New time-changes of unipotent flows on quotients of Lorentz groups

Siyuan Tang

ABSTRACT. We study the cocompact lattices $\Gamma \subset SO(n,1)$ so that the Laplace-Beltrami operator Δ on $SO(n)\backslash SO(n,1)/\Gamma$ has eigenvalues in $(0, \frac{1}{4})$, and then show that there exist time-changes of unipotent flows on $SO(n,1)/\Gamma$ that are not measurably conjugate to the unperturbed ones.

A main ingredient of the proof is a stronger version of the branching of the complementary series. Combining it with a refinement of the works of Ratner and Flaminio-Forni is adequate for our purpose.

CONTENTS

1. Introduction

1.1. **Main results.** Let G be a semisimple Lie group, Γ be a lattice of G. Let $X = G/\Gamma$ be the homogeneous space equipped with the Haar measure μ. Then classical unipotent flows u^t on X have been studied by an extensive literature. Besides, one can build new parabolic flows in terms of u^t via perturbations. Perhaps the simplest perturbations are *time-changes*, i.e. flows that move points along the same orbits, but with different speeds. Time-changes preserve certain ergodic and spectral properties. For instance, [FU12] and [DA12] showed that the sufficiently regular time-changes of horocycle flows have the *Lebesgue spectrum*. Later [Sim18] extended this result to the case of semisimple unipotent flows. In particular, we know that all time-changes of unipotent flows satisfying a mild differentiability condition are *strongly mixing* (the horocycle case was first discovered by [Mar77]).

One may then ask whether time-changes produce genuinely new flows, i.e. a time-change of the unipotent flow is actually not *measurably conjugated* to the unperturbed one. However, the question is in general difficult and the answer is known only in a few cases. For horocycle flows, [Rat86] and [FF03] showed that sufficiently regular time-changes which are measurably conjugate to the (unperturbed) horocycle flow are **rare** (in fact, they form a countable codimension subspace). However, no similar results are known for other unipotent flows. On the other hand, it is worth mentioning that [Rav19] provides **new** examples of parabolic perturbations on $SL(3, \mathbf{R})/\Gamma$ for which one can study ergodic theoretical properties. They are not (but related to) time-changes. Whether the perturbations produce genuinely new flows in that setting remains open.

In this paper, we manage to generalize [Rat86] and [FF03] to $G = SO(n, 1)$ setting. More precisely, let $\mathfrak{g} = \mathfrak{so}(n, 1)$ be the corresponding Lie algebra, $U \in \mathfrak{g}$ be a nilpotent element. Then it induces a unipotent flow $\phi_t^U(x) = \exp(tU)x = u^t x$ on X. Let τ be a positive integrable function on X with $\int_X \tau(x)d\mu(x) = 1$. Then define a cocycle $\xi : X \times \mathbf{R} \to \mathbf{R}$ by

$$\xi(x, t) := \int_0^t \tau(\phi_s^U(x))ds = \int_0^t \tau(u^s x)ds.$$

Then the flow $\phi_t^{U,\tau} : X \to X$ obtained from the unipotent flow u^t by the time-change τ is given by the relation

$$\phi_{\xi(x,t)}^{U,\tau}(x) := u^t x.$$

Besides, we require that the time-changes have the effective mixing property. More precisely, let $\mathbf{K}(X)$ be the set of all positive integrable functions α on X such that α, α^{-1} are bounded and satisfies

$$\left| \int_X \alpha(x)\alpha(u^t x)d\mu(x) - \left(\int_X \alpha(x)\mu(x) \right)^2 \right| \leq D_\alpha |t|^{-\sigma_\alpha}$$

for some $D_\alpha, \sigma_\alpha > 0$. In other words, elements $\alpha \in \mathbf{K}(X)$ have polynomial decay of correlations. Note that [KM99] has shown that sufficiently regular functions on X are in $\mathbf{K}(X)$.

First of all, the construction of time-changes naturally connects it to *cohomological properties*. We say that two functions g_1, g_2 on X are *measurable (respectively L^2, smooth, etc.) cohomologous over the flow u^t* if there exists a measurable (respectively L^2, smooth, etc.) function f on X, called the *transfer function*, such that

$$\int_0^T g_1(u^t x) - g_2(u^t x) dt = f(u^T x) - f(x). \tag{1.1}$$

An elementary argument establishes that the flows generated by cohomologous time-changes are always measurably conjugated. On the contrary, we deduce the following generalization of [Rat86]:

Theorem 1.1. *Let $\tau \in \mathbf{K}(X)$. Suppose that there is a measurable conjugacy map $\psi : (X, \mu) \to (X, \mu_\tau)$ such that*

$$\psi(\phi_t^U(x)) = \phi_t^{U,\tau}(\psi(x))$$

for $t \in \mathbf{R}$ and μ-a.e. $x \in X$, where $d\mu_\tau = \tau d\mu$. Then $\tau(x)$ and $\tau(cx)$ are measurably cohomologous for all $c \in C_G(U)$. Besides, if $\tau(x)$ and $\tau(cx)$ are indeed L^1-cohomologous for all $c \in C_G(U)$, then 1 and τ are measurably cohomologous. Here $C_G(U) := \{g \in G : \operatorname{Ad} g.U = U\}$ denotes the centralizer of U in G.

Remark 1.2. It is possible to extend the result to two time-changes $\tau_1 \in \mathbf{K}(G/\Gamma_1)$, $\tau_2 \in \mathbf{K}(G/\Gamma_2)$, similar to [Rat86]. More precisely, we can assume that there is a measurable conjugacy map $\psi : (G/\Gamma_1, \mu_{\tau_1}) \to (G/\Gamma_2, \mu_{\tau_2})$ such that

$$\psi(\phi_t^{U,\tau_1}(x)) = \phi_t^{U,\tau_2}(\psi(x)).$$

Then we shall again obtain cohomologous results for τ_1, τ_2 and, with further regularity assumptions for τ_1, τ_2, we have Γ_1 and Γ_2 are conjugate. The proofs will be a bit more complicated and we do not need it here.

Thus, in order to find a time-changed flow that is not measurably conjugated to the unperturbed one, we should study the cohomological equation (1.1). Or equivalently, the differential equation

$$g(x) = Uf(x) \tag{1.2}$$

once f is differentiable along U-direction. [FF03] studied the equation (1.2) on the irreducible unitary representations of $SO(2,1)$ by classifying the U-invariant distributions. Flaminio-Forni realized that the invariant distributions are the **only** obstructions to the existence of smooth solutions of equation (1.2). Besides, acting by the geodesic flow on the U-invariant distributions, Flaminio-Forni established precise asymptotics for the ergodic averages along the orbits of the horocycle flow on $SO(2,1)/\Gamma$ when Γ is a cocompact lattice. However, it seems difficult to generalize these ideas to $SO(n,1)$ for $n \geq 3$. One reason is that the equation (1.2) in the $SO(2,1)$-representations is an *ordinary difference equation* (OdE), but in the $SO(n,1)$-representations becomes a *partial difference equation* (PdE) when $n \geq 3$. However, if we pay our attention to certain complementary series, then one may possibly restrict the $SO(n,1)$-representations to the subgroup $SO(n-1,1)$, and hence [FF03] may apply.

Let $G = SO(n,1)$, $H = SO(n-1,1)$, $\Gamma \subset G$ be a cocompact lattice such that the Laplace-Beltrami operator Δ has eigenvalues in $(0, \frac{1}{4})$. See [Ran74], [SWY80], [Bro88] for the existence of these lattices (see Section 4.2). Then [Zha15] (see also [Muk68], [SV12], [SZ16]) has shown that $L^2(G/\Gamma)$ contains a G-complementary series π_ν so that it further contains a H-complementary series $\pi^\flat_{\nu-\frac{1}{2}}$ as a direct summand. (See Section 4 for further explanation for the notation.) Here we prove that the corresponding Sobolev spaces also have this property (see Theorem 4.8):

Theorem 1.3. *Let $n \geq 3$, $\rho^\flat < \nu < \rho$, $s \geq 0$, $G = SO(n,1)$ and $H = SO(n-1,1)$. Then $(\pi^\flat_{\nu-\frac{1}{2}}, W^s_H(\mathcal{H}^\flat_{\nu-\frac{1}{2}}))$ is a direct summand of $(\pi_\nu, W^s_G(\mathcal{H}_\nu))$ restricted to H, where $\rho = \frac{n-1}{2}$ and $\rho^\flat = \frac{n-2}{2}$ denote the half sum of the positive roots in $\mathfrak{g}$ and $\mathfrak{h}$, respectively.*

Thus, by repeatedly using Theorem 1.3, we are able to study a complementary series of $SO(2,1)$ as a direct summand of $L^2(G/\Gamma)$. Then by applying a similar argument of [FF03], we can study the ergodic average $\frac{1}{T}\int_0^T g(\phi^U_t(x))dt$ in terms of U-invariant distributions. Then we get

Theorem 1.4. *Let the notation and assumptions be as above. Then there is a sufficiently regular function g on $X = G/\Gamma$ of integral zero $\mu(g) = 0$ that is not measurably cohomologous to 0, i.e. there are no measurable functions f satisfying*

$$\int_0^T g(\phi^U_t(x))dt = f(\phi^U_T(x)) - f(x).$$

Moreover, if there are some $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$ such that $\phi^Z_\lambda g$ is not L^2-cohomologous to g, then $\phi^Z_\lambda g$ is not measurably cohomologous to g.

Theorem 1.1 and 1.4 yield

Corollary 1.5. *There is a cocompact lattice $\Gamma \subset G = SO(n,1)$, and a time-change of a unipotent flow on $X = G/\Gamma$ that is not measurably conjugate to the unperturbed unipotent flows.*

Proof. Let g be given by Theorem 1.4. Via *Sobolev embedding* (Lemma 5.2), it is possible to choose g to be continuous. After multiplying a constant if necessary, we can take $\tau = 1 + g$ to be positive and integrable. Now assume that there is a measurable conjugacy map $\psi : (X, \mu) \to (X, \mu_\tau)$ such that

$$\psi(\phi^U_t(x)) = \phi^{U,\tau}_t(\psi(x))$$

for $t \in \mathbf{R}$ and μ-a.e. $x \in X$. Then by Theorem 1.1, $\tau(x)$ and $\tau(cx)$ are measurably cohomologous for all $c \in C_G(U)$. If there exists $c \in C_G(U)$ such that $\tau(x)$ and

$\tau(cx)$ are not L^1-cohomologous, then $g(x) = \tau(x) - 1$ and $g(cx) = \tau(cx) - 1$ are not L^1- (and hence are not L^2-) cohomologous. Thus, by Theorem 1.4, $\tau(x)$ and $\tau(cx)$ are not measurably cohomologous either, which leads to a contradiction. Thus, we conclude that $\tau(x)$ and $\tau(cx)$ are indeed L^1-cohomologous for all $c \in C_G(U)$. Via Theorem 1.1 again, we see that $\tau = 1 + g$ and 1 are measurably cohomologous, but it again violates Theorem 1.4. $\square$

Thus, we conclude that sufficiently regular time-changes on X which are measurably conjugate to the unperturbed unipotent flow are rare, in the sense that the complement of the set of these time-changes has at least finite codimension.

Besides, Theorem 1.4 also implies that the *central limit theorem* does not hold for unipotent flows on $X = G/\Gamma$ (see Corollary 5.20):

Corollary 1.6. *There is a cocompact lattice $\Gamma \subset G = SO(n, 1)$, and a function g on $X = G/\Gamma$ such that, as $T \to \infty$, any weak limit of the probability distributions*

$$\frac{\frac{1}{T}\int_0^T g(\phi_t^U(x))dt}{\left\|\frac{1}{T}\int_0^T g(\phi_t^U(\cdot))dt\right\|_{L^2}}$$

has a nonzero compact support.

1.2. **Structure of the paper.** In Section 2 we recall basic definitions, including some basic material on the Lie algebra $\mathfrak{so}(n, 1)$ (in Section 2.1), as well as time-changes (Section 2.2). In Section 3, we deduce Theorem 1.1. This requires studying the shearing property of u_X^t for nearby points. More precisely, we provide a quantitative estimate (Proposition 3.3) for the difference of nearby points in terms of the length of unipotent orbits. Then, the estimate can deduce extra equivariant properties (Lemma 3.21 and 3.24). Then combining Ratner's theorem, we obtain Theorem 3.1 which states that the measurable conjugacies are almost algebraic. In particular, we obtain the cohomologous relations. In Section 4, we state a number of results of the representation theory, which will be used as tools to study the cohomological equations. In particular, we prove a Sobolev version of the branching of complementary series (Theorem 1.3 or Theorem 4.8). Finally, in Section 5, we apply Flaminio-Forni argument (Theorem 5.3) to find a required time-change function τ (Theorem 1.4) in Corollary 5.21. Then combining Theorem 1.1, we conclude that τ is a nontrivial time-change (Corollary 1.5). Besides, we present the central limit theorem of unipotent flows does not hold (Corollary 1.6) in Corollary 5.20.

Acknowledgements. The paper was written under the guidance of Prof. David Fisher for my PhD thesis, and I am sincerely grateful for his help. I would also like to thank Prof. Livio Flaminio, Prof. Giovanni Forni and Prof. Adam Kanigowski for useful conversations. Besides, I would like to express my deep appreciation to the referee, who examined the paper with great care, pointing out many inaccuracies, and suggesting improvements in various aspects of writing.

2. Preliminaries

2.1. **Definitions.** Let $G := SO(n,1)$; more precisely we define

$$G := \left\{ g \in SL_{n+1}(\mathbf{R}) : \begin{bmatrix} I_n & \\ & -1 \end{bmatrix} g^T \begin{bmatrix} I_n & \\ & -1 \end{bmatrix} = g^{-1} \right\}$$

where I_n is the $n \times n$ identity matrix. The corresponding Lie algebra is given by

$$\begin{aligned} \mathfrak{g} &= \left\{ v \in \mathfrak{sl}_{n+1}(\mathbf{R}) : \begin{bmatrix} I_n & \\ & -1 \end{bmatrix} v^T \begin{bmatrix} I_n & \\ & -1 \end{bmatrix} = -v \right\} \\ &= \left\{ \begin{bmatrix} \mathbf{l} & \\ & 0 \end{bmatrix} : \mathbf{l} \in \mathfrak{so}(n) \right\} \oplus \left\{ \begin{bmatrix} 0 & \mathbf{p} \\ \mathbf{p}^T & 0 \end{bmatrix} : \mathbf{p} \in \mathbf{R}^n \right\}. \end{aligned}$$

Let E_{ij} be the $(n \times n)$-matrix with 1 in the (i,j)-entry and 0 otherwise. Let $e_k \in \mathbf{R}^n$ be the k-th standard basis (vertical) vector. Set

$$Y_k := \begin{bmatrix} 0 & e_k \\ e_k^T & 0 \end{bmatrix}, \quad \Theta_{ij} := \begin{bmatrix} E_{ji} - E_{ij} & 0 \\ 0 & 0 \end{bmatrix}.$$

Then Y_i, Θ_{ij} form a basis of $\mathfrak{g} = \mathfrak{so}(n,1)$. Let $\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{p}$ be the corresponding Cartan decomposition. Let $\mathfrak{a} = \mathbf{R}Y_n \subset \mathfrak{p}$ be a maximal abelian subspace of $\mathfrak{p}$. Then the root space decomposition of $\mathfrak{g}$ is given by

$$\mathfrak{g} = \mathfrak{g}_{-1} \oplus \mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{g}_1. \tag{2.1}$$

Denote by $\mathfrak{n} := \mathfrak{g}_1$ the sum of the positive root spaces. Let ρ be the half sum of positive roots. Sometimes, we adopt the convention by identifying $\mathfrak{a}^*$ with $\mathbf{C}$ via $\lambda \mapsto \lambda(Y_n)$. Thus, $\rho = \rho(Y_n) = (n-1)/2$. We write

$$a^t := \exp(tY_n)$$

for the geodesic flow.

Let $\Gamma \subset G$ be a lattice, $X := G/\Gamma$, μ be the Haar probability measure on X. Fix a nilpotent $U \in \mathfrak{g}_{-1}^\flat$. Then U defines a unipotent flow

$$\phi_t^U(x) = \exp(tU)x = u^t x$$

on G/Γ and satisfies

$$[Y_n, U] = -U.$$

Then using the *Killing form*, there exists $\tilde{U} \in \mathfrak{g}$ such that $\{U, Y_n, \tilde{U}\}$ spans a $\mathfrak{sl}_2$-triple. Denote

$$\tilde{u}^t := \exp(t\tilde{U}).$$

For convenience, we choose

$$U := \begin{bmatrix} 0 & e_{n-1} & e_{n-1} \\ -e_{n-1}^T & 0 & 0 \\ e_{n-1}^T & 0 & 0 \end{bmatrix}, \quad \tilde{U} := \begin{bmatrix} 0 & -e_{n-1} & e_{n-1} \\ e_{n-1}^T & 0 & 0 \\ e_{n-1}^T & 0 & 0 \end{bmatrix}. \tag{2.2}$$

Then $\langle u^t, a^t, \tilde{u}^t\rangle$ generates $SO(2,1) \subset SO(n,1)$. Further, if we consider $\mathfrak{g}$ as a $\mathfrak{sl}_2(\mathbf{R})$-representation via the adjoint map, then by the complete reducibility of $\mathfrak{sl}_2(\mathbf{R})$, there is a orthogonal decomposition

$$\mathfrak{g} = \mathfrak{sl}_2(\mathbf{R}) \oplus V^\perp$$

where $V^\perp$ consists of irreducible representations with highest weights 2 and 0. For elements $g \in \exp \mathfrak{g}$, we decompose

$$g = h\exp(v), \quad h \in SO(2,1), \quad v \in V^\perp.$$

Moreover, it is convenient to think about $h \in SO(2,1)$ as a 2×2 matrix with determinant 1. Thus, consider the isogeny $\iota : SL_2(\mathbf{R}) \to SO(2,1) \subset G$ induced by $\mathfrak{sl}_2(\mathbf{R}) \to \mathrm{Span}\{U, Y_n, \tilde{U}\} \subset \mathfrak{g}$. This is a two-to-one immersion. In the following, for $h \in SO(2,1)$ and v in an irreducible representation, we often write

$$h = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad v = b_0 v_0 + \cdots + b_\varsigma v_\varsigma$$

where v_i are weight vectors in $\mathfrak{g}$ of weight i. Notice that h should more appropriately be written as $\iota(h)$.

For the centralizer $C_{\mathfrak{g}}(U)$, we have the corresponding decomposition:

$$C_{\mathfrak{g}}(U) = \mathbf{R}U \oplus V_C^\perp \tag{2.3}$$

where $V_C^\perp$ consists of highest weight vectors other than U (see also Lemma 3.9). More precisely, under the setting (2.2), one may calculate

$$C_{\mathfrak{g}}(U) = \left\{ \begin{bmatrix} \mathbf{c} & \\ & 0 \end{bmatrix} : \mathbf{c} \in \mathfrak{so}(n-2) \right\} \oplus \left\{ \begin{bmatrix} 0 & \mathbf{u} & \mathbf{u} \\ -\mathbf{u}^T & 0 & 0 \\ \mathbf{u}^T & 0 & 0 \end{bmatrix} : \mathbf{u} \in \mathbf{R}^{n-1} \right\}. \tag{2.4}$$

Note that the first summand consists of semisimple elements, and the second summand consists of nilpotent elements.

2.2. **Time-changes.** Let $\phi_t^{U,\tau}$ be a *time change* for the unipotent flow ϕ_t^U, $t \in \mathbf{R}$. More precisely, we assume

- $\tau : X \to \mathbf{R}^+$ is a integrable nonnegative function on X satisfying
$$\int_X \tau(x) d\mu(x) = 1,$$
- $\xi : X \times \mathbf{R} \to \mathbf{R}$ is the cocycle defined by
$$\xi(x,t) := \int_0^t \tau(\phi_s^U(x)) ds = \int_0^t \tau(u^s x) ds, \tag{2.5}$$
- $\phi_t^{U,\tau} : X \to X$ is given by the relation
$$\phi_{\xi(x,t)}^{U,\tau}(x) := u^t x.$$

Remark 2.1. Note that $\phi_t^{U,1} = \phi_t^U$. Besides, one can check that $\phi_t^{U,\tau}$ preserves the probability measure on X defined by $d\mu_\tau := \tau d\mu$. On the other hand, if τ is smooth, then the time-change $\phi_t^{U,\tau}$ is the flow on X generated by the smooth vector field $U_\tau := U/\tau$ (see [FU12]).

The unipotent flows ϕ_t^U, as well as their time-changes $\phi_t^{U,\tau}$ are *parabolic flows*, in the sense that nearby orbits diverge polynomially in time. We shall quantitatively study it via the effective ergodicity of the unipotent flows.

Besides, the construction of time-changes naturally connects to cohomological properties. We say that two functions g_1, g_2 on X are *measurable (respectively L^2, smooth, etc.) cohomologous over the flow u^t* if there exists a measurable (respectively L^2, smooth, etc.) function f on X, called the *transfer function*, such that

$$\int_0^T g_1(u^t x) - g_2(u^t x) dt = f(u^T x) - f(x) \tag{2.6}$$

for μ-a.e. $x \in X$. We also say that g is *measurably (respectively L^2, smooth, etc.) trivial* if g and 0 are cohomologous. For the related discussion, see [AFRU19] and references therein. Then conjugacies naturally arise from cohomologous time-changes. More precisely, one may verify that two time-changes τ_1, τ_2 are cohomologous via a transfer function f iff the map $\psi_f : X \to X$ defined by

$$\psi_f : x \mapsto \phi^U_{z(x)}(x)$$

where $z : X \times \mathbf{R} \to \mathbf{R}$ is defined by the relation

$$f(x) = \xi_2(x, z_f(x)) = \int_0^{z_f(x)} \tau_2(\phi_s^U(x)) ds,$$

is an invertible conjugacy between ϕ_t^{U,τ_1} and ϕ_t^{U,τ_2}, i.e.

$$\psi_f(\phi_t^{U,\tau_1}(x)) = \phi_t^{U,\tau_2}(\psi_f(x)).$$

On the other hand, if f is differentiable along U-direction, then differentiate (2.6) along U and we get the *cohomological equation*

$$g_1(x) - g_2(x) = Uf(x).$$

We shall discuss it further in Section 5.

In [Rat86], Ratner considered a particular class $\mathbf{K}(X)$ of time changes. More precisely, $\mathbf{K}(X)$ consists of all positive integrable functions α on X such that α, α^{-1} are bounded and satisfies

$$\left| \int_X \alpha(x)\alpha(u^t x) d\mu(x) - \left(\int_X \alpha(x)\mu(x) \right)^2 \right| \leq D_\alpha |t|^{-\sigma_\alpha}$$

for some $D_\alpha, \sigma_\alpha > 0$. This is the effective mixing property of the unipotent flow ϕ_t^U. Note that [KM99] (see also [Ven10]) have shown that there is $\kappa > 0$ such that

$$\left| \langle \phi_t^U(f), g \rangle - \left(\int_X f(x)\mu(x) \right) \left(\int_X g(x)\mu(x) \right) \right| \ll (1+|t|)^{-\kappa} \|f\|_{W^s} \|g\|_{W^s}$$

for $f, g \in C^\infty(X)$, where $s \geq \dim(K)$ and W^s denotes the Sobolev space on $X = G/\Gamma$ that will be defined later (Section 4.3).

3. Measurable conjugacies and transfer functions

In this section, we shall use the shearing properties of unipotent flows and show that any measurable conjugacy between unipotent flows and their time-changes is almost algebraic. More precisely, we deduce

Theorem 3.1. *Let the notation and assumptions be as above. Let $\phi_t^{U,\tau}$ be a time change for the unipotent flow u^t with $\tau \in \mathbf{K}(X)$. Suppose that there is a measurable conjugacy map $\psi : (X, \mu) \to (X, \mu_\tau)$ such that*

$$\psi(\phi_t^U(x)) = \phi_t^{U,\tau}(\psi(x))$$

for $t \in \mathbf{R}$ and μ-a.e. $x \in X$. Then there exists a measurable map $\varpi : X \times C_G(U) \to C_G(U)$ such that

$$\psi(cx) = \varpi(x, c)\psi(x) \tag{3.1}$$

for $c \in C_G(U)$, μ-almost all $x \in X$. Besides, $\varpi(x, c) = u^{\alpha(x,c)}\beta(c)$ where $\alpha(x, c) \in \mathbf{R}$ and $\beta(c) \in \exp V_C^\perp$. Moreover, if $\alpha(\cdot, c) \in L^1(X)$ for all $c \in C_G(U)$, then there are points $x_0, y_0 \in X$, an automorphism Φ of G that fixes $SO(2,1)$ (i.e. $\Phi(g) = g$ for $g \in SO(2,1)$) and a map $c : X \to C_G(U)$ such that

$$\psi(gx_0) = c(gx_0)\Phi(g)y_0 \tag{3.2}$$

for any $g \in G$. Similarly, $c(x) = u^{a(x)}b$ where $a(x) \in \mathbf{R}$ and $b \in \exp V_C^\perp$.

The proof of Theorem 3.1 use the strategy similar to Ratner's theorem. We consider the unipotent orbits of nearby points and look at their images under the given measurable conjugacy ψ. By Lusin's theorem, the images inherit a similar behavior as in the range. Using this phenomenon, we can study the difference of the nearby points under ψ and so obtain the extra equivariant properties of ψ. The difficulty is to pull the information of unipotent orbits from the homogeneous spaces back to Lie groups. It requires us to observe the long unipotent orbits with "gaps" and make use of their polynomial growth nature (see Proposition 3.3).

Note that combining (3.1) (respectively (3.2)) with Corollary 3.23 (respectively Corollary 3.27), we obtain a criterion for the solutions of the cohomological equation (Theorem 1.1):

Corollary 3.2. *Let the notation and assumptions be as above. Then $\tau(x)$ and $\tau(cx)$ are measurably cohomologous for all $c \in C_G(U)$. Besides, if $\tau(x)$ and $\tau(cx)$ are indeed L^1-cohomologous for all $c \in C_G(U)$, then 1 and τ are measurably cohomologous.*

3.1. **Shearing properties.** We shall study the *shearing property* of the unipotent flow φ_t^U. Roughly speaking, it states that if two points start out so close together that we cannot tell them apart, then the first difference we see, which is often called the *fastest relative motion*, will fall in the centralizer $C_G(U)$. It is only much later that we will detect any other difference between their paths.

In the following, we shall prove Proposition 3.3, which provide a quantitative estimate of the difference of two nearby points via the shearing property stated above. The philosophy of Proposition 3.3 is:

> For certain $\lambda > 0$, $t : [0,\infty) \to [0,\infty)$, if 99% of $s \in [0,\lambda]$ $d_X(u^{t(s)}y, u^s x) < \epsilon$, then the only possible situation is that there is a big **interval** $I \subset [0,\lambda]$ (say of 90% length) so that $d_X(u^{t(s)}y, u^s x) < \epsilon$ for all $s \in I$.

Roughly speaking, it connects to the fact that polynomials do not have extreme oscillations. The $SL(2,\mathbf{R})$ version of this property has already been established by Ratner [Rat86]. The method is also inspired by the proof of *Ratner's theorem.* See [Ein06], [EMV09] and references therein.

Proposition 3.3 (Shearing)**.** *Let the notation and assumptions be as above. Given $\eta \in (0,1)$ and $m > 1$, there are*

- $\rho = \rho(\eta) > 0$,
- $\theta = \theta(\rho) > 0$,

such that for any sufficiently small $\sigma \in (0,\sigma_\rho)$, there are

- *a compact $K = K(\rho,\sigma) \subset X$ with $\mu(K) > 1-\sigma$,*
- *$\epsilon = \epsilon(K,m) \in (0,1)$ close to 0*

satisfying the following property: Let $x \in K$, $y \in B_X(x,\epsilon)$, and a subset $A \subset \mathbf{R}^+$ satisfy the following conditions

(i) if $s \in A$, then

$$u^s x \in K \quad \textit{and} \quad d_X(u^{t(s)}y, u^s x) < \epsilon$$

for some increasing function $t : [0,\infty) \to [0,\infty)$,

(ii) we have the Hölder inequality:

$$|(t(s') - t(s)) - (s' - s)| \leq |s' - s|^{1-\eta} \tag{3.3}$$

for all $s, s' \in A$ with $s' > s$, $\max\{(s'-s),(t(s')-t(s))\} \geq m$.

Then for any $\lambda \in A$ satisfying $\mathrm{Leb}(A \cap [0,\lambda]) > (1-\theta)\lambda$, there is $s_\lambda \in A \cap [0,\lambda]$ such that

$$u^{t(s_\lambda)}y = h_\lambda \exp(v_\lambda) u^{s_\lambda} x \tag{3.4}$$

where $h_\lambda \in SO(2,1)$ and $v_\lambda \in V^\perp$ satisfy

$$h_\lambda = \begin{bmatrix} 1 + O(\lambda^{-2\rho}) & O(\lambda^{-1-2\rho}) \\ O(\epsilon) & 1 + O(\lambda^{-2\rho}) \end{bmatrix}$$

$$v_\lambda = O(\lambda^{-\frac{1+2\rho}{2}\varsigma})v_0 + O(\lambda^{-\frac{1+2\rho}{2}(\varsigma-1)})v_1 + \cdots + O(\epsilon)v_\varsigma.$$

3.2. **Quantitative estimates.** In order to prove Proposition 3.3, we shall develop a collection α of finitely many subintervals of $[0,\lambda]$ through the assumptions. Then we shall show that there is a big interval from the collection α. In the following, we first verify a combinatorial result that helps us to find the big interval in α.

Let I be an interval in $\mathbf{R}$ and let J_i, J_j be disjoint subintervals of I, $J_i = [x_i, y_i]$, $y_i < x_j$ if $i < j$. Denote

$$d(J_i, J_j) \coloneqq \mathrm{Leb}[y_i, x_j] = x_j - y_i.$$

For a collection β of finitely many intervals, we define

$$|\beta| \coloneqq \mathrm{Leb}\left(\bigcup_{J\in\beta} J\right).$$

Besides, for a collection β of finitely many intervals, an interval I, let

$$\beta \cap I \coloneqq \{I \cap J : J \in \beta\}.$$

Proposition 3.4 (Existence of large intervals, Solovay, [Rat79])**.** *Given $\rho \in (0,1)$, there is $\theta = \theta(\rho) \in (0,1)$ such that if I is an interval of length $\lambda > 1$ and $\mathcal{G} \cup \mathcal{B} = \{J_1, \dots, J_n\}$ is a partition of I into good and bad intervals such that*

(1) for any two good intervals $J_i, J_j \in \mathcal{G}$, we have

$$d(J_i, J_j) \geq [\min\{\mathrm{Leb}(J_i), \mathrm{Leb}(J_j)\}]^{1+\rho}, \tag{3.5}$$

(2) $\mathrm{Leb}(J) \leq \frac{3}{4}\lambda$ for any good interval $J \in \mathcal{G}$,
(3) $\mathrm{Leb}(J) \geq 1$ for any bad interval $J \in \mathcal{B}$,

then the measure of bad intervals $\mathrm{Leb}(\bigcup_{J\in\mathcal{B}} J) \geq \theta\lambda$.

Remark 3.5. The idea of Proposition 3.4 is to consider the arrangement of intervals in α. It turns out that under the assumptions, the worst arrangement would be like the complement of a Cantor set. A careful calculation of the quantities under this situation leads to the result.

Proof. Assume that $\left(\frac{4}{3}\right)^{k-1} \leq \lambda \leq \left(\frac{4}{3}\right)^{k}$ for some $k \geq 1$. Let $\mathcal{G}_n \coloneqq \{J \in \mathcal{G} : \left(\frac{3}{4}\right)^{n+1}\lambda \leq |J| \leq \left(\frac{3}{4}\right)^{n}\lambda\}$, $\mathcal{G}_{\leq n} \coloneqq \bigcup_{i=1}^{n} \mathcal{G}_i$, and $\mathcal{B}_{\leq n}$ be the collection of remaining intervals forming $I \setminus \bigcup_{J\in\mathcal{G}_{\leq n}} J$. Then given $n \in \mathbf{N}$, $J \in \mathcal{B}_{\leq n}$, by (3.5), we have

$$\begin{aligned}\frac{|\mathcal{B}_{\leq n+1} \cap J|}{\mathrm{Leb}(J)} &= \frac{|\mathcal{B}_{\leq n+1} \cap J|}{|\mathcal{G}_{n+1} \cap J| + |\mathcal{B}_{\leq n+1} \cap J|} = \left(1 + \frac{|\mathcal{G}_{n+1} \cap J|}{|\mathcal{B}_{\leq n+1} \cap J|}\right)^{-1} \\ &\geq \left(1 + \frac{l\left(\frac{3}{4}\right)^{n+1}\lambda}{(l-1)\left(\frac{3}{4}\right)^{(n+2)(1+\rho)}\lambda^{1+\rho}}\right)^{-1} = \left(1 + C\left(\frac{3}{4}\right)^{(k-n)\rho}\right)^{-1}\end{aligned}$$

where $l \geq 2$ is the number of intervals in $\mathcal{G}_{n+1} \cap J$, and $C > 0$ is some constant depending on ρ. One can also show that when $k = 0, 1$, we have a similar relation.

By summing over $J \in \mathcal{B}_{\leq n}$, we obtain

$$\frac{|\mathcal{B}_{\leq n+1}|}{|\mathcal{B}_{\leq n}|} \geq \left(1 + C\left(\frac{3}{4}\right)^{(k-n)\rho}\right)^{-1}.$$

Note that by (2), $|\mathcal{B}_{\leq 0}| = \lambda$, and by (3), $\mathcal{B}_{\leq n} = \mathcal{B}_{\leq n+1}$ for all $n \geq k$. We calculate

$$|\mathcal{B}| = |\bigcap_{k\geq 0} \mathcal{B}_{\leq k}| = \lim_{k\to\infty} |\mathcal{B}_{\leq k}| = \prod_{n=0}^{\infty} \frac{|\mathcal{B}_{\leq n+1}|}{|\mathcal{B}_{\leq n}|} \cdot \lambda \geq \prod_{n=0}^{k} \left(1 + C\left(\frac{3}{4}\right)^{(k-n)\rho}\right)^{-1} \cdot \lambda.$$

Take

$$\theta := \prod_{m=0}^{\infty} \left(1 + C\left(\frac{3}{4}\right)^{m\rho}\right)^{-1} \leq \prod_{n=0}^{k} \left(1 + C\left(\frac{3}{4}\right)^{(k-n)\rho}\right)^{-1}$$

and the proposition follows. □

In light of (3.5), we say that two intervals $I, J \subset \mathbf{R}$ have an *effective gap* if

$$d(I, J) \geq [\min\{\mathrm{Leb}(I), \mathrm{Leb}(J)\}]^{1+\rho}$$

for some $\rho > 0$. Later, we shall obtain some quantitative results relative to the effective gap.

Lemma 3.6. *For sufficiently small vector $v \in \mathfrak{g}$, we have*

$$\log \exp(gvg^{-1}) = gvg^{-1}$$

for all $g \in G$, where log *denotes the principal logarithm.*

Proof. According to [Hig08], for any square complex matrix v, $\log \exp(v) = v$ iff $|\operatorname{Im} \lambda_i| < \pi$ for every eigenvalue λ_i of v. Then the consequence follows from the fact that $\operatorname{Ad} g$ does not change the eigenvalues of v. □

Lemma 3.7. *Let $\mathfrak{g}$ be a Lie algebra, and $\mathfrak{g} = V_1 \oplus V_2$ be a decomposition of vector spaces. Then the multiplication map $\zeta : V_1 \oplus V_2 \to G$ defined by*

$$(\alpha, \beta) \mapsto \exp(\alpha)\exp(\beta)$$

induces a diffeomorphism on small neighborhoods $U_1 \subset V_1$ and $U_2 \subset V_2$ of 0.

Proof. Note that $d\zeta_{(0,0)} : (\alpha, \beta) \mapsto \alpha + \beta$. Then the consequence follows from the inverse function theorem. □

In the following, $A \ll B$ means there is a constant $C > 0$ such that $A \leq CB$. Besides, we write $A \ll_\kappa B$ if the constant $C(\kappa)$ depends on some coefficient κ.

Lemma 3.8. *Fix numbers $\epsilon > 0$, $\eta \in (0, 1]$, a real polynomial $p(x) = v_0 + v_1 x + \cdots + v_k x^k \in \mathbf{R}[x]$. Assume further that there exist intervals $[0, \bar{l}_1] \cup [l_2, \bar{l}_2] \cup \cdots \cup [l_m, \bar{l}_m]$ such that*

$$|p(t)| \ll \max\{\epsilon, t^{1-\eta}\} \quad \textit{iff} \quad t \in [0, \bar{l}_1] \cup [l_2, \bar{l}_2] \cup \cdots \cup [l_m, \bar{l}_m] \tag{3.6}$$

Then $\bar{l}_1$ has the lower bound l depending on $\max_i |v_i|$, ϵ, η and the implicit constant such that $l \nearrow \infty$ as $\max_i |v_i| \searrow 0$ for fixed ϵ, η. Besides, $m \leq k$ and we have

(1) $|v_i| \ll_{k,\eta} \bar{l}_1^{1-i-\eta}$ *for all* $1 \leq i \leq k$;

(2) Fix $\rho \in (0,1)$. *For* $1 \leq j \leq k-1$, *sufficiently large* $\bar{l}_j$, *assume that the intervals* $[0, \bar{l}_j]$ *and* $[l_{j+1}, \bar{l}_{j+1}]$ *do not have an effective gap:*

$$l_{j+1} - \bar{l}_j \leq \min\{\bar{l}_j, \bar{l}_{j+1} - l_{j+1}\}^{1+\rho}. \tag{3.7}$$

Then there exists $\xi(\rho,k) \in (0,1)$ *with* $\xi(\rho,k) \to 1$ *as* $\rho \to 0$ *such that*

$$|v_i| \ll_{k,\eta} \bar{l}_j^{\xi(\rho,k)(1-i-\eta)}$$

for all $1 \leq i \leq k$.

Proof. The number m of intervals in (3.6) can be bounded by k via an elementary study of polynomials.

(1) Let $F(x) := v_1(\bar{l}_1 x)^\eta + \cdots + v_k(\bar{l}_1 x)^{k-1+\eta}$ for $x \in [0,1]$. Then we have

$$\begin{pmatrix} v_1 \bar{l}_1^{\eta} \\ v_2 \bar{l}_1^{1+\eta} \\ \vdots \\ v_k \bar{l}_1^{k-1+\eta} \end{pmatrix} = \begin{bmatrix} (1/k)^\eta & (1/k)^{1+\eta} & \cdots & (1/k)^{k-1+\eta} \\ (2/k)^\eta & (2/k)^{1+\eta} & \cdots & (2/k)^{k-1+\eta} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}^{-1} \begin{pmatrix} F(1/k) \\ F(2/k) \\ \vdots \\ F(1) \end{pmatrix}.$$

By (3.6), we know that $|F(1/k)|, |F(2/k)|, \cdots, |F(1)| \ll 1$. Thus, we obtain $|v_i| \ll_{k,\eta} \bar{l}_1^{1-i-\eta}$ for all $1 \leq i \leq k$.

(2) This follows easily by induction. Assume that the statement holds for $j-1$. For j, the only difficult situation is when $\bar{l}_j \leq l_{j+1} - \bar{l}_j$ and $\bar{l}_{j+1} - l_{j+1} \leq l_{j+1} - \bar{l}_j$. If this is the case, then

$$\bar{l}_{j+1} = (\bar{l}_{j+1} - l_{j+1}) + (l_{j+1} - \bar{l}_j) + \bar{l}_j \leq 3\bar{l}_j^{1+\rho}.$$

Thus, by induction hypothesis, we get

$$|v_i| \ll \bar{l}_j^{\xi(\rho,j)(1-i-\eta)} \ll \bar{l}_{j+1}^{\frac{\xi(\rho,j)}{1+\rho}(1-i-\eta)}$$

for all $1 \leq i \leq k$. □

Lemma 3.9. *By the weight decomposition, an irreducible* $\mathfrak{sl}_2(\mathbf{R})$*-representation* V_ς *is the direct sum of weight spaces, each of which is* 1 *dimensional. More precisely, there exists a basis* $v_0, \ldots, v_\varsigma \in V_\varsigma$ *such that*

$$U.v_i = (i+1)v_{i+1}, \quad Y_n.v_i = \frac{\varsigma - 2i}{2} v_i.$$

Thus, if V_ς is an irreducible representation of $\mathfrak{sl}_2(\mathbf{R})$ with the highest weight $\varsigma \leq 2$, then for any $v = b_0 v_0 + \cdots + b_\varsigma v_\varsigma \in V_\varsigma$, we have

$$\begin{aligned} \exp(tU).v &= \sum_{n=0}^{\varsigma} \sum_{i=0}^{n} b_i \binom{n}{i} t^{n-i} v_n, \\ \exp(\omega Y_n).v &= \sum_{n=0}^{\varsigma} b_n e^{(\varsigma - 2n)\omega/2} v_n. \end{aligned} \tag{3.8}$$

In the following, we consider the decomposition $\mathfrak{g} = \mathfrak{sl}_2(\mathbf{R}) \oplus V^\perp$, where $\mathfrak{sl}_2(\mathbf{R}) = \mathrm{Span}\{U, Y_n, \tilde{U}\} \subset \mathfrak{g}$ is the $\mathfrak{sl}_2$-triple. We shall study $\mathfrak{sl}_2(\mathbf{R})$ and $V^\perp$ separately. We can first assume that $V^\perp = V_\varsigma$ is irreducible.

By Lemma 3.7, for sufficiently small $\epsilon > 0$, $g \in B_G(e, \epsilon)$, we have

$$g = h \exp v \tag{3.9}$$

for some $h \in B_{SO(2,1)}(e, \epsilon)$ and $v \in B_{V^\perp}(0, \epsilon)$. Now we discuss a necessary condition for $h \in B_{SO(2,1)}(e, \epsilon)$ in a small neighborhood of the identity. Recall that we consider h as a (2×2)-matrix $h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO(2,1)$. Then one may obtain that a necessary condition for $h \in B_{SO(2,1)}(e, \epsilon)$ is that $|b|, |c| < \epsilon$, $1 - \epsilon < |a|, |d| < 1 + \epsilon$.

Next, let $t(s) \in \mathbf{R}^+$ be a function of $s \in \mathbf{R}^+$. Then by (3.8) , we have

$$\begin{aligned} u^t g u^{-s} =& u^t h \exp v u^{-s} \\ =& (u^t h u^{-s})(u^s \exp(v) u^{-s}) \\ =& (u^t h u^{-s}) \exp(\mathrm{Ad}\, u^s.v) \\ =& (u^t h u^{-s}) \exp\left(\sum_{n=0}^{\varsigma} \sum_{i=0}^{n} b_i \binom{n}{i} s^{n-i} v_n \right). \end{aligned}$$

Moreover, by Lemma 3.6, 3.7, one can show for $G = SO(n,1)$ that $u^t g u^{-s} \ll \epsilon$ iff

$$u^t h u^{-s} \ll \epsilon, \quad \mathrm{Ad}\, u^s.v = \sum_{n=0}^{\varsigma} \sum_{i=0}^{n} b_i \binom{n}{i} s^{n-i} v_n \ll \epsilon. \tag{3.10}$$

Thus, we split the elements close to the identity into two parts, namely the $SO(2,1)$-part and the $V^\perp$-part.

As shown in (3.10), we shall consider the elements of the form $u^t h u^{-s} \in B_{SO(2,1)}(e, \epsilon)$. A direct calculation shows

$$u^t h u^{-s} = \begin{bmatrix} 1 & \\ t & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & \\ -s & 1 \end{bmatrix} = \begin{bmatrix} a - bs & b \\ c + (a-d)s - bs^2 + (t-s)(a-bs) & d + bt \end{bmatrix}.$$

If we further require $|s - t| \ll_\eta \max\{\epsilon, s^{1-\eta}\}$ (cf. (3.3)), then we see that

$$\begin{aligned} & |-bs^2 + (a-d)s + c + (-bs + a)(t-s)| < \epsilon \\ \Rightarrow\ & |-bs^2 + (a-d)s| - |c| - |(-bs+a)(t-s)| < \epsilon \\ \Rightarrow\ & |-bs^2 + (a-d)s| < 2\epsilon + 2|t-s| \\ \Rightarrow\ & |-bs^2 + (a-d)s| \ll_\eta \max\{\epsilon, s^{1-\eta}\}. \end{aligned} \tag{3.11}$$

By Lemma 3.8, we immediately obtain

Lemma 3.10 (Estimates for $SO(2,1)$-coefficients)**.** *Given $\eta \in (0,1)$, a sufficiently small $\epsilon > 0$, a matrix $h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in B_{SO(2,1)}(e, \epsilon)$, then the solutions $s \in [0, \infty)$ of* *the following inequality*

$$|-bs^2+(a-d)s| \ll_\eta \max\{\epsilon, s^{1-\eta}\} \tag{3.12}$$

consist of at most two intervals, say $[0,\bar{l}_1(h)]\cup[l_2(h),\bar{l}_2(h)]$ *where* $\bar{l}_1$ *has the lower bound* $l(\epsilon,\eta)$ *such that* $l(\epsilon,\eta)\nearrow\infty$ *as* $\epsilon\searrow 0$ *for fixed* η*. Moreover, we have*

(1) $|b|\ll_\eta \bar{l}_1^{-1-\eta}$ *and* $|a-d|\ll_\eta \bar{l}_1^{-\eta}$*;*

(2) If we further assume that the intervals $[0,\bar{l}_1]$ *and* $[l_2,\bar{l}_2]$ *do not have an effective gap (3.7), i.e.* $l_2-\bar{l}_1\le\min\{\bar{l}_1,\bar{l}_2-l_2\}^{1+\rho}$*, then*

$$|b|\ll_\eta \bar{l}_2^{\xi(\rho)(-1-\eta)},\quad |a-d|\ll_\eta \bar{l}_2^{\xi(\rho)(-\eta)}.$$

Next, we study the situation when $Adu^s.v\ll\epsilon$. Again by Lemma 3.8, we have

Lemma 3.11 (Estimates for $V^\perp$-coefficients)**.** *Fix* $v=b_0v_0+\cdots+b_\varsigma v_\varsigma\in B_{V_\varsigma}(0,\epsilon)$*. Assume that*

$$Adu^s.v\ll\epsilon_0 \quad \textit{iff} \quad s\in[0,\bar{l}_1(v)]\cup\cdots\cup[l_m(v),\bar{l}_m(v)]$$

where $\bar{l}_1$ *has the lower bound* $l(\epsilon,\eta)$ *such that* $l(\epsilon,\eta)\nearrow\infty$ *as* $\epsilon\searrow 0$ *for fixed* η*. Then* $m=m(v)$ *is bounded by a constant depending on* ς*. Moreover, for* $1\le j\le\varsigma-1$*, the intervals* $[0,\bar{l}_j]$ *and* $[l_{j+1},\bar{l}_{j+1}]$ *do not have an effective gap (3.7), i.e.* $l_{j+1}-\bar{l}_j\le\min\{\bar{l}_j,\bar{l}_{j+1}-l_{j+1}\}^{1+\rho}$*, then we have*

$$|b_i|\ll_{\varsigma,\eta}\bar{l}_j^{\xi(\rho,\varsigma)(-\varsigma+i)}.$$

For $g=h\exp(v)\in G$, we conclude from (3.11), Lemma 3.10 and 3.11 that

$$u^thu^{-s}\ll\epsilon \quad \text{implies} \quad s\in[0,\bar{l}_1(h)]\cup[l_2(h),\bar{l}_2(h)]$$
$$Adu^s.v\ll\epsilon \quad \text{iff} \quad s\in[0,\bar{l}_1(v)]\cup\cdots\cup[l_{m(v)}(v),\bar{l}_{m(v)}(v)].$$

Write $l_1(h)=l_1(v)=0$ and we shall consider the family of intervals

$$\{[l_k(g),\bar{l}_k(g)]\}_k\coloneqq\{[l_i(h),\bar{l}_i(h)]\cap[l_j(v),\bar{l}_j(v)]\}_{i,j} \tag{3.13}$$

where the intervals $[l_i(h),\bar{l}_i(h)]$, $[l_j(v),\bar{l}_j(v)]$ are given by Lemma 3.10, 3.11 respectively, and $\bar{l}_k(g)<l_{k+1}(g)$ for all k. Thus, in particular, $l_1(g)=0$ and $[0,\bar{l}_1(g)]=[0,\bar{l}_1(h)]\cap[0,\bar{l}_1(v)]$.

Now assume that there exists k such that $[0,\bar{l}_k(g)]$ and $[l_{k+1}(g),\bar{l}_{k+1}(g)]$ do not have an effective gap (3.7), i.e.

$$l_{k+1}(g)-\bar{l}_k(g)\le\min\{\bar{l}_k(g),\bar{l}_{k+1}(g)-l_{k+1}(g)\}^{1+\rho}.$$

Then clearly, the corresponding "$SO(2,1)$-part" and "$V^\perp$-part" should not have effective gaps either. More precisely, for the $SO(2,1)$-part, we define

$$i_{\ge k}\coloneqq\min\{i\in\{1,2\}:\bar{l}_k(g)\le\bar{l}_i(h)\},\quad i_{\le k+1}\coloneqq\max\{i\in\{1,2\}:l_{k+1}(g)\ge l_i(h)\}.$$

Thus, we know

$$[0,\bar{l}_k(g)]\subset[0,\bar{l}_{i_{\ge k}}(h)],\quad [l_{k+1}(g),\bar{l}_{k+1}(g)]\subset[l_{i_{\le k+1}}(h),\bar{l}_{i_{\le k+1}}(h)]$$

and hence $[0,\bar{l}_{i_{\geq k}}(h)]$ and $[l_{i_{\leq k+1}}(h),\bar{l}_{i_{\leq k+1}}(h)]$ do not have an effective gap (3.7). Similarly, for the $V^{\perp}$-part, we define

$$j_{\geq k} \coloneqq \min\{j : \bar{l}_k(g) \leq \bar{l}_j(v)\}, \quad j_{\leq k+1} \coloneqq \max\{j : l_{k+1}(g) \geq l_j(v)\}.$$

Then we know

$$[0,\bar{l}_k(g)] \subset [0,\bar{l}_{j_{\geq k}}(v)], \quad [l_{k+1}(g),\bar{l}_{k+1}(g)] \subset [l_{j_{\leq k+1}}(v),\bar{l}_{j_{\leq k+1}}(v)]$$

and hence $[0,\bar{l}_{j_{\geq k}}(v)]$ and $[l_{j_{\leq k+1}}(v),\bar{l}_{j_{\leq k+1}}(v)]$ do not have an effective gap (3.7). Further, one observes

$$\begin{aligned}
[0,\bar{l}_k(g)] =&[0,\bar{l}_{i_{\geq k}}(h)] \cap [0,\bar{l}_{j_{\geq k}}(v)]\\
[l_{k+1}(g),\bar{l}_{k+1}(g)] =&[l_{i_{\leq k+1}}(h),\bar{l}_{i_{\leq k+1}}(h)] \cap [l_{j_{\leq k+1}}(v),\bar{l}_{j_{\leq k+1}}(v)].
\end{aligned}$$

Now recall by the definition (3.13) that the number of intervals in $\{[l_k(g),\bar{l}_k(g)]\}_k$ is bounded by a constant $C(\varsigma) > 0$ because the numbers of intervals $\{[l_i(h),\bar{l}_i(h)]\}_i$, $\{[l_j(v),\bar{l}_j(v)]\}_j$ are. Since $\varsigma \leq 2$ when $\mathfrak{g} = \mathfrak{so}(n,1)$, we see that $C(\varsigma)$ is uniformly bounded for all ς. Thus, we conclude that the number of intervals in $\{[l_k(g),\bar{l}_k(g)]\}_k$ is uniformly bounded for all $g \in G$. Then, combining with Lemma 3.11 and 3.10, we obtain

Lemma 3.12 (Estimates for G-coefficients)**.** *Let $g = h\exp v \in B_G(e,\epsilon)$ be as above, where*

$$h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO(2,1), \quad v = b_0v_0 + \cdots + b_\varsigma v_\varsigma \in V_\varsigma.$$

Next, let $t(s) \in \mathbf{R}^+$ be a function of $s \in \mathbf{R}^+$ which satisfies $|s-t(s)| \ll_\eta \max\{\epsilon, s^{1-\eta}\}$. Then there exist intervals $\{[l_k(g),\bar{l}_k(g)]\}_k$ such that

$$u^tgu^{-s} \ll \epsilon \quad \text{implies} \quad s \in \bigcup_k [l_k(g),\bar{l}_k(g)]. \tag{3.14}$$

where $\bar{l}_1$ has the lower bound $l(\epsilon,\eta)$ such that $l(\epsilon,\eta) \nearrow \infty$ as $\epsilon \searrow 0$ for fixed η. Besides, $k \leq C$ for some constant $C = C(\mathfrak{g}) > 0$, and

1. *$|b| \ll_\eta \bar{l}_1(g)^{-1-\eta}$, $|a-d| \ll_\eta \bar{l}_1(g)^{-\eta}$, $|b_i| \ll_{\varsigma,\eta} \bar{l}_1(g)^{-\varsigma+i}$ for all $0 \leq i \leq \varsigma$;*
2. *If we further assume that the intervals $[0,\bar{l}_k(g)]$ and $[l_{k+1}(g),\bar{l}_{k+1}(g)]$ do not have an effective gap (3.7). Then there exists $\xi = \xi(\rho) \in (0,1)$ with $\xi \to 1$ as $\rho \to 0$ such that*

$$|b| \ll_\eta \bar{l}_k(g)^{-\xi(1+\eta)}, \quad |a-d| \ll_\eta \bar{l}_k(g)^{-\xi\eta}, \quad |b_i| \ll_{\varsigma,\eta} \bar{l}_k(g)^{-\xi(\varsigma-i)}$$

for all $1 \leq i \leq \varsigma$.

Remark 3.13. Since $\varsigma \leq 2$ for $\mathfrak{g} = \mathfrak{so}(n,1)$, we might obtain Lemma 3.12 via an explicit discussion of intervals (and this is simpler at first glance). However, as in general V is not irreducible (for $n \geq 4$), it would be convenient to repeat the above argument to conclude Lemma 3.12.

3.3. **Proof of Proposition 3.3.** Now we start to prove Proposition 3.3. We shall adopt a similar strategy as in [Rat86]. More precisely, we shall construct a collection β_ρ of disjoint subintervals of $\mathbf{R}^+$ so that its union covers $A \cap [0, \lambda]$ and every pair has an effective gap (3.5). Then, we apply Proposition 3.4 to obtain a large interval in β_ρ. We first specify the quantities claimed in Proposition 3.3.

- (Choice of ρ) Choose a small $\rho > 0$ that satisfies

$$\frac{1+2\rho}{\xi(2\rho)} < 1+\eta, \quad 1+2\delta < 1+2\rho < 2\xi(2\rho) \tag{3.15}$$

  where $\xi(2\rho)$ was defined in Lemma 3.12, and $\delta \coloneqq 3\rho/4$.
- (Choice of θ) Let $\theta = \theta(\rho)$ be as in Proposition 3.4.
- (Choice of σ_ρ) Then $\sigma_\rho > 0$ can be chosen as

$$\sigma_\rho < \frac{\rho}{4+6\rho}. \tag{3.16}$$

- (Choice of Δ, K_1; injectivity radius) Let $\pi : G \to X$ be the natural quotient map. Since Γ is discrete, there is a compact subset $K_1 \subset X$, $\mu(K_1) > 1 - \frac{1}{2}\sigma$ and $\Delta \in (0, 1)$ such that for any $g_x \in \pi^{-1}(K_1)$, $g_y \in G$ satisfying

$$d(g_x, g_y) < 2\Delta, \quad d(u^s g_x, u^t g_y \gamma) < 2\Delta \text{ with } e \neq \gamma \in \Gamma, \tag{3.17}$$

  we must have $\max\{|t|, |s|\} \geq m$. Here d denotes the metric on X, and m is given by the assumption of Proposition 3.3. In particular, it implies that for any $g_x \in \pi^{-1}(K_1)$, $g_y \in G$ satisfying

$$d(g_x, g_y) < 2\Delta, \quad d(g_x, g_y\gamma) < 2\Delta \tag{3.18}$$

  for some $\gamma \in \Gamma$, then $\gamma = e$.
- (Choice of K_2, K, T_0, T; ergodicity of a^T) Since the diagonal action a^T is ergodic on (X, μ), there is a compact subset $K_2 \subset G/\Gamma$, $\mu(K_2) > 1 - \frac{1}{2}\sigma$ and $T_0 = T_0(K_2) > 0$ such that the relative length measure K_2 on $[x, a^T x]$ (and $[a^{-T}x, x]$) is greater than $1 - \sigma$ for any $x \in K_2$, $|T| \geq T_0$. Assume that

$$K \coloneqq K_1 \cap K_2 \tag{3.19}$$

  Note that $\mu(K) > 1 - \sigma$. The choice will be used in (3.33).
- (Choice of ϵ) Let $0 < \epsilon < \Delta$ be so small that for $g \in B_G(e, \epsilon)$

$$\bar{l}_1(g) \geq l(\epsilon, \eta) > \max\{e^{(1+2\delta)^{-1}T_0}, m\} \tag{3.20}$$

  where $\bar{l}_1, l$ are defined in Lemma 3.12, and $\delta \coloneqq 3\rho/4$.

Thus, $0 < \rho, \xi, \theta, \epsilon < 1$ and $K \subset X$ have been chosen. Next let us describe some notation that will be used later. Let $x \in X$, $y \in B_X(x, \epsilon)$. We say that $(g_x, g_y) \in G \times G$ *covers* (x, y) if $d_G(g_x, g_y) < \epsilon$ and $\pi(g_x) = x$, $\pi(g_y) = y$.

Definition 3.14 (ϵ-block)**.** Suppose that $x \in X$, $y \in B(x, \epsilon)$, (g_x, g_y) covers (x, y), and $r \in (0, \infty]$ satisfies

$$d_G(u^r g_x, u^{t(r)} g_y) < \epsilon.$$

Then we define the *ϵ-block of g_x, g_y of length r* by

$$\mathrm{BL}(g_x, g_y) \coloneqq \{(u^s g_x, u^{t(s)} g_y) \in G \times G : 0 \leq s \leq r\}.$$

Similarly, we define the *ϵ-block of x, y of length r* by

$$\mathrm{BL}(x, y) \coloneqq \pi\, \mathrm{BL}(g_x, g_y) = \{(u^s x, u^{t(s)} y) \in X \times X : 0 \leq s \leq r\}.$$

We also write

$$\mathrm{BL}(x, y) = \{(x, y), (u^r x, u^{t(r)} y)\} = \{(x, y), (\overline{x}, \overline{y})\}$$

emphasizing that (x, y) is the first and $(\overline{x}, \overline{y})$ is the last pair of the block $\mathrm{BL}(x, y)$.

Construction of β_0. Let $x \in X$, $y \in B(x, \epsilon)$ and assume that $A \subset \mathbf{R}^+$ satisfies (i), (ii) as in Proposition 3.3 (and assume without loss of generality that $0 \in A$). For $\lambda \in A$ denote $A_\lambda \coloneqq A \cap [0, \lambda]$ and assume that

$$\mathrm{Leb}(A_\lambda) > (1 - \theta)\lambda. \tag{3.21}$$

Now we construct a collection β_0 of ϵ-blocks. Let $x_1 \coloneqq x$, $y_1 \coloneqq y$. Suppose that $(g_{x_1}, g_{y_1}) \in G \times G$ covers (x_1, y_1) and

$$\overline{s}_1 \coloneqq \sup\{s \in A_\lambda \cap [0, \bar{l}_1(g_{y_1} g_{x_1}^{-1})] : d_G(u^{t(s)} g_{y_1}, u^s g_{x_1}) < \epsilon\}.$$

Let BL_1 be the ϵ-block of x_1, y_1 of length $\overline{s}_1$, $\mathrm{BL}_1 = \{(x_1, y_1), (\overline{x}_1, \overline{y}_1)\}$. To define BL_2, we take

$$s_2 \coloneqq \inf\{s \in A_\lambda : s > \overline{s}_1\}$$

and apply the above procedure to $x_2 \coloneqq u^{s_2} x_1, \quad y_2 \coloneqq u^{t(s_2)} y_1$ (Note that by (3.14), $s_2 > \overline{s}_1$). This process defines a collection $\beta_0 = \{\mathrm{BL}_1, \ldots, \mathrm{BL}_n\}$ of ϵ-blocks on the orbit intervals $[x_1, u^\lambda x_1]$, $[y_1, u^{t(\lambda)} y_1]$ (see Figure 1):

$$x_i = u^{s_i} x_1, \quad \overline{x}_i = u^{\overline{s}_i} x_1, \quad y_i = u^{t_i} y_1, \quad \overline{y}_i = u^{\overline{t}_i} y_1.$$

Note also that by the assumption of A, we have $x_i, \overline{x}_i \in K$ for all i.

Remark 3.15. Notice that any $\mathrm{BL}_i = \{(x_i, y_i), (\overline{x}_i, \overline{y}_i)\} \in \beta_0$ has length $\leq \bar{l}_1(g_{y_i} g_{x_i}^{-1})$. Write $g_{y_i} = h \exp(v) g_{x_i}$, where $h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO(2,1)$. Then by Lemma 3.12, we immediately conclude that

$$|b| \ll_\eta \bar{l}_1(h)^{-1-\eta} \leq \bar{l}_1(g_{y_i} g_{x_i}^{-1})^{-1-\eta} \leq |\,\mathrm{BL}_i\,|^{-1-\eta}$$
$$|a - d| \ll_\eta \bar{l}_1(h)^{-\eta} \leq \bar{l}_1(g_{y_i} g_{x_i}^{-1})^{-\eta} \leq |\,\mathrm{BL}_i\,|^{-\eta}$$

where $|\,\mathrm{BL}_i\,|$ denotes the length of the ϵ-block BL_i.

For a collection β_0 of ϵ-blocks, a shifting problem may occur.

Definition 3.16 (Shifting)**.** For integers $i < j$, assume that $(g_{x_i}, g_{y_i}) \in G \times G$, $g_{y_i} \in B_G(g_{x_i}, \epsilon)$ covers (x_i, y_i). Then there is a unique $\gamma \in \Gamma$ such that

$$d_G(g_{x_j}, g_{y_j} \gamma) < \epsilon \tag{3.22}$$

where $g_{y_j} \coloneqq u^{t_j - t_i} g_{y_i}$, $g_{x_j} \coloneqq u^{s_j - s_i} g_{x_i}$. We write

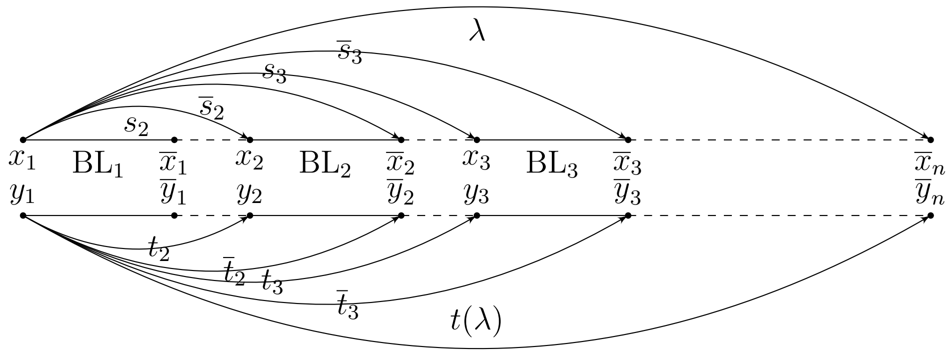

FIGURE 1. A collection of ϵ-blocks $\{\mathrm{BL}_1, \dots, \mathrm{BL}_n\}$. The solid straight lines are the unipotent orbits in the ϵ-blocks and the dashed lines are the rest of the unipotent orbits. The bent curves indicate the length defined by the letters.

- (Shifting) $(x_i, y_i) \overset{\Gamma}{\sim} (x_j, y_j)$ if $\gamma \neq e$ in (3.22),
- (Non-shifting) $(x_i, y_i) \overset{e}{\sim} (x_j, y_j)$ if $\gamma = e$ in (3.22).

Construction of β_ρ. Now we construct a new collection $\beta_\rho = \{\overline{\mathrm{BL}}_1, \dots, \overline{\mathrm{BL}}_k\}$ by the following procedure. The idea is to connect ϵ-blocks in $\beta_0 = \{\mathrm{BL}_1, \dots, \mathrm{BL}_n\}$ so that each pair of new blocks must have an effective gap. Take $\mathrm{BL}_1 \in \beta_0$, $g_{y_1} = h\exp(v) g_{x_1}$ and

$$h = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SO(2,1), \quad v = b_0 v_0 + \cdots + b_\varsigma v_\varsigma \in V_\varsigma.$$

Then by Lemma 3.12, one can write $u^{t(s)} g u^{-s} \in B_G(e, \epsilon)$ for

$$s \in \bigcup_k [l_k(g), \bar{l}_k(g)] \tag{3.23}$$

where $k \leq C$ is uniformly bounded for all $g \in G$. Then consider the following two cases:

(i) There is no $j \in \{2, \dots, n\}$ such that $(x_1, y_1) \overset{e}{\sim} (x_j, y_j)$.
(ii) There is $j \in \{2, \dots, n\}$ such that $(x_1, y_1) \overset{e}{\sim} (x_j, y_j)$.

In case (i), we set $\overline{\mathrm{BL}}_1 = \mathrm{BL}_1$. Then by Remark 3.15, we have

$$|b| \ll \bar{l}_1(g_{y_1} g_{x_1}^{-1})^{-1-\eta}, \quad |a - d| \leq \bar{l}_1(g_{y_1} g_{x_1}^{-1})^{-\eta} \tag{3.24}$$

In case (ii), suppose that $g_{x_j} = u^{s_j} g_{x_1}$, $g_{y_j} = u^{t_j} g_{y_1}$. Clearly, by the construction, $\bar{s}_j > \bar{l}_1(g_{y_1} g_{x_1}^{-1})$. On the other hand, by (3.23), we get

$$\bar{s}_j \in \bigcup_k [l_k(g_{y_1} g_{x_1}^{-1}), \bar{l}_k(g_{y_1} g_{x_1}^{-1})]$$

and $k \le C$ is uniformly bounded for all $g \in G$. Assume that $j_{\max}$ is the maximal j among $\overline{s}_j \in [l_2(g_{y_1}g_{x_1}^{-1}), \overline{l}_2(g_{y_1}g_{x_1}^{-1})]$. Whether $[0, \overline{l}_1(g_{y_1}g_{x_1}^{-1})]$ and $[l_2(g_{y_1}g_{x_1}^{-1}), \overline{l}_2(g_{y_1}g_{x_1}^{-1})]$ have an effective gap leads to a dichotomy of choices:

$$\overline{\mathrm{BL}}_1 = \begin{cases} \text{remains unchanged} & , \text{ if } l_2(g_{y_1}g_{x_1}^{-1}) - \overline{l}_1(g_{y_1}g_{x_1}^{-1}) > \overline{l}_1(g_{y_1}g_{x_1}^{-1})^{1+2\rho} \\ \{(x_1, y_1), (\overline{x}_{j_{\max}}, \overline{y}_{j_{\max}})\} & , \text{ otherwise} \end{cases}.$$

If the first case occurs, we will not change $\overline{\mathrm{BL}}_1$ anymore. If the second case occurs, i.e. we redefine $\overline{\mathrm{BL}}_1 = \{(x_1, y_1), (\overline{x}_{j_{\max}}, \overline{y}_{j_{\max}})\}$, then we repeat the construction for the new $\overline{\mathrm{BL}}_1$ again:

Suppose that there is $\overline{s}_j > \overline{l}_2(g_{y_1}g_{x_1}^{-1})$. Then assume $j_{\max}$ to be the maximal j among $\overline{s}_j \in [l_3(g_{y_1}g_{x_1}^{-1}), \overline{l}_3(g_{y_1}g_{x_1}^{-1})]$. Then again, we set

$$\overline{\mathrm{BL}}_1 = \begin{cases} \text{remains unchanged} & , \text{ if } l_3(g_y g_x^{-1}) - \overline{l}_3(g_{y_1}g_{x_1}^{-1}) > \overline{l}_2(g_{y_1}g_{x_1}^{-1})^{1+2\rho} \\ \{(x_1, y_1), (\overline{x}_{j_{\max}}, \overline{y}_{j_{\max}})\} & , \text{ otherwise} \end{cases}$$

and so on.

The process will stop since the number of intervals is uniformly bounded for all $g \in G$. Now $\overline{BL}_1 \in \beta_\rho$ has been constructed. By the choice of $\overline{\mathrm{BL}}_1$ and Lemma 3.12, we conclude that

$$|b| \ll_\eta |\,\mathrm{BL}_1\,|^{-\xi(1+\eta)}, \quad |a-d| \ll_\eta |\,\mathrm{BL}_1\,|^{-\xi\eta}, \quad |b_i| \ll_{\varsigma,\eta} |\,\mathrm{BL}_1\,|^{-\xi(\varsigma-i)} \tag{3.25}$$

for all $1 \le i \le \varsigma$.

Next, we repeat the above argument to construct $\overline{\mathrm{BL}}_{m+1}$. More precisely, suppose that $\overline{\mathrm{BL}}_m = \{(x_{j_{m-1}+1}, y_{j_{m-1}+1}), (\overline{x}_{j_m}, \overline{y}_{j_m})\} \in \beta_\rho$ has been constructed. To define $\overline{\mathrm{BL}}_{m+1}$, we repeat the above argument to $\mathrm{BL}_{j_m+1} \in \beta_0$. Thus, β_ρ is completely defined. Further, one may conclude some basic properties of β_ρ:

Lemma 3.17. *For any $\overline{\mathrm{BL}}_i = \{(x_i', y_i'), (\overline{x}_i', \overline{y}_i')\}$ in the collection $\beta_\rho = \{\overline{\mathrm{BL}}_1, \ldots, \overline{\mathrm{BL}}_k\}$ of ϵ-blocks, we have*

$$y_i' = h_i \exp(v_i) x_i' \tag{3.26}$$

where

$$h_i = \begin{bmatrix} 1 + O(r_i^{-2\rho}) & O(r_i^{-1-2\rho}) \\ O(\epsilon) & 1 + O(r_i^{-2\rho}) \end{bmatrix}, \quad v_i = O(r_i^{-\xi\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma$$

for some $r_i \ge \max\{e^{(1+2\delta)^{-1}T_0}, |\overline{\mathrm{BL}}_i|\}$ where T_0 is given by (3.19).

Proof. (3.26) follows immediately from (3.24), (3.25), (3.20), (3.15). □

Lemma 3.18. *For any $\overline{\mathrm{BL}}' \ne \overline{\mathrm{BL}}'' \in \beta_\rho$, we have*

$$d(\overline{\mathrm{BL}}', \overline{\mathrm{BL}}'') > \max\{e^{(1+2\delta)^{-1}T_0}, [\min\{|\overline{\mathrm{BL}}'|, |\overline{\mathrm{BL}}''|\}]^{1+\rho}\} \tag{3.27}$$

where the distance of blocks is defined by the distance of the intervals provided by the x-coordinate, $\delta := 3\rho/4$ and $|\,\mathrm{BL}\,|$ denotes the length of the ϵ-block BL.

Proof. Suppose $\overline{\mathrm{BL}}', \overline{\mathrm{BL}}'' \in \beta_\rho$ and write

$$\overline{\mathrm{BL}}' = \{(x', y'), (\overline{x}', \overline{y}')\}, \quad \overline{\mathrm{BL}}'' = \{(x'', y''), (\overline{x}'', \overline{y}'')\}, \quad x'' = u^s \overline{x}', \quad y'' = u^t \overline{y}'.$$

If $\overline{\mathrm{BL}}' \overset{e}{\sim} \overline{\mathrm{BL}}''$, then by the above construction, we know

$$d(\overline{\mathrm{BL}}', \overline{\mathrm{BL}}'') \geq |\overline{\mathrm{BL}}'|^{1+2\rho}$$

and so (3.27) holds in this situation. It remains to show that if $\overline{\mathrm{BL}}' \overset{\Gamma}{\sim} \overline{\mathrm{BL}}''$, (3.27) also holds. Suppose that $\overline{\mathrm{BL}}' \overset{\Gamma}{\sim} \overline{\mathrm{BL}}''$, and $g_{x''} = u^s g_{\overline{x}'}$. It follows that

$$g_{y''} = u^t g_{\overline{y}'} \gamma \quad \text{ for some } e \neq \gamma \in \Gamma. \tag{3.28}$$

Then via (3.26) and (3.25), there exists $r \geq \max\{e^{(\frac{1}{2}+\delta)^{-1}T_0}, \mathrm{Leb}(\overline{\mathrm{BL}}')\}$ such that

$$\begin{aligned} g_{\overline{y}'} &= h^{\overline{y}'}_{\overline{x}'} \exp(v^{\overline{y}'}_{\overline{x}'}) u^{-s} g_{x''} \\ g_{\overline{y}'}\gamma &= u^{-t} h^{y''}_{x''} \exp(v^{y''}_{x''}) g_{x''} \end{aligned} \tag{3.29}$$

where $g_{\overline{y}'} g_{\overline{x}'}^{-1} = h^{\overline{y}'}_{\overline{x}'} \exp(v^{\overline{y}'}_{\overline{x}'})$, $g_{y''} g_{x''}^{-1} = h^{y''}_{x''} \exp(v^{y''}_{x''})$ can be estimated by Lemma 3.17

$$h^{\overline{y}'}_{\overline{x}'} = \begin{bmatrix} 1+O(\epsilon) & O(r^{-1-2\rho}) \\ O(\epsilon) & 1+O(\epsilon) \end{bmatrix}, \quad v^{\overline{y}'}_{\overline{x}'} = O(r^{-\xi_\varsigma}) v_0 + \cdots + O(\epsilon) v_\varsigma$$

$$h^{y''}_{x''} = \begin{bmatrix} 1+O(\epsilon) & O(r^{-1-2\rho}) \\ O(\epsilon) & 1+O(\epsilon) \end{bmatrix}, \quad v^{y''}_{x''} = O(r^{-\xi_\varsigma}) v_0 + \cdots + O(\epsilon) v_\varsigma.$$

Now assume that one of s, t is not greater than $r^{1+\rho}$. Then by (3.3) and (3.17), we know

$$0 < s, t \leq O(r^{1+\rho}). \tag{3.30}$$

Since $r > e^{(1+2\delta)^{-1}T_0}$, let $e^{\omega_0} := r^{1+2\delta}$ and we know $\omega_0 > T_0$. Since $x'' \in K \subset K_2$, it follows from the choice of K_2 and T_0 that the relative length measure of K on $[x'', a^{\omega_0} x'']$ is greater than $1-\sigma$. This implies that there is ω satisfying

$$(1-\sigma)\omega_0 < \omega \leq \omega_0$$

such that $a^\omega x'' \in K$ and therefore

$$a^\omega g_{x''} \in \pi^{-1}(K) \tag{3.31}$$

where a^ω denotes the diagonal action. On the other hand, by (3.29), we have

$$\begin{aligned} a^\omega g_{\overline{y}'} &= (a^\omega h^{\overline{y}'}_{\overline{x}'} a^{-\omega}) \exp(\mathrm{Ad}\, a^\omega . v^{\overline{y}'}_{\overline{x}'})(a^\omega u^{-s} a^{-\omega}) a^\omega g_{x''} \\ a^\omega g_{\overline{y}'}\gamma &= (a^\omega u^{-t} a^{-\omega})(a^\omega h^{y''}_{x''} a^{-\omega}) \exp(\mathrm{Ad}\, a^\omega . v^{y''}_{x''}) a^\omega g_{x''} \end{aligned} \tag{3.32}$$

Notice that by the choice of ω, we have

$$e^{\omega/2} \in [r^{(1-\sigma)(\frac{1}{2}+\delta)}, r^{\frac{1}{2}+\delta}].$$

Then according to (3.32) and Lemma 3.9, we get

$$a^\omega \begin{bmatrix} 1+O(\epsilon) & O(r^{-1-2\rho}) \\ O(\epsilon) & 1+O(\epsilon) \end{bmatrix} a^{-\omega} = \begin{bmatrix} 1+O(\epsilon) & O(r^{2\delta-2\rho}) \\ O(\epsilon) & 1+O(\epsilon) \end{bmatrix}$$

$$\begin{aligned}(3.33)\qquad &\operatorname{Ad} a^{\omega}.(O(r^{-\xi\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma)\\ &= \begin{cases} O(r^{1+2\delta-2\xi})v_0 + O(\epsilon)v_1 + O(r^{-(1-\sigma)(1+2\delta)})v_2 & \text{, if } \varsigma = 2\\ O(r^{\frac{1+2\delta-2\xi}{2}})v_0 + O(r^{-(1-\sigma)(\frac{1}{2}+\delta)})v_1 & \text{, if } \varsigma = 1\\ O(\epsilon)v_0 & \text{, if } \varsigma = 0\end{cases}\end{aligned}$$

$$a^{\omega}u^{-t}a^{-\omega} = u^{-te^{-\omega}}, \quad a^{\omega}u^{-s}a^{-\omega} = u^{-se^{-\omega}}.$$

Thus, by (3.33) and (3.15), we can require T_0 sufficiently large, and then r will be large so that

$$a^{\omega}h_{\overline{x}'}^{\overline{y}'}a^{-\omega},\ \exp(\operatorname{Ad} a^{\omega}.v_{\overline{x}'}^{\overline{y}'}),\ a^{\omega}h_{x''}^{y''}a^{-\omega},\ \exp(\operatorname{Ad} a^{\omega}.v_{x''}^{y''}) \in B_G(e,\epsilon).$$

On the other hand, by (3.16), we have

$$1 + \rho - (1-\sigma)(1+2\delta) = 1 + \rho - (1-\sigma)(1 + \frac{3}{2}\rho) < -\frac{1}{4}\rho$$

and then by (3.30)

$$|-te^{-\omega}| = O(r^{-\frac{1}{4}\rho}) < \Delta, \quad |-se^{-\omega}| = O(r^{-\frac{1}{4}\rho}) < \Delta.$$

It follows from (3.19), (3.32) that

$$d_G(a^{\omega}g_{\overline{y}'}, a^{\omega}g_{x''}) < 2\Delta \ \text{ and } \ d_G(a^{\omega}g_{\overline{y}'}\gamma, a^{\omega}g_{x''}) < 2\Delta.$$

Then by (3.18) and (3.31), we conclude $\gamma = e$, which contradicts (3.28). Thus, both s, t are greater than $r^{1+\rho}$, and (3.27) follows. □

Proof of Proposition 3.3. Let $I = A_\lambda = A \cap [0,\lambda]$, $\mathcal{G}$ be the subintervals of I obtained by taking the x-coordinate of ϵ-blocks in β_ρ. Note that according to the hypotheses,

$$\operatorname{Leb}\left(\bigcup_{J\in\mathcal{G}} J\right) \geq \operatorname{Leb}(A) \geq (1-\theta)\lambda.$$

Then by (3.27), we can use Proposition 3.4 and obtain a good interval $J_\lambda = [x(\lambda), \overline{x}(\lambda)] \in \mathcal{G}$ satisfying

$$\operatorname{Leb}(J_\lambda) > \frac{3}{4}\lambda. \tag{3.34}$$

Correspondingly, there is a ϵ-block $\overline{\mathrm{BL}}(\lambda) = \{(x(\lambda), y(\lambda)), (\overline{x}(\lambda), \overline{y}(\lambda))\} \in \beta_\rho$ such that its x-coordinate has length greater than $\frac{3}{4}\lambda$. Then by (3.26) (and also the choice of $\xi(2\rho)$ in (3.15)), we get

$$y(\lambda) = h_\lambda \exp(v_\lambda)x(\lambda). \tag{3.35}$$

where

$$h_\lambda = \begin{bmatrix} 1 + O(r^{-2\rho}) & O(r^{-1-2\rho})\\ O(\epsilon) & 1 + O(r^{-2\rho})\end{bmatrix}, \quad v_\lambda = O(r^{-\frac{1+2\rho}{2}\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma$$

for some $r \geq \max\{e^{(\frac{1}{2}+\delta)^{-1}T_0}, \operatorname{Leb}(\overline{\mathrm{BL}}(\lambda))\} > \frac{3}{4}\lambda$. Therefore, we complete the proof of Proposition 3.3. □

3.4. **Proof of Theorem 3.1.** Now we are in the position to verify the equivariant properties of ψ (Theorem 3.1) via Proposition 3.3. We first consider the central direction $c \in C_G(U)$ in Lemma 3.21. Then with the "a-adjustment", we study the opposite unipotent direction $\tilde{u}$ in Lemma 3.25. Finally, since the central and the opposite unipotent directions generate the whole Lie group, we can obtain the rigidity of ψ with the help of Ratner's theorem.

Recall that in (2.5), we have defined the cocycle ξ. Now define $z : X \times \mathbf{R} \to \mathbf{R}$ by the relation $t = \xi(x, z(x,t))$, i.e.

$$t = \int_0^{z(x,t)} \tau(\phi_s^U(x))ds. \tag{3.36}$$

Then by the conjugate assumption in Theorem 3.1, we know that

$$\psi(u^t x) = \phi_t^{U,\tau}(\psi(x)) = u^{z(\psi(x),t)}\psi(x).$$

Moreover, using ergodic theorem, we get

$$|t - z(\psi(x), t)| = o(t) \tag{3.37}$$

for μ-almost all $x \in X$. Further, according to Lemma 3.1 [Rat86], when $\tau \in \mathbf{K}(X)$, we have the effective ergodicity: given $\sigma > 0$, there is $K = K(\sigma) \subset X$ with $\mu(K) > 1 - \sigma$ and t_K such that

$$|t - z(\psi(x), t)| = O(t^{1-\eta}) \tag{3.38}$$

for some $\eta > 0$, all $t \geq t_K$ and $x \in K$.

Proposition 3.19 (Lusin's theorem)**.** *Let $(X, \mathscr{B}, \mu)$ be the completion of μ on the Borel σ-algebra $\mathcal{B}_X$. Let $\psi : X \to X$ be measurable. Then given $\sigma \in (0,1)$, there is $K = K(\sigma) \in \mathscr{B}$, $\mu(K) > 1 - \sigma$ such that ψ is uniformly continuous on K.*

By (3.38), there are $P_n \subset X$ with $\mu(P_n) > 1 - 2^{-n}$ and λ_n such that

$$|t - z(\psi(x), t)| = O(t^{1-\eta}) \tag{3.39}$$

for some $\eta > 0$, all $t \geq \lambda_n$ and $x \in P_n$. Now let $a_n \coloneqq a^{(1+\gamma)\log\lambda_n}$ for some $\gamma \in (0, 2\rho)$, and let

$$\Psi_n(x) \coloneqq a_n\psi(a_{-n}x). \tag{3.40}$$

Our goal is to show that Ψ_n, after passing to a subsequence, has a pointwise limit Ψ as $n \to \infty$. By the ergodic theorem (3.37), Ψ will be u^t-equivariant, and then *Ratner theorem* applies. First, by an elementary argument, we have

Lemma 3.20. *For μ-almost all $x \in X$, there exists a subsequence $\{n(x,l)\}_{l\in\mathbf{N}} \subset \mathbf{N}$ and $y(x) \in X$ such that*

$$\lim_{l\to\infty} \Psi_{n(x,l)}(x) = y(x).$$

Proof. Write $X = \bigcup_{n=1} K_n$, where K_n are compact and $\mu(K_n) \nearrow 1$ as $n \to \infty$. We claim that $\mu(\Omega) = 1$, where

$$\Omega := \bigcup_{n\geq 1} \bigcap_{k\geq 1} \bigcup_{m\geq k} \Psi_m^{-1}(K_n).$$

But this follows from a direct calculation (recall that $d\mu_\tau := \tau d\mu$)

$$\begin{aligned}\mu\left(\bigcup_{n\geq 1} \bigcap_{k\geq 1} \bigcup_{m\geq k} \Psi_m^{-1}(K_n)\right) \geq& \mu\left(\bigcap_{k\geq 1} \bigcup_{m\geq k} \Psi_m^{-1}(K_n)\right)\\ =& \lim_{k\to\infty} \mu\left(\bigcup_{m\geq k} \Psi_m^{-1}(K_n)\right) \geq \mu_\tau(a_m^{-1}K_n).\end{aligned}$$

Note that $\mu(a_m^{-1}K_n) = \mu(K_n) \nearrow 1$ as $n \to \infty$. Then since μ_τ and μ are equivalent, $\mu_\tau(a_m^{-1}K_n) \nearrow 1$ as $n \to \infty$ and the claim follows. For almost all $x \in \Omega$, there exists $n \geq 1$ such that $\Psi_m(x) \in K_n$ for infinitely many m. $\square$

However, the subsequence $\{n(x,l)\}_{l\in\mathbf{N}} \subset \mathbf{N}$ obtained from Lemma 3.20 relies on $x \in X = G/\Gamma$. To get rid of this, we manage to verify that the limit of Ψ_n has some equivariant properties. We first prove (3.1) of Theorem 3.1:

Lemma 3.21. *There is a measurable map $\varpi : X \times C_G(U) \to C_G(U)$ such that*

$$\psi(cx) = \varpi(x,c)\psi(x) \tag{3.41}$$

for $c \in C_G(U)$, μ-almost all $x \in X$.

Proof. By Proposition 3.19, for some $\sigma \in (0,1)$ close to 0, there is $K_1 \subset X$ such that $\mu(K_1) > 1-\sigma$ and $\psi|_{K_1}$ is uniformly continuous. On the other hand, by (3.38), there are $K_2 \subset X$ with $\mu(K_2) > 1-\sigma$ and t_{K_2} such that

$$|t - z(y,t)| = O(t^{1-\eta}) \tag{3.42}$$

for some $\eta > 0$, all $t \geq t_{K_2}$ and $y \in K_2$. Then letting $m = t_{K_2}$ and according to Proposition 3.3, we obtain quantities ρ, θ, ϵ and a compact set K_3 with $\mu(K_3) > 1-\sigma$. Let $K := K_1 \cap K_2 \cap K_3$ and σ so small that $\sigma \ll \theta$. Then $\mu(K) > 1 - 3\sigma$. We shall study the unipotent orbits on K.

Next, fix a sufficiently small $\delta > 0$ so that $d(\psi(x), \psi(y)) < \epsilon$ whenever $d(x,y) < \delta$ and $x, y \in K$. Given a u^t-generic point $x \in X$, there is $A_x \subset \mathbf{R}^+$ such that $\psi(u^s cx) \in B(\psi(u^s x), \epsilon)$ for all $s \in A_x$. More precisely, given $c \in B_G(e,\delta) \cap C_G(U)$, by ergodic theorem, there is $\lambda_0 \gg m$ such that

$$u^{z(\psi(cx),s)}\psi(cx) \in B(u^{z(\psi(x),s)}\psi(x), \epsilon)$$

for $s \in A_x$ and $\mathrm{Leb}(A_x \cap [0,\lambda]) \geq (1-\sigma)\lambda$ whenever $\lambda \geq \lambda_0$. Then via Proposition 3.3 (see also (3.34)) for any $\lambda_k \geq \lambda_0$, there is an interval $J_k = \{(x_k, y_k), (\overline{x}_k, \overline{y}_k)\} \subset [0, \lambda_k]$ with $(x_k, y_k) \overset{e}{\sim} (\overline{x}_k, \overline{y}_k)$ and $x_k = u^{z(\psi(x),s_k)}\psi(x)$, $y_k = u^{z(\psi(cx),s_k)}\psi(cx)$ for some $s_k \in \mathbf{R}^+$ such that

$$y_k = h_k \exp(v_k) x_k$$

where

$$h_k = \begin{bmatrix} 1+O(\lambda_k^{-2\rho}) & O(\lambda_k^{-1-2\rho}) \\ O(\epsilon) & 1+O(\lambda_k^{-2\rho}) \end{bmatrix}, \quad v_k = O(\lambda_k^{-\frac{1+2\rho}{2}\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma$$

and $|J_k| \geq \frac{3}{4}\lambda_k$. Then we can choose an increasing sequence of λ_k so that $J_n \cap J_m \neq \emptyset$ for any $n, m \in \mathbf{N}$. But it forces $J_k \subset J_{k+1}$. Thus, write $s_x = \inf \bigcup_k J_k$ and we conclude

$$u^{z(\psi(cx),s_x)}\psi(cx) = h_x \exp(v_x) u^{z(\psi(x),s_x)}\psi(x), \quad h_x = \begin{bmatrix} 1 & \\ O(\epsilon) & 1 \end{bmatrix}, \quad v_x \in C_G(U).$$

Thus, for $c \in B_G(e,\delta) \cap C_G(U)$, we set

$$\varpi(x,c) \coloneqq u^{-z(\psi(cx),s_x)} h_x \exp(v_x) u^{z(\psi(x),s_x)}.$$

For general $c \in C_G(U)$, we can define it by iteration, since

$$\varpi(x,c^k) = \prod_{j=0}^{k-1} \varpi(c^j x, c). \tag{3.43}$$

The consequence follows. □

Now we explore further properties of ϖ. First of all, ϖ satisfies

$$u^{z(\psi(cx),t)}\varpi(x,c)\psi(x) = \varpi(u^t x, c) u^{z(\psi(x),t)}\psi(x) \tag{3.44}$$

for μ-a.e. $x \in X$. Moreover

Lemma 3.22. *For $t \in \mathbf{R}$, we have*

$$u^{z(\psi(cx),t)}\varpi(x,c) = \varpi(u^t x, c) u^{z(\psi(x),t)} \tag{3.45}$$

for μ-a.e. $x \in X$.

Proof. First of all, by (3.36), we have the cocycle identity

$$z(\psi(x), T+t) = z(\psi(u^t x), T) + z(\psi(x), t) \tag{3.46}$$

for all $t, T \in \mathbf{R}$ and μ-a.e. $x \in X$. Let $g_x \in G$ be a representative of $\psi(x)$, i.e. $g_x\Gamma = \psi(x)$. Then for any $t \in \mathbf{R}$, we deduce from (3.44) that

$$g_x^{-1} Y_c(x,t) g_x \in \Gamma$$

for μ-a.e. $x \in X$, where

$$Y_c(x,t) \coloneqq (\varpi(u^t x, c) u^{z(\psi(x),t)})^{-1} u^{z(\psi(cx),t)} \varpi(x,c).$$

It follows that there is a $\gamma \in \Gamma$ such that

$$|\{t \in \mathbf{R} : g_x^{-1} Y_c(x,t) g_x = \gamma \text{ for } \mu\text{-a.e. } x \in X\}| > 0.$$

Fix arbitrary $t_0 \in \mathbf{R}$ such that $g_x^{-1} Y_c(x,t_0) g_x = \gamma$. Then we have

$$|\{T \in \mathbf{R} : Y_c(x, T+t_0) = Y_c(x,t_0) \text{ for } \mu\text{-a.e. } x \in X\}| > 0. \tag{3.47}$$

On the other hand, note that by (3.46), we have

$$Y_c(x, T + t_0)(Y_c(x, t_0))^{-1} = Y_c(u^{t_0} x, T).$$

Then after replacing x by $u^{t_0}x$, we can assume that (3.47) holds for $t_0 = 0$.

Now suppose that there are $T, t \in \mathbf{R}$ such that

$$Y_c(x, T) = Y_c(x, t) = Y_c(x, 0) \equiv e \tag{3.48}$$

for μ-a.e. $x \in X$. Then we claim that $Y_c(x, T + t) = e$ as well. In fact, replacing x by $u^t x$ and using (3.46), (3.48), we get

$$\begin{aligned} e =& Y_c(u^t x, T) = (\varpi(u^{T+t}x, c)u^{z(\psi(u^t x),T)})^{-1}u^{z(\psi(cu^t x),T)}\varpi(u^t x, c) \\ =&(\varpi(u^{T+t}x, c)u^{z(\psi(u^t x),T)+z(\psi(x),t)})^{-1}u^{z(\psi(cu^t x),T)}(\varpi(u^t x, c)u^{z(\psi(x),t)}) \\ =&(\varpi(u^{T+t}x, c)u^{z(\psi(x),T+t)})^{-1}u^{z(\psi(cu^t x),T)}(u^{z(\psi(cx),t)}\varpi(x, c)) \\ =&(\varpi(u^{T+t}x, c)u^{z(\psi(x),T+t)})^{-1}u^{z(\psi(cx),T+t)}\varpi(x, c) = Y_c(x, T + t) \end{aligned}$$

for μ-a.e. $x \in X$. Thus, we see that $\{T \in \mathbf{R} : Y_c(x, T) = e \text{ for } \mu\text{-a.e. } x \in X\}$ is a group with positive Lebesgue measure, which can only be the whole $\mathbf{R}$. $\square$

In light of (3.45), we consider the orthogonal decomposition (2.3) and write

$$\varpi(x, c) = u^{\alpha(x,c)}\beta(x, c) \tag{3.49}$$

where $\alpha(x, c) \in \mathbf{R}$ and $\beta(x, c) \in \exp V_C^{\perp}$. Then by (3.45), we have

$$z(\psi(cx), t) + \alpha(x, c) = \alpha(u^t x, c) + z(\psi(x), t), \quad \beta(x, c) = \beta(u^t x, c) \tag{3.50}$$

for all $t \in \mathbf{R}$. Via the ergodicity of the unipotent flow u^t, we conclude that

$$\beta(x, c) \equiv \beta(c)$$

for all $c \in C_G(U)$. Besides, $\beta : C_G(U) \to \exp V_C^{\perp}$ must be surjective. This is because ψ is bijective and so for a.e. $x \in X$, $\varpi(x, \cdot) : C_G(U) \to C_G(U)$ is surjective.

On the other hand, consider

$$F(x, c_1, c_2) := (\varpi(x, c_1c_2))^{-1}\varpi(c_2x, c_1)\varpi(x, c_2)$$

for $x \in X$, $c_1, c_2 \in C_G(U)$. By (3.45), one can show that

$$F(u^t x, c_1, c_2) = F(x, c_1, c_2)$$

for $c_1, c_2 \in C_G(U)$, a.e. $x \in X$. Then by the ergodicity, $F(x, c_1, c_2) \equiv F(c_1, c_2)$. Besides, by (3.41), we know that

$$\psi(x) = F(x, c_1, c_2)\psi(x).$$

Since ψ is bijective, we conclude that

$$\varpi(x, c_1c_2) = \varpi(c_2x, c_1)\varpi(x, c_2)$$

for a.e. x, $c_1, c_2 \in C_G(U)$. In particular, we have $\beta(c_1c_2) = \beta(c_1)\beta(c_2)$. Further, we always have $\beta(u^t) \equiv e$. Therefore, we can restrict our attention to $\exp V_C^{\perp}$ and conclude that $d\beta : V_C^{\perp} \to V_C^{\perp}$ is an automorphism.

Lemma 3.21 can be interpreted by the language of cohomology. More precisely, Lemma 3.21 implies the time change τ and $\tau \circ c$ are measurably cohomologous.

Corollary 3.23. *Let $\tau \in \mathbf{K}(X)$. Suppose that there is a measurable conjugacy map $\psi : (X, \mu) \to (X, \mu_\tau)$ such that*

$$\psi(\phi_t^U(x)) = \phi_t^{U,\tau}(\psi(x))$$

for $t \in \mathbf{R}$ and μ-a.e. $x \in X$. Then $\tau(x)$ and $\tau(cx)$ are (measurably) cohomologous for all $c \in C_G(U)$. Besides, $\alpha(\cdot, c) \in L^1(X)$ for some $c \in C_G(U)$ iff the transfer function is in L^1.

Proof. By (3.50), we have

$$\begin{aligned}
&\int_0^{z(\psi(x),t)} \tau(u^s\psi(x)) - \tau(u^s\beta(c)\psi(x))ds \\
=&\int_0^{z(\psi(cx),t)} \tau(u^s\psi(cx))ds - \int_0^{z(\psi(x),t)} \tau(u^s\beta(c)\psi(x))ds \\
=&\int_0^{z(\psi(cx),t)} \tau(u^{\alpha(x,c)+s}\beta(c)\psi(x))ds - \int_0^{z(\psi(x),t)} \tau(u^s\beta(c)\psi(x))ds \\
=&\int_0^{\alpha(x,c)+z(\psi(cx),t)} \tau(u^s\beta(c)\psi(x))ds - \int_0^{\alpha(x,c)} \tau(u^s\beta(c)\psi(x))ds - \int_0^{z(\psi(x),t)} \tau(u^s\beta(c)\psi(x))ds \\
=&\int_0^{z(\psi(x),t)+\alpha(u^tx,c)} \tau(u^s\beta(c)\psi(x))ds - \int_0^{z(\psi(x),t)} \tau(u^s\beta(c)\psi(x))ds - \int_0^{\alpha(x,c)} \tau(u^s\beta(c)\psi(x))ds \\
=&\int_0^{\alpha(u^tx,c)} \tau(u^s\beta(c)\psi(u^tx))ds - \int_0^{\alpha(x,c)} \tau(u^s\beta(c)\psi(x))ds.
\end{aligned}$$

Thus, we can take the transfer function as

$$g_c(y) := \int_0^{\alpha(\psi^{-1}(y),c)} \tau(u^s\beta(c)y)ds.$$

Then $\tau(x)$ and $\tau(\beta(c)x)$ are (measurably) cohomologous for all $c \in C_G(U)$. Since $d\beta : V_C^\perp \to V_C^\perp$ is surjective, this is equivalent to say that $\tau(x)$ and $\tau(cx)$ are (measurably) cohomologous for all $c \in C_G(U)$. Since τ is bounded, we conclude that $g_c \in L^1(X)$ iff $\alpha(\cdot, c) \in L^1(X)$. $\square$

If $\tau(x)$ and $\tau(\beta(c)x)$ are cohomologous with a L^1 transfer function, then we are able to do more via the *ergodic theorem*.

Lemma 3.24. *If $\alpha(\cdot, c) \in L^1(X)$ for all $c \in C_G(U)$, then for any $c \in \exp(V_C^\perp \cap \mathfrak{g}_{-1})$ (see (2.1)), there exists $C \in \exp(V_C^\perp \cap \mathfrak{g}_{-1})$ such that*

$$\lim_{l\to\infty} d(\Psi_n(cx), C\Psi_n(x)) = 0$$

for μ-almost all $x \in X$, where Ψ_n is given by (3.40).

Proof. Fix an (orthonormal) basis $\{U, V_1, \dots, V_{n-2}\} \subset \mathfrak{g}_{-1}$. For $c_i = \exp V_i$, $\phi_t^{V_i}(x) = \exp(tV_i)x = c_i^t x$ defines an (ergodic) unipotent flow. Thus, if $\alpha(\cdot, c_i)$ is integrable, via (3.43), (3.49) and ergodic theorem, we obtain

$$\left|\frac{1}{k}\alpha(x, c_i^k) - \int \alpha(y, c_i)d\mu(y)\right| \to 0 \tag{3.51}$$

for μ-almost all $x \in X$. Thus, by Lemma 3.21 and (3.49), one can calculate

$$\begin{aligned}\Psi_n(c_i x) &= a_n\psi(a_{-n}c_i x) = a_n\psi(a_{-n}c_i a_n a_{-n}x) = a_n\psi(c_i^{\lambda_n^{1+\gamma}} a_{-n}x)\\ &= a_n u^{\alpha(a_{-n}x, c_i^{\lambda_n^{1+\gamma}})}\beta(c_i^{\lambda_n^{1+\gamma}})\psi(a_{-n}x) = u^{\lambda_n^{-(1+\gamma)}\alpha(a_{-n}x, c_i^{\lambda_n^{1+\gamma}})} \cdot a_n\beta(c_i^{\lambda_n^{1+\gamma}})a_{-n} \cdot \Psi_n(x).\end{aligned}$$

Since $V_i \in V_C^\perp \cap \mathfrak{g}_{-1}$ is nilpotent (recall (2.4)), the fact that $d\beta : V_C^\perp \to V_C^\perp$ is an automorphism implies $d\beta(V_i) \in \mathfrak{g}_{-1}$ is also nilpotent. Write $\beta(c_i) = \exp(v_i)$, where $v_i \in \mathfrak{g}_{-1}$. Then

$$a_n\beta(c_i^{\lambda_n^{1+\gamma}})a_{-n} = a_n\beta(c_i)^{\lambda_n^{1+\gamma}}a_{-n} = \exp(v_i).$$

Next, by (3.51), we can enlarge λ_n so that $\mu(W_n) > 1 - 2^{-n}$, where

$$W_n := \left\{y \in X : \left|\frac{1}{\lambda_n^{1+\gamma}}\alpha(y, c^{\lambda_n^{1+\gamma}}) - \int \alpha(\cdot, c)\right| < \frac{1}{n}\right\}.$$

It follows that $\mu(\bigcup_{m\geq 1}\bigcap_{n\geq m} a_n W_n) = 1$. Then for any $x \in \bigcup_{m\geq 1}\bigcap_{n\geq m} a_n W_n$, there exists a number $m > 0$ such that for any $n \geq m$, we have

$$\left|\frac{1}{\lambda_n^{1+\gamma}}\alpha(a_n^{-1}x, c^{\lambda_n^{1+\gamma}}) - \int \alpha(\cdot, c)\right| < \frac{1}{n}.$$

Thus, we conclude that for μ-almost all $x \in X$,

$$\lim_{n\to\infty} d_G(u^{\lambda_n^{-(1+\gamma)}\alpha(a_{-n}x, c^{\lambda_n^{1+\gamma}})}, u^{\int \alpha(\cdot, c)}) = 0.$$

The consequence follows. □

Let $\mathfrak{sl}_2(\mathbf{R}) = \mathrm{Span}\{U, Y_n, \tilde{U}\} \subset \mathfrak{g}$ be a $\mathfrak{sl}_2$-triple, $\tilde{u} = \exp(\tilde{U})$. It is again convenient to consider $u, a, \tilde{u} \in SO(2,1)$ as (2×2)-matrices. Then we have

Lemma 3.25. *Let the notation and assumption be as above. For $\delta > 0$, let $\tilde{u}^p \in B_G(e, \delta)$ for $p \in \mathbf{R}$. Then for sufficiently small $\delta > 0$ and for μ-almost all $x \in X$, there exists an element $C_{\tilde{u}}(x, p) \in C_G(\mathfrak{sl}_2(\mathbf{R}))$ such that*

$$\lim_{n\to\infty} d(\Psi_n(\tilde{u}^p x), C_{\tilde{u}}(x,p)\tilde{u}^p\Psi_n(x)) = 0.$$

Proof. Recall we have defined P_n in (3.39). Note that

$$\mu\left(\bigcup_{k\geq 1}\bigcap_{n\geq k} a_n P_n\right) = 1.$$

Suppose that $x, \tilde{u}^p x \in \bigcup_{k\geq 1}\bigcap_{n\geq k} a_n P_n$. Let $t(\lambda_n) := \frac{\lambda_n}{1-p\lambda_n^{-\gamma}}$ and consider

$$\begin{aligned}d(u^t a_n^{-1}\tilde{u}^p x, u^{\lambda_n} a_n^{-1}x) &= d(u^t \tilde{u}^{p\lambda_n^{-1-\gamma}} a_n^{-1}x, u^{\lambda_n} a_n^{-1}x)\\ &= d\left(\begin{bmatrix} 1-\lambda_n^{-\gamma}p & \lambda_n^{-1-\gamma}p \\ 0 & 1+\lambda_n^{-1-\gamma}pt\end{bmatrix} u^{\lambda_n} a_n^{-1}x, u^{\lambda_n}a_n^{-1}x\right) \leq \delta.\end{aligned}$$

Then by the continuity of ψ, we get

$$d(u^{z(\psi(a_n^{-1}\tilde{u}^p x),t)}\psi(a_n^{-1}\tilde{u}^p x), u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x)) < \epsilon. \tag{3.52}$$

Similarly, letting $\tilde{t}(\lambda_n) := \frac{z(\psi(a_n^{-1}x),\lambda_n)}{1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma}}$, we get

$$\begin{aligned}&d(u^{\tilde{t}}\tilde{u}^{p\lambda_n^{-1-\gamma}}\psi(a_n^{-1}x), u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x))\\ =&d(u^{\tilde{t}}\tilde{u}^{p\lambda_n^{-1-\gamma}}u^{-z(\psi(a_n^{-1}x),\lambda_n)}u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x), u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x))\\ =&d\left(\begin{bmatrix} 1-\lambda_n^{-1-\gamma}pz(\psi(a_n^{-1}x),\lambda_n) & \lambda_n^{-1-\gamma}p \\ 0 & 1+\lambda_n^{-1-\gamma}pt\end{bmatrix} u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x), u^{z(\psi(a_n^{-1}x),\lambda_n)}\psi(a_n^{-1}x)\right) < \delta.\end{aligned} \tag{3.53}$$

Combining (3.52) with (3.53), we obtain

$$d(u^{z(\psi(a_n^{-1}\tilde{u}^p x),t)}\psi(a_n^{-1}\tilde{u}^p x), u^{\tilde{t}}\tilde{u}^{p\lambda_n^{-1-\gamma}}\psi(a_n^{-1}x)) \ll \epsilon.$$

In order to apply Proposition 3.3, we need to consider

$$\begin{aligned}&|z(\psi(a_n^{-1}\tilde{u}^p x),t)-\tilde{t}|\\ \leq&|z(\psi(a_n^{-1}\tilde{u}^p x),t)-t| + |t-\tilde{t}|\\ =&O(t^{1-\eta}) + \left|\frac{\lambda_n}{1-p\lambda_n^{-\gamma}} - \frac{z(\psi(a_n^{-1}x),\lambda_n)}{1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma}}\right|\\ \leq&O(\lambda_n^{1-\eta}) + \left|\frac{\lambda_n}{1-p\lambda_n^{-\gamma}} - \frac{\lambda_n}{1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma}}\right|\\ &+\left|\frac{\lambda_n}{1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma}} - \frac{z(\psi(a_n^{-1}x),\lambda_n)}{1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma}}\right|\\ \leq&O(\lambda_n^{1-\eta}) + \left|\frac{p\lambda_n^{-\gamma}(z(\psi(a_n^{-1}x),\lambda_n)-\lambda_n)}{(1-p\lambda_n^{-\gamma})(1-pz(\psi(a_n^{-1}x),\lambda_n)\lambda_n^{-1-\gamma})}\right| + O(\lambda_n^{1-\eta})\\ =&O(\lambda_n^{1-\eta}) + o(\lambda_n^{1-\eta}) + O(\lambda_n^{1-\eta}).\end{aligned}$$

Thus, via Proposition 3.3, we conclude that

$$u^{t_{\lambda_n}}\psi(a_n^{-1}\tilde{u}^p x) = h_n \exp(v_n) u^{s_{\lambda_n}}\tilde{u}^{p\lambda_n^{-1-\gamma}}\psi(a_n^{-1}x)$$

where

$$h_n = \begin{bmatrix} 1+O(\lambda_n^{-2\rho}) & O(\lambda_n^{-1-2\rho}) \\ O(\epsilon) & 1+O(\lambda_n^{-2\rho})\end{bmatrix}, \quad v_n = O(\lambda_n^{-\frac{1+2\rho}{2}\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma$$

for some $t_{\lambda_n}, s_{\lambda_n} \ll \lambda_n$.

It follows that

$$\begin{aligned} u^{t_{\lambda_n}\lambda_n^{-1-\gamma}}\Psi_n(\tilde{u}^p x) =& a_n h_n \exp(v_n) u^{s_{\lambda_n}} \tilde{u}^{p\lambda_n^{-1-\gamma}} \psi(a_n^{-1}x) \\ =& a_n h_n a_n^{-1} \exp(\operatorname{Ad} a_n.v_n) u^{s_{\lambda_n}\lambda_n^{-1-\gamma}} \tilde{u}^p \Psi_n(x). \end{aligned}$$

Then, one can calculate

$$a_n h_n a_n^{-1} = \begin{bmatrix} 1+O(\lambda_n^{-2\rho}) & O(\lambda_n^{\gamma-2\rho}) \\ O(\lambda_n^{-1-\gamma}) & 1+O(\lambda_n^{-2\rho}) \end{bmatrix}, \quad \operatorname{Ad} a_n.v_n = O(\lambda_n^{\frac{\gamma-2\rho}{2}\varsigma})v_0 + \cdots + O(\epsilon)v_\varsigma.$$

Thus, letting $n \to \infty$, the consequence follows. $\square$

It is worth noting that $\{c, \tilde{u}^p : c \in \exp \mathfrak{g}_{-1}, p \in \mathbf{R}\}$ already generates the whole group $G = SO(n,1)$. Thus, using Lemma 3.21, 3.25 and *Fubini's theorem*, we get

Corollary 3.26. *There exists a sufficiently small $\delta > 0$, a map $f(g) \in G$ such that for μ-almost all $x \in X$, we have*

$$\lim_{n\to\infty} d(\Psi_n(gx), f(g)\Psi_n(x)) = 0 \tag{3.54}$$

for almost all $g \in B_G(e,\delta)$.

Now fix $x \in X$ so that Corollary 3.26 and Lemma 3.20 apply. Then by Lemma 3.20, we can fix a universal subsequence $\{n(l)\}_{l\in\mathbf{N}} \subset \mathbf{N}$ and $y \in X$ such that

$$\lim_{l\to\infty} \Psi_{n(l)}(x) = y.$$

Write $\Psi(x) \coloneqq y$. Then, (3.54) implies that $\Psi_{n(l)}(gx) \to f(g)y \eqqcolon \Psi(gx)$ as $l \to \infty$ for $g \in B_G(e,\delta)$. Finally, since u^t is ergodic, we have $\mu(u^{\mathbf{R}}B_G(e,\delta)x) = 1$ and

$$\Psi(u^t gx) \coloneqq \lim_{l\to\infty} \Psi_{n(l)}(u^t gx) = \lim_{l\to\infty} u^{\lambda_{n(l)}^{-1-\gamma} z(\psi(gx), \lambda_{n(l)}^{1+\gamma} t)} \Psi_{n(l)}(gx) = u^t \Psi(gx)$$

is well defined for $u^t g \in u^{\mathbf{R}} B_G(e,\delta)$. In other words, we obtain a (surjective) u^t-equivariant map $\Psi : X \to X$. Next, consider the graph map $\overline{\Psi} : X \to X \times X$ defined by

$$\overline{\Psi} : x \mapsto (x, \Psi(x)).$$

Then $\overline{\Psi}_*\mu$ is a $(u^t \times u^t)$-invariant and ergodic measure supported on $\operatorname{graph}(\Psi)$. By *Ratner's theorem*, we conclude that there is a subgroup $S \leq G \times G$ and a point $(x_0, y_0) \in X \times X$ such that

$$\operatorname{graph}(\Psi) = \operatorname{supp}(\overline{\Psi}_*\mu) = S.(x_0, y_0).$$

It is then not hard to see that S is the graph of an automorphism $\Phi : G \to G$ (cf. [Mor05]). Thus, we see that

$$\Psi(gx_0) = \Phi(g)y_0$$

is an affine map. By Lemma 3.25, we know that $\Phi(\tilde{u}^p) = C_{\tilde{u}}(p)\tilde{u}^p$ for some $C_{\tilde{u}}(p) \in C_G(U)$. On the other hand, the JacobsonMorozov theorem asserts that all $\mathfrak{sl}_2$-triples are conjugate under the action of the group $C_G(U)$. Since Φ fixes u^t, we conclude that Φ fixes $SO(2,1)$.

On the other hand, since

$$\lim_{l\to\infty} d(\Psi_{n(l)}(gx), \Psi(gx)) = 0$$

for $g \in B_G(e,\delta)$, μ-almost all $x \in X$. Thus, for sufficiently large $l \in \mathbf{N}$, most points $x \in X$, we have

$$\epsilon > d(\Psi_{n(l)}(u^s x), \Psi(u^s x)) = d(u^{\lambda_{n(l)}^{-1-\gamma} z(\psi(gx), \lambda_{n(l)}^{1+\gamma} s)} \Psi_{n(l)}(x), u^s \Psi(x))$$

for most of the time $s \in \mathbf{R}$. Applying Proposition 3.3 to $t(s) = \lambda_{n(l)}^{-1-\gamma} z(\psi(gx), \lambda_{n(l)}^{1+\gamma} s)$ (similar to the proof of Lemma 3.21), there exists $c(x) \in C_G(U)$ such that

$$\Psi_{n(l)}(x) = c(x)\Psi(x).$$

It follows that there exists a function $c : X \to \mathbf{C}$ such that

$$\psi(gx_0) = c(gx_0)\Phi(g)y_0.$$

Therefore, we have proved Theorem 3.1.

Similar to Corollary 3.23, by (3.2), we have

Corollary 3.27. *Let $\tau \in \mathbf{K}(X)$. Suppose that there is a measurable conjugacy map $\psi : (X,\mu) \to (X,\mu_\tau)$ such that*

$$\psi(\phi_t^U(x)) = \phi_t^{U,\tau}(\psi(x))$$

for $t \in \mathbf{R}$ and μ-a.e. $x \in X$. Assume further that $\tau(x)$ and $\tau(cx)$ are L^1-cohomologous for all $c \in C_G(U)$. Then 1 and τ are cohomologous.

Proof. Write $c(x) = u^{a(x)}b$, i.e. by (3.2), $\psi(gx_0) = u^{a(gx_0)} b\Phi(g)y_0$. Note that $a(gx_0) + z(\psi(gx_0), t) = t + a(u^t gx_0)$. It follows that

$$\begin{aligned}
&\int_0^t 1 - \tau(u^s b\Phi(g)y_0)ds \\
&= \int_0^{z(\psi(gx_0),t)} \tau(u^s u^{a(gx_0)} b\Phi(g)y_0)ds - \int_0^t \tau(u^s b\Phi(g)y_0)ds \\
&= \int_0^{z(\psi(gx_0),t)+a(gx_0)} \tau(u^s b\Phi(g)y_0)ds - \int_0^{a(gx_0)} \tau(u^s b\Phi(g)y_0)ds - \int_0^t \tau(u^s b\Phi(g)y_0)ds \\
&= \int_0^{t+a(u^t gx_0)} \tau(u^s b\Phi(g)y_0)ds - \int_0^{a(gx_0)} \tau(u^s b\Phi(g)y_0)ds - \int_0^t \tau(u^s b\Phi(g)y_0)ds \\
&= \int_0^{a(u^t gx_0)} \tau(u^s b\Phi(u^t g)y_0)ds - \int_0^{a(gx_0)} \tau(u^s b\Phi(g)y_0)ds.
\end{aligned}$$

Then 1 and $\tau(b\Phi(g)y_0)$ are cohomologous. Because τ and $\tau \circ b$ are cohomologous by assumption, the consequence follows. $\square$

4. Restriction of representations

4.1. **Unitary representations of $SO(n,1)$.** Now we adopt the standard notation in [Kna01] Chapter 7 to develop the unitary representation of G. Let $\sigma_{\mathbf{n}}$ be an irreducible unitary representation of $M = SO(n-1)$, where $\mathbf{n}$ indicates the highest weight. Besides, we require $\mathbf{n} = (n_i)_{1\le i\le\lfloor\frac{n-1}{2}\rfloor}$ satisfies

$$\begin{array}{ll} 0 \le n_1 \le \cdots \le n_{k-1} & , \text{ if } n = 2k \\ |n_1| \le n_2 \le \cdots \le n_k & , \text{ if } n = 2k+1 \end{array}.$$

Then for $\nu \in \mathbf{C}$, let $(\mathcal{H}_{\mathbf{n},\nu}, \pi_{\mathbf{n},\nu})$ be the induced representation of G from MAN given by

$$\{f : G \to \mathbf{C} \big| f(gme^{tY_n}n) = e^{-(\nu+\rho)t}\sigma_{\mathbf{n}}(m)^{-1}f(g),\ me^{tY_n}n \in MAN,\ f|_K \in L^2(K)\}$$

where $K = SO(n)$ is a maximal compact subgroup of G, with the group operation

$$(\pi_{\mathbf{n},\nu}(g)f)(x) = f(g^{-1}x).$$

It is possible to show that

$$\begin{array}{ll} \pi_{\mathbf{n},\nu} \text{ is unitary equivalent to } \pi_{\mathbf{n},-\nu} & , \text{ if } n = 2k \\ \pi_{\mathbf{n},\nu} \text{ is unitary equivalent to } \pi_{\mathbf{n}_1,-\nu} & , \text{ if } n = 2k+1 \end{array} \tag{4.1}$$

where $\mathbf{n_1} = (-n_1, n_2, \ldots, n_k)$.

Note that f in $\pi_{\mathbf{n},\nu}$ are invariant under M. Thus, $\mathcal{H}_{\mathbf{n},\nu}$ can be realized on $L^2(K/M) = L^2(S^{n-1})$. The natural L^2-norm on $L^2(S^{n-1})$ can define a unitary representation for π_ν only when $\nu = it$ for $t \in \mathbf{R}$. It is tempered, and called the *principal series*. However, it is still possible to unitarize the representations for $\nu \in (-\rho, 0) \cup (0, \rho)$ by other norms (see Theorem 4.4). They are called the *complementary series* and not tempered.

For a fixed $(\mathcal{H}_{\mathbf{n},\nu}, \pi_{\mathbf{n},\nu})$, the K-restricted representation of $K = SO(n)$ is a direct sum of K-irreducible representations $\mathcal{H}_{\mathbf{m}}$. Thus, we have

$$\mathcal{H}_{\mathbf{n},\nu} = \bigoplus_{\mathbf{m}} \mathcal{W}_{\mathbf{m}} \tag{4.2}$$

where $\mathbf{m} = (m_i)_{1\le i\le\lceil\frac{n-1}{2}\rceil}$ indicates the highest weight and satisfies

$$\begin{array}{ll} |m_1| \le n_1 \le m_2 \le n_2 \le \cdots \le m_{k-1} \le n_{k-1} \le m_k < \infty & , \text{ if } n = 2k \\ |n_1| \le m_1 \le n_2 \le m_2 \le \cdots \le m_{k-1} \le n_k \le m_k < \infty & , \text{ if } n = 2k+1 \end{array}.$$

There is a standard orthonormal basis for $\mathcal{W}_{\mathbf{m}}$ (and hence for $\mathcal{H}_{\mathbf{n},\nu}$), called the *Gelfand-Tsetlin basis*. See [GT50], [Hir62a], [Ram13] for more details. However, we do not need it here.

In the following, we are mainly interested in the case $\sigma_{\mathbf{n}} = 1$ and hence $\mathbf{n} = 0$. (It follows that $\mathbf{m} = (0, \ldots, 0, m_k)$ and so we consider $\mathbf{m}$ as an integer.) In this case, the representations are *spherical* (or *class one*) and we shall denote $(\mathcal{H}_{0,\nu}, \pi_{0,\nu})$ by $(\mathcal{H}_\nu, \pi_\nu)$. For more information about the general cases, one may see [Hir62b], [Thi74], and so on.

In order to make the restriction map clear, we review some facts about spherical harmonics (see [JW77], [Vil78], also [Zha15]). We identify $\mathfrak{p}$ with $\mathbf{R}^n$, and consider the adjoint action of K on $\mathfrak{p}$. We fix a K-invariant inner product on $\mathfrak{p}$ so that $Y_1, \ldots, Y_n$ form an orthonormal basis. Then the homogeneous space $K/M \cong S^{n-1}$. Let $\hat{K}$ be the unitary dual of K, i.e. the set of equivalent classes of irreducible finite dimensional representations of K. If $(\pi_\gamma, V_\gamma) \in \gamma \in \hat{K}$, let $V_\gamma^M := \{v \in V_\gamma : \pi(M)v = v\}$ be the space of M-fixed vectors. General representation theory, namely *Frobenius reciprocity* and *Peter-Weyl theorem*, implies that

$$L^2(S^{n-1}) = \bigoplus_{\gamma \in \hat{K}} n_\gamma V_\gamma$$

where $n_\gamma = \dim V_\gamma^M$.

However, we can explore further properties of n_γ and V_γ^M. Let $x_1, \ldots, x_n$ be the standard coordinates for $\mathfrak{p} = \mathbf{R}^n$. Let $r^2 = \sum_{i=1}^n x_i^2$ and $\Delta_n = \sum_{i=1}^n \partial^2/\partial x_i^2$ be the standard *Laplacian*. Let $\mathcal{P}^p$ be the space of all homogeneous polynomials of degree p in the variables $x_1, \ldots, x_n$ and let $W^p = \ker \Delta_n|_{\mathcal{P}^p}$ be the *spherical harmonics*. Clearly, W^p is a K-representation, and W^0 is the trivial representation. Besides, it is known that

$$L^2(S^{n-1}) = \bigoplus_{p \geq 0} W_p$$

where we are identifying elements of W^p and their restrictions to the unit sphere $S^{n-1} \subset \mathbf{R}^n$. Moreover, it is proved for $p \geq 1$ that

$$\begin{array}{ll} W_p = \mathbf{C}\chi_p \oplus \mathbf{C}\chi_{-p} & , \text{ if } n = 2 \\ W_p \text{ is irreducible} & , \text{ if } n \geq 3 \end{array}$$

where χ_p is the character on S^1 of degree p. Thus, we conclude that

$$\mathcal{W}_{\mathbf{m}} = \begin{cases} \mathbf{C}\chi_{\mathbf{m}} & , \text{ if } n = 2 \\ W_{\mathbf{m}} & , \text{ if } n \geq 3 \end{cases} \tag{4.3}$$

to align the notation. The subspace $(\mathcal{W}_{\mathbf{m}})^M$ of M-fixed vectors is 1-dimensional

$$(\mathcal{W}_{\mathbf{m}})^M = \mathbf{C}\phi_{\mathbf{m}}$$

where $\phi_{\mathbf{m}}$ is a generator normalized by $\phi_{\mathbf{m}}(Y_n) = 1$. They depend only on the last variable $x_n \in S^{n-1}$ of $x = (x_1, \ldots, x_n)$. In the following, we put the upper-index the dimension n as we shall treat it as a variable, such as $\phi_{\mathbf{m}} = \phi_{\mathbf{m}}^n$.

Lemma 4.1 (Theorem 3.1 [JW77])**.** *The polynomials $\phi_{\mathbf{m}}^n$ is given as follows:*

$$x_n = \cos\xi, \quad \phi_{\mathbf{m}}^n(x_n) := \cos^{\mathbf{m}}\xi F(-\frac{\mathbf{m}}{2}, -\frac{\mathbf{m}-1}{2}, \frac{n-1}{2}, -\tan^2\xi)$$

where $F(a, b, c, x)$ is the Gauss hypergeometric function ${}_2F_1$,

$$F(a, b, c, x) = \sum_{m=0}^{\infty} \frac{(a)_m (b)_m}{(c)_m} \frac{x^m}{m!}$$

and $(a)_m = \prod_{j=0}^{m-1}(a+j)$ *is the Pochammer symbol.*

Next, we introduce the notation on the subgroup $H = SO(n-1,1) \subset G$. In the following, we shall use superscript $\flat$ to indicate the corresponding H-data, as we obtained for G, and many of them are obtained by restriction of $H \subset G$. For example, we write

$$H = K^\flat A N^\flat$$

for the *Iwasawa decomposition* of H (note that as $H \subset G$, we may require the maximal abelian subgroups A of H and G coincide). Besides, $K^\flat = K \cap H$ and $N^\flat = N \cap H$. Here, for simplicity, we choose H so that Y_1 is invariant under $K^\flat$.

Again, we are able to construct the unitary representation of H. For $\nu \in \mathbf{C}$, let $(\pi_\nu^\flat, \mathcal{H}_\nu^\flat)$ be the induced representation of H from $M^\flat A N^\flat$ given by

$$\{f : H \to \mathbf{C} \Big| f(gme^{tY_n}n) = e^{-(\nu+\rho^\flat)t} f(g),\ me^{tY_n}n \in M^\flat A N^\flat,\ f|_{K^\flat} \in L^2(K^\flat)\}.$$

Similarly as for G, the complementary series of H are defined for $\nu \in (-\rho^\flat, 0) \cup (0, \rho^\flat)$. Now for $\nu \in (\rho^\flat, \rho)$, we can define the restriction map $\mathrm{Res} : \mathcal{H}_{-\nu} \to \mathcal{H}^\flat_{\frac{1}{2}-\nu}$ by

$$\mathrm{Res} : f \mapsto f|_H. \tag{4.4}$$

One important consequence is that Res is H-equivariant, i.e.

$$\mathrm{Res}(\pi_{-\nu}(h)f) = \pi^\flat_{\frac{1}{2}-\nu}(h)\,\mathrm{Res}(f) \tag{4.5}$$

for all $h \in H$ and $f \in \mathcal{H}_{-\nu}$. When we realize them as elements in L^2, then the restriction map (4.4) becomes $\mathrm{Res} : L^2(K/M) \to L^2(K^\flat/M^\flat)$ by

$$\mathrm{Res} : f \mapsto f|_{Y_1=0}.$$

It is known that

$$L^2(K^\flat/M^\flat) = L^2(S^{n-2}) = \bigoplus_l \mathcal{V}_l \tag{4.6}$$

where $\mathcal{V}_l$ is the space of harmonic polynomials in $n-1$ variables of degree $\mathbf{m}$ defined in (4.3). Then we have

Lemma 4.2 (Lemma 3.3 [Zha15])**.** *The branching of* $\mathcal{W}_\mathbf{m}$ *and* $\mathrm{Res}(\mathcal{W}_\mathbf{m})$ *under* $K^\flat$ *is given by*

$$\mathcal{W}_\mathbf{m} = \bigoplus_{|l| \le \mathbf{m}} \widetilde{\mathcal{V}}_l, \quad \mathrm{Res}(\mathcal{W}_\mathbf{m}) = \bigoplus_{\substack{|l| \le \mathbf{m} \\ \mathbf{m}-l \ \text{even}}} \mathcal{V}_l \tag{4.7}$$

where $\widetilde{\mathcal{V}}_l \subset L^2(S^{n-1})$ *denotes the* $K^\flat$*-irreducible representation of highest weight* l *in* $L^2(K/M)$*. Further, the isomorphism* $\mathcal{V}_l \to \widetilde{\mathcal{V}}_l$ *is given by*

$$h(x_2, \ldots, x_n) \mapsto h(x_2, \ldots, x_n)\phi^{n+2s}_{p-s}(x_1)$$

where ϕ *is given in (4.1).*

4.2. **Casimir and Laplace operators.** In this section, we review the Casimir operators and Laplace-Beltrami operators on $SO(n,1)$. See [Ram13] and the references therein. The *Casimir operator* for $SO(n,1)$ is

$$\Box_n := -\sum_{k=1}^{n} Y_k^2 + \sum_{1\le i<j\le n} \Theta_{ij}^2.$$

It is in the center of the universal enveloping algebra of $\mathfrak{g}$, and therefore acts as a scalar $c_n(\mathbf{n},\nu)$ in any irreducible unitary representation $\mathcal{H}_{\mathbf{n},\nu}$. By [Thi74] Theorem 3 (or [Thi73] Lemma 6), we know that

$$\begin{aligned} c_n(\mathbf{n},\nu) =& \rho_n^2 - \nu^2 - \langle \mathbf{n}, \mathbf{n} + 2\rho_{M_n}\rangle \\ =& \begin{cases} \rho_n^2 - \nu^2 - \sum_{i=1}^{k-1} n_i(n_i + 2i - 1) & , \text{ if } n = 2k \\ \rho_n^2 - \nu^2 - \sum_{i=1}^{k} n_i(n_i + 2i - 2) & , \text{ if } n = 2k+1 \end{cases} \end{aligned} \tag{4.8}$$

where ρ_n and ρ_{M_n} are the half-sum of positive roots of $SO(n,1)$ and $M_n = SO(n-1)$ respectively. Similarly, the *Casimir operator* of $K_n = SO(n)$ is given by

$$\Box_{K_n} := \sum_{1\le i<j\le n} \Theta_{ij}^2.$$

It again acts as a scalar in any irreducible unitary representation $\mathcal{H}_{\mathbf{m}}$. As $\mathbf{m}$ indicates the highest weight of $\mathcal{H}_{\mathbf{m}}$, we conclude from the standard representation theory (e.g. [Hum12] Section 23) that the scalar is

$$\begin{aligned} c_{K_n}(\mathbf{m}) =& \langle \mathbf{m}, \mathbf{m} + 2\rho_{K_n}\rangle \\ =& \begin{cases} -\sum_{i=1}^{k} m_i(m_i + 2i - 2) & , \text{ if } n = 2k \\ -\sum_{i=1}^{k} m_i(m_i + 2i - 1) & , \text{ if } n = 2k+1 \end{cases} \end{aligned} \tag{4.9}$$

where ρ_{K_n} is the half sum of the positive roots of K_n.

Note that now the *Laplace-Beltrami operator* Δ is then defined by

$$\Delta := \Box_n - 2\Box_{K_n}.$$

Since then Δ commute with K_n, we can define the Laplace-Beltrami operator and Sobolev norms on $K\backslash G/\Gamma$, after making a standard identification between $L^2(K\backslash G/\Gamma)$ and the subspace $L^2(G/\Gamma)^K$ of K-invariant elements of $L^2(G/\Gamma)$.

Also, recall that a *spherical representation* of $G = SO(n,1)$ (e.g. [Cor90] Section 4) is a representation which contains a nontrivial K-fixed vector. Now define the spherical part $L^2(G/\Gamma)^{\mathrm{sph}}$ to be the minimal subrepresentation containing the K-fixed part $L^2(G/\Gamma)^K$. Then the spherical part $L^2(G/\Gamma)^{\mathrm{sph}}$ decomposes discretely or continuously into irreducible spherical unitary representations of G:

$$L^2(G/\Gamma)^{\mathrm{sph}} = \int \pi_\lambda d\mu(\lambda).$$

Harish-Chandra (e.g. [Sha00]) has shown that the spherical representation π_λ occurs in the decomposition, correspond to the L^2-spectrum of the Laplacian Δ acting on the locally symmetric space $K\backslash G/\Gamma$. In particular, the complementary series π_ν

$(0 < \nu < \rho_n)$ lies in the support of μ iff $\rho_n^2 - \nu^2$ lies in the spectrum of Δ. In other words, the complementary series occurs iff the spectrum satisfies

$$\mathrm{Spec}(\Delta) \cap (0, \rho_n^2) \neq \emptyset.$$

It is easy to see from many points of view that the smallest nonzero eigenvalue in $\mathrm{Spec}(\Delta)$ can be made arbitrarily small, even for cocompact lattice Γ. For instance, by [Mil76], there exists a hyperbolic manifold X with positive first Betti number. Then let X^k be the cyclic covering of degree k induced by a fixed surjective homomorphism

$$\pi_1(X) \xrightarrow{\varphi} \mathbf{Z} \to \mathbf{Z}/k$$

where φ is independent of k. Then there are constants $c_1(X)$ and $c_2(X)$ such that the smallest nonzero eigenvalue $\lambda(X)$ in $\mathrm{Spec}(\Delta_X)$ satisfies

$$c_1(X)k^{-2} \leq \lambda(X^k) \leq c_2(X)k^{-2}$$

as what we wanted. See [Ran74], [SWY80], [Bro88] for more details. Thus, we take it for granted that there exist cocompact lattices Γ for which $L^2(G/\Gamma)$ contains complementary series with spectral parameter $\nu \in (\rho_{n-1}, \rho_n)$ as a direct summand.

Although not needed in our proof, it is worth mentioning other results for the study of $\mathrm{Spec}(\Delta)$. For example, Lax and Phillips have shown that for *geometrically finite* discrete subgroup Γ, the spectrum $\mathrm{Spec}(\Delta)$ of Δ on $\mathbf{H}^n/\Gamma$ has at most finitely many L^2-eigenvalues in the interval $[0, \rho_n^2)$ [LP82] and purely absolutely continuous spectrum of infinite multiplicity in $[\rho_n^2, \infty)$ [LP84]. On the other hand, let $\widehat{G}^{\mathrm{sph}}$ be the spherical unitary dual of $G = SO(n,1)$, that is

$$\widehat{G}^{\mathrm{sph}} = \{\pi_\lambda \bmod \pm 1 : \lambda \in i\mathbf{R} \cup [-\rho_n, \rho_n]\}.$$

Then let $\widehat{G}^{\mathrm{sph}}_{\mathrm{Aut}}$ be its automorphic dual, consisting of all π_λ which occur in $L^2(G/\Gamma)$ where Γ varies over all congruence subgroups of $G(\mathbf{Z})$. We have the following *generalized Ramanujan conjecture* for G.

Conjecture 4.3 (Generalized Ramanujan conjecture)**.** *Let $G = SO(n,1)$. Then*

$$\widehat{G}^{\mathrm{sph}}_{\mathrm{Aut}} = i\mathbf{R} \cup \{\rho_n, \rho_n - 1, \ldots, \rho_n - \lfloor \rho_n \rfloor\}.$$

For $n = 2$, it reduces to the *Selberg's 1/4 conjecture*. See [Sar05] for more details.

4.3. **Hilbert and Sobolev structures.** As mentioned in Section 4.1, for $\nu \in (-\rho, 0) \cup (0, \rho)$, one may define a spherical complementary series $(\mathcal{H}_\nu, \pi_\nu)$. Besides, the elements $\mathcal{H}_\nu$ can be realized on $L^2(K/M) = L^2(S^{n-1})$. Then the norms $\|\cdot\|_{\mathcal{H}_\nu}$ can be obtained by

Theorem 4.4 (Theorem 6.2 [JW77], [Kos69])**.** *For $\nu \in (-\rho, \rho)$, $w = \sum_{\mathbf{m}} w_{\mathbf{m}} \in L^2(S^{n-1}) = \bigoplus_{\mathbf{m}} \mathcal{W}_{\mathbf{m}}$, the norm $\|\cdot\|_{\pi_\nu}$ on $\mathcal{H}_\nu$ is given by*

$$\|w\|^2_{\mathcal{H}_\nu} = \sum_{\mathbf{m}} d_{\mathbf{m}}(\nu)\|w_{\mathbf{m}}\|^2 \tag{4.10}$$

where $\|w_{\mathbf{m}}\|^2$ *is the* L^2*-norm, and*

$$d_{\mathbf{m}}(-\nu) = \frac{(\rho+\nu)_{\mathbf{m}}}{(\rho-\nu)_{\mathbf{m}}} = \frac{\Gamma(\rho+\nu+\mathbf{m})}{\Gamma(\rho+\nu)\Gamma(\rho-\nu+\mathbf{m})}.$$

Remark 4.5. Via *Stirlings formula*, we can estimate

$$d_{\mathbf{m}}(-\nu) \asymp_{n,\nu} (1+\mathbf{m})^{2\nu}. \tag{4.11}$$

where $A \asymp B$ means there is a constant $C>0$ such that $C^{-1}B \le A \le CB$. On the other hand, the norm clearly indicates that $\langle w_{\mathbf{m}_1}, w_{\mathbf{m}_2}\rangle_{\mathcal{H}_{-\nu}} = 0$ for $w_{\mathbf{m}_1} \in \mathcal{W}_{\mathbf{m}_1}$ and $w_{\mathbf{m}_2} \in \mathcal{W}_{\mathbf{m}_2}$. Thus we still have the orthogonal decomposition (4.2):

$$\mathcal{H}_{-\nu} = \bigoplus_{\mathbf{m}} \mathcal{W}_{\mathbf{m}}.$$

Having been introduced the norm (Hilbert structure) $\|\cdot\|_{\mathcal{H}_\nu}$ on $\mathcal{H}_\nu$, we can then discuss the Sobolev structure on it. Let $\mathcal{H}$ be a unitary representation of G. As in [FF03] and other related results, the Laplace-Beltrami operator Δ_G gives unitary representation spaces a Sobolev structure. The *Sobolev space of order* $s \ge 0$ is the Hilbert space $W^s_G(\mathcal{H}) \subset \mathcal{H}$ that is the maximal domain given by the inner product

$$\langle f, g\rangle_{W^s_G(\mathcal{H})} \coloneqq \langle (1+\Delta_G)^s f, g\rangle$$

for $f, g \in \mathcal{H}$. Besides, the space of smooth vectors is given by

$$C^\infty(\mathcal{H}) = \bigcap_{s\ge 0} W^s_G(\mathcal{H}).$$

Denote by $\mathcal{E}'(\mathcal{H}) \coloneqq (C^\infty(\mathcal{H}))'$ its distributional dual. Note that, when $\mathcal{H} = L^2(G/\Gamma)$, $W^s_G(G/\Gamma) \coloneqq W^s_G(L^2(G/\Gamma))$ coincides with the natural Sobolev structure on G/Γ and hence $C^\infty(G/\Gamma)$ is the space of infinite differentiable functions on G/Γ. On the other hand, for $s>0$, the distributional dual of $W^s_G(\mathcal{H})$ is the Sobolev space $W^{-s}_G(\mathcal{H}) = (W^s_G(\mathcal{H}))' \subset \mathcal{E}'(\mathcal{H})$.

For an irreducible unitary representation $\mathcal{H}_{\mathbf{n},\nu}$ of $G = SO(n,1)$, the Sobolev inner product can be computed via (4.8), (4.9): for $f = \sum_{\mathbf{m}} f_{\mathbf{m}}, g = \sum_{\mathbf{m}} g_{\mathbf{m}} \in W^s_G(\mathcal{H}_{\mathbf{n},\nu})$, we have

$$\begin{aligned}\langle f, g\rangle_{W^s_G(\mathcal{H}_{\mathbf{n},\nu})} =& \langle (I+\Delta_G)^s f, g\rangle_{\mathcal{H}_{\mathbf{n},\nu}}\\ =& \sum_{\mathbf{m}} \langle (I+\Delta_G)^s f_{\mathbf{m}}, g\rangle_{\mathcal{H}_{\mathbf{n},\nu}}\\ =& \sum_{\mathbf{m}} (1+c_n(\mathbf{n},\nu)+c_{K_n}(\mathbf{m}))^s \langle f_{\mathbf{m}}, g_{\mathbf{m}}\rangle_{\mathcal{H}_{\mathbf{n},\nu}}.\end{aligned} \tag{4.12}$$

Remark 4.6. Again, (4.12) indicates that $\langle w_{\mathbf{m}_1}, w_{\mathbf{m}_2}\rangle_{W^s_G(\mathcal{H}_{\mathbf{n},\nu})} = 0$ for $w_{\mathbf{m}_1} \in \mathcal{W}_{\mathbf{m}_1}$ and $w_{\mathbf{m}_2} \in \mathcal{W}_{\mathbf{m}_2}$. Thus, we still have the orthogonal decomposition (cf. (4.2)):

$$W^s_G(\mathcal{H}_{\mathbf{n},\nu}) = \bigoplus_{\mathbf{m}} \mathcal{W}_{\mathbf{m}}.$$

It is easy to estimate the coefficients

Lemma 4.7. *Let the notation and assumptions be as above. Then*

$$1+c_n(\mathbf{n},\nu)+c_{K_n}(\mathbf{m})\asymp_{\mathbf{n},\nu}1+\|\mathbf{m}\|_\infty^2$$

where $\|\mathbf{m}\|_\infty$ is the maximal number of $\mathbf{m}=(m_1,\dots,m_k)$.

4.4. **Norms estimate.** In this section, we shall show that certain G-complementary series contains a H-complementary series as a discrete component. The result for $G=SO(3,1)$, $H=SO(2,1)$ has already been shown by [Muk68]. Here we adopt the method as in [Zha15] (or [SV12]), and make a slight generalization. More precisely, the idea in [Zha15] is to estimate the operator norm of the projection with respect to the norms on Hilbert spaces. In the following, instead of thinking about the Hilbert norm, we make a more precise estimate for the Sobolev norm. We include the proofs to keep the paper as self-contained as possible.

Theorem 4.8. *Let $n\geq 3$, $\rho^\flat<\nu<\rho$, $s\geq 0$, $G=SO(n,1)$ and $H=SO(n-1,1)$. Then $(\pi^\flat_{\nu-\frac{1}{2}},W^s_H(\mathcal{H}^\flat_{\nu-\frac{1}{2}}))$ is a direct summand of $(\pi_\nu,W^s_G(\mathcal{H}_\nu))$ restricted to H.*

In the following, we replace π_ν and $\pi^\flat_{\nu-\frac{1}{2}}$ by the unitarily equivalent representations $\pi_{-\nu}$ and $\pi^\flat_{\frac{1}{2}-\nu}$ via (4.1) (or Section 6 [JW77]), for the sake of introducing the restriction map. First of all, we estimate the operator norm of the restriction map $\mathrm{Res}:W^s_G(\mathcal{H}_{-\nu})\to W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})$ for $\nu\in(\rho^\flat,\rho)$. By (4.2), (4.6), (4.12), we have

$$W^s_G(\mathcal{H}_{-\nu})=\bigoplus_{\mathbf{m}}\mathcal{W}_{\mathbf{m}},\quad W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})=\bigoplus_{l}\mathcal{V}_l. \tag{4.13}$$

Via (4.7), for $|l|\leq\mathbf{m}$, we consider the orthogonal projections

$$P_{\mathbf{m},l}:\mathcal{W}_{\mathbf{m}}\to\widetilde{\mathcal{V}}_l,\quad \overline{P}_{\mathbf{m},l}:\mathrm{Res}(\mathcal{W}_{\mathbf{m}})\to\mathcal{V}_l,\quad \mathrm{Res}_{\mathbf{m},l}:=\overline{P}_{\mathbf{m},l}\,\mathrm{Res}:\mathcal{W}_{\mathbf{m}}\to\mathcal{V}_l.$$

Then $\mathrm{Res}=\sum_{\mathbf{m}}\sum_{|l|\leq\mathbf{m}}\mathrm{Res}_{\mathbf{m},l}$. Using the orthogonality (4.13), we can deduce an estimate for the operator norms via an elementary argument:

Lemma 4.9. *The operator norm $\|\cdot\|_{\mathrm{op}}$ of $\mathrm{Res}:W^s_G(\mathcal{H}_{-\nu})\to W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})$ satisfies*

$$\|\mathrm{Res}\|^2_{\mathrm{op}}=\sup_l\sum_{\mathbf{m}\geq|l|}\|\mathrm{Res}_{\mathbf{m},l}\|^2_{\mathrm{op}}.$$

Proof. This is exactly Lemma 3.2 [Zha15]. Recall Lemma 4.2 and fix arbitrarily $w = \sum_{\mathbf{m}} \sum_{\mathbf{m} \geq |l|} P_{\mathbf{m},l} w_{\mathbf{m}} \in W^s_G(\mathcal{H}_{-\nu})$. Then by *Cauchy-Schwarz inequality*, we have

$$\begin{aligned}
\| \operatorname{Res} w\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} &= \sum_l \| \sum_{\mathbf{m} \geq |l|} \operatorname{Res}_{\mathbf{m},l} P_{\mathbf{m},l} w\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} \\
&\leq \sum_l \left(\sum_{\mathbf{m} \geq |l|} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}} \| P_{\mathbf{m},l} w\|_{W^s_G(\mathcal{H}_{-\nu})} \right)^2 \\
&\leq \sum_l \left(\sum_{\mathbf{m} \geq |l|} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}} \right) \left(\sum_{\mathbf{m} \geq |l|} \| P_{\mathbf{m},l} w\|^2_{W^s_G(\mathcal{H}_{-\nu})} \right) \\
&= \left(\sup_l \sum_{\mathbf{m} \geq |l|} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}} \right) \cdot \|w\|^2_{W^s_G(\mathcal{H}_{-\nu})}.
\end{aligned}$$

On the other hand, since Res is H- (or $K^\flat$-) equivariant, so is Res^*. But then each $\operatorname{Res}^*_{\mathbf{m},l}$ is a scalar constant of an isometry operator by *Schur's lemma*. Thus, for any $v \in \mathcal{V}_l$, we have

$$\begin{aligned}
\| \operatorname{Res}^* v\|^2_{W^s_G(\mathcal{H}_{-\nu})} &= \| \sum_{|l| \leq \mathbf{m}} \operatorname{Res}^*_{\mathbf{m},l} v\|^2_{W^s_G(\mathcal{H}_{-\nu})} = \sum_{|l| \leq \mathbf{m}} \| \operatorname{Res}^*_{\mathbf{m},l} v\|^2_{W^s_G(\mathcal{H}_{-\nu})} \\
&= \sum_{|l| \leq \mathbf{m}} \| \operatorname{Res}^*_{\mathbf{m},l} \|^2_{\mathrm{op}} \|v\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} = \left(\sum_{|l| \leq \mathbf{m}} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}} \right) \cdot \|v\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})}.
\end{aligned}$$

The consequence follows. ◻

Thus, we want to estimate the operator norm of $\operatorname{Res}_{\mathbf{m},l}$. With the help of the harmonic analysis, we can obtain the operator norm in L^2-sense.

Lemma 4.10 (Proposition 3.4 [Zha15])**.** *Let the notation and assumptions be as above. Then for $\mathbf{m} - l$ even, the $(L^2(S^{n-1}), L^2(S^{n-2}))$-norm of $\operatorname{Res}_{\mathbf{m},l} : \mathcal{W}_{\mathbf{m}} \to \mathcal{V}_l$ is given by*

$$\| \operatorname{Res}_{\mathbf{m},l} \|^2_{L^2} = \frac{(2\mathbf{m}+n-2)\Gamma(\frac{n}{2})\Gamma(\frac{n+\mathbf{m}+l-2}{2})\Gamma(\frac{\mathbf{m}-l+1}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\Gamma(\frac{\mathbf{m}-l+2}{2})\Gamma(\frac{n+\mathbf{m}+l-1}{2})} \asymp_{n,\nu} \frac{\mathbf{m}+1}{(\mathbf{m}+l+1)^{\frac{1}{2}}(\mathbf{m}-l+1)^{\frac{1}{2}}}.$$

Then by Theorem 4.4 (and (4.11)), the $(\mathcal{H}_{-\nu}, \mathcal{H}^\flat_{\frac{1}{2}-\nu})$-norm of $\operatorname{Res}_{\mathbf{m},l} : \mathcal{W}_{\mathbf{m}} \to \mathcal{V}_l$ is given by

$$\| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathcal{H}} = \frac{d^\flat_l(\frac{1}{2}-\nu)}{d_{\mathbf{m}}(-\nu)} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{L^2} \asymp_{n,\nu} \frac{(1+l)^{2\nu-1}}{(1+\mathbf{m})^{2\nu}} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{L^2}$$

where $d^\flat_l(\frac{1}{2}-\nu) = \frac{(\rho-\frac{1}{2}+\nu)_l}{(\rho+\frac{1}{2}-\nu)_l}$ denotes the coefficients in (4.10) for $\mathcal{H}^\flat_{\frac{1}{2}-\nu}$. Finally, by (4.12) (and Lemma 4.7), the $(W^s_G(\mathcal{H}_{-\nu}), W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu}))$-norm of $\operatorname{Res}_{\mathbf{m},l} : \mathcal{W}_{\mathbf{m}} \to \mathcal{V}_l$ is given by

$$\| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}^2 \asymp_{n,\nu} \frac{(1+l)^{2s}}{(1+\mathbf{m})^{2s}} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathcal{H}}^2.$$

Thus, we conclude

Proposition 4.11. *There is a constant $C = C(n, \nu, s) > 1$ such that*

$$C^{-1} \le \inf_l \sum_{\mathbf{m} \ge |l|} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}^2 \le \sup_l \sum_{\mathbf{m} \ge |l|} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}^2 \le C$$

Proof. By Lemma 4.2 and the above estimates, for $\nu \in (\rho^\flat, \rho)$, $2k = \mathbf{m} - l$, we have

$$\begin{aligned}
\sum_{\mathbf{m} \ge |l|} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}^2 &= \sum_{\substack{\mathbf{m} \ge |l| \\ \mathbf{m}-l \text{ even}}} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}^2 \\
&\asymp_{n,\nu} \sum_{\substack{\mathbf{m} \ge |l| \\ \mathbf{m}-l \text{ even}}} \frac{(1+l)^{2s}}{(1+\mathbf{m})^{2s}} \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathcal{H}}^2 \\
&\asymp_{n,\nu} \sum_{\substack{\mathbf{m} \ge |l| \\ \mathbf{m}-l \text{ even}}} \frac{(1+l)^{2\nu+2s-1}}{(1+\mathbf{m})^{2\nu+2s}} \| \operatorname{Res}_{\mathbf{m},l} \|_{L^2}^2 \\
&\asymp_{n,\nu} \sum_{\substack{\mathbf{m} \ge |l| \\ \mathbf{m}-l \text{ even}}} \frac{(1+l)^{2\nu+2s-1}}{(1+\mathbf{m})^{2\nu+2s}} \frac{\mathbf{m}+1}{(\mathbf{m}+l+1)^{\frac{1}{2}}(\mathbf{m}-l+1)^{\frac{1}{2}}} \\
&= \sum_{k \ge 0} \frac{(1+l)^{2\nu+2s-1}}{(1+l+2k)^{2\nu+2s}} \frac{1+l+2k}{(1+2l+2k)^{\frac{1}{2}}(1+2k)^{\frac{1}{2}}}. \qquad (4.14)
\end{aligned}$$

It remains to show that (4.14) is controlled by constants independent of $l \ge 0$. This can be done by the standard integral test. More precisely, the series is controlled by the first term

$$\frac{(1+l)^{2\nu+2s-1}}{(1+l)^{2\nu+2s}} \frac{1+l}{(1+2l)^{\frac{1}{2}}} = \frac{1}{\sqrt{2l+1}}$$

and the integral

$$\int_0^\infty \frac{(1+l)^{2\nu+2s-1}}{(1+l+2k)^{2\nu+2s}} \frac{1+l+2k}{(1+2l+2k)^{\frac{1}{2}}(1+2k)^{\frac{1}{2}}} dk.$$

The first term is bounded above by a constant independent of l. For the integral, we change the variable from k to xl and obtain

$$\begin{aligned}
&\overset{k=xl}{=\!=} \int_0^\infty \frac{(1+l)^{2\nu+2s-1}}{(1+l+2lx)^{2\nu+2s}} \frac{1+l+2lx}{(1+2l+2lx)^{\frac{1}{2}}(1+2lx)^{\frac{1}{2}}} l dx \\
&\asymp_{n,\nu,s} \int_0^\infty \frac{1}{(1+2x)^{2\nu+2s-1}} \frac{1}{\sqrt{(2+2x) \cdot 2x}} dx.
\end{aligned}$$

Since $\nu > \rho^\flat \geq \frac{1}{2}$, the latter integral is finite (and independent of l). This is already enough for Proposition 4.11. $\square$

Finally, using the *open mapping theorem*, we are able to prove Theorem 4.8:

Proof of Theorem 4.8. By (4.1), we can replace π_ν and $\pi^\flat_{\nu-\frac{1}{2}}$ by the unitarily equivalent representations $\pi_{-\nu}$ and $\pi^\flat_{\frac{1}{2}-\nu}$. Since $\sup_l \sum_{\mathbf{m}\geq|l|} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}}$ is bounded above and equal to the operator norm of Res by Lemma 4.9, we conclude that the restriction map $\operatorname{Res} : W^s_G(\mathcal{H}_{-\nu}) \to W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})$ is continuous.

On the other hand, since $\inf_l \sum_{\mathbf{m}\geq|l|} \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}}$ is bounded below, we may then deduce that $\operatorname{Res} : W^s_G(\mathcal{H}_{-\nu}) \to W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})$ is surjective. More precisely, assume that $v = \sum_l v_l \in W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})$. Then fix l, and we conclude from Lemma 4.2 that $\operatorname{Res}_{\mathbf{m},l}|_{\widetilde{\mathcal{V}}_l}$ are isomorphisms for all $\mathbf{m} \geq |l|$ and $\mathbf{m}-l$ even. For simplicity, the sums $\sum$ are all over $\mathbf{m}$ with $\mathbf{m} \geq |l|$ and $\mathbf{m}-l$ even. Let $u_l = \sum c^{\mathbf{m}}_l u^{\mathbf{m}}_l$ where

$$u^{\mathbf{m}}_l := (\operatorname{Res}_{\mathbf{m},l}|_{\widetilde{\mathcal{V}}_l})^{-1} v_l, \quad c^{\mathbf{m}}_l := \frac{\|v_l\|_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})}}{\sum \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}}} \cdot \frac{\| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}}{\|u^{\mathbf{m}}_l\|_{W^s_G(\mathcal{H}_{-\nu})}}.$$

Then

$$\operatorname{Res} u_l = \sum \operatorname{Res}_{\mathbf{m},l} c^{\mathbf{m}}_l u^{\mathbf{m}}_l = \sum c^{\mathbf{m}}_l v_l. \tag{4.15}$$

Since $\operatorname{Res}_{\mathbf{m},l}|_{\widetilde{\mathcal{V}}_l}$ is a $K^\flat$-equivariant isomorphism, by *Schur's lemma*, we know that $\operatorname{Res}_{\mathbf{m},l}|_{\widetilde{\mathcal{V}}_l}$ is a scalar constant of an isometry operator. Then we get

$$\begin{aligned}
\| \operatorname{Res} u_l\|_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} &= \sum \| \operatorname{Res}_{\mathbf{m},l} c^{\mathbf{m}}_l u^{\mathbf{m}}_l\|_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} \\
&= \sum \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}} \|c^{\mathbf{m}}_l u^{\mathbf{m}}_l\|_{W^s_G(\mathcal{H}_{-\nu})} \\
&= \sum \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}} \cdot \frac{\|v_l\|_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})}}{\sum \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}}} \cdot \frac{\| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}}}{\|u^{\mathbf{m}}_l\|_{W^s_G(\mathcal{H}_{-\nu})}} \cdot \|u^{\mathbf{m}}_l\|_{W^s_G(\mathcal{H}_{-\nu})} \\
&= \|v_l\|_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})}. \qquad (4.16)
\end{aligned}$$

Thus, $\operatorname{Res} u_l = v_l$ by (4.15) and (4.16). On the other hand, one has

$$\begin{aligned}
\| \operatorname{Res} u_l\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} &= \Big(\sum \| \operatorname{Res}_{\mathbf{m},l} \|_{\mathrm{op}} \|c^{\mathbf{m}}_l u^{\mathbf{m}}_l\|_{W^s_G(\mathcal{H}_{-\nu})}\Big)^2 \\
&= \sum \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}} \sum \|c^{\mathbf{m}}_l u^{\mathbf{m}}_l\|^2_{W^s_G(\mathcal{H}_{-\nu})} \\
&= \sum \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}} \|u_l\|^2_{W^s_G(\mathcal{H}_{-\nu})} \\
&\geq \inf_l \Big(\sum \| \operatorname{Res}_{\mathbf{m},l} \|^2_{\mathrm{op}}\Big) \|u_l\|^2_{W^s_G(\mathcal{H}_{-\nu})}. \qquad (4.17)
\end{aligned}$$

Combining (4.17) with (4.16), we get

$$\begin{aligned}\|v\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} = \sum_l v_l \|v_l\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} &= \sum_l \|\operatorname{Res} u_l\|^2_{W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu})} \\ &\geq \inf_l \left(\sum \|\operatorname{Res}_{\mathbf{m},l}\|^2_{\mathrm{op}}\right) \sum_l \|u_l\|^2_{W^s_G(\mathcal{H}_{-\nu})}.\end{aligned}$$

Thus, $u := \sum_l u_l \in W^s_G(\mathcal{H}_{-\nu})$ is well defined and satisfies $\operatorname{Res} u = v$, which proves the surjectivity.

Now the orthogonal decomposition implies $W^s_G(\mathcal{H}_{-\nu}) = \ker(\operatorname{Res}) \oplus \ker(\operatorname{Res})^\perp$. It induces a continuous H-equivariant bijection

$$\ker(\operatorname{Res})^\perp \to W^s_H(\mathcal{H}^\flat_{\frac{1}{2}-\nu}).$$

The consequence then follows from the *open mapping theorem.* □

5. Effective estimates for ergodic averages

In Section 4.2, we see that it is possible to find a cocompact lattice $\Gamma \subset G$ such that $L^2(G/\Gamma)$ contains a complementary series $\mathcal{H}_*$ of $G = SO(n,1)$ with spectral parameter $\tilde{\nu} \in (\rho_{n-1}, \rho_n)$ as a direct summand. We write the orthogonal decomposition

$$\mathcal{H} = \mathcal{H}_* \oplus \mathcal{H}_*^\perp.$$

Let $H = SO(2,1)$. When we study the H-action on $\mathcal{H}_*$, by repeatedly using Theorem 4.8, there is an H-complementary series $\mathcal{H}_\nu$ with $\nu = \tilde{\nu} - \rho_{n-1} \in (0, \frac{1}{2})$ such that for any $r \geq 0$, we have

$$W^r_G(\mathcal{H}_*) = W^{r,\nu}_1 \oplus W^{r,\nu}_2.$$

where the restriction map $\operatorname{Res} : W^{r,\nu}_1 \to W^r_H(\mathcal{H}_\nu)$ is H-equivariant isomorphism. Then, for $r \in \mathbf{R}$, we further have the following decomposition

$$W^r_G(G/\Gamma) = W^{r,\nu}_1 \oplus W^{r,\nu}_2 \oplus W^r_G(\mathcal{H}_*^\perp). \tag{5.1}$$

Remark 5.1. It needs not be true that $W^{r,\nu}_1 \subset W^{t,\nu}_1$ for $r > t$.

Later, we want to make sure that some specific elements in $W^r(G/\Gamma)$ are bounded on G/Γ. As G/Γ is compact, we only need to verify that they are continuous, which can be done by *Sobolev embedding theorem*.

Lemma 5.2 (Sobolev embedding theorem)**.** *For $r > r_0 := \dim(G/\Gamma)/2$, there is a constant $C = C(G/\Gamma) > 0$ such that*

$$|f(x)| < C\|f\|_{W^r_G}$$

for any $f \in W^r_G(G/\Gamma)$ and $x \in G/\Gamma$.

Proof. This is the standard Sobolev embedding theorem, e.g. [Aub82]. □

5.1. **Spectral decomposition of unipotent orbits.** Assume that G/Γ is compact. Recall from Section 4.3 that we let $C^\infty(G/\Gamma)$ be the space of infinite differentiable functions on G/Γ, and $\mathcal{E}'(G/\Gamma) = (C^\infty(G/\Gamma))'$ be its distributional dual. On the other hand, recall from Section 4.1 that we choose $\mathfrak{a} = \mathbf{R}Y_n \subset \mathfrak{g}$. Recall that we fix a nilpotent $U \in \mathfrak{g}^\flat_{-1}$. Thus, U defines a unipotent flow $\phi^U_t(x) \coloneqq \exp(tU)x$ on G/Γ and satisfies

$$[Y_n, U] = -U.$$

In this section, we want to study the ergodic average

$$S_{x,T}(f) \coloneqq \frac{1}{T}\int_0^T f(\phi^U_t(x))dt \tag{5.2}$$

of unipotent flows for functions $f \in W^{r_0,\nu}_1 \subset L^2(G/\Gamma)$. The proof relies on the characterization of the space of *invariant distributions* for unipotent flows. Specifically, we make use of the argument in [FF03]. See also [Mie06], [Ram13], [Wan15] for related discussions.

The space of U*-invariant distributions* for a given H-unitary representation $\mathcal{H}$ is then defined by

$$\mathcal{I}_U(\mathcal{H}) \coloneqq \{\mathcal{D} \in \mathcal{E}'(\mathcal{H}) : \mathcal{L}_U\mathcal{D} = 0\}.$$

Similarly, we define The space of U*-invariant distributions of order* s to be

$$\mathcal{I}^r_U(\mathcal{H}) \coloneqq \{\mathcal{D} \in W^{-r}_H(\mathcal{H}) : \mathcal{L}_U\mathcal{D} = 0\}.$$

Clearly, the necessary condition for $g \in W^r(\mathcal{H})$ having the form $g = Uf$ for some $f \in W^{r+1}_H(\mathcal{H})$ is $g \in \ker \mathcal{I}^r_U(\mathcal{H}) = \{g \in \mathcal{H} : \mathcal{D}(g) = 0 \text{ for any } \mathcal{D} \in \mathcal{I}^r_U(\mathcal{H})\}$, since

$$\mathcal{D}(g) = \mathcal{D}(Uf) = -\mathcal{L}_U\mathcal{D}(f) = 0$$

for any $\mathcal{D} \in \mathcal{I}^r_U(\mathcal{H})$. On the other hand, Flaminio and Forni [FF03] have characterized the spaces of U-invariant distributions for all $SO(2,1)$-irreducible unitary representations, and shown that they are the **only** obstructions to the existence of smooth solutions of the cohomological equation $Uf = g$. Here we need the results for the complementary series:

Theorem 5.3 ($SO(2,1)$-complementary series, [FF03])**.** *For $\nu \in (0, \frac{1}{2})$, let $(\mathcal{H}_\nu, \pi_\nu)$ be a complementary series of $H = SO(2,1)$. Then the space $\mathcal{I}_U(\mathcal{H}_\nu)$ has dimension 2 and it is generated by two Y_n-eigenvectors $\mathcal{D}^\pm_\nu$ of eigenvalues $-(1 \pm 2\nu)/2$ and Sobolev order $(1 \pm 2\nu)/2$, respectively. In other words*

$$\mathcal{L}_{Y_n}\mathcal{D}^\pm_\nu = -\frac{1 \pm 2\nu}{2}\mathcal{D}^\pm_\nu, \quad \mathcal{D}^\pm_\nu \in W_H^{-\frac{1\pm 2\nu}{2}}(\mathcal{H}_\nu).$$

Besides, let $s > (1+2\nu)/2$ and $t < s-1$. Then there is a constant $C(\nu, s, t) > 0$ such that, for all $g \in \ker \mathcal{I}^s_U(\mathcal{H}_\nu)$, the cohomological equation has a solution $f \in W^t_H(\mathcal{H}_\nu)$ which satisfies the Sobolev estimate

$$\|f\|_{W^t_H(\mathcal{H}_\nu)} \le C(\nu, s, t)\|g\|_{W^s_H(\mathcal{H}_\nu)}.$$

On the other hand, let $g \in W^s(\mathcal{H}_\nu)$, $s > (1+2\nu)/2$. If the equation $Uf = g$ has a solution $f \in W^t(\mathcal{H}_\nu)$ with $t \geq (2\nu-1)/2$, then $\mathcal{D}^{\pm}_\nu(g) = 0$.

Thus, for $r_0 := \dim(G/\Gamma)/2 > 1$ and H-complementary series $\mathcal{H}_\nu$, $\mathcal{I}^{r_0}_U(\mathcal{H}_\nu) \subset W_H^{-r_0}(\mathcal{H}_\nu)$ is closed. Then the orthogonal decomposition is of the form

$$W_H^{-r_0}(\mathcal{H}_\nu) = \mathcal{I}^{r_0}_U(\mathcal{H}_\nu) \oplus \mathcal{I}^{r_0}_U(\mathcal{H}_\nu)^{\perp}. \tag{5.3}$$

Combining (5.3) with (5.1), we get

$$W_G^{-r_0}(G/\Gamma) = (\mathcal{I}^{r_0}_1 \oplus \mathcal{I}^{r_0}_2) \oplus W_2^{-r_0,\nu} \oplus W_G^{-r_0}(\mathcal{H}^{\perp}_*). \tag{5.4}$$

where $\mathcal{I}^{r_0}_1 := \mathrm{Res}^{-1}\,\mathcal{I}^{r_0}_U(\mathcal{H}_\nu)$, $\mathcal{I}^{r_0}_2 := \mathrm{Res}^{-1}\,\mathcal{I}^{r_0}_U(\mathcal{H}_\nu)^{\perp}$.

Remark 5.4. The spaces $\mathcal{I}^{r_0}_1$, $W_2^{-r_0,\nu}$, $W_G^{-r_0}(\mathcal{H}^{\perp}_*)$ are $\phi^{Y_n}_t$-invariant. However, $\mathcal{I}^{r_0}_2$ is not $\phi^{Y_n}_t$-invariant.

According to the previous results, ϕ^X_t has a spectral decomposition on the space $\mathcal{I}^{r_0}_U(\mathcal{H}_\nu)$. More precisely, for all $t \in \mathbf{R}$, we have

$$\phi^{Y_n}_t(\mathcal{D}^{\pm}_\nu) = e^{-\frac{1\pm 2\nu}{2}t}\mathcal{D}^{\pm}_\nu.$$

Now we consider the ergodic average $S_{x,T}$ defined in (5.2) as a distribution in $W_G^{-r_0}(G/\Gamma)$. By (5.4), we can write

$$S_{x,T} = c_+(x,T)\mathcal{D}^+_\nu + c_-(x,T)\mathcal{D}^-_\nu + \mathcal{R}(x,T) + \mathcal{C}(x,T) \tag{5.5}$$

where $\mathcal{D}^{\pm}_\nu \in \mathcal{I}^{r_0}_1$, $\mathcal{R}(x,T) \in \mathcal{I}^{r_0}_2$, $\mathcal{C}(x,T) \in W_2^{-r_0,\nu} \oplus W_G^{-r_0}(\mathcal{H}^{\perp}_*)$.

Remark 5.5. The distributions $\mathcal{D}^{\pm}_\nu \in \mathcal{I}^{r_0}_1$ should more appropriately be written as $\mathcal{D}^{\pm}_\nu \circ \mathrm{Res}$, as $\mathcal{D}^{\pm}_\nu \in W_H^{-\frac{1\pm 2\nu}{2}}(\mathcal{H}_\nu)$ has already been given in Theorem 5.3. We abuse notation if it makes no confusion.

Thus, we can analyze the ergodic average $S_{x,T}$ via the distributions. The method has already been used to study the ergodic averages for horocycle flows in [FF03]. We adopt the same strategy here and provide proofs for the sake of completeness.

Remark 5.6. It is possible to obtain a more explicit decomposition than (5.5). For instance, [Muk68] provides the full decomposition of the complementary series $\mathcal{H}_\nu$ of $G = SO(3,1)$ under $H = SO(2,1)$. If $0 < \nu \leq \frac{1}{2}$, it is a sum of two direct integrals of spherical principal series, and if $\frac{1}{2} < \nu < 1$, it contains one extra discrete component, the complementary series $\mathcal{H}_{\nu-\frac{1}{2}}$, as we have shown. However, the relations of Sobolev structures on the principal series are not quite clear. Thus, it seems that we cannot apply the Flaminio-Forni argument to get further information.

In the following, we want to apply Theorem 5.3 with different Sobolev orders, and hence the space $W_1^{r,\nu}$ is no longer convenient, as indicated in Remark 5.1. Thus, for $s \geq 0$, we introduce

$$W_1^{s,r_0,\nu} := \mathrm{Res}^{-1}(W_H^{r_0+s}(\mathcal{H}_\nu)).$$

In particular, $W_1^{0,r_0,\nu} = W_1^{r_0,\nu}$. Then we have $W_1^{s,r_0,\nu} \subset W_1^{t,r_0,\nu}$ whenever $s > t$. As in (5.4), we have

$$W_1^{-s,-r_0,\nu} = \mathcal{I}_1^{r_0+s} \oplus \mathcal{I}_2^{r_0+s} \tag{5.6}$$

where $\mathcal{I}_1^{r_0+s} := \operatorname{Res}^{-1} \mathcal{I}_U^{r_0+s}(\mathcal{H}_\nu)$, $\mathcal{I}_2^{r_0+s} := \operatorname{Res}^{-1} \mathcal{I}_U^{r_0+s}(\mathcal{H}_\nu)^\perp$.

Next, we collect some basic results with respect to the decomposition (5.5). First, we observe that the norms of $S_{x,T}$ in $W_1^{-s,-r_0,\nu} \cong W_H^{-r_0-s}(\mathcal{H}_\nu)$ are equivalent to their coefficients in the decomposition.

Lemma 5.7 ($W_1^{-s,-r_0,\nu}$-norm estimates)**.** *For $s \geq 0$, we have*

$$|c_+(x,T)|^2 + |c_-(x,T)|^2 + \|\mathcal{R}(x,T)\|^2_{W_1^{-s,-r_0,\nu}} \asymp_s \|S_{x,T}\|^2_{W_1^{-s,-r_0,\nu}}.$$

Proof. It follows directly from the orthogonal decomposition, and the fact that $\{\mathcal{D}_\nu^\pm\} \subset \mathcal{I}_U^s(\mathcal{H}_\nu)$ is a basis by Theorem 5.3. □

Combining Lemma 5.7 with Sobolev embedding theorem (Lemma 5.2), we obtain a uniform upper bound for the coefficients:

Corollary 5.8. *For $s \geq 0$, there exists a constant $C = C(s) > 0$ such that*

$$|c_+(x,T)|^2 + |c_-(x,T)|^2 + \|\mathcal{R}(x,T)\|^2_{W_1^{-s,-r_0,\nu}} \leq C$$

for all x, T.

Proof. Note that $|S_{x,T}(f)| \leq \max_{x\in G/\Gamma} |f(x)|$ for any $f \in W_1^{s,r_0,\nu} \subset W_1^{r_0,\nu}$. □

5.2. **Estimates for coefficients via Gottschalk-Hedlund.** Based on the study of the cohomological equation $Uf = g$, we can obtain a better bound for $\mathcal{R}(x,T)$. Recall that the restriction map $\operatorname{Res} : W_1^{s,r_0,\nu} \to W_H^{r_0+s}(\mathcal{H}_\nu)$ is H-equivariant. Thus, the cohomological equation $Uf = g$ on $W_H^{r_0+s}(\mathcal{H}_\nu)$ is equivalent to $\operatorname{Res}(Uf) = U\operatorname{Res}(f) = \operatorname{Res}(g)$ on $W_1^{s,r_0,\nu}$.

Lemma 5.9 (Pointwise bound for $\mathcal{R}(x,T)$)**.** *For $s > 1$, there exists a constant $C = C(\nu, s) > 0$ such that*

$$\|\mathcal{R}(x,T)\|_{W_1^{-s,-r_0,\nu}} \leq CT^{-1}.$$

Proof. The orthogonal decomposition (5.6) implies

$$W_1^{s,r_0,\nu} = \ker(\mathcal{I}_1^{r_0+s}) \oplus \ker(\mathcal{I}_2^{r_0+s}). \tag{5.7}$$

Then, for any $g \in W_1^{s,r_0,\nu}$, there is a unique orthogonal decomposition $g = g_1 + g_2$ where $g_1 \in \ker(\mathcal{I}_1^{r_0+s})$ and $g_2 \in \ker(\mathcal{I}_2^{r_0+s})$. Since $\mathcal{R}(x,T) \in \mathcal{I}_2^{r_0+s}$, we have

$$\mathcal{R}(x,T)(g) = \mathcal{R}(x,T)(g_1 + g_2) = \mathcal{R}(x,T)(g_1) = S_{x,T}(g_1). \tag{5.8}$$

Now since $g_1 \in \ker(\mathcal{I}_1^{r_0+s})$, by Theorem 5.3, there exists a function $f_1 \in W_1^{t,r_0,\nu}$ with $t \in (0, s-1)$, such that $Uf_1 = g_1$ and

$$\|f_1\|_{W_1^{t,r_0,\nu}} \ll_{\nu,s,t} \|g_1\|_{W_1^{s,r_0,\nu}}.$$

By the Sobolev embedding theorem (Lemma 5.2), we conclude that

$$\max_{x\in G/\Gamma}|f_1(x)| \ll \|f_1\|_{W_1^{t,r_0,\nu}} \ll_{\nu,s,t} \|g_1\|_{W_1^{s,r_0,\nu}}.$$

It follows that

$$|S_{x,T}(g_1)| = \frac{1}{T}|f_1\circ\phi_T^U(x) - f_1(x)| \ll_{\nu,s,t} \frac{1}{T}\|g_1\|_{W_1^{s,r_0,\nu}}. \tag{5.9}$$

Therefore, using (5.8), (5.9), we make an appropriate choice of $t(s)\in(0,s-1)$ and then there exists $C=C(\nu,s)>0$ such that

$$|\mathcal{R}(x,T)(g)| \ll_{\nu,s,t(s)} \frac{1}{T}\|g_1\|_{W_1^{s,r_0,\nu}} \le \frac{C}{T}\|g\|_{W_1^{s,r_0,\nu}}.$$

The consequence follows. □

We also need a L^2-bound for $\mathcal{R}(x,T)$ in order to get the lower bound for ergodic averages. The proof for the L^2-bound (Lemma 5.10) is completely similar to the pointwise bound (Lemma 5.9).

Lemma 5.10 (L^2-bound for $\mathcal{R}(x,T)$). *For $s>1$, there exists a constant $C=C(\nu,s)>0$ such that for $g\in W_1^{s,r_0,\nu}$, we have*

$$\|\mathcal{R}(\cdot,T)(g)\|_{L^2(G/\Gamma)} \le \frac{C}{T}\|g\|_{W_1^{s,r_0,\nu}}.$$

Proof. As in the proof of Lemma 5.9, we write

$$g = g_1+g_2 \in \ker(\mathcal{I}_1^{r_0+s})\oplus\ker(\mathcal{I}_2^{r_0+s}).$$

Then we have $Uf_1=g_1$ and

$$\begin{aligned}\|\mathcal{R}(\cdot,T)(g)\|_{L^2} =&\|S_{\cdot,T}(g_1)\|_{L^2} \le \frac{2}{T}\|f_1\|_{L^2}\\ \le&\frac{2}{T}\max_{x\in G/\Gamma}|f_1(x)| \ll_{\nu,s,t(s)} \|g_1\|_{W_1^{s,r_0,\nu}} \le \frac{C}{T}\|g\|_{W_1^{s,r_0,\nu}}.\end{aligned}$$

This proves Lemma 5.10. □

The following Gottschalk-Hedlund theorem is a useful criterion for L^2-solutions for the cohomological equation for ergodic measurable flows ϕ_t.

Lemma 5.11 (Gottschalk-Hedlund). *If an L^2-function f is a solution of the equation*

$$\left.\frac{df\circ\phi_t}{dt}\right|_{t=0} = g \tag{5.10}$$

then the one-parameter family of functions G_T defined by

$$G_T(x) \coloneqq \int_0^T g(\phi_t(x))dt$$

is equibounded in L^2 by $2\|f\|$. Conversely, if the family G_T is equibounded, then the cohomological equation has an L^2-solution.

Proof. The L^2-norm of G_T is clearly bounded by $2\|f\|$ if f is a solution of (5.10). On the other hand, if the family of functions $\{G_T\}_{T\geq 0}$ is equibounded in L^2, then the family of functions $\{f_T\}_{T\geq 0}$ defined by

$$f_T(x) := -\frac{1}{T}\int_0^T \int_0^t G(\phi_s(x))dsdt$$

is equibounded in L^2. Then by ergodic theorem, G has zero ergodic average, and any weak limit $f \in L^2$ of $\{f_T\}_{T\geq 0}$ is a L^2-solution of (5.10). $\square$

The following results provide an important information about L^2-bounds for the ergodic averages.

In the following we shall use *Hahn-Banach theorem* to construct functions dual to $\mathcal{D}_\nu^\pm$ in order to estimate the coefficients. More precisely, there is a 1-dimensional space $(\mathcal{D}_\nu^+)' \subset W_1^{s,r_0,\nu}$ such that $g \in (\mathcal{D}_\nu^+)'$ satisfies

$$\mathcal{D}_\nu^+(g) \neq 0, \quad \mathcal{D}_\nu^-(g) = \mathcal{R}(x,T)(g) = \mathcal{C}(x,T)(g) = 0$$

and $(\mathcal{D}_\nu^-)'$ can be similarly defined.

Lemma 5.12 (L^2-bound for $c_\pm(x,T)$)**.** *For $s > 1$, $\mathcal{D}_\nu^\pm \in \mathcal{I}_1^{r_0+s}$, there exists a constant $C(\mathcal{D}_\nu^\pm) > 0$ such that*

$$\|c_\pm(\cdot,T)\|_{L^2} \leq C(\mathcal{D}_\nu^\pm). \tag{5.11}$$

On the other hand, $c_\pm(x,T)$ satisfies the L^2 lower bound

$$\sup_{T\in\mathbf{R}^+} T\|c_\pm(\cdot,T)\|_{L^2} = \infty. \tag{5.12}$$

Moreover, if $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$, and $g \in (D_\nu^\pm)'$ such that the equation

$$g - \phi_\lambda^Z g = Uf \tag{5.13}$$

has no L^2-solutions f, then we have

$$\sup_{T\in\mathbf{R}^+} T\|c_\pm(\cdot,T) - \phi_\lambda^Z c_\pm(\cdot,T)\|_{L^2} = \infty. \tag{5.14}$$

Proof. We only consider the coefficient $c_+(x,T)$. Then there is a unique function $g \in (D_\nu^+)' \subset W_1^{s,r_0,\nu}$ (cf. (5.7)) such that

$$D_\nu^+(g) = 1, \quad D_\nu^-(g) = \mathcal{R}(x,T)(g) = \mathcal{C}(x,T)(g) = 0$$

for all x, T. It follows that

$$\|c_+(\cdot,T)\|_{L^2} = \|S_{\cdot,T}(g)\|_{L^2} \leq \|g\|_{L^2}.$$

This proves (5.11). On the other hand, since $D_\nu^+(\mathrm{Res}(g)) \neq 0$ (see Remark 5.5), by Theorem 5.3, the equation $U\,\mathrm{Res}(f) = \mathrm{Res}(g)$ has no solutions $\mathrm{Res}(f) \in \mathcal{H}_\nu$. Thus, we conclude that $Uf = g$ does not have L^2-solutions f. Then by Gottschalk-Hedlund theorem (Lemma 5.11), the family of functions

$$Tc_+(x,T) = TS_{x,T}(g) = \int_0^T g(\phi_t^U(x))dt$$

is not equibounded in $L^2(G/\Gamma)$. This proves (5.12). Similarly, if (5.13) has no L^2-solutions f, then the family of functions

$$\begin{aligned} TS_{x,T}(g-\phi^Z_\lambda g) =& \int_0^T g(\phi^U_t(x)) - g(\phi^U_t\phi^Z_\lambda(x))dt \\ =& T(c_+(x,T)-c_+(\phi^Z_\lambda(x),T))\mathcal{D}(g) \end{aligned}$$

is not in L^2-equibounded by Gottschalk-Hedlund theorem again. This proves (5.14). □

5.3. **Estimates for coefficients via geodesic renormalization.** Recall that by the choice of U, Y_n, we have the renormalization

$$\phi^{Y_n}_t \circ \phi^U_s = \phi^U_{se^{-t}} \circ \phi^{Y_n}_t.$$

It follows that

$$\phi^{Y_n}_t(S_{x,T}) = S_{\phi^{Y_n}_{-t}(x),e^tT}.$$

We shall use (5.15) to study the asymptotic behavior of $S_{x,T}$. Recall that in the decomposition (5.4), $\mathcal{I}^s_U(\mathcal{H}_\nu)^\perp$ is not $\phi^{Y_n}_t$-invariant. We need to show that the remainder term $\mathcal{R}(x,T)\in\mathcal{I}^s_U(\mathcal{H}_\nu)^\perp$ is still negligible under Y_n-action.

It is convenient to discretize the geodesic flow. More precisely, fix $\sigma\in[1,2]$, $x\in G/\Gamma$, $T\geq 0$. For any $l\in\mathbf{N}$, we consider

$$\phi^{Y_n}_{l\sigma}(S_{x,T}) = S_{\phi^{Y_n}_{l\sigma}(x),e^{l\sigma}T}. \tag{5.15}$$

Similar to (5.5), the ergodic average $\phi^{Y_n}_{l\sigma}(S_{x,T})$ has the decomposition

$$\phi^{Y_n}_{l\sigma}(S_{x,T}) = c^T_+(x,l)\mathcal{D}^+_\nu + c^T_-(x,l)\mathcal{D}^-_\nu + \mathcal{R}^T(x,l) + \mathcal{C}^T(x,l). \tag{5.16}$$

We prove pointwise and L^2-bounds for the functions $c^T_\pm(x,l)$, $\mathcal{R}^T(x,l)$. By the identity (5.15) and the definition (5.5), we have

$$c^T_\pm(x,l) = c_\pm(\phi^{Y_n}_{l\sigma}(x),e^{l\sigma}T), \quad \mathcal{R}^T(x,l)=\mathcal{R}(\phi^{Y_n}_{l\sigma}(x),e^{l\sigma}T). \tag{5.17}$$

Note that $\mathcal{R}$-component is not $\phi^{Y_n}_t$-invariant, but we still have $\phi^{Y_n}_t\mathcal{R}^T(x,l)\in W_1^{-s,-r_0,\nu}$. Now we estimate the remainder term $\mathcal{R}^T(x,l)$ after pushforward by one geodesic step $\phi^{Y_n}_\sigma$. Let $r^T_\pm(x,l) \coloneqq c_\pm(\phi^{Y_n}_\sigma\mathcal{R}^T(x,l))$ be its $\mathcal{D}^\pm_\nu$-component. Then, similar to (5.5), we have

$$\begin{aligned} \phi^{Y_n}_\sigma\mathcal{R}^T(x,l) =& c_+(\phi^{Y_n}_\sigma\mathcal{R}^T(x,l))\mathcal{D}^+_\nu + c_-(\phi^{Y_n}_\sigma\mathcal{R}^T(x,l))\mathcal{D}^-_\nu + \mathcal{R}(\phi^{Y_n}_\sigma\mathcal{R}^T(x,l)) \\ =& r^T_+(x,l)\mathcal{D}^+_\nu + r^T_-(x,l)\mathcal{D}^-_\nu + \mathcal{R}^T(x,l+1). \end{aligned} \tag{5.18}$$

Moreover, we get

$$c^T_\pm(x,l+1) = c_\pm(\phi^{Y_n}_\sigma(c^T_\pm(x,l)+\mathcal{R}^T(x,l))) = c^T_\pm(x,l)e^{\frac{1\pm2\nu}{2}\sigma} + r^T_\pm(x,l). \tag{5.19}$$

We want to have an effective estimate for the coefficients $c^T_\pm(x,l)$ of $\mathcal{D}$-components. To solve the recurrence relation (5.19), we need the following elementary result.

Lemma 5.13. *Let $A : V \to V$ be a linear map. Let $\{R_l\} \subset V$. The solution x_l of the following difference equation*

$$x_{l+1} = A(x_l) + R_l$$

has the form

$$x_l = A^l(x_0) + \sum_{j=0}^{l-1} A^{l-j-1} R_j.$$

Thus, it remains to estimate the remainder terms $r^T_{\pm}(x, l)$.

Lemma 5.14 (Pointwise bound for $r^T_{\pm}(x, l)$)**.** *For fixed $\sigma \in [1, 2]$, $s > 1$, there exists a constant $C = C(\nu, s) > 0$ such that*

$$|r^T_+(x, l)|^2 + |r^T_-(x, l)|^2 \leq C(e^{l\sigma}T)^{-2}$$

for all $x \in G/\Gamma, T \in \mathbf{R}^+, l \in \mathbf{N}$.

Proof. Let $C(s) \coloneqq \max_{\sigma\in[1,2]} \|\phi^{Y_n}_\sigma\|$, where $\|\cdot\|$ denotes the operator norm. Similar to Lemma 5.7, using (5.18), we have the estimate:

$$|r^T_{\pm}(x, l)|^2 = |c_{\pm}(\phi^{Y_n}_\sigma \mathcal{R}^T(x, l))|^2 \ll_s \|\phi^{Y_n}_\sigma \mathcal{R}^T(x, l)\|^2_{W_1^{-s,-r_0,\nu}} \ll_s C(s)\|\mathcal{R}^T(x, l)\|^2_{W_1^{-s,-r_0,\nu}}.$$

Then by (5.17) and Lemma 5.9, we have

$$\|\mathcal{R}^T(x, l)\|^2_{W_1^{-s,-r_0,\nu}} = \|\mathcal{R}(\phi^{Y_n}_{l\sigma}(x), e^{l\sigma}T)\|^2_{W_1^{-s,-r_0,\nu}} \ll_{\nu,s} (e^{l\sigma}T)^{-2}.$$

The consequence follows. □

As in Lemma 5.12, we can also estimate the L^2-norm of $r^T_{\pm}(x, l)$.

Lemma 5.15 (L^2-bound for $r^T_{\pm}(x, l)$)**.** *For fixed $\sigma \in [1, 2]$, $s > 1$, there exists a constant $C = C(\nu, s) > 0$ such that*

$$\|r^T_{\pm}(\cdot, l)\|_{L^2} \leq C(e^{l\sigma}T)^{-1}$$

for any $T \in \mathbf{R}^+, l \in \mathbf{N}$.

Proof. Again, we choose the function $g \in (\mathcal{D}^+_\nu)' \subset W_1^{s,r_0,\nu}$ such that

$$\mathcal{D}^+_\nu(g) = 1, \quad \mathcal{D}^-_\nu(g) = \mathcal{R}^T(x, l+1)(g) = 0.$$

Recall that by the definition

$$r^T_+(x, l) = \mathcal{R}^T(x, l)(\phi^{Y_n}_{-\sigma} g) = \mathcal{R}(\phi^{Y_n}_{l\sigma}(x), e^{l\sigma}T)(\phi^{Y_n}_{-\sigma} g).$$

Then by Lemma 5.10, we have

$$\|r^T_+(\cdot, l)\|_{L^2} = \|\mathcal{R}^T(\cdot, l)(\phi^{Y_n}_{-\sigma} g)\|_{L^2} \ll_{\nu,s} \|g\|_{W_1^{s,r_0,\nu}}(e^{l\sigma}T)^{-1}.$$

The argument for $r^T_-(x, l)$ is similar. □

Now we are in the position to estimate the coefficients $c^T_{\pm}(x, l)$ of $\mathcal{D}$-components.

Lemma 5.16 (Pointwise bound for $c_\pm^T(x,l)$)**.** *Let $\sigma \in [1,2]$, $s>1$, $T \in \mathbf{R}^+$. Then there exists a constant $C = C(\nu, s, T) > 0$ such that*

$$|c_\pm^T(x,l)| \le Ce^{-\frac{1\pm 2\nu}{2}l\sigma}$$

for all $x \in G/\Gamma, l \in \mathbf{N}$.

Proof. Choose $V = \mathbf{C}$, $A = e^{-\frac{1\pm 2\nu}{2}\sigma}$, $x_l = |c_\pm^T(x,l)|$ and $R_l = |r_\pm^T(x,l)|$. Then by (5.19) and Lemma 5.13, we obtain

$$|c_\pm^T(x,l)| \le |c_\pm^T(x,0)|e^{-\frac{1\pm 2\nu}{2}l\sigma} + \sum_{j=0}^{l-1} |r_\pm^T(x,j)|e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma}. \tag{5.20}$$

By Corollary 5.8, we have

$$|c_\pm^T(x,0)|^2 = |c_\pm(x,T)|^2 \le C(s). \tag{5.21}$$

On the other hand, by Lemma 5.14, we have

$$|r_\pm^T(x,j)|^2 \le C(\nu,s)(e^{j\sigma}T)^{-2}.$$

It follows that

$$\begin{aligned}\sum_{j=0}^{l-1} |r_\pm^T(x,j)|e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} &\le C(\nu,s)\sum_{j=0}^{l-1}(e^{j\sigma}T)^{-1}e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} \\ &= C(\nu,s)e^{\frac{1\pm 2\nu}{2}\sigma}T^{-1}e^{-\frac{1\pm 2\nu}{2}l\sigma}\sum_{j=0}^{l-1} e^{\frac{-1\pm 2\nu}{2}j\sigma} \le C(\nu,s)T^{-1}e^{-\frac{1\pm 2\nu}{2}l\sigma}. \end{aligned} \tag{5.22}$$

Recall that $\nu \in (0,1/2)$. Then by (5.20), (5.21), (5.22), we conclude

$$|c_\pm^T(x,l)| \le C(\nu,s)T^{-1}e^{-\frac{1\pm 2\nu}{2}l\sigma}$$

The proves Lemma 5.16. □

Lemma 5.17 (L^2-bound for $c_\pm^T(x,l)$)**.** *Let $\sigma \in [1,2]$, $s>1$, $T \in \mathbf{R}^+$. Then there exists a constant $C = C(\nu,s,T) > 0$ such that*

$$\|c_\pm^T(\cdot,l)\|_{L^2} \le Ce^{-\frac{1\pm 2\nu}{2}l\sigma} \tag{5.23}$$

for all $x \in G/\Gamma, l \in \mathbf{N}$. On the other hand, there exist $C_0 = C_0(\nu,s) > 0$, $T_0 = T_0(\nu,s) > 0$ such that

$$\|c_\pm^{T_0}(\cdot,l)\|_{L^2} \ge C_0 e^{-\frac{1\pm 2\nu}{2}l\sigma} \tag{5.24}$$

for all $l \in \mathbf{N}$. Further, if $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$, and $g \in (D_\nu^\pm)'$ such that the equation

$$g - \phi_\lambda^Z g = Uf \tag{5.25}$$

has no L^2-solutions f, then there exist $C_1 = C_1(\nu,s) > 0$, $T_1 = T_1(\nu,s) > 0$ such that

$$\|c_\pm^{T_1}(\cdot,l) - \phi_\lambda^Z c_\pm^{T_1}(\cdot,l)\|_{L^2} \ge C_1 e^{-\frac{1\pm 2\nu}{2}l\sigma} \tag{5.26}$$

for all $l \in \mathbf{N}$.

Proof. By (5.19) and Lemma 5.13, we obtain

$$\|c_\pm^T(\cdot,l)\|_{L^2} \leq \|c_\pm^T(\cdot,0)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}l\sigma} + \sum_{j=0}^{l-1} \|r_\pm^T(\cdot,j)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma}. \tag{5.27}$$

Similar to the pointwise upper bound (Lemma 5.16), we can apply Lemma 5.15 (cf. (5.22)), and obtain

$$\sum_{j=0}^{l-1} \|r_\pm^T(\cdot,j)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} \leq C(\nu,s)T^{-1} e^{-\frac{1\pm 2\nu}{2}l\sigma}. \tag{5.28}$$

Combining with Lemma 5.12, we obtain the L^2-upper bound (5.23).

On the other hand, using again (5.19) and Lemma 5.13, and then (5.28), we obtain a lower bound

$$\begin{aligned}\|c_\pm^T(\cdot,l)\|_{L^2} \geq & \|c_\pm^T(\cdot,0)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}l\sigma} - \sum_{j=0}^{l-1} \|r_\pm^T(\cdot,j)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} \\ \geq & (\|c_\pm^T(\cdot,0)\|_{L^2} - C(\nu,s)T^{-1}) e^{-\frac{1\pm 2\nu}{2}l\sigma}.\end{aligned} \tag{5.29}$$

By (5.12), there exists $T_0 = T_0(\nu,s) > 0$ such that

$$T_0\|c_\pm^{T_0}(\cdot,0)\|_{L^2} = T_0\|c_\pm(\cdot,T_0)\|_{L^2} > 2C(\nu,s).$$

It follows that

$$\|c_\pm^{T_0}(\cdot,l)\|_{L^2} \geq C(\nu,s)T_0^{-1} e^{-\frac{1\pm 2\nu}{2}l\sigma}.$$

This proves (5.24).

Finally, using again (5.19) and Lemma 5.13, and then (5.28), we obtain a lower bound

$$\begin{aligned}&\|c_\pm^T(\cdot,l) - \phi_\lambda^Z c_\pm^T(\cdot,l)\|_{L^2} \\ \geq & \|c_\pm^T(\cdot,0) - \phi_\lambda^Z c_\pm^T(\cdot,0)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}l\sigma} - \sum_{j=0}^{l-1} (\|r_\pm^T(\cdot,j)\|_{L^2} + \|\phi_\lambda^Z r_\pm^T(\cdot,j)\|_{L^2}) e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} \\ = & \|c_\pm^T(\cdot,0) - \phi_\lambda^Z c_\pm^T(\cdot,0)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}l\sigma} - \sum_{j=0}^{l-1} 2\|r_\pm^T(\cdot,j)\|_{L^2} e^{-\frac{1\pm 2\nu}{2}(l-j-1)\sigma} \\ \geq & (\|c_\pm^T(\cdot,0) - \phi_\lambda^Z c_\pm^T(\cdot,0)\|_{L^2} - 2C(\nu,s)T^{-1}) e^{-\frac{1\pm 2\nu}{2}l\sigma}.\end{aligned} \tag{5.30}$$

Then, with the assumption (5.25), (5.14) shows that there exists $T_1 = T_1(\nu,s) > 0$ such that

$$T_1\|c_\pm(\cdot,T_1) - \phi_\lambda^Z c_\pm(\cdot,T_1)\|_{L^2} > 3C(\nu,s).$$

Then (5.30) shows

$$\|c_\pm^{T_1}(\cdot,l) - \phi_\lambda^Z c_\pm^{T_1}(\cdot,l)\|_{L^2} \geq C(\nu,s)T_1 e^{-\frac{1\pm 2\nu}{2}l\sigma}.$$

It implies (5.26). □

Now we recover the continuous time by replacing $e^{l\sigma}$ by T. Note that based on our assumption, we have $[1,\infty) \subset \{e^{l\sigma} : l \in \mathbf{N}, \sigma \in [1,2]\}$. Therefore, the above results can be translated to:

Proposition 5.18 (Estimates for $c_\pm(x,T)$)**.** *Let $s > 1$. Then there exists $T_0 = T_0(\nu, s) > 0$ such that*

$$|c_\pm(x,T)| \ll_{\nu,s} T^{-\frac{1\pm\nu}{2}}, \quad \|c_\pm(\cdot,T)\|_{L^2} \gg_{\nu,s} T^{-\frac{1\pm 2\nu}{2}}$$

for all $x \in G/\Gamma$, $T \geq T_0$. Besides, if $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$, and $g \in (D_\nu^\pm)'$ such that the equation

$$g - \phi_\lambda^Z g = Uf$$

has no L^2-solutions f, then

$$\|c_\pm(\cdot,T) - \phi_\lambda^Z c_\pm(\cdot,T)\|_{L^2} \gg_{\nu,s} T^{-\frac{1\pm 2\nu}{2}}$$

for all $T \geq T_0$.

The following corollary, given by an elementary integral argument, is an important criterion for the existence of measurable solutions of the cohomological equation $Uf = g$:

Corollary 5.19. *There exists $T_0 = T_0(\nu, s) > 0$, $\gamma = \gamma(\nu, s) > 0$ such that for any $T \geq T_0$, there exists a measurable set $A_T \subset G/\Gamma$ of measure at least γ, such that*

$$|c_\pm(x,T)| \geq_{\nu,s} T^{-\frac{1\pm 2\nu}{2}} \tag{5.31}$$

for all $x \in A_T$. Besides, if $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$, and $g \in (D_\nu^\pm)'$ such that the equation

$$g - \phi_\lambda^Z g = Uf \tag{5.32}$$

has no L^2-solutions f, then

$$|c_\pm(x,T) - c_\pm(\phi_\lambda^Z(x),T)| \geq_{\nu,s} T^{-\frac{1\pm 2\nu}{2}} \tag{5.33}$$

for all $x \in A_T$.

Proof. By Proposition 5.18, we have

$$|c_\pm(x,T)| \ll_{\nu,s} T^{-\frac{1\pm\nu}{2}} \ll_{\nu,s} \|c_\pm(\cdot,T)\|_{L^2}$$

for $T \geq T_0$. More precisely, there is $C = C(\nu, s) > 0$ such that

$$|c_\pm(x,T)| \leq C\|c_\pm(\cdot,T)\|_{L^2} \tag{5.34}$$

for $T \geq T_0$. Now let

$$A_T \coloneqq \{x \in G/\Gamma : |c_\pm(x,T)| > \frac{1}{2}\|c_\pm(\cdot,T)\|_{L^2}\}.$$

Then, we have

$$\|c_\pm(\cdot,T)\|_{L^2} = \int_{A_T\cup((G/\Gamma)\setminus A_T)} |c_\pm(x,T)|^2 d\mu(x) \leq (C\mu(A_T) + \frac{1}{2})\|c_\pm(\cdot,T)\|_{L^2}$$

for $T \geq T_0$. It follows that for $T \geq T_0$

$$\mu(A_T) \geq \frac{1}{2C} =: \gamma.$$

This proves (5.31).

Next, assume that (5.32) holds. Then using Proposition 5.18 again, we have

$$|c_\pm(x,T) - c_\pm(\phi^Z_\lambda(x),T)| \ll_{\nu,s} T^{-\frac{1\pm\nu}{2}} \ll_{\nu,s} \|c_\pm(\cdot,T) - \phi^Z_\lambda c_\pm(\cdot,T)\|_{L^2}$$

for $T \geq T_0$. A similar argument as above proves (5.33). □

Now for $s > 1$, we consider $g^+ \in (D^+_\nu)' \subset W^{s,r_0,\nu}_1$ satisfying

$$D^+_\nu(g^+) \neq 0, \quad D^-_\nu(g^+) = \mathcal{R}(x,T)(g^+) = \mathcal{C}(x,T)(g^+) = 0.$$

Then by (5.5),the ergodic average $S_{x,T}$ of g^+ is

$$S_{x,T}(g^+) = \frac{1}{T}\int_0^T g^+(\phi^U_t(x))dt = c_+(x,T)\mathcal{D}^+_\nu(g^+). \tag{5.35}$$

Then there are several interesting consequences related to these functions. The following result implies that the *central limit theorem* does no hold for unipotent flow on G/Γ.

Corollary 5.20. *As $T \to \infty$, any weak limit of the probability distributions*

$$\frac{\frac{1}{T}\int_0^T g^+(\phi^U_t(x))dt}{\left\|\frac{1}{T}\int_0^T g^+(\phi^U_t(\cdot))dt\right\|_{L^2}}$$

has a nonzero compact support.

Proof. By (5.34), the distributions are uniformly bounded above by C for sufficiently large T. On the other hand, (5.31) shows that the distributions are bounded below on a measurable set of positive measure γ. The consequence follows. □

Moreover, the functions g^+ is not *measurably trivial*, in the sense that there are no measurable functions f satisfy

$$\int_0^T g^+(\phi^U_t(x))dt = f(\phi^U_T(x)) - f(x). \tag{5.36}$$

Hence, we finally arrive at Theorem 1.4:

Corollary 5.21. *The functions $g^\pm \in (D^\pm_\nu)' \subset W^{s,r_0,\nu}_1$ are not measurably trivial. Moreover, if there are some $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$ such that $\phi^Z_\lambda g^\pm$ are not L^2-cohomologous to $g^\pm$, then $\phi^Z_\lambda g^\pm$ are not measurably cohomologous to $g^\pm$.*

Proof. Assume by contradiction that there is a measurable function f satisfies (5.36). Then by *Luzin's theorem*, for given $\gamma > 0$, there exists a constant $C = C(\gamma) > 0$ such that for any $T > 0$, there exists a measurable set $B_T = B_T(\gamma) \subset G/\Gamma$ of measure $\mu(B_T) < \gamma$ such that

$$\left|\int_0^T g^+(\phi^U_t(x))dt\right| \leq C$$

for all $x \in B_T^c$. On the other hand, (5.35) and (5.31) imply that there exists a measurable set $A_T \subset G/\Gamma$ of measure $\mu(A_T) \geq \gamma$ such that

$$\left|\int_0^T g^+(\phi_t^U(x))dt\right| = |Tc_+(x,T)| \geq_{\nu,s} T^{\frac{1-2\nu}{2}}$$

for all $x \in A_T$. It is a contradiction. Thus, $g^\pm$ are not measurably trivial.

Similarly, if there are some $Z \in C_{\mathfrak{g}}(U)$, $\lambda \in \mathbf{R}$ such that $\phi_\lambda^Z g^+$ is not L^2-cohomologous to g^+, then $\phi_\lambda^Z g^\pm$ are not measurably cohomologous to $g^\pm$. Then (5.33) and (5.35) imply that there exists a measurable set $A_T \subset G/\Gamma$ of measure $\mu(A_T) \geq \gamma$ such that

$$\begin{aligned}&\left|\int_0^T g^+(\phi_t^U(x)) - \phi_\lambda^Z g^+(\phi_t^U(x))dt\right|\\ =&T|c_+(x,T) - c_+(\phi_\lambda^Z(x),T)| \geq_{\nu,s} T^{\frac{1-2\nu}{2}}.\end{aligned}$$

Again, it contradicts *Luzin's theorem*. □

References

[AFRU19] Artur Avila, Giovanni Forni, Davide Ravotti, and Corinna Ulcigrai. Mixing for smooth time-changes of general nilflows. *arXiv preprint arXiv:1905.11628*, 2019.

[Aub82] Thierry Aubin. *Nonlinear analysis on manifolds. Monge-Ampere equations*, volume 252. Springer Science & Business Media, 1982.

[Bro88] Robert Brooks. Injectivity radius and low eigenvalues of hyperbolic manifolds. *Journal für die reine und angewandte Mathematik*, 1988(390):117–129, 1988.

[Cor90] Kevin Corlette. Hausdorff dimensions of limit sets i. *Inventiones mathematicae*, 102(1):521–541, 1990.

[DA12] Rafael Tiedra De Aldecoa. Spectral analysis of time changes of horocycle flows. *Journal of Modern Dynamics*, 6(2):275, 2012.

[Ein06] Manfred Einsiedler. Ratner's theorem on $\mathrm{SL}(2,\mathbf{R})$-invariant measures. *arXiv preprint math/0603483*, 2006.

[EMV09] Manfred Einsiedler, Grigory Margulis, and Akshay Venkatesh. Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces. *Inventiones mathematicae*, 177(1):137–212, 2009.

[FF03] Livio Flaminio and Giovanni Forni. Invariant distributions and time averages for horocycle flows. *Duke Mathematical Journal*, 119(3):465–526, 2003.

[FU12] Giovanni Forni and Corinna Ulcigrai. Time-changes of horocycle flows. *Journal of Modern Dynamics*, 6(2):251, 2012.

[GT50] I. M. Gelfand and M. L. Tsetlin. Finite-dimensional representations of groups of orthogonal matrices. In *Doklady Akad. Nauk SSSR (NS)*, volume 71, pages 1017–1020, 1950.

[Hig08] Nicholas J Higham. *Functions of matrices: theory and computation*, volume 104. Siam, 2008.

[Hir62a] Takeshi Hirai. On infinitesimal operators of irreducible representations of the lorentz group of n-th order. *Proceedings of the Japan Academy*, 38(3):83–87, 1962.

[Hir62b] Takeshi Hirai. On irreducible representations of the lorentz group of n-th order. *Proceedings of the Japan Academy*, 38(6):258–262, 1962.

[Hum12] James E Humphreys. *Introduction to Lie algebras and representation theory*, volume 9. Springer Science & Business Media, 2012.

[JW77] Kenneth D Johnson and Nolan R Wallach. Composition series and intertwining operators for the spherical principal series. i. *Transactions of the American Mathematical Society*, 229:137–173, 1977.

[KM99] Dmitry Y Kleinbock and Gregory A Margulis. Logarithm laws for flows on homogeneous spaces. *Inventiones mathematicae*, 138(3):451–494, 1999.

[Kna01] Anthony W Knapp. *Representation theory of semisimple groups: an overview based on examples*, volume 36. Princeton university press, 2001.

[Kos69] Bertram Kostant. On the existence and irreducibility of certain series of representations. *Bulletin of the American Mathematical Society*, 75(4):627–642, 1969.

[LP82] Peter D Lax and Ralph S Phillips. The asymptotic distribution of lattice points in euclidean and non-euclidean spaces. *Journal of Functional Analysis*, 46(3):280–350, 1982.

[LP84] Peter D Lax and Ralph S Phillips. Translation representations for automorphic solutions of the wave equation in non-euclidean spaces. i. *Communications on pure and applied mathematics*, 37(3):303–328, 1984.

[Mar77] Brian Marcus. Ergodic properties of horocycle flows for surfaces of negative curvature. *Annals of mathematics*, pages 81–105, 1977.

[Mie06] David Johnathan Mieczkowski. The cohomological equation and representation theory. 2006.

[Mil76] John J Millson. On the first betti number of a constant negatively curved manifold. *Annals of Mathematics*, pages 235–247, 1976.

[Mor05] Dave Witte Morris. *Ratner's theorems on unipotent flows*. University of Chicago Press, 2005.

[Muk68] N Mukunda. Unitary representations of the lorentz groups: Reduction of the supplementary series under a noncompact subgroup. *Journal of Mathematical Physics*, 9(3):417–431, 1968.

[Ram13] Felipe A Ramírez. Invariant distributions and cohomology for geodesic flows and higher cohomology of higher-rank anosov actions. *Journal of Functional Analysis*, 265(6):1002–1063, 2013.

[Ran74] Burton Randol. Small eigenvalues of the laplace operator on compact riemann surfaces. *Bulletin of the American Mathematical Society*, 80(5):996–1000, 1974.

[Rat79] Marina Ratner. The cartesian square of the horocycle flow is not loosely bernoulli. *Israel Journal of Mathematics*, 34(1):72–96, 1979.

[Rat86] Marina Ratner. Rigidity of time changes for horocycle flows. *Acta mathematica*, 156(1):1–32, 1986.

[Rav19] Davide Ravotti. Parabolic perturbations of unipotent flows on compact quotients of $\mathrm{SL}(3, \mathbf{R})$. *Communications in Mathematical Physics*, 371(1):331–351, 2019.

[Sar05] Peter Sarnak. Notes on the generalized ramanujan conjectures. *Harmonic analysis, the trace formula, and Shimura varieties*, 4:659–685, 2005.

[Sha00] Yehuda Shalom. Rigidity, unitary representations of semisimple groups, and fundamental groups of manifolds with rank one transformation group. *Annals of Mathematics*, 152(1):113–182, 2000.

[Sim18] Lucia D Simonelli. Absolutely continuous spectrum for parabolic flows/maps. *Discrete & Continuous Dynamical Systems-A*, 38(1):263, 2018.

[SV12] Birgit Speh and Tyakal Nanjundiah Venkataramana. On the restriction of representations of $\mathrm{SL}(2, \mathbf{C})$ to $\mathrm{SL}(2, \mathbf{R})$. In *Representation Theory, Complex Analysis, and Integral Geometry*, pages 231–249. Springer, 2012.

[SWY80] Richard Schoen, Scott Wolpert, and Shing-Tung Yau. Geometric bounds on the low eigenvalues of a compact surface. In *Proc. Sympos. Pure Math*, volume 36, pages 279–285, 1980.

[SZ16] Birgit Speh and Genkai Zhang. Restriction to symmetric subgroups of unitary representations of rank one semisimple lie groups. *Mathematische Zeitschrift*, 283(1-2):629–647, 2016.

[Thi73] Ernest Thieleker. On the quasi-simple irreducible representations of the lorentz groups. *Transactions of the American Mathematical Society*, 179:465–505, 1973.

[Thi74] Ernest A Thieleker. The unitary representations of the generalized lorentz groups. *Transactions of the American mathematical Society*, 199:327–367, 1974.

[Ven10] Akshay Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Annals of Mathematics*, pages 989–1094, 2010.

[Vil78] Naum IAkovlevich Vilenkin. *Special functions and the theory of group representations*, volume 22. American Mathematical Soc., 1978.

[Wan15] Zhenqi Jenny Wang. Cohomological equation and cocycle rigidity of parabolic actions in some higher-rank lie groups. *Geometric and Functional Analysis*, 25(6):1956–2020, 2015.

[Zha15] Genkai Zhang. Discrete components in restriction of unitary representations of rank one semisimple lie groups. *Journal of Functional Analysis*, 269(12):3689–3713, 2015.

Department of Mathematics, IU, Bloomington, IN 47401
E-mail address: `1992.siyuan.tang@gmail.com, siyutang@indiana.edu`