ON MIXING PROPERTIES OF COMPACT GROUP EXTENSIONS OF HYPERBOLIC SYSTEMS.

DMITRY DOLGOPYAT

Abstract. We study compact group extensions of hyperbolic diffeomorphisms. We relate mixing properties of such extensions with accessibility properties of their stable and unstable laminations. We show that generically the correlations decay faster than any power of time. In particular, this is always the case for ergodic semisimple extensions as well as for stably ergodic extensions of Anosov diffeomorphisms of infranilmanifolds.

1. Introduction

1.1. **Overview.** This paper treats compact group extensions of hyperbolic systems. These systems have attracted much attention in the past because they provide one of the simplest examples of weakly hyperbolic systems. Due to the major developments in 60' and 70' the theory of uniformly hyperbolic systems (i.e., Anosov and Axiom A diffeomorphisms) is quite well understood (see [3, 7]). It is also now generally accepted that the hyperbolic structure is the main cause of the chaotic behavior in deterministic systems. Thus it is important to understand how much the assumptions of uniform hyperbolicity can be weakened so that the same conclusions remain valid. One direction of research which experiences a new wave of interest now is the theory of partially hyperbolic or slightly less generally transversely hyperbolic systems. In this case our diffeomorphism preserves some foliation and is hyperbolic in the transverse direction, at least, when restricted to the non-wandering set. The systems we deal with can be specified by the requirement that the foliation involved has compact leaves and the maps between leaves are isometries. If G is a compact group the diffeomorphisms with this property form an open set in the space of G–equivariant dynamical systems and they play the same role in the equivariant theory as Axiom A play in the space of all diffeomorphisms.

Thus the systems under consideration are the simplest partially hyperbolic systems since we have very strong control over what happens in the center. Besides harmonic analysis can be used to study such systems. These reasons make compact group extensions over hyperbolic systems an attractive object of investigation. In fact qualitative properties of these systems are well understood now. The progress here can be summarized as follows. First, Brin in a series of papers [10, 11, 12] applied the general theory of partially hyperbolic systems [13] to show that, in the volume–preserving case, such systems are generically ergodic and weak mixing. It then follows from the general theory of compact group extensions [43] that they are also Bernoulli. Quite recently Burns and Wilkinson [17] used new advances in partially hyperbolic theory [28, 40, 41] to show that generically ergodicity of such systems persists under small not necessary equivariant perturbations. In another direction Field, Parry and Pollicott generalized Brin's theory to the non-volume preserving context. By contrast not much is known about quantitative properties of such systems. This paper is a first step in this direction.

To explain our results we need to introduce some notation. Let F be a topologically mixing Axiom A diffeomorphism on a compact manifold Y. Let f be a Holder continuous function and μ_f be a Gibbs measure with potential f. Also, let G be a compact connected and simply connected Lie group and X be a transitive G–space. Write $M = Y \times X$. Let $\tau : Y \to G$ be a smooth function. Consider the skew action

$$T(y, x) = (F(y), \tau(y)x). \tag{1}$$

It preserves measure $\mu = \mu_f \times$Haar. If A and B are functions on M let

$$\rho_{A,B}(n) = \int A(y, x)B(T^n(y, x))d\mu(y, x).$$

Denote by

$$\bar{\rho}_{A,B}(n) = \rho_{A,B}(n) - \int A(y, x)d\mu(y, x) \int B(y, x)d\mu(y, x)$$

the correlation function. Call T **rapidly mixing** ($T \in \mathcal{RM}$) if $\bar{\rho}$ is a continuous map from $C^\infty(M) \times C^\infty(M)$ to rapidly decreasing sequences, that is given k there are constants C, r such that

$$|\bar{\rho}_{A,B}| \leq C||A||_{C^r(M)}||B||_{C^r(M)}n^{-k}. \tag{2}$$

On the first glance this definition depends also on the Gibbs potential f but we will show that it is not the case. One may think that better bounds should hold for generic extensions. However the decay of correlation this definition requires is fast enough to imply good stochastic behavior. As an example in Subsection 6.1 we derive the Central Limit

Theorem from it. On the other hand (2) is mild enough so that it can be verified in many cases.

As in qualitative theory, accessibility properties of the system under consideration play important role in our analysis. Let Ω_F be the non–wandering set of F and $\Omega = \Omega_F \times Y$ be the non-wandering set of T. Let m', m'' be points in Ω. We say that m'' is accessible from m' if there is a chain of points $m' = m_0, m_1 \dots m_n = m''$ such that m_{j+1} belongs to either stable or unstable manifold of m_j. (We call such a chain n-legged.) Given m, the set of points in the same fiber which are accessible from m lie on an orbit of a group Γ_t which we call Brin transitivity group. As usual different choices of reference point give conjugated groups. The Brin transitivity group can be obtained as follows. Let $\bar{T}$ be the principal extension associated to T (that is $\bar{T}$ acts by (1) on $Y \times G$). Let $\Gamma(n, R)$ be the set of points which can be accessed from (y, id) by n-legged chains such that the distance between m_{j+1} and m_j inside the corresponding stable (unstable) manifold is at most R. Then if n, R are large enough, $\Gamma(n, R)$ generates Γ_t. It was shown by Brin that T is mixing if and only if Γ_t acts ergodically on X. Here we prove the following refinement.

Theorem 1.1. *Let n, R be so large that $\Gamma(n, R)$ generates Γ_t. Then $T \in \mathcal{RM}$ if and only if $\Gamma(n, R)$ is Diophantine.*

Here as usual Diophantine condition means the absence of resonances. More exactly we call a subset $S \subset G$ Diophantine for the action of G on X if for large k, S does not have non–constant almost invariant vectors in $C^k(X)$. See Appendix A for details.

It can be shown that a generic pair of elements of G is Diophantine. (The exceptional set is a union of a countable number of positive codimension submanifolds. In case G is semisimple it is a finite union of algebraic subvarieties. See [25].) From this we can deduce that in a generic family, the condition of Theorem 1.1 is satisfied on the set of full measure. A drawback of this result is that it does not tell how the constant C from (2) varies along the family. Thus one may wonder how large the interior of $\mathcal{RM}$ is. This question is easier if Ω is large [38, 27] (since then Γ_t is also large) or if G is semisiple.

Let $\mathcal{ERG}$ be the set of ergodic group extensions.

Corollary 1.2. *If f is an Anosov diffeomorphism of an infranilmanifold then* $\mathrm{Int}(\mathcal{RM}) = \mathrm{Int}(\mathcal{ERG})$.

This is a direct consequence of Theorem 1.1 and [17]. This result is quite satisfying because one would not expect good mixing properties from a diffeomorphism which can be well–approximated by non–ergodic ones.

Corollary 1.3. *If G is semisimple then* $\mathrm{Int}(\mathcal{RM}) = \mathrm{Int}(\mathcal{ERG}) = \mathcal{ERG}$.

In general we can reduce the problem to an Abelian extension. Let T_a be the factor of T on $Y \times (X/[G,G])$.

Corollary 1.4. *If $T \in \mathcal{ERG}$ then $T \in \mathcal{RM}$ if and only if $T_a \in \mathcal{RM}$.*

Still in the general case of compact extensions of Axiom A diffeomorphism we do not know how large the interior of rapidly mixing diffeomorphisms is. To get some insight into this we study two related classes of dynamical systems. These are compact group extensions of subshifts of a finite type and of expanding maps of Riemannian manifolds. Heuristically the subshifts of finite type are less rigid than Axiom A diffeos because any subshift of a finite type has an Axiom A realization but small perturbations of the subshift correspond to piecewise Holder perturbations of diffeomorphisms. Similarly natural extensions of expanding maps have Axiom A realizations but the unstable foliation will be more smooth than in the general case. So they are more rigid. Nonetheless, in both cases we show that the interior of rapidly mixing maps is dense. In the second case even the interior of the exponentially mixing maps is dense. This suggests that the same result might be true in the context of compact extensions of Axiom A diffeomorphisms.

1.2. **Organization of the paper.** Let us describe the structure of the paper. Section 2 is preliminary. Here we recall necessary facts about Axiom A diffeomorphisms and symbolic dynamics. We also present Brin's theory of compact extensions and its generalization by Field, Parry and Pollicott. In Section 3 we study compact group extensions of expanding maps. First, we describe the Lie algebra of Brin transitivity group. We then proceed to show that if this algebra equals the whole Lie algebra of G (infinitesimal complete non-integrability) then the system is exponentially mixing. Under some technical assumptions we establish the converse of this statement. Also we show that if this condition is not satisfied the map can be made non-ergodic by an arbitrary small perturbation. We conclude Section 3 by showing that infinitesimal complete non-integrability is generic. Section 4 treats symbolic dynamical systems. We show that, in the absence of resonances, our skew extension is rapidly mixing. (See Appendix A for the detailed discussion of the notion of resonances we use). We also describe the reduction of a general extension to the semisimple and abelian cases. We conclude Section 4 by showing that rapid mixing is generic. In Section 5 we apply the results of the previous section to study extensions of Axiom A diffeomorphisms and prove Theorem 1.1 and Corollaries 1.2–1.4. Section 6 contains some applications of our estimates. Some open questions are collected in Section 7.

For the reader familiar with the concepts of Section 2, Sections 3 and 4–6 constitute blocks which could be read separately. Roughly speaking the difference between Section 3 and Section 4 is that in the former we work with Lie algebras while in the latter we work with Lie groups. The unavailability of the differential calculus accounts for the fact that results of Section 4 are weaker than results of Section 3.

Some of the arguments of this paper are similar to [20]–[22]. The main difference which appear here as compared to [20]–[22] is that we have to work with arbitrary finite dimensional representations rather than one-dimensional ones. Still we show that most of the results of [20]–[22] can be generalized to the setting of the present paper.

Notation. if W is a subset of G we denote by $< W >$ the smallest Lie subgroup of G containing W.

Acknowledgment. It is a pleasure for me to thank W. Parry, M. Pollicott, M. Ratner, K. Schmidt and A. Wilkinson for useful discussions. This work is supported by the Miller Institute of Basic Research in Science. I am also grateful for the referee who found 341 errors and misprints in the original version of this paper.

2. Preliminaries.

2.1. **Subshifts of finite type.** In this section we recall how to reduce the study of Axiom A diffeomorphisms to symbolic systems. First, we recall some facts about subshifts of finite type we. For proofs and more information on the subject see [7, 37].

For a $n \times n$ matrix $\mathbf{A}$ whose entries are zeroes and ones we denote by $\Sigma_{\mathbf{A}} = \{\{\omega_i\}_{i=-\infty}^{+\infty} : \mathbf{A}_{\omega_i \omega_{i+1}} = 1\}$ the configuration space of a subshift of a finite type. Usually we omit $\mathbf{A}$ and write Σ instead of $\Sigma_{\mathbf{A}}$. The shift σ acts on Σ by $(\sigma\omega)_i = \omega_{i+1}$. The one-sided shift $(\Sigma^+_{\mathbf{A}}, \sigma)$ is defined in the same way but the index set is the set of non-negative integers. For $\theta < 1$ we consider the distance $d_\theta(\omega^1, \omega^2) = \theta^k$ where $k = \max\{j : \omega^1_i = \omega^2_i \text{ for } |i| \leq j\}$. If $\mathcal{X}$ is a metric space we denote by $C_\theta(\Sigma, \mathcal{X})$ the space of $d_\theta-$Lipschitz functions from Σ to $\mathcal{X}$. $C^+_\theta(\Sigma, \mathcal{X})$ is defined similarly to Σ^+ instead of Σ. There is a natural embedding of $C^+_\theta(\Sigma, \mathcal{X})$ to $C_\theta(\Sigma, \mathcal{X})$ corresponding to the projection $\Sigma \to \Sigma^+$. We use the notation $L(h)$ for the Lipschitz constant of h. If $\mathcal{X}$ is a Banach space we write $h_n(\omega) =$

$\sum_{i=0}^{n-1} h(\sigma^i\omega)$. Functions f_1 and f_2 are called cohomologous ($f_1 \sim f_2$) if there is a function f_3 such that $f_1(\omega) = f_2(\omega) + f_3(\omega) - f_3(\sigma\omega) + \text{Const}$. For any $f \in C_\theta(\Sigma, \mathcal{X})$ there exists a function $\tilde{f} \in C^+_{\sqrt{\theta}}(\Sigma, \mathcal{X})$ such that $f \sim \tilde{f}$. If $\bar{\omega}, \tilde{\omega}$ are points in Σ and $\bar{\omega}_0 = \tilde{\omega}_0$ we denote by $[\bar{\omega}, \tilde{\omega}]$ their local product. That is, $[\bar{\omega}, \tilde{\omega}]_j = \bar{\omega}_j$ if $j \le 0$ and $[\bar{\omega}, \tilde{\omega}]_j = \tilde{\omega}_j$ if $j \ge 0$.

We assume that σ is topologically mixing (that is all entries of some power of $\mathbf{A}$ are positive). The **pressure functional** on $C_\theta(\Sigma, \mathbb{R})$ is defined by

$$Pr(f) = \sup_{\tilde{\nu}} \int f(\omega)\, d\tilde{\nu} + h_{\tilde{\nu}}(\sigma)$$

where the supremum is taken over the set of $\sigma-$invariant probability measures and $h_{\tilde{\nu}}(\sigma)$ is the measure theoretic entropy of σ with respect to $\tilde{\nu}$. μ_f is called the equilibrium state or the **Gibbs measure** with potential f if $\int f(\omega)\, d\mu_f + h_{\mu_f}(\sigma) = Pr(f)$. For $C_\theta(\Sigma, \mathbb{R})$ potentials, Gibbs measures exist and are unique. It is clear that cohomologous functions have the same Gibbs measure. Take $f \in C^+_\theta(\Sigma, \mathbb{R})$ and let μ_f be its Gibbs measure. To describe ν it is enough to specify its projection to Σ^+. To this end consider the transfer operator $\mathcal{L}_f : C_\theta(\Sigma^+) \to C_\theta(\Sigma^+)$

$$(\mathcal{L}_f h)(\omega) = \sum_{\sigma\varpi = \omega} e^{f(\varpi)} h(\varpi).$$

The structure of the spectrum of the transfer operator is described by the Ruelle-Perron-Frobenius Theorem. Namely, the leading eigenvalue of $\mathcal{L}_f$ is simple and if h_f is the corresponding eigenfunction and ν_f is the corresponding eigenmeasure then $\mu_f = h_f \nu_f$.

A function f is called **normalized** if $\mathcal{L}_f 1 = 1$. Given f there is unique normalized $\tilde{f}$ such that $f \sim \tilde{f}$. Let f be normalized and $w = w_1 w_2 \dots w_n$ be an admissible word (that is $\mathbf{A}_{w_i w_{i+1}} = 1$). The map $\varpi(\omega) = w\omega$ is defined on a subset of $\Sigma^+_{\mathbf{A}}$. On this subset the following equation holds:

$$\frac{d\mu_f(\varpi(\omega))}{d\mu(\omega)} = \exp\left[f_n(\varpi(\omega))\right].$$

Gibbs measures are exponentially mixing in the sense that $\forall A, B \in C_\theta(\Sigma)$

$$|\mu_f\left(A\left(\bar{B} \circ \sigma^n\right)\right) - \mu_f(A)\mu_f(\bar{B})| \le \text{Const}\xi^n ||A||_\theta ||B||_\theta \tag{3}$$

for some $\xi < 1$.

2.2. **Brin groups.** Here we review Brin theory of compact group extensions ([13, 10, 11]). We include some proofs to make this paper more self-contained as well as because later on we shall use the similar methods to obtain a quantitative version of the results of this subsection. For different expositions of Brin's theory see [17, 38].

Let $\sigma : \Sigma \to \Sigma$ be a topologically mixing subshift of a finite type. We consider on Σ a Gibbs measure μ_f with potential $f \in C_\theta(\Sigma, \mathbb{R})$. Let G be a compact connected Lie group, X be a transitive G–space and dx be the G-invariant probability measure. We assume that (G, X) is a presentation in the sense that no normal subgroup of G acts transitively on X.

Let $\mathcal{M} = \Sigma \times X$. We denote by $C_{k,\theta}(\Sigma)$ ($C_{k,\theta}(\Sigma^+)$) the space $C_\theta(\Sigma, C_k(X))$ ($C_\theta(\Sigma^+, C_k(X))$). Let $\tau \in C_\theta(\Sigma, G)$ be a Holder continuous function. Form a skew product $T : \mathcal{M} \to \mathcal{M}$

$$T(\omega, x) = (\sigma\omega, \tau(\omega)x)$$

and let $d\mu = d\mu_f dx$. For $\omega \in \Sigma$ introduce stable and unstable sets:

$$W^s(\omega) = \{\varpi : \exists n_0 : \varpi_i = \omega_i \quad \text{for} \quad i \geq n_0\},$$

$$W^u(\omega) = \{\varpi : \exists n_0 : \varpi_i = \omega_i \quad \text{for} \quad i \leq n_0\},$$

Define $\tau_n(\omega) = \tau(\sigma^{n-1}\omega) \dots \tau(\sigma\omega)\tau(\omega)$. For $\varpi \in W^s(\omega)$, let

$$\Delta_s(\omega, \varpi) = \lim_{N\to\infty} \tau_N^{-1}(\varpi)\tau_N(\omega) \tag{4}$$

and for $\varpi \in W^u(\omega)$, let

$$\Delta_u(\omega, \varpi) = \lim_{N\to\infty} \tau_N(\sigma^{-N}\varpi)\tau_N^{-1}(\sigma^{-N}\omega). \tag{5}$$

Now set

$$W^s(\omega, x) = \{(\varpi, y) : \varpi \in W^s(\omega), y = \Delta_s(\omega, \varpi)x\}, \tag{6}$$

$$W^u(\omega, x) = \{(\varpi, y) : \varpi \in W^u(\omega), y = \Delta_u(\omega, \varpi)x\}, \tag{7}$$

It is easy to see that $\mathrm{dist}(T^n(\omega, x), T^n(\varpi, y)) \to 0$ as $n \to +\infty$ exponentially fast if $(\varpi, y) \in W^s(\omega, x)$ and $\mathrm{dist}(T^n(\omega, x), T^n(\varpi, y)) \to 0$ as $n \to -\infty$ exponentially fast if $(\varpi, y) \in W^u(\omega, x)$. By a **t-chain** in Σ we mean a set of points $\omega^0, \omega^1, \dots \omega^n$ such that for all i either $\omega^{i+1} \in W^s(\omega^i)$ or $\omega^{i+1} \in W^s(\omega^i)$. An **e-chain** is defined by also allowing that $\omega^{i+1} = \sigma^n\omega^i$. We can also define e- and t-chains in $\mathcal{M}$. As usual we say that $(\omega^0, x_0)(\omega^1, x_1) \dots (\omega^n, x_n)$ covers $W = (\omega^0\omega^1 \dots \omega^n)$. By (6) and (7) for any such chain we have $x_n = g(W)x_0$, where g does not depend on x_0. We also say that any chain connects its endpoints. If an (e- or t-) chain W has $\omega^0 = \omega^n = \omega$ we say that W is a closed chain at ω.

Definition. The ergodicity group $\Gamma_e(\omega)$ is the subgroup of G generated (set-theoretically) by $g(W)$ for all e-chains W at ω.

Definition. The transitivity group $\Gamma_t(\omega)$ is the subgroup of G generated by $g(W)$ for all closed t-chains W at ω.

We refer to Γ_e and Γ_t as **the Brin groups.** Note that the Brin groups can be defined (and used) in much more general framework of extensions of partially hyperbolic systems (see [34]). It is interesting to see how much of the theory described below works in that setting.

Proposition 2.1. *For any e-chain* $W = (\omega^0, \omega^1 \dots \omega^n)$

$$\Gamma_t(\omega^n) = g(W)\Gamma_t(\omega^0)g^{-1}(W),$$

$$\Gamma_e(\omega^n) = g(W)\Gamma_e(\omega^0)g^{-1}(W).$$

PROOF: It is enough to consider two-point chains $W = (\omega, \tilde{\omega})$, the general case follows by induction. If $\tilde{\omega} \in W^s(\omega) \bigcup W^u(\omega)$ we note that if V is a closed t-chain (e-chain) at ω then $\tilde{\omega} V \tilde{\omega}$ is a closed is a closed t-chain (e-chain) at $\tilde{\omega}$. Thus $\Gamma_*(\tilde{\omega}) \supset g(W)\Gamma_*(\omega)g^{-1}(W)$. Similarly $\Gamma_*(\omega) \supset g^{-1}(W)\Gamma_*(\tilde{\omega})g(W)$. If $\tilde{\omega} = \sigma^n \omega$ then V is a closed chain at ω iff $\sigma^n W$ is a closed chain at $\tilde{\omega}$. $\square$

As any two points in Σ can be connected by a t-chain, we get the following consequence of the preceding result.

Proposition 2.2. *(i)* $\forall \omega^1, \omega^2$ $\Gamma_t(\omega^1)$ *is conjugated to* $\Gamma_t(\omega^2)$ *and* $\Gamma_e(\omega^1)$ *is conjugated to* $\Gamma_e(\omega^2)$*;*
(ii) $\forall \omega$ $\Gamma_t(\omega)$ *is normal in* $\Gamma_e(\omega)$.

If we make a change of coordinates

$$(\omega', x') = (\omega, \alpha(\omega)x) \tag{8}$$

then in the new coordinates $T(\omega', x') = (\sigma\omega', \tau'(\omega))$ where

$$\tau'(\omega) = \alpha(\sigma\omega)\tau(\omega)\alpha^{-1}(\omega).$$

Γ's are transformed according to the following rule.

Proposition 2.3. *In the coordinates* $(\omega', x') = (\omega, \alpha(\omega)x)$

$$\Gamma'_e(\omega) = \alpha(\omega)\Gamma_e(\omega)\alpha^{-1}(\omega)$$

$$\Gamma'_t(\omega) = \alpha(\omega)\Gamma_t(\omega)\alpha^{-1}(\omega)$$

Definition. T is called **reduced** if $\forall \omega^1, \omega^2$ there is a t-chain W connecting ω^1 to ω^2 such that $g(W) = id$ (the identity element in G).

Proposition 2.4. *If T is reduced then*
(i) $\Gamma_t(\omega)$ and $\Gamma_e(\omega)$ do not depend on ω;
(ii) If W is a t-chain (e-chain) then $g(W) \in \Gamma_t$ ($g(W) \in \Gamma_e$);
(iii) Γ_e/Γ_t is cyclic.

PROOF: (i) follows immediately from Proposition 2.2.

(ii) Let $W = (\omega^0 \dots \omega^n)$ be a t-chain (e-chain) and let $V = (\tilde{\omega}^0 \dots \tilde{\omega}^m)$ be a t-chain with $\tilde{\omega}^0 = \omega^n$, $\tilde{\omega}^m = \omega^0$, $g(V) = id$, then

$$g((\omega^0 \dots \omega^n \tilde{\omega}^1 \dots \tilde{\omega}^m)) = g(W).$$

(iii) By (ii), if $\omega^2 \in W^s(\omega^1)$ then $\Delta_s(\omega^1, \omega^2) \in \Gamma_t$, $\Delta_s(\sigma\omega^1, \sigma\omega^2) \in \Gamma_t$. But $\Delta_s(\sigma\omega^1, \sigma\omega^2) = \tau(\omega^1)\Delta_s(\omega^1, \omega^2)\tau^{-1}(\omega^2)$. Therefore if $\omega^2 \in W^s(\omega^1)$ then $\tau(\omega^1) \equiv \tau(\omega^2) \bmod \Gamma_t$. The same is true if $\omega^2 \in W^u(\omega^1)$ and hence if ω^2 can be connected to ω^1 by a t-chain. ∎

Proposition 2.5. *Every T can be reduced by a change of coordinates (8).*

PROOF: Clearly T is reduced if every ω can be connected to some fixed $\omega^0 \in \Sigma$ by a t-chain W with $g(W) = id$. Now choose any chain $W(\omega)$ connecting ω to ω^0 such that $g(W(\omega))$ is continuous and set $\alpha(\omega) = g(W(\omega))$ in (8) □.

Proposition 2.6. *(i) T is ergodic iff $\bar{\Gamma}_e$ acts transitively on X;*
(ii) T is weak mixing iff $\bar{\Gamma}_t$ acts transitively on X.

PROOF: We prove weak mixing criterion. Ergodicity is similar but easier. By Proposition 2.5 we may assume that T is reduced.

(a) Let $\bar{\Gamma}_t$ be transitive and $h(\omega, x)$ be an eigenfunction of T. It follows from [38] that we can assume that h is continuous. Then it is easy to see that it is constant along $W^s(\cdot)$ and $W^u(\cdot)$. Thus $\forall \omega, x$ $\forall g \in \bar{\Gamma}_t$ $h(\omega, gx) = h(\omega, x)$. Since $\bar{\Gamma}_t$ is transitive, $h(\omega, x)$ depends only on the base point and since σ is weak-mixing, h is constant.

(b) Assume that $\bar{\Gamma}_t$ is not transitive. If $\bar{\Gamma}_e$ is not transitive then any $\bar{\Gamma}_e$–invariant function on X lifts to a T invariant function on $\mathcal{M}$ so we may assume that $\bar{\Gamma}_e$ is transitive. Let $\mathcal{A}$ be the algebra of the sets of the form $\Sigma \times Z$ where Z is Γ_t invariant. Then T preserves $\mathcal{A}$ and the action of T on $\mathcal{A}$ is a factor of a group shift on $K = \bar{\Gamma}_e/\bar{\Gamma}_t$. Thus it has pure point spectrum and so T is not weak-mixing. □

By a theorem of Rudolph ([43]) any weak-mixing compact group extension of Bernoulli shift is Bernoulli shift, therefore, we get

Corollary 2.7. *If G is compact then T is Bernoulli iff $\bar{\Gamma}_t$ acts transitively on X.*

Remark. Since in our case τ is Holder continuous we do not have to use the deep result of [43] to obtain the last statement. In fact straightforward arguments of [36], [14], [42] would suffice. The later approach is similar to one used in the present paper to derive estimates on correlation function.

It is known that if G is semisimple then ergodicity implies weak mixing. Note that this is a consequence of the following statement (we need part (a) here while part (b) will be used later on).

Proposition 2.8. *(a) Let X be a transitive space of a compact connected semisimple Lie group G, $H_1 \subset H_2$ be subgroups of G. Assume that H_1 is normal in H_2 and $H_2/H_1 = \mathbb{T}^d \times F$ where F is a finite group. If H_2 acts transitively on X then so does H_1.*

(b) Let X be a transitive space of a compact connected Lie group G, $H \subset G$ be a closed subgroup. Then H is transitive on X iff it is transitive on $X/[G, G]$ and $X/\text{Center}(G)$.

Proof: (a) We may assume that H_1 is connected by passing to its identity component. Also since X is connected we may assume that so is H_2 and hence that $F = \{\text{id}\}$. Now $\mathbb{T}^d = H_2/H_1$ acts on $Y = H_1\backslash X$ and because this action is transitive $Y \equiv \mathbb{T}^m$ for some m. So X and hence G fiber over $\mathbb{T}^m$. Therefore $m = 0$ as claimed;

(b) By the same argument as before we can neglect finite covers and assume that $\text{Center}(G) \bigcap [G, G] = \{\text{id}\}$. Since X is a transitive G-space it equals G/Γ for some subgroup Γ of G. Take $g \in G$ As H is transitive on $X/[G, G]$, $\exists h \in H, \gamma \in \Gamma, g' \in [G, G]$ such that $hg\gamma = g'$. Since H is transitive on $X/\text{Center}(G)$, we can apply (a) to conclude that $[H, H]$ also acts transitively on $G/\Gamma\text{Center}(G)$. Equivalently the left action of $\Gamma\text{Center}(G)$ on $[H, H]\backslash G$ is transitive. Hence Γ acts transitively on $[H, H]\text{Center}(G)\backslash G$. Again by (a) $[\Gamma, \Gamma]$ acts transitively on $[H, H]\text{Center}(G)/G$. Thus $\exists h' \in [H, H], \gamma' \in [\Gamma, \Gamma]$ such that $h'g'\gamma' \in \text{Center}(G)$. But also $h'g'\gamma' \in [G, G]$ thus $h'g'\gamma' = \text{id}$. But $h'g'\gamma' = h'hg\gamma\gamma'$, so $Hg\Gamma = G$. $\square$

2.3. **One sided subshifts.**

Here we discuss the reduction of two–sided subshifts to one–sided ones.

First, we show that by a change of variables we can obtain that that $\tau(\omega)$ depends only on the future. Given two sequences ω^1, ω^2 such that $\omega_0^1 = \omega_0^2$ let $[\omega^1, \omega^2]$ denote their local product, that is, $[\omega^1, \omega^2]_j = \omega_j^1$ for $j \leq 0$ and $[\omega^1, \omega^2]_j = \omega_j^2$ for $j \geq 0$. For each $a \in \{1 \dots n\}$ choose a sequence $\hat{\omega}(a)$ such that $\hat{\omega}(a)_0 = a$. Let $\phi(\omega) = [\hat{\omega}(\omega_0), \omega]$. Make change of variables (8) with

$$\alpha(\omega) = \Delta_s^{-1}(\phi(\omega), \omega).$$

It is easy to see that in the new variables local stable manifolds are flat, that is if $\omega_j^1 = \omega_j^2$ for $j \geq 0$ then $\Delta_s'(\omega^1, \omega^2) = \text{id}$. Thus

$$\text{id} = \Delta_s'(\omega^1, \omega^2) = \tau'(\omega^1)\Delta_s'(\sigma\omega^1, \sigma\omega^2)\left(\tau'(\omega^2)\right)^{-1} = \tau'(\omega^1)\left(\tau'(\omega^2)\right)^{-1}.$$

Hence $\tau'(\omega^1) = \tau'(\omega^2)$, i.e. τ' depends only on the future coordinates.

Now we show that we also can assume that $A, B \in C_{k,\theta}(\Sigma^+)$. In fact suppose that for all such functions $|\bar{\rho}_{A,B}(N)| \to 0$. Take $A, B \in C_{k,\theta}(\Sigma)$. For any cylinder $\mathcal{C}_{n,j} = \mathcal{C}_{-n,0}(\omega_j)$ choose a sequence $\xi_{n,j} \in \mathcal{C}_{n,j}$. If $H \in C_{k,\theta}(\Sigma)$ denote $H^{(n)}(\omega, x) = H([\xi_{n,j(\omega)}, \omega], \tau_n^{-1}(\omega)\tau_n(\xi_{n,j(\omega)})x)$, then

$$\|H - H^{(n)}\|_0 \leq \text{Const}\|H\|_{k,\theta}\theta^n,$$

where $||H||_{k,\theta}$ denotes the norm of H as the element of $C_\theta(\Sigma, C^k(X))$. Also,

$$\|H^{(n)} \circ T^n\|_{k,\theta} \leq \text{Const}\|H^{(n)}\|_{\theta,0}(\frac{1}{\theta})^n$$

and $H^{(n)} \circ T^n \in C_\theta(\Sigma^+)$. So

$$\begin{aligned}\bar{\rho}_{A,B}(N) = \bar{\rho}_{A^{(n)},B^{(n)}}(N) + O(\theta^n) &= \bar{\rho}_{A^{(n)}\circ T^n, B^{(n)}\circ T^n}(N) + O(\theta^n) \quad (9)\\ &= o_{N\to 0}(1) + O_{n\to\infty}(\theta^n).\end{aligned}$$

Thus $\bar{\rho}_{A,B}(N) \to 0, N \to \infty$.

Let T be reduced and $A \in C_{k,\theta}(\Sigma)$. Given $A \in C_{k,\theta}(\mathcal{M})$ let

$$B(\omega, x) = \sum_n [A(T^n(\omega, x)) - A(T^n(\phi(\omega), \Delta_s(\omega, \phi(\omega))x))].$$

Then $B \in C_{k,\theta}(\mathcal{M})$ since the derivative with respect to the second variable of the n-th term of this sum is exponentially small and

$$A - B + B \circ T \in C_{k,\theta}(\Sigma^+) \quad (10)$$

(see [47, 37] for more details).

2.4. **An expression for the correlation function.** In this section we provide an expression for the correlation function we shall use later on. By the preceding section we can assume that (Σ, σ) is an one-sided subshift of finite type. If ω, ϖ are two-sided sequences such that $\omega_i = \varpi_i$ for $i \geq 0$ then $\Delta_s(\omega, \varpi) = \text{id}$. We also assume that the potential f of Gibbs measure μ is normalized, that is

$$\sum_{\sigma\omega=\varpi} e^{f(\omega)} = 1. \quad (11)$$

Let Δ be the Laplace operator of some G-invariant Riemann metric on X. Let $\mathbb{H}_\lambda$ be the space of functions satisfying $\Delta f = \lambda^2 f$, $\int f(x)\,dx = 0$

and $\mathbb{H}_0$ be the space of constants. In this subsection we provide a formula for correlation function

$$\rho_{A,B}(n) = \int A(q)\overline{B(T^n q)}\, d\mu(q).$$

We have

$$\rho_{A,B}(n) = \int\int A(\omega, x)\overline{B(\sigma^n\omega, \tau_n(\omega)x)}d\mu dx.$$

Let us make the change of variables $\varpi = \sigma^n\omega$, $y = \tau_n(\omega)x$ then by the definition of Gibbs measures

$$\rho_{A,B}(n) = \int\int \overline{B(\varpi, y)} \sum_{\sigma^n\omega=\varpi} e^{f_n}(\omega)A(\omega, \tau_n^{-1}(\omega)y)d\mu(\varpi)dy.$$

Regard now A, B as functions $\Sigma \to L^2(X)$. Denote $(\pi(g)f)(x) = f(g^{-1}x)$, then we can rewrite the above expression as

$$\rho_{A,B}(n) = \int (\mathcal{L}_\pi^n A)(\omega, x)B(\omega, x)\, d\mu(\omega, x)$$

where

$$(\mathcal{L}_\pi \vec{H})(\varpi) = \sum_{\sigma\omega=\varpi} e^{f(\omega)}\pi(\tau(\omega))\vec{H}(\omega).$$

Finally decompose $A = \int A(\omega, x)\, dx + \sum_\lambda A_\lambda$, $B = \int B(\omega, x)\, dx + \sum_\lambda B_\lambda$ where $A_\lambda(\omega, \cdot), B_\lambda(\omega, \cdot) \in \mathbb{H}_\lambda$. and write

$$(\mathcal{L}_\lambda \vec{H})(\varpi) = \sum_{\sigma\omega=\varpi} e^{f(\omega)}\pi_\lambda(\tau(\omega))\vec{H}(\omega),$$

where π_λ is the restriction of π to $\mathbb{H}_\lambda$. Then using (3) we get

$$\rho_{A,B}(n) = \int A(q)\, d\mu(q) \int \overline{B(q)}\, d\mu(q) + \bar{\rho}_{A,B}(n) + O(\xi^n),$$

$$\bar{\rho}_{A,B}(n) = \sum_\lambda \int\int (\mathcal{L}_\lambda^n A_\lambda, B_\lambda)(\varpi)d\mu(\varpi) + O(\xi^n). \tag{12}$$

2.5. **Axiom A diffeomorphisms.** Recall that $F : Y \to Y$ satisfies Axiom A if there is an F–invariant splitting $T_{\Omega_F}Y = E_s \oplus E_u$ and constants $C, \xi < 1$ such that

(a) for any $v \in E_s(x)$, $n \geq 0$ $||dF^n(v)|| \leq C\xi^n||v||$,

(b) for any $v \in E_u(x)$, $n \geq 0$ $||dF^{-n}(v)|| \leq C\xi^n||v||$.

We suppose that the restriction of F to Ω_F is topologically mixing. We shall use the following statement (see [9]).

Proposition 2.9. *There exists a subshift of a finite type Σ, $\theta < 1$ and a surjective d_θ–Lipschitz map $p : \Sigma \to \Omega_F$ such that $p \circ \sigma = F \circ p$ and if μ_f is the Gibbs measure with potential f on Y then $p_* \mu_f$ is the Gibbs measure on Σ with potential $f \circ p$.*

Thus if $T_Y : Y \times X \to Y \times X$ is a compact extension with skewing function τ, we can associate to it the extension $T_\Sigma : \Sigma \times X \to \Sigma \times X$ given by $T_\Sigma(\omega, x) = (\sigma\omega, \tau(p\omega)x)$. Now if $P = (p, \mathrm{id})$ then $P \circ T_\Sigma = T_Y \circ P$. This allows us to reduce the study of T_Y to that of T_Σ.

3. Expanding maps.

3.1. **Content of this section.** In this section we study compact group extensions of expanding maps. We assume that σ is an expanding map of a compact connected Riemannian manifold M. In this section we use notation which is slightly different from one used in the rest of the paper. Namely we denote by x points in M. Let $(\tilde{M}, \sigma)$ be the natural extension of (M, σ). Points in $\tilde{M}$ will be denoted by $q = (x, \vec{y})$.

The structure of expanding maps is given by the following result of Gromov and Shub [30, 44].

Proposition 3.1. *(*[30]*) Let σ be an expanding map of a compact connected boundaryless manifold M. Then there exist a nilpotent simply-connected Lie group N and a subgroup Γ of $\mathrm{Aff}(N)$ acting discretely on N such that $M = N/\Gamma$. Moreover there exist an expanding automorphism $\alpha \in \mathrm{Aut}(N)$ and a homeomorphism $\xi : M \to M$ such that $\alpha(\Gamma) = \Gamma$ and $\sigma = \xi\alpha\xi^{-1}$.*

In particular the universal cover $\hat{M}$ of M is $\mathbb{R}^d$ and the action of σ on the first cohomology group of M has no non-trivial fixed points.

Given $\tau \in C^\infty(M, G)$ we define skew extension $T : M \times X \to M \times X$ by $T(x, \eta) = (\sigma x, \tau(x)\eta)$. Recall the classical fact that expanding maps always have a unique absolutely continuous invariant measure (see [19], for example). Multiplying this measure by the Haar measure on X we obtain a smooth invariant measure for the compact extension.

The Brin groups for compact extensions of σ are defined using the stable and unstable sets exactly as it was done in Section 2. We will also consider infinitesimal analogues of the Brin groups. In the next subsection we introduce the notion of infinitesimal complete non-integrability which is an infinitesimal analogue of the property that transitivity group is whole of G. The results of this section then could be formulated as follows.

Theorem 3.2. (Mixing) *Infinitesimal complete non-integrability implies exponential mixing with respect to the smooth invariant measure.*

Definition. We say that T is stably ergodic if for all pairs $(\tilde{\sigma}, \tilde{\tau})$ C^2–close to (σ, τ) $T_{\tilde{\sigma},\tilde{\tau}}$ is ergodic.

Theorem 3.3. (Characterization) *If $X = G$ then the following properties are equivalent:*
(a) T is stably ergodic;
(b) T is exponentially mixing;
(c) T is infinitesimally non-integrable.

Theorem 3.4. (Prevalence) *Infinitesimal complete non-integrability is generic among compact extensions of expanding maps in the sense that the complimentary subset is a positive codimension submanifold.*

3.2. **Infinitesimal transitivity group.** Here we describe an infinitesimal version of Γ_t. For $x \in M$ let $\mathfrak{h}(x)$ be the span of

$$\partial_{\vec{e}}^{x'}[\Delta_u((x, \vec{y}), (x', \vec{y}')) - \Delta_u((x, \vec{\tilde{y}}), (x', \vec{\tilde{y}}'))]_{x'=x}$$

for all $\vec{e}, \vec{y}, \vec{\tilde{y}}$ ($\vec{y}', \vec{\tilde{y}}'$ are chosen so that $(x', \vec{y}') \in W^u(x, \vec{y})$ and $(x', \vec{\tilde{y}}') \in W^u(x, \vec{\tilde{y}})$). Here $\partial_{\vec{e}}^{x'}$ means the derivative with respect to x' applied to $\vec{e}$.

The plane field $\mathfrak{h}(x)$ is a lower-semicontinuous, that is given x_0 there is a continuous plane field $\tilde{\mathfrak{h}}(x)$ with $\tilde{\mathfrak{h}}(x_0) = \mathfrak{h}(x_0)$ and $\tilde{\mathfrak{h}}(x) \subset \mathfrak{h}(x)$. As

$$\Delta_u(\sigma q, \sigma q') = \tau(x)\Delta_u(q, q')\tau^{-1}(x'),$$

we have

$$\mathfrak{h}(\sigma x) \supset \mathrm{Ad}(\tau(x))\mathfrak{h}(x).$$

So the ergodicity of σ implies

Proposition 3.5. *The conjugacy class of $\mathfrak{h}(x)$ is constant almost everywhere.*

Let $\mathfrak{h}$ is a representative of this class, $F = \{x : \mathfrak{h}(x)$ is conjugated to $\mathfrak{h}\}$, $\tilde{F} = \{x : \dim\mathfrak{h}(x) = \dim\mathfrak{h}\}$. Then $F \subset \tilde{F}$, F has full measure and $\tilde{F}$ is open (by semicontinuity). Also for $x \in \tilde{F}$

$$\mathfrak{h}(\sigma x) = \mathrm{Ad}(\tau(x))\mathfrak{h}(x). \tag{13}$$

Proposition 3.6. $\mathfrak{h}$ *is Holder continuous on $\tilde{F}$.*

Proof: For fixed $y, \tilde{y}, \vec{e}$

$$V(x, y, \tilde{y}, \vec{e}) = \partial_{\vec{e}}^{x}[\Delta_u((x, y)(x', y')) - \Delta_u((x, \tilde{y})(x', \tilde{y}'))]$$

is Holder continuous in x by the general theory of partially hyperbolic systems [33] (or by differentiating the product formula for Δ_u (5) term by term). If $x_0 \in \tilde{F}$ and $\mathfrak{h}(x_0)$ is generated by $\{V_j = V_j(x_0, y_j, \tilde{y}_j, \vec{e}_j)\}$ then for x near x_0 $\mathfrak{h}(x)$ will be generated by $V(x, y_j, \tilde{y}_j, \vec{e}_j)$. □

Lemma 3.7. *Let $W = (q_0, q_1 \dots q_m)$ be an e-chain with $x_0, x_m \in \tilde{F}$, (we write $q_j = (x_j, \vec{y}_j)$) then*

$$\mathrm{Ad}g(W)\mathfrak{h}(x_0) = \mathfrak{h}(x_m). \tag{14}$$

PROOF: Since $\mathfrak{h}(x)$ is continuous on $\tilde{F}$ and $g(W)$ depends continuously on W it suffices to prove this statement for a dense set of chains, so we may assume that $x_i \in F$. Therefore it is enough to verify this statement for $m = 1$. The case when $q_1 = \sigma^n q_2$ follows from (13). Also if $q_1 \in W^s(q_0)$, then $\sigma^n x_0 = \sigma^n x_1$ for some n, so

$$\mathfrak{h}(x_1) = \mathrm{Ad}(\tau_n^{-1}(x_1))\mathfrak{h}(\sigma^n x_1) = \mathrm{Ad}(\tau_n^{-1}(x_1))\mathfrak{h}(\sigma^n x_0) =$$

$$\mathrm{Ad}(\tau_n^{-1}(x_1)\mathrm{Ad}(\tau_n(x_0))\mathfrak{h}(x_0) =$$

$$\mathrm{Ad}(\tau_n^{-1}(x_1)\tau_n(x_0))\mathfrak{h}(x_0) = \mathrm{Ad}(g(W))\mathfrak{h}(x_0).$$

So it remains to consider the case $q_1 \in W^u(q_0)$. Again it suffices to consider a dense set of pairs. By the above proposition we can find an open subset $U \subset N$ and a Holder function $\alpha : U \times U \to GL(\mathfrak{g})$ such that $\forall \tilde{x}, \tilde{\tilde{x}} \in U$ $\mathfrak{h}(\tilde{x}) = \alpha(\tilde{x}, \tilde{\tilde{x}})\mathfrak{h}(\tilde{\tilde{x}})$. Moreover we can assume that α is close to id by shrinking U if necessary. Now, we may assume that $y_n^0, y_n^1 \in U$ for infinitely many n since this condition is satisfied on a dense set of pairs. Then

$$\mathfrak{h}(x_0) = \mathrm{Ad}(\tau_n(y_n^0))\mathfrak{h}(y_n^0) = \mathrm{Ad}(\tau_n(y_n^0))\alpha(y_n^0, y_n^1)\mathfrak{h}(y_n^0) =$$

$$\mathrm{Ad}(\tau_n(y_n^0))\alpha(y_n^0, y_n^1)\mathrm{Ad}(\tau_n^{-1}(y_n^1)\mathfrak{h}(x_1).$$

Passing to the limit as $n \to \infty$ we obtain the statement required. □

Now semicontinuity implies

Corollary 3.8. *If in the previous lemma $x_0 \in \tilde{F}$, $x_m \in M$ then*

$$\mathrm{Ad}g(W)\mathfrak{h}(x_0) \supset \mathfrak{h}(x_m).$$

In particular, $\tilde{F} = F$.

Let $\bar{\mathfrak{h}}(x)$ be the subalgebra generated by $\mathfrak{h}(x)$ and $H(x)$ be the corresponding subgroup.

Corollary 3.9. $\Gamma_e(x) \subset \mathrm{Norm}(H(x))$.

Corollary 3.10. *If either T is reduced or $\bar{\Gamma}_e = G$, then $H(x) \equiv H$ almost surely and always $H(x) \subset H$. Also $\mathfrak{h}$ is an ideal and hence $\bar{\mathfrak{h}} = \mathfrak{h}$.*

Lemma 3.11. *If $\bar{\Gamma}_e = G$ then $\overline{\Gamma_t} = \bar{H}$.*

Proof: By definition $\mathfrak{h} \subset L(\overline{\Gamma_t})$ without any assumptions, so $H \subset \overline{\Gamma_t}$. Locally we can always make a change of variables (8) so that in a neighborhood $U(x_0)$, $\forall x, x'$ $\exists y, y'$ such that $\Delta_u((x,y),(x',y')) = \mathrm{id}$. (Under this change of variables $\mathfrak{h}(x)$ gets replaced by a conjugated subspace but by our assumption $\mathfrak{h}(x)$ is an ideal and so this change of variables does not affect $\mathfrak{h}(x)$.)

Thus if q is close to q', then $\Delta_u(q,q') \in H$. (To see this, join q and q' by a smooth curve $\gamma(\xi)$, then

$$\frac{\partial}{\partial \xi}\left[\Delta_u^{-1}(\gamma(0),\gamma(\xi_0))\Delta_u(\gamma(0),\gamma(\xi))\right] \in \mathfrak{h}.)$$

As M is connected, $\Delta_u(q,q')$ is always in H. As in the proof of Proposition 2.4 we get $\tau(x) \equiv \tau(x') (\mathrm{mod}\ \ H)$. Thus if $q' \in W^s(q)$ then $\Delta_s(q,q') = \tau^{-n}(x')\tau_n(x)$ for some n and so $\Delta_s(q',q)$ belongs to H. $\square$

Let $\mathfrak{g}$ denote the Lie algebra of G.

Definition. Call T **infinitesimally completely non-integrable** if $\mathfrak{h} = \mathfrak{g}$.

3.3. **Complete non-integrability and stable ergodicity.** Here we begin the proof of Theorem 3.3. In this subsection we work with principal extensions that is we assume that $X = G$. First we record the following consequence of Lemma 3.11.

Corollary 3.12. *If G is semisimple, then T is ergodic if and only if $H = G$.*

Lemma 3.13. *If $G = \mathbb{T}^d$, then T is stably ergodic if and only if $H = \mathbb{R}^d$.*

Proof: (a) If $H = \mathbb{R}^d$ then $\bar{\Gamma}_t = \mathbb{T}^d$ and so T is ergodic. Also if $\tilde{T}$ is close to T then by the semicontinuity of $\mathfrak{h}$, $H(\tilde{T}) = \mathbb{R}^d$ as well and so $\tilde{T}$ is ergodic;

(b) Let $H \neq \mathbb{R}^d$. We want to show that T is not stably ergodic. We represent $\mathbb{T}^d$ as $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Without the loss of generality we may assume that

$$H \bigcap \mathbb{Z}^d = \{0\} \tag{15}$$

since this could always be ensured by a passage to a factor group. Denote $E' = \mathbb{R}^d/H$. Denote

$$\Theta(x,\vec{y},\vec{e}) = \sum_{j=1}^{\infty} \left(\partial_{(d\sigma^{-j}\vec{e})}\,\tau\right)(y_j) \tag{16}$$

Here in $(d\sigma^{-j})\vec{e}$ we take the local branch of σ^{-j} corresponding to y_j. Let Θ' be the projection of Θ to E'. By assumption Θ' does not depend on $\vec{y}$ and so it defines a 1-form on M. Being the uniform sum of closed forms (locally we can invert σ and write $\Theta' = \lim_{n\to\infty} \sum_{j=1}^{n} d(\tau' \circ \sigma^{-j})$), Θ' is closed (τ' denotes the image of τ in E'). Also from the last identity it is clear that

$$\sigma^*(\Theta') = \Theta' + d(\tau'). \tag{17}$$

Hence cohomology class of Θ' is σ–invariant. Since $1 \notin \mathrm{Sp}(\sigma^*)$ (see Proposition 3.1 and the discussion thereafter), Θ' is closed, $\Theta' = d\alpha'$. Hence the previous equation reads $d(\alpha' \circ \sigma - \alpha') = \tau'$. Let $\alpha(x)$ be some preimage of α' in $\mathbb{R}^d$. Let x_0 be the fixed point of σ. Let us make a change of variables $(x,t) \to (x, t-\alpha(x))$. After this change τ is replaced by τ^* where $\tau^*(x) - \tau^*(x_0) \in H/\mathbb{Z}^d$. In particular by (15) τ^* is homotopic to a constant map, so it can be written in the form $\tau^* = \pi(\hat{\tau}^*)$ where $\hat{\tau}^* \colon M \to \mathbb{R}^d$. Now by a small perturbation we can pass from τ to $\tilde{\tau}$, where $\tilde{\tau} = \pi(\hat{\tilde{\tau}})$, where $\hat{\tilde{\tau}} \in \hat{\tilde{\tau}}(q_0) + \tilde{H}$, where $\hat{\tilde{\tau}}(q_0)$ has rational component and $\dim(\tilde{H}) = \dim(H)$ and $\tilde{H}$ is generated by rational vectors. But then $T(\tilde{\tau})$ is not ergodic. □

Lemma 3.14. *Complete uniform non-integrability is equivalent to stable ergodicity.*

PROOF: (a) If $\mathfrak{h} = \mathfrak{g}$ then T is stably ergodic as in the proof of Lemma 3.13.

(b) Let T be ergodic. Then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ and since $\bar{\Gamma}_t = \mathbf{H}$ we see that if $\mathfrak{h} \neq \mathfrak{g}$, then $\mathfrak{h}/[\mathfrak{g},\mathfrak{g}] \neq \mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$. But in this case the maximal abelian subextension T_a of T is not stably ergodic. So arbitrary close to T_a there is a non-ergodic extension $\tilde{T}_a$. But since T_a lifts to $M \times G$ so does $\tilde{T}_a$. □

3.4. **Decay of correlations.** In this subsection we prove Theorem 3.2. Let μ be the smooth invariant measure for σ. Let $\theta = \frac{1}{\min_x ||d\sigma(x)||}$. Denote by λ_0 the minimal eigenvalue of Δ on X.

In view of (12) we have to find bounds for the transfer operator

$$(\mathcal{L}_\lambda \vec{H})(x) = \sum_{\sigma y = x} e^{f(y)} \pi_\lambda(\tau(y)) \vec{H}(y)$$

where $e^{f(y)} = \frac{d\mu(y)}{d\mu(x)}$. We need an auxiliary estimate.

Proposition 3.15. *Given a branch $y = \sigma^{-n}x$ the following estimate holds*

$$\left\| \frac{d\tau_n(y)}{dx} \right\|_1 \leq \frac{\theta}{1-\theta} ||\tau||_1.$$

PROOF:

$$\left\|\frac{d\tau_n(y)}{dx}\right\| \leq \sum_{j=0}^{n-1}\left\|\frac{d\tau}{d\sigma^j y}\right\|\left\|\frac{d\sigma^j y}{dx}\right\| \leq \sum_{j=0}^{n-1}||\tau||_1\theta^j \leq \frac{\theta}{1-\theta}||\tau||_1. \quad \square$$

Let

$$\kappa = 4\theta\left(\frac{||f||}{\lambda_0}+\frac{1}{1-\theta}\right). \tag{18}$$

Introduce a norm $\|\vec{H}\|_\lambda = \max(\|\vec{H}\|_0, \frac{\kappa\|D\vec{H}\|_0}{\lambda})$. The following estimate is analogous to [22].

Proposition 3.16. *If* $n\ \frac{\theta^n}{\lambda_0} < \frac{1}{4}$ *then* $\forall\lambda \neq 0\ \|\mathcal{L}_\lambda^n \vec{H}\|_\lambda \leq \|H\|_\lambda$.

PROOF: We have

$$||\mathcal{L_\lambda}^n\vec{H}||_0 \leq \sum_{\sigma^n y=x} e^{f_n(y)}||\vec{H}||_0 \leq ||\vec{H}||_0||\mathcal{L}_0^n 1|| = ||\vec{H}||_0.$$

Now let us estimate the derivative.

$$\left|\frac{d(\mathcal{L_\lambda}^n\vec{H})}{dx}\right| \leq \sum_{\sigma^n y=x} e^{f_n(y)}\left|\frac{df}{dy}\right|\left|\frac{dy}{dx}\right|||\vec{H}||+$$

$$\sum_{\sigma^n y=x} e^{f_n(y)}\pi_\lambda(\frac{d\tau_n}{dx})||\vec{H}|| + \sum_{\sigma^n y=x} e^{f_n(y)}\left\|\frac{d\vec{H}}{dy}\right\|\left|\frac{dy}{dx}\right| \leq$$

$$\theta^n||f||||\vec{H}||_0 + \frac{\theta}{1-\theta}\lambda||\tau||||\vec{H}||_0 + \theta^n\left\|\frac{d\vec{H}}{dx}\right\|_0 \tag{19}$$

(here the inequality $||\pi_\lambda(Z)|| \leq \lambda||Z||$ was used to estimate the second term.) Thus

$$\frac{d(\mathcal{L_\lambda}^n\vec{H})}{dx} \leq \left(\theta^n||f|| + \lambda\frac{\theta}{1-\theta}||\tau||\right)||\vec{H}||_0 + \theta^n||\frac{d\vec{H}}{dx}||_0 \quad \square.$$

We need an auxiliary estimate.

Lemma 3.17. *There are constants* $n_1, \varepsilon_0, \varepsilon_1, \varepsilon_2$, *an open set* $U \subset M$ *and vectorfields* $e_1(x), e_2(x)\dots e_l(x)$ *such that* $\frac{1}{2} \leq \|e_j(x)\| \leq 1$ *and for any* $N \geq n_1$ *there are inverse branches*

$$y_{11}(x), y_{12}(x), y_{21}(x), y_{22}(x)\dots y_{l1}(x), y_{l2}(x)$$

of σ^N *such that* $\forall\vec{H}\exists j : \forall\vec{H}^* : \|\vec{H}^* - \vec{H}\| \leq \varepsilon_0$ *the following is true. Let*

$$\Theta_{jk}^N(x) = \pi_\lambda[\partial_{e_j}^{x'}\left(\tau_N(y_{jk}(x)\tau_N^{-1}(y_{jk}(x'))\right)_{x'=x}]$$

then

$$\varepsilon_1\lambda \leq \|[\Theta_{j1}^N(x) - \Theta_{j2}^N(x)]\vec{H}^*\| \leq \varepsilon_2\lambda$$

.

PROOF: Fix $x_0 \in F$. By assumption $\exists Z_1, Z_2 \dots Z_l \in \mathfrak{h}(x_0)$ which span $L(G)$. Let $\mathbf{D}$ denote a Casimir operator. As $\pi_\lambda(\mathbf{D})\vec{H} = \lambda^2\vec{H}$ $\exists j : \|\pi_\lambda(Z_j)\vec{H}\| \geq \varepsilon_3\lambda$. Since always $\|\pi_\lambda(Z)\tilde{\vec{H}}\| \leq \lambda\|L\|\|\tilde{\vec{H}}\|$ we have $\|\pi_\lambda(\tilde{Z})\vec{H}^*\| \geq \frac{\varepsilon_3}{2}\lambda$ for $\tilde{Z}$ close to Z_j and $\vec{H}^*$ close to $\vec{H}$. Now $\exists \vec{y}^{j1}, \vec{y}^{j2}, e_j$ such that

$$\lim_{N\to\infty} \Theta_{j1}^N(x_0) - \Theta_{j1}^N(x_0) = Z_j.$$

Thus if N is large, $e_j(x)$ is close to e_j and x close to x_0 then $\Theta_{j1}^N(x) - \Theta_{j2}(x)$ is close to Z_j. □

We need more notation. Let n_2 be a number such that $\sigma^{n_2}U = M$. Define

$$\hat{C} = \sup_j ||d\sigma^{n_2} e_j||_{C^2}, \tag{20}$$

$$E = 2\left(\frac{\theta}{1-\theta} + 2\right), \tag{21}$$

$$N = \frac{16E+8}{\varepsilon_1} \tag{22}$$

$$\delta = \frac{4}{N^2\hat{C}} \tag{23}$$

Let ε_4 be a number such that

$$\varepsilon_4 < \frac{1}{4} \tag{24}$$

and if $\vec{Z}_1, \vec{Z}_2$ are two vectors such that

$$||\vec{Z}_1|| \geq \frac{3}{64}||\vec{Z}_2||$$

and

$$||\vec{Z}_1 + \vec{Z}_2|| \geq ||\vec{Z}_1|| + (1-\varepsilon_4)||\vec{Z}_2||$$

then

$$\left\|\frac{\vec{Z}_1}{||\vec{Z}_1||} - \frac{\vec{Z}_2}{||\vec{Z}_2||}\right\| \leq \delta \tag{25}$$

Set $n_0 = n_2 + n_3$. where a number $n_3 \geq n_1$ is such that the following inequalities hold

$$\frac{\theta^{n_0}}{\lambda_0} < \frac{1}{4}, \tag{26}$$

$$N\theta^{n_0} \leq \frac{1}{8\kappa}, \tag{27}$$

$$N\delta\theta^{n_0} \leq \frac{\ln 2}{2}, \tag{28}$$

$$N\delta\theta^{n_0} \leq \frac{1}{32\kappa}. \tag{29}$$

Let $w(x)$ be a branch of $\sigma^{-n_2} : M \to U$ and set $z_{jk} = y_{jk}^{(n_3)} \circ w$.

We now follow a construction of [22]. Let us recall it. Divide M into "cubes" of diameter $\frac{\delta}{\lambda}$: $M = \bigcup_t \mathcal{C}_t(\lambda)$. (Here by cube in M we mean an image of a cube in $\mathbb{R}^d$ under the covering map.)

We now want to improve upon the estimate of Proposition 3.16. Let

$$\mathcal{K}_A = \{R : M \to \mathbb{R} : \|\partial_x \ln R\| \leq A\}.$$

Lemma 3.18. *There exist $\varepsilon, \bar{n}$ so that for given λ there are linear operators $\mathcal{N}_1(\lambda), \mathcal{N}_2(\lambda) \dots \mathcal{N}_{l(\lambda)}(\lambda)$ preserving $\mathcal{K}_{2\lambda}$ and such that*
a) For $R \in \mathcal{K}_{2\lambda}$

$$\int |\mathcal{N}_j R|^2 \, d\nu \leq (1-\varepsilon) \int R^2 \, d\nu;$$

b) If $|\vec{H}(x)| \leq R(x)$, $\|D\vec{H}(x)\| \leq 2\kappa\lambda R(x)$ for some $R \in \mathcal{K}_{2\lambda}$ then there exist $j = j(H,R)$ such that

$$|\mathcal{L}_\lambda{}^{\bar{n}} \vec{H}(x)| \leq (\mathcal{N}_j(R))(x) \tag{30}$$

and

$$\|D(\mathcal{L}_\lambda{}^{\bar{n}} \vec{H})(x)\| \leq 2\kappa\lambda(\mathcal{N}_j R)(x). \tag{31}$$

To prove this lemma we need several auxiliary estimates.

Take a cutoff function $\phi_t(x)$ satisfying
(a) $\operatorname{supp} \phi_t \in C_t$;
(b) $\phi_t(x) = 1$ if $x \in C_t$ and $\operatorname{dist}(x, \partial C_t) > \frac{\delta}{8\lambda}$;
(c) $\|\phi_t\|_1 \leq C\lambda$.

Set $\phi_{tjk} = \phi_t \circ z_{jk}^{-1}$. If J is a set of indices let $\phi_J = \sum_{(tjk)\in J} \phi_{tjk}$. Set $\mathcal{N}^{(J,\varepsilon_4)} R = \mathcal{L}_f^{n_0}((1 - \varepsilon_4 \phi_J)R)$. Call J $N-$dense if $\forall t \exists t' \in J$ such that $\operatorname{dist}(\mathcal{C}_t, \mathcal{C}_t') < \frac{N\delta}{\lambda}$. The following result is essentially proven in [22].

Proposition 3.19.

$$(1) \quad \mathcal{N}^{(J,\varepsilon_4)} : \mathcal{K}_{2\lambda} \to \mathcal{K}_{2\lambda}; \tag{32}$$

(2) If $\|\vec{H}\|(x) \leq R(x)$, $\|D\vec{H}\|(x) \leq 2\kappa\lambda R(x)$ then

$$\|D(\mathcal{L}_\lambda{}^{n_0} \vec{H})\|(x) \leq 2\kappa\lambda \mathcal{N}^{(J,\varepsilon_4)} R(x); \tag{33}$$

(3) $\exists \varepsilon_5$ such that if J is $N-$dense and $R \in \mathcal{K}_{2\lambda}$, then

$$\int (\mathcal{N}^{(J,\varepsilon_4)} R)^2 d\mu \leq (1-\varepsilon_5) \int R^2 d\mu.$$

PROOF: (1) and (3) are established in [22] (note that (1) and (3) deal with functions $M \to \mathbb{R}$ rather than $M \to H_\lambda(X)$.) (2) follows from (19) and the second condition of Proposition 32 by the same calculation as in Proposition 3.16. $\square$

We want to find $N-$dense J so that $\mathcal{N}^{(J,\varepsilon_4)}R$ satisfies (31). Let

$$\rho^\epsilon_{jk}(x) = \frac{\sum_{jk} e^{f_{n_0}(z_jk)}\pi_\lambda(\tau_{n_0}z_{jk})H(z_{jk})}{\left(\sum_{jk} e^{f_{n_0}(z_jk)}R(z_{jk})\right) - \epsilon e^{f_{n_0}(z_{j_0k_0})}R(z_{j_0k_0})}.$$

Call $\mathcal{C}_t$ good if $\exists j_0(t), k_0(t)$ so that on $\mathcal{C}_t$ $\|\rho^{\varepsilon_4}_{j_0k_0}\| \leq 1$.

Let $\mathcal{C} = \bigcup_{\mathrm{dist}(\mathcal{C}_{t'}\mathcal{C}_t)\leq \frac{N\delta}{|\lambda|}} \mathcal{C}_{t'}$.

Definition. If Φ is a function on a set U let

$$\mathrm{Osc}_U\Phi = \max_U \Phi - \min_U \Phi.$$

Proposition 3.20. *Let $\vec{H}, R$ satisfy $\|\vec{H}\| \leq R$, $\|\vec{H}'\| \leq 2\kappa\lambda R$, $R \in \mathcal{K}_{2\lambda}$.*
a) $\forall x, x' \in \mathcal{C}$, $\forall j, k$

$$\frac{1}{2} \leq \frac{R(z_{jk}(x))}{R(z_{jk}(x'))} \leq 2; \tag{34}$$

b) Fix j, k. Then either

$$\|\vec{H}(z_{jk}(x)\| \leq \frac{3}{4}R(z_{jk}(x))\forall x \in \mathcal{C} \tag{35}$$

$$\textit{or} \ \ \|\vec{H}(z_{jk}(x))\| \geq \frac{1}{4}R(z_{jk}(x).\forall x \in \mathcal{C} \tag{36}$$

(c) Moreover if (36) holds then

$$\|\vec{H}(z_{jk}(x)) - \vec{H}(z_{jk}(x'))\| \leq \delta\|\vec{H}(z_{jk}(x))\|. \tag{37}$$

and

$$\left\|\frac{\vec{H}(z_{jk}(x))}{||\vec{H}(z_{jk}(x))||} - \frac{\vec{H}(z_{jk}(x'))}{||\vec{H}(z_{jk}(x'))||}\right\| \leq 2\delta \tag{38}$$

PROOF: (a) We have $|\frac{d}{dz}\ln R| \leq 2\lambda\theta^{n_0}$. Thus

$$\mathrm{Osc}_{\mathcal{C}}(\ln R) \leq 2\lambda\theta^{n_0}\frac{N\delta}{\lambda} = 2N\delta\theta^{n_0}.$$

By (28) the oscillation of $\ln R$ on $\mathcal{C}$ is less than $\ln 2$ as claimed;

(b) Suppose there is a point $\tilde{x}$ such that

$$||\vec{H}(z_{jk}(\tilde{x})|| \leq \frac{1}{4}R(z_{jk}(\tilde{x})).$$

Then $\forall x \in \mathcal{C}$

$$||\frac{d\vec{H}}{dx}(x)|| \leq 2\kappa\lambda \cdot 2R(z_{jk}(\tilde{x}))|\frac{dz}{dx}| \leq 4\kappa\lambda R(z_{jk}(\tilde{x}))\theta^{n_0}.$$

Thus $\forall x \in \mathcal{C}$

$$||\vec{H}(z_{jk}(x))|| \leq ||\vec{H}(z_{jk}(\tilde{x}))|| + \frac{N\delta}{\lambda}4\kappa\lambda R(z_{jk}(\tilde{x}))\theta^{n_0} \leq \tag{39}$$

$$\left(\frac{1}{4} + 4\kappa N\delta\theta^{n_0}\right) R(z_{jk}(\tilde{x})) \leq \frac{3}{8}R(z_{jk}(\tilde{x}))$$

(the last inequality holds since (29) implies that $4\kappa N\delta\theta^{n_0} < \frac{1}{8}$.)

But

$$\frac{3}{8}R(z_{jk}(\tilde{x})) \leq \frac{3}{4}R(z_{jk}(x))$$

by (a);

(c) (37) follows from (39) and (27). (38) follows from (37) because

$$\left\|\frac{\vec{H}(z_{jk}(x))}{||\vec{H}(z_{jk}(x))||} - \frac{\vec{H}(z_{jk}(x'))}{||\vec{H}(z_{jk}(x'))||}\right\| \leq$$

$$\left\|\frac{\vec{H}(z_{jk}(x))}{||\vec{H}(z_{jk}(x))||} - \frac{\vec{H}(z_{jk}(x'))}{||\vec{H}(z_{jk}(x))||}\right\| + \left\|\frac{\vec{H}(z_{jk}(x'))}{||\vec{H}(z_{jk}(x))||} - \frac{\vec{H}(z_{jk}(x'))}{||\vec{H}(z_{jk}(x'))||}\right\| \leq$$

$$\delta + \frac{||\vec{H}(z_{jk}(x'))||}{||\vec{H}(z_{jk}(x))||||\vec{H}(z_{jk}(x'))||}||\vec{H}(z_{jk}(x')) - \vec{H}(z_{jk}(x))|| \leq 2\delta. \quad \square$$

Lemma 3.21. *$\forall t$ $\exists t'$ such that $\mathcal{C}_{t'}$ is good and* $\mathrm{dist}(\mathcal{C}_t, \mathcal{C}_{t'}) \leq \frac{N\delta}{\lambda}$.

Proof: If for some j_0, k_0 alternative (35) holds then $\|\rho_{j_0,k_0}^{\frac{1}{4}}(x)\| \leq 1$, so we may assume that (36) is always true. We assume that no $\mathcal{C}_s \subset \mathcal{C}$ is good and get a contradiction. So suppose that $\forall s, j, k$ $\exists x(s, j, k) \in C_s$ such that $\rho_{jk}^{\varepsilon_4}(x(s, j, k)) > 1$. Take some $x_0 \in \mathcal{C}$ and choose j_0, k_0 such that $R(z_{j_0k_0}(x_0))$ is the smallest. (34) implies that $\forall x, j, k$

$$R(z_{j_0k_0}(x)) \leq 4R(z_{jk}(x)).$$

Let $\vec{M}_{jk}(x) = \frac{\vec{H}(z_{jk}(x))}{\|\vec{H}(z_{jk}(x))\|}$, $\vec{K}(x) = \pi_\lambda(\tau_{n_0}(z_{j_0k_0}(x)))\vec{M}_{j_0k_0}(x)$. (35) and (25) now give

$$\|\pi_\lambda(\tau_{n_0}(z_{jk}(x(t, j_0, k_0)))\vec{M}_{jk}(x(t, j_0, k_0)) - \vec{K}(x(t, j_0, k_0))\| \leq \delta$$

Proposition 3.15 and (38) now imply that $\forall j, k$

$$\mathrm{Osc}_{\mathcal{C}_s}(\pi_\lambda(\tau_{n_0}(z_{jk}(x)\vec{M}_{jk}(x) \leq \frac{E\delta}{2}$$

so $\forall x \in C_s$

$$\|\pi_\lambda(\tau_{n_0}(z_{jk}(x))\vec{M}_{jk}(x) - \vec{K}(x)\| \leq E\delta \tag{40}$$

where $E = 2(\frac{\theta}{1-\theta}+1)$ by (37) and Proposition 3.15. Since s is arbitrary this holds for all $x \in \mathcal{C}$. Hence $\forall x, x' \in \mathcal{C}$ $\forall j$

$$\|\vec{K}(x') - \pi_\lambda(\tau_{n_0}(z_{j1}(x'))\tau_{n_0}^{-1}(z_{j1}(x)))\vec{K}(x)\| \leq$$

$$\|\vec{K}(x') - \pi_\lambda(\tau_{n_0}(z_{j1}(x'))\vec{M}_{j1}(x')\| + \|\vec{M}_{j1}(x') - \pi_\lambda(\tau_{n_0}^{-1}(z_{j1}(x)))\vec{K}(x)\| \leq$$

The first term here can be bounded by $E\delta$ while the second one is less than

$$\|\vec{M}_{j1}(x') - \vec{M}_{j1}(x)\| + \|\vec{M}_{j1}(x) - \tau_{n_0}^{-1}(z_{j1}(x)))\vec{K}(x)\| \leq (E+1)\delta.$$

Hence

$$\|\vec{K}(x') - \pi_\lambda(\tau_{n_0}(z_{j1}(x'))\tau_{n_0}^{-1}(z_{j1}(x)))\vec{K}(x)\| \leq (2E+1)\delta.$$

By the same token

$$\|\vec{K}(x') - \pi_\lambda(\tau_{n_0}(z_{j2}(x'))\tau_{n_0}^{-1}(z_{j2}(x)))\vec{K}(x)\| \leq (2E+1)\delta.$$

Therefore

$$\|\left[\pi_\lambda(\tau_{n_0}(z_{j1}(x'))\tau_{n_0}^{-1}(z_{j1}(x))) - \pi_\lambda(\tau_{n_0}(z_{j2}(x'))\tau_{n_0}^{-1}(z_{j2}(x)))\right]\vec{K}(x)\| \\ \leq (4E+2)\delta. \tag{41}$$

Now let j, e_j be as in Lemma 3.17 with $\vec{H} = \vec{K}(x_0)$. Set $\tilde{e}_j = d\sigma^{n_2}e_j$ and let x be obtained from x_0 by shifting along the flowlines of $\tilde{e}_j$ on distance $\frac{N\delta}{\lambda}$. Let $x(t)$ be this flowline. Let $\vec{H}(t) =$

$$\left[\pi_\lambda(\tau_{n_0}(z_{j1}(x))\tau_{n_0}^{-1}(z_{j1}(x_0))) - \pi_\lambda(\tau_{n_0}(z_{j2}(x))\tau_{n_0}^{-1}(z_{j2}(x_0)))\right]\vec{K}(x_0).$$

Then $\vec{H}(0) = 0$,

$$(\partial_t\vec{H})(0) = \left[\Theta_{j1}^{n_3} - \Theta_{j2}^{n_3}\right](w(x_0))\vec{K}(x_0)$$

and $(\partial_t{}^2\vec{H})(t) = \pi_\lambda(Y(t))\vec{K}(x_0)$ where $Y(t)$ is a second order differential operator and $\|Y(t)\| \leq$ Const. So

$$\|(\partial_t{}^2\vec{H})(t)\| \leq \hat{C}\lambda^2. \tag{42}$$

Hence

$$\vec{H}\left(\frac{N\delta}{\lambda}\right) = \frac{N\delta}{\lambda}\left[\Theta_{j1}^{n_3} - \Theta_{j2}^{n_3}\right](w(x_0))\vec{K}(x_0) + r \tag{43}$$

where $r < \hat{C}(N\delta^2)$. Hence

$$\left\|\vec{H}\left(\frac{N\delta}{\lambda}\right)\right\| \geq N\delta\varepsilon_1 - \hat{C}(N\delta)^2. \tag{44}$$

Now by (22) $N\delta\varepsilon_1 \geq (16E+8)\delta$ whereas by (23) $\hat{C}N^2\delta^2 = 4\delta$. Thus (44) contradicts (41), which proves Lemma 3.21. □

PROOF OF LEMMA 3.18: Set $J = \{(j_0(t), k_0(t), t) : \mathcal{C}_t \text{ is good }\}$. Then $\mathcal{N}^{J,\varepsilon_4}$ satisfies conditions (a) and (b) of Lemma 3.18 and so this lemma is established. □

PROOF OF THE THEOREM 3.2: Define recursively $R_0 \equiv \|A_\lambda\|_\lambda \cdot 1$, $R_{s+1} = \mathcal{N}^{J(R_s, {\mathcal{L}_\lambda}^{n_0 s} A_\lambda), \varepsilon_4} R_s$ then $\|{\mathcal{L}_\lambda}^{n_0 s} A_\lambda\|(x) \leq R_s(x)$ and so

$$\int ({\mathcal{L}_\lambda}^{n_0 s} A_\lambda) B_\lambda d\mu \leq \left(\int R_s^2 d\mu\right)^{\frac{1}{2}} \left(\int |B_\lambda|^2 d\mu\right)^{\frac{1}{2}} \leq$$
$$(1-\varepsilon)^s \|A_\lambda\| \|B_\lambda\|. \quad \square$$

3.5. **Characterization of exponential mixing.** In this subsection we again assume that $X = G$. We will finish the proof of Theorem 3.3 by establishing the following result.

Proposition 3.22. *If $X = G$ and T is exponentially mixing, then it is completely uniformly non-integrable.*

PROOF: Assume T is ergodic but $\mathfrak{h} \neq \mathfrak{g}$. We must show that T does not mix exponentially. Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$, where $\mathfrak{g}_1$ is the center of $\mathfrak{g}$, $\mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}]$. Since $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ and H contains $[G, G]$ by Corollary 2.6, we see that $\mathfrak{h} = \tilde{\mathfrak{g}} \oplus \mathfrak{g}_2$ where $\tilde{\mathfrak{g}} \neq \mathfrak{g}_1$. In this case we show that T has poor ergodic properties even on $G/[G, G]$ so we can assume from the beginning that $G = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ and $\tilde{\mathfrak{g}} \bigcap \mathbb{Z}^d = \{0\}$. We can regard $\tilde{\mathfrak{g}}$ as a subspace of $\mathbb{R}^d$. Let $P : M \times \mathbb{T}^d \to \mathbb{T}^d$ be the natural projection. The proof of Lemma 3.13 shows that we can obtain $\tau(x) = a_0 + \alpha(x)$, $\alpha(x) \in \tilde{\mathfrak{g}}$ by a coordinate change. Let $A(x, t) = \phi_1(t)$, $B(x, t) = \phi_2(t)$, where $\phi_1, \phi_2 \geq 0$, $\int \phi_1(t)\, dt = \int \phi_2(t)\, dt = 1$ $\operatorname{supp} \phi_1 \subset \mathcal{B}(x_1, \varepsilon)$, $\operatorname{supp} \phi_2 \subset \mathcal{B}(x_2, \varepsilon)$, $\|\phi_1\|, \|\phi_2\| \leq \varepsilon^{-N}$. If T were exponentially mixing, there would be a constant C such that $\int A(q) B(T^{C \ln \frac{1}{\varepsilon}} q)\, d\mu(q) > 0$ and therefore $P\mathcal{B}(x_1, \varepsilon) \bigcap \mathcal{B}(x_2, \varepsilon) \neq \emptyset$. Thus $P(T^{C \ln \frac{1}{\varepsilon}} P^{-1}(x_1))$ is a 2ε–net in $\mathbb{T}^d$ i.e $P(T^{C \ln \frac{1}{\varepsilon}} \mathcal{B}(x_1, 2\varepsilon)) = \mathbb{T}^d$. However the pullback of this set to $\mathbb{R}^d$ is contained in 2ε neighborhood of the ball in $C a_0 \ln \frac{1}{\varepsilon} + \tilde{\mathfrak{g}}$ centered at $C a_0 \ln \frac{1}{\varepsilon} + x_0$ and of radius $\mathrm{Const}\, a_0 \ln \frac{1}{\varepsilon}$. So its volume tends to 0 as $\varepsilon \to 0$, a contradiction. Hence T does not mix exponentially. □

3.6. **Prevalence of complete non-integrability.** In this subsection we prove Theorem 3.4.

Proposition 3.23. *If*

$$\bar{\Gamma}_e/\mathrm{Center}(G) = G/\mathrm{Center}(G) \tag{45}$$

and

$$\mathfrak{h}/[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \tag{46}$$

then $\mathfrak{h} = \mathfrak{g}$.

Proof: By (45), $[\bar{\Gamma}_e, \bar{\Gamma}_e] = [G, G]$ so $\mathfrak{h}$ is $[G, G]$ invariant. Applying Lemma 3.11 to $G/\text{Center}(G)$ we obtain $h/\text{Center}(\mathfrak{g}) = \mathfrak{g}/\text{Center}(\mathfrak{g})$ and by $[G, G]$ invariance, $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{h}$. This together with (46) implies that $\mathfrak{h} = \mathfrak{g}$. □

Proof of Theorem 3.4: By the above proposition we need to show that both (45) and (46) are violated at most on a manifold of a positive codimension.

(45): We can assume without loss of generality that G is semisimple. Let q_1 and q_2 be periodic points of periods n_1 and n_2 respectively and W be some t-chain joining q_1 and q_2. Then

$$\bar{\Gamma}_e(q_1) \supset < \{\tau_{n_1}(q_1), g(W)\tau_{n_2}(q_2)g^{-1}(W)\} > .$$

But the set of pairs $(g_1, g_2) \in G \times G$ such that $< g_1, g_2 > \neq G$ is an algebraic submanifold of positive codimension [25]. (Recall that $< g_1, g_2 >$ denotes the subgroup generated by g_1 and g_2.) Thus (45) is true generically.

(46): Here we can assume without loss of generality that $G = \mathbb{T}^d$. Denote

$$V(\vec{e}, vy, \vec{\tilde{y}}, x) = \partial_{\vec{e}}^{x'}[\Delta_u((x, \vec{y}), (x', \vec{y}')) - \Delta_u((x, \vec{\tilde{y}}), (x', \vec{\tilde{y}}'))]_{x'=x}$$

(see Subsection 3.2.) To show that generically $\mathfrak{h}(q) = \mathbb{R}^d$ it suffices to show that (always) for any $x \in M$

$$\text{Range}\left(\frac{dV}{d\tau}(\vec{e}, \vec{y}, \vec{\tilde{y}}, x)\right) = \mathbb{R}^d.$$

But

$$\frac{dV}{d\tau}(\vec{e}, \vec{y}, \vec{\tilde{y}}, x))(\delta\tau) = \Theta(\delta\tau, x, \vec{y}, \vec{e}) - \Theta(\delta\tau, x, \vec{\tilde{y}}, \vec{e})$$

where $\Theta(\tau, \dots)$ is defined by (16). Now let U be a small ball in M and y, $\tilde{y}$ be two sequences such that the preimages of x corresponding to y visit U exactly once (say $x_j = \sigma^{-j}x$) and no preimage of x corresponding to $\tilde{y}$ visits U. Let $\delta\tau = \phi(x)\vec{v}$, where $\text{supp}\,\phi \subset U$ and $\vec{v} \in \mathbb{R}^d$. Then $\frac{dV}{d\tau}(\vec{e}, \vec{y}, \vec{\tilde{y}}, x))(\delta\tau) = (\partial_{\sigma^{-j}\vec{e}}\phi)\vec{v}$ and such vectors span $\mathbb{R}^d$. □

4. Subshifts of finite type

4.1. **Content of this section.** In this section we study mixing rates of compact group extensions of one–sided subshifts of finite type.

The key notion of this section is that of Diophantine subset discussed in Appendix A. To state our results we need some auxiliary notation. If ω^0 is a two-sided sequence, let $\Gamma_t(\omega^0, l_1, l_2)$ be the set $\{g(W)\}$ for all t-chains $W = (\omega^1, \omega^1 \dots \omega^l)$, $\omega^l = \omega^0$ such that $l \leq l_1$ and if

$\omega^{j+1} \in W^s(\omega^j)$ then $(\sigma^{l_2}\omega^{j+1})_+ = (\sigma^{l_2}\omega^j)_+$ and if $\omega^{j+1} \in W^u(\omega^j)$ then $(\sigma^{-l_2}\omega^{j+1})_- = (\sigma^{-l_2}\omega^j)_-$, where ω_+ (ω_-) denotes $\{\omega_j\}_{j\geq 0}$ ($\{\omega_j\}_{j\leq 0}$).

Theorem 4.1. (Mixing) *If for some ω^0, l_1, l_2 $\Gamma_t(\omega^0, l_1, l_2)$ is Diophantine, then $\forall A, B \in C_{\theta,k}(\Sigma^+)$*

$$|\bar{\rho}_{A,B}(n)| \leq \mathrm{Const}\|A\|_k\|B\|_0 \left(\frac{1}{n}\right)^{\beta(k)}, \tag{47}$$

where $\beta(k) \to \infty$ as $k \to \infty$.

Theorem 4.2. (Characterization) *If (47) holds, then $\Gamma_t(\omega^0, l_1, l_2)$ is Diophantine for any ω^0 for large l_1, l_2.*

Theorem 4.3. (Prevalence) *The set of τ's such that $\Gamma_t(\omega, l_1, l_2)$ is Diophantine for large l_1, l_2, contains an open and dense subset of $C_\theta(\Sigma^+, G)$.*

Remark. This result can be rephrased by saying that a generic skew product over a one-sided subshift of finite type is stably rapidly mixing.

4.2. **Decay of correlations.** PROOF OF THEOREM 4.1. Without loss of generality we may assume that $\int A(x)dx = 0$. Let

$$\|\vec{H}\|_\lambda = \max\left(\|\vec{H}\|_0, \mathrm{Const}\frac{L(\vec{H})}{|\lambda|}\right)$$

where $L(\vec{H})$ denotes the Lipschitz constant of $\vec{H}$ as an element of $C_\theta(\Sigma^+, L^2(G))$ and Const is chosen in such a way that $\|\mathcal{L}_\lambda^n\|_\lambda \leq 1$ for large n (cf the proof of Proposition 3.16) We need the following estimate.

Proposition 4.4. $\|\mathcal{L}_\lambda^n\|_\lambda \leq \mathrm{Const}|\lambda|^{\beta_1}\left(1 - |\lambda|^{-\beta_2}\right)^n.$

Corollary 4.5. *If $A \in C_{k,\theta}$ and $\int A(\omega, x)d\mu(\omega)dx = 0$, then*

$$\|\mathcal{L}_\pi^n A\|_0 \leq \mathrm{Const}\|A\|_k n^{-\beta(k)}, \tag{48}$$

$\beta(k) \to \infty$ as $k \to \infty$.

Clearly this corollary proves Theorem 4.1. $\square$Let us first derive Corollary 4.5 from Proposition 4.4 and then return to the proof of the proposition.

PROOF OF THE COROLLARY: We have $L(A_\lambda) \leq \|A\|_k\mathrm{Const}|\lambda|$ and $\|A_\lambda\|_0 \leq \mathrm{Const}\|A\|_k|\lambda|^{-\beta_3}$ where $\beta_3 \to \infty$ as $k \to \infty$. Using the bound (see [37] or Proposition 3.16)

$$L(\mathcal{L}_\lambda^m\vec{H}) \leq \mathrm{Const}|\lambda|(|H\|_0 + \theta^n L(\vec{H}))$$

we obtain

$$L\left(\mathcal{L}_\lambda^{\mathrm{Const}\ln|\lambda|}\vec{H}\right) \leq |\lambda|^{-\beta_3+2}$$

if Const is large enough. Hence

$$\|\mathcal{L}_\lambda^n \vec{H}\| \leq \mathrm{Const}|\lambda|^{\beta_1}\left(1-|\lambda|^{-\beta_2}\right)^{n-\mathrm{Const}\ln|\lambda|}$$

Also we always have $\|\mathcal{L}_\lambda^n A_\lambda\|_0 \leq \|A_\lambda\|_0 \leq \|A\|_k ||\lambda||^{-\beta_3}$. Thus

$$\|\mathcal{L}_\pi^n A\|_0 \leq \sum_{|\lambda|\leq n^{\frac{1}{2}\beta_2}} \|\mathcal{L}_\lambda^n A_\lambda\|_0 + \sum_{|\lambda|>n^{\frac{1}{2}\beta_2}} \|A_\lambda\|_0$$

The first term is at most $\mathrm{Const}\|A\|_k e^{-\sqrt{n}}$ while the second does not exceed

$$\|A\|_k \sum_{|\lambda|>n^{\frac{1}{2}\beta_2}} |\lambda|^{-\beta_3} \leq \mathrm{Const}\|A\|_k n^{-\beta(k)}. \quad \square$$

Proof of the Lemma: Set $m(\lambda) = [C_1 \ln|\lambda|]$, $\tilde{m}(\lambda) = m(\lambda) + l_2$ where the restrictions on C_1 will be clear later (see Lemma 4.7).

Lemma 4.6. *If $\forall \vec{H}$ such that $\|\vec{H}\|_\lambda \leq 1$ $\exists \beta_4 > 0, \omega \in \Sigma$ and $m \leq \tilde{m}(\lambda)$ so that $\|\mathcal{L}_\lambda^m \vec{H}(\omega)\|_0 < 1 - |\lambda|^{-\beta_4}$ then the statement of Proposition 4.4 is true.*

Proof: Repeat the proof of Lemma 3 from [20]. (See also the proof of (1)⇒(4) in Theorem A.2.) $\square$

So we have to prove that for β_4 large enough the conditions of the Lemma are satisfied. So take some $\vec{H}$ with $||H||_\lambda \leq 1$. So we assume that

$$\|(\mathcal{L}_\lambda{}^m \vec{H})(\omega)\|_0 \geq 1 - |\lambda|^{-\beta_4} \tag{49}$$

for $m \leq \tilde{m}(\lambda)$ and get a contradiction. Consider two points $\bar{\omega}, \tilde{\omega}$ such that $\sigma^m \bar{\omega} = \sigma^m \tilde{\omega} = \omega$. Consider $(\mathcal{L}_\lambda^m \vec{H})(\omega)$.

Among other terms it contains

$$e^{f_m(\bar{\omega})}\pi_\lambda(\tau_m(\bar{\omega}))\vec{H}(\bar{\omega}) + e^{f_m(\tilde{\omega})}\pi_\lambda(\tau_m(\tilde{\omega}))\vec{H}(\tilde{\omega}).$$

(49) implies that

$$||e^{f_m(\bar{\omega})}\pi_\lambda(\tau_m(\bar{\omega}))\vec{H}(\bar{\omega}) + e^{f_m(\tilde{\omega})}\pi_\lambda(\tau_m(\tilde{\omega}))\vec{H}(\tilde{\omega})|| \geq$$

$$(1-\lambda^{-\beta_4})||e^{f_m(\bar{\omega})}\pi_\lambda(\tau_m(\bar{\omega}))\vec{H}(\bar{\omega})|| + ||e^{f_m(\tilde{\omega})}\pi_\lambda(\tau_m(\tilde{\omega}))\vec{H}(\tilde{\omega})||.$$

Therefore $e^{f_m(\bar{\omega})}\pi_\lambda(\tau_m(\bar{\omega}))\vec{H}(\bar{\omega})$ and $e^{f_m(\tilde{\omega})}\pi_\lambda(\tau_m(\tilde{\omega}))\vec{H}(\tilde{\omega})$ are almost collinear. That is

$$\|\pi_\lambda(\tau_m(\bar{\omega}))\vec{H}(\bar{\omega}) - \pi_\lambda(\tau_m(\tilde{\omega}))\vec{H}(\tilde{\omega})\| \leq C\lambda^{-\beta_5}, \tag{50}$$

$\beta_5 \to \infty$ as $\beta_4 \to \infty$. Denote by $\omega(m, j)$ an one-sided sequence

$$(\omega(m, j))_i = \omega^j_{i-m}.$$

Let $\vec{H}^j = \vec{H}(\omega(m(\lambda), j))$, $\vec{K}^j = \pi_\lambda(\tau_{m(\lambda)}(\omega(m(\lambda), j)))\vec{H}^j$. By assumption $\exists W$ such that $g(W)$ satisfies (60) with $\vec{H} = \vec{K}^0$.

Lemma 4.7. *If $\vec{H}$ satisfies (50) then*

$$\|\vec{K}^{j+1} - \pi_\lambda(g(\omega^j, \omega^{j+1}))\vec{K}^j\| \le C\lambda^{-\beta_6} \tag{51}$$

where $\beta_6 \to \infty$ as $C_1 \to \infty$.

Proof: Consider the following two cases:

(1) If $\omega^{j+1} \in W^s(\omega^j)$, then applying (50) with $\tilde{\omega} = \omega(m(\lambda), j)$, $\bar{\omega} = \omega(m(\lambda), j)$, $m = \tilde{m}(\lambda)$ we get

$$\|\vec{K}^{j+1} - \pi_\lambda(\tau^{-1}_{l_2}(\omega^{j+1}_+)\tau_{l_2}(\omega^j_+))\vec{K}^j\| \le C\lambda^{-\beta_5};$$

(2) If $\omega^{j+1} \in W^s(\omega^j)$, then

$$\|\vec{H}^{j+1} - \vec{H}\| \ \le \mathrm{Const}\theta^{m(\lambda)-l_2}\|H\|_\lambda \le \overline{\mathrm{Const}}\theta^{m(\lambda)}\lambda.$$

Using the relation between $\vec{H}^j$ and $\vec{K}^j$ we get

$$\|\vec{K}^j - \pi_\lambda\left(\tau^{-1}_{m(\lambda)}(\omega(m(\lambda), j+1))\tau_{m(\lambda)}(\omega(m(\lambda, j)))\right)\vec{K}^j\| \le \mathrm{Const}\theta^{m(\lambda)}\lambda$$

but

$$\left|\tau^{-1}_{m(\lambda)}(\omega(m(\lambda), j+1))\tau_{m(\lambda)}(\omega(m(\lambda, j))) - \Delta_u(\omega^j, \omega^{j+1})\right| \le \mathrm{Const}\theta^{m(\lambda)}$$

while $\|\pi_\lambda(g) - 1\| \le \lambda$ which completes the proof. □

Adding (51) for all j and using $\vec{K}^l = \vec{K}^0$ we get

$$\|\pi_\lambda(g(W))\vec{K}^0 - \vec{K}^0\| \le C(\beta_6)\lambda^{-\beta_6}l_1$$

where β_6 can be made as large as we wish by choosing C_1 large which contradicts to the Diophantine condition. Hence Proposition 4.4 is established. □

Corollary 4.8. *For G semisimple ergodicity implies rapid mixing.*

Proof: Γ_t acts transitively on X. Since

$$\bigcup_{l_1, l_2} \Gamma_t(\omega^0, l_1, l_2)$$

generate $\Gamma_t(\omega^0)$ Corollary A.5 shows that $\Gamma_t(\omega^0, l_1, l_2)$ is Diophantine for large l_1, l_2. □

4.3. **Characterization of rapid mixing.** PROOF OF THEOREM 4.2. By the results of Section 2, we can assume without the loss of generality that T is reduced. Because T is mixing the only way $\Gamma_t(\omega^0, l_1, l_2)$ can fail to be Diophantine is that $p(\Gamma_t(\omega^0, l_1, l_2))$ is not Diophantine on $[\bar{\Gamma}_t, \bar{\Gamma}_t]\backslash X$. Hence we may assume from the beginning that $G = X = \mathbf{T}^d$. Fix l_1, l_2. If $\Gamma_t(\omega^0, l_1, l_2)$ is not Diophantine $\forall\beta\ \exists\vec{m} : \forall W$

$$|\exp(2\pi i(\vec{m}, g(W))) - 1| \leq \frac{1}{|\vec{m}|^\beta}.$$

In particular (Proposition 2.4) $\forall\omega^1, \omega^2$

$$|\exp(2\pi i(\vec{m}, \tau(\omega^1) - \tau(\omega^2))) - 1| \leq \frac{1}{|\vec{m}|^\beta}.$$

Thus $\exists\varphi_0$ such that

$$|\exp(2\pi i(\vec{m}, \tau_n(\omega))) - \exp(2\pi i\varphi_0)| \leq \frac{n}{|\vec{m}|^\beta}.$$

If β is large enough this is incompatible to (47) with $A(\omega, t) = B(\omega, t) = e^{2\pi i(\vec{m}, t)}$. □

This statement has several nice corollaries.

Definition. Call T **locally transitive** if there are ω^0, l_1, l_2 such that $\Gamma_t(\omega^0, l_1, l_2)x_0 = X$ for any $x_0 \in X$.

Let T_a be the maximal abelian subextension of T.

Corollary 4.9. *If T is ergodic then it is rapidly mixing if and only if its maximal abelian subextension is rapidly mixing.*

PROOF: This follows immediately from Theorems 4.1, 4.2 and Corollary A.7. □

Corollary 4.10. *The property of rapid mixing does not depend on the Gibbs measure in the base.*

Corollary 4.11. *If T_a is locally transitive then $T \in \mathcal{RM}$.*

4.4. **Prevalence of rapid mixing.** This subsection complements the results of Appendix A in the following way. In the appendix we show that Diophantineness is generic in measure theoretic sense. However, for toral action, the opposite property is topologically generic. Therefore, even though most of finite sets are Diophantine, the corresponding constants behave rather irregularly which makes this result of limited value. However, our condition in Theorem 4.1 involves much larger group (namely, the group of all closed t-chains). This explains why Theorem 4.3 holds.

Proof of Theorem 4.3: Clearly it is enough to consider the action of G on itself by translations. We again reduce the problem to the toral case. Indeed it follows from the results of Appendix A, that for an open and dense set of extensions $\bar{\Gamma}_t \supset [G,G]$. Hence we can factor it out and end up with toral extensions as claimed. (We could also appeal here to [27] and Corollary 4.9.) So let $G = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. In this case $\Gamma_t(\omega)$ does not depend on ω by Proposition 2.2 so we will move the base point freely. We will consider the simplest t-chains of the form $W = (\omega^1, \omega^2, \omega^3, \omega^4)$ such that $\omega^1_0 = \omega^2_0 = \omega^3_0 = \omega^4_0$ and $\omega^1_+ = \omega^2_+$, $\omega^3_+ = \omega^4_+$, $\omega^1_- = \omega^4_-$, $\omega^2_- = \omega^3_-$. So we let

$$\varphi(\omega^1,\omega^2,\omega^3,\omega^4) = g(\omega^1,\omega^2,\omega^3,\omega^4) =$$

$$\sum_{j=-\infty}^{\infty} \left[\tau(\sigma^j\omega^1) - \tau(\sigma^j\omega^2) + \tau(\sigma^j\omega^3)\tau(\sigma^j\omega^4)\right]. \tag{52}$$

We will write φ_τ if it is not clear which skewing function is used. Recall that we consider the case $\tau \in C^+(\Sigma)$. The following bound is immediate.

Proposition 4.12. *If $\forall j,k\ d(\omega^j,\omega^k) \le \theta^N$ then*

$$\varphi(\omega^1,\omega^2,\omega^3,\omega^4) \le \mathrm{Const}\theta^{2N}\|\tau\|.$$

Proof: As $\omega^1_+ = \omega^2_+$, $\omega^3_+ = \omega^4_+$, all positive terms in (52) vanish. The first non-negative term corresponds to $j = -N$. $\square$

Fix some element α of our alphabet. If n_0 is large enough, we find $(d+2)$ periodic points $\omega^1 \dots \omega^{d+2}$ of prime period n_0 such that $\omega^j_0 = \alpha$ and their orbits do not intersect. Let $\alpha\kappa^1, \alpha\kappa^2, \dots \alpha\kappa^{d+2}$ be the corresponding words of length n_0. Let $\xi_+ = (\alpha\kappa^{d+1})_0^\infty$, $\xi_+ = (\alpha\kappa^{d+2})^0_{-\infty}$. Finally denote by $\mathcal{C}_{jN}$ the cylinders $\mathcal{C}_{jN} = \mathcal{C}_{(\alpha\kappa^j)^{2N+2}}$. We consider perturbations of τ of the form

$$\tilde{\tau} = \tau + \sum_{j=1}^{d} e_j \sum_{l=1}^{\infty} \varepsilon_{jl}\theta^{2ln_0}\chi_{\mathcal{C}_{jl}} \tag{53}$$

with $|\varepsilon_{jl}| \le \epsilon$. We prove that for any ϵ we can make $\tilde{\tau}$ satisfy the conditions of the proposition. We will choose parameters ε_{jl} by induction. Assume that we already defined $\{\varepsilon_{jl}\}_{l<N}$. Let

$$\omega^{jN1} = \xi_-(\alpha\kappa^j)^{N+1}|(\alpha\kappa^j)^{N+1}\xi_+, \quad \omega^{jN2} = \xi_-(\alpha\kappa^j)^{N}|(\alpha\kappa^j)^{N+1}\xi_+,$$
$$\omega^{jN3} = \xi_-(\alpha\kappa^j)^{N}|(\alpha\kappa^j)^{N}\xi_+, \quad \omega^{jN4} = \xi_-(\alpha\kappa^j)^{N+1}|(\alpha\kappa^j)^{N}\xi_+$$

(here | is used to mark the place before the zeroth letter). We have

$$\varphi_{\tilde{\tau}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}) =$$
$$\varphi_\tau(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}) + \varphi_{\tau(1,N-1)}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}) +$$

$$\varphi_{\tau^{(N,N)}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})+\varphi_{\tau^{(N+1,\infty)}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})$$

where the second term corresponds to the summation from 1 to $N-1$ in (53), the third one corresponds to the N-th term and the last one corresponds to the remainder. Now

$$\varphi_{\tau^{(N+1,\infty)}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})=0$$

as no ω^{jNk} contains $(\alpha\kappa^l)^{2N+4}$. Also

$$\varphi_{\tau^{(N,N)}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})=\varepsilon_{jN}\theta^{2Nn_0}e_j.$$

Let

$$\eta_{jN}=\frac{\varphi_{\tau}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})+\varphi_{\tau^{(1,N-1)}}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})}{\theta^{2Nn_0}}.$$

By Proposition 4.12 η_{jN} is less then some constant E. So

$$\varphi_{\tilde\tau}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})=\theta^{2Nn_0}\left(\eta_{jN}+\varepsilon_{jN}e_j\right).$$

The next statement follows immediately by compactness arguments.

Proposition 4.13. *Let d,ϵ be fixed. There is a constant δ such that we always can choose $\varepsilon_{jN}\in[-\varepsilon,\varepsilon]$ so that*

$$|\mathrm{Vol}(\eta_{1N}+\varepsilon_{1N}e_1,\dots,\eta_{dN}+\varepsilon_{dN}e_d)|\geq 2\delta.$$

Thus $\tilde\tau$ and its small perturbations will satisfy

$$\left|\mathrm{Vol}\left(\frac{\varphi_{\bar\tau}(\omega^{1N1},\omega^{1N2},\omega^{1N3},\omega^{1N1})}{\theta^{2Nn_0}},\dots,\frac{\varphi_{\bar\tau}(\omega^{dN1},\omega^{dN2},\omega^{dN3},\omega^{dN1})}{\theta^{2Nn_0}}\right)\right| \\ \geq\delta \tag{54}$$

Also if $\bar\tau$ is close to $\tilde\tau$ then

$$|\varphi_{\bar\tau}(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})|\leq 2E \tag{55}$$

We claim that (54) and (55) guarantee that the set

$$\{\varphi(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})\}_{jN=11}^{d\infty}$$

is Diophantine. Indeed take some $\vec m$. Let $K=\max m_j$. Take minimal N such that $E\theta^{2Nn_0}\leq\frac{1}{100dK}$. Then $|(\vec m,\varphi(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}))|\leq\frac{1}{50}$. So in order for $\exp[2\pi i(\vec m,\varphi(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}))]$ to be close to 1, this product has to be small. However this is impossible for all j. Indeed, if $|(\vec m,\varphi(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4}))|\leq\tilde\delta$ then all the vectors $\frac{\varphi(\omega^{jN1},\omega^{jN2},\omega^{jN3},\omega^{jN4})}{\theta^{2Nn_0}}$ are confined to the cylinder whose base has radius $2E$ and is perpendicular to $\vec m$ and whose height is $200Ed\tilde\delta$. If $\tilde\delta$ is small enough this is incompatible with (54). This completes the proof. □

Remark. In the toral case we consider the same perturbation as in [39] but we analyze its effect more carefully. In fact, for extensions over symbolic systems it is not true that stable ergodicity implies polynomial decay of correlations. For example, let Σ be the full two shift, $G = \mathbb{T}^1$ and $\tau(\omega) = \sum_j \theta^{n_j} \chi_{\mathcal{C}_{(1)^{n_j}}}(\omega)$. Then if n_j grow very fast, we can very well approximate τ by locally constant functions, so the decay of correlations in this example can be arbitrary slow. Instead, stably ergodic systems have the property that $\exists \gamma_n = \gamma_n(\tau) \to 0, k = k(\tau)$ such that if $\tilde{\tau}$ is close to τ then

$$|\bar{\rho}_{A,B}(n, \tilde{T})| \leq \|A\|_k \|B\|_k \gamma_n.$$

Nor it is true that rapid mixing is stable. For example, consider the set of T's with skewing function locally constant with fixed number of domains of the constancy (still $G = \mathbb{T}^1$).Then almost all T's in this set are rapidly mixing but the set of the transformations not having this property contains a countable intersection of open dense sets.

5. Axiom A.

Here we finish the proof of the theorems given in the introduction. Theorem 1.1 follows immediately from Theorems 4.1 and 4.2 via the reduction described in subsection 2.5. Likewise Corollary 1.3 follows from Corollaries 4.8 and A.6(b) and Corollary 1.4 follows from Corollary 4.9. To prove Corollary 1.2 more arguments are needed since a perturbation inside subshifts of finite type can be done more easily than for Anosov diffeomorphisms. We shall use the following result of Burns and Wilkinson:

Proposition 5.1. *([17], Theorems 9.1 and 12.1) Let $F : Y \to Y$ be an Anosov diffeomorphism of an infranilmanifold and T be a compact group extension of F with $X = G$. If T is stably ergodic then it is locally transitive.*

Proof of Corollary 1.2: If $T \in \mathrm{Int}(\mathcal{ERG})$ then, in particular, $T \in \mathcal{ERG}$ and so $T \in \mathcal{RM}$ if and only if $T_e \in \mathcal{RM}$. So we can assume from the beginning that $G = \mathbb{T}^d$. Then we can also suppose that $X = G$. So, after all reductions we have $T(y, x) = (F(y), \tau(y)x)$ where F is Anosov, Y is infranilmanifold and $X = G = \mathbb{T}^d$. In this case Proposition 5.1 shows that T is locally transitive and we are done by Corollary 4.11. ∎

6. Applications

Here we derive some consequences from our bounds for correlation decay. More applications will be presented elsewhere [23].

6.1. **Central Limit Theorem.**

Corollary 6.1. *Under the conditions of Theorem 1.1 there exists k such that $\forall A \in C^k(M)$ the sequence $\{A(T^n(y,x))\}$ satisfies Central Limit Theorem (CLT).*

PROOF: By (10) it is enough to prove CLT for extensions of subshifts of finite type satisfying Theorem 4.1 and $A \in C_{\theta,k}(\Sigma^+)$. Recall [35] that if T is an endomorphism of a measure space (M, ν) then the following conditions suffice for CLT

$$(a) \qquad \sum_n \Big| \int \bar{A}(m)\bar{A}(T^m)d\nu(m) \Big| \leq \infty \quad \text{and}$$

(b) $\sum_n (U^{*n}A)(m)$ converges uniformly, where U^* is the dual to

$$(UA)(m) = A(Tm).$$

In our case (a) follows by Theorem 4.1 and (b) follows by Corollary 4.5 since in our situation $U = \mathcal{L}_\pi$. $\square$

6.2. **Equidistribution of the leaves.**

Here we provide estimate for equidistribution of the images of local unstable manifolds under the conditions of Theorem 4.1. But first we should pass to functions in $C_\theta(\Sigma)$ rather then $C_{k,\theta}(\Sigma^+)$.

Corollary 6.2. *Under the conditions of Theorem 4.1 for any pair $A, B \in C_{k,\theta}(\Sigma)$*

$$|\bar{\rho}_{A,B}(N)| \leq C\|A\|_{k,\theta}\|B\|_{0,\theta}N^{-\theta(k)},$$

$\tilde{\beta}(k) \to \infty$ as $k \to \infty$.

PROOF: Plug the estimate of Theorem 4.1 in equation (9) . $\square$

Corollary 6.3. *Under the same conditions*

$$|\bar{\rho}_{A,B}(N)| \leq C\|A\|_{0,\theta}\|B\|_{k,\theta}N^{-\tilde{\beta}(k)},$$

PROOF: Replace T by T^{-1}. $\square$

Now we provide quantitative version of the K–property. Let

$$W^u_{loc}(\tilde{\omega}) = \{\omega : \omega_- = \tilde{\omega}_-\}.$$

On $\mathcal{C}_{\tilde{\omega}_0}$ write ([32]) $d\mu(\omega) = J(\omega)d\mu_+(\omega_+)d\mu_-(\omega_-)$. Denote $\int\limits_{W^u_{loc}(\tilde{\omega})} H = \int\limits_{W^u_{loc}(\tilde{\omega})} H(\omega)J(\omega)\, d\mu_+(\omega)$.

Proposition 6.4. *If $\alpha \in C_\theta(\Sigma, X)$ then*

$$\int_{W^u_{loc}} B(T^N(\omega, \alpha(\omega))) \to \int B(\omega, x) d\mu(\omega) dx.$$

PROOF: Let ϕ be a cutoff function concentrated on $[-1,1]^d$. Let $I_{\mathcal{C}}$ denote the indicator of $\mathcal{C}$. Set

$$A^{(n,\varepsilon)}(\omega, x) = \frac{I_{\mathcal{C}_{-n,0}(\tilde{\omega})}(\omega) \phi(\frac{(\exp^{-1}_{\alpha(\omega)} x}{\varepsilon}))}{J(\omega) \left(\int_X \phi(\frac{(\exp^{-1}_{\alpha(\omega)} x}{\varepsilon})) dx\right) \mu(\mathcal{C}_{-n,0}(\tilde{\omega}))},$$

Then $\|A\|_{\theta,0} \leq \text{Const} K^n \varepsilon^{-d}$ and $\int A^{(n,\varepsilon)}(\omega, x) d\mu(\omega) dx = 1 + O(\theta^n)$. Also if $(\omega, x) \in \operatorname{supp} A^{(n,\varepsilon)}$ then

$$|B(T^N(\omega, x)) - B(T^N([\tilde{\omega}, \omega], \alpha([\tilde{\omega}, \omega])| \leq \text{Const}(\theta^n + \varepsilon).$$

Therefore $\rho_{A^{(n,\varepsilon)},B}(N) = \int_{W^u_{loc}} B(T^N(\omega, \alpha(\omega)))(1 + O(\theta^n + \varepsilon))$. On the other hand

$$\rho_{A^{(n,\varepsilon)},B}(N) = \int B(\omega, x) d\mu(\omega) dx (1 + O(\theta^n + \varepsilon)) + O(K^n \varepsilon^{-d} \gamma_N)$$

where

$$\gamma_N = \sup \frac{|\bar{\rho}_{A,B}(N)|}{\|A\|_{\theta,0} \|B\|_{\theta,k}} \tag{56}$$

Comparing these two estimates we get

$$| \int_{W^u_{loc}} B(T^N(\omega, \alpha(\omega))) - \int B(\omega, x) d\mu(\omega) dx| \leq$$

$$\text{Const}(\theta^n + \varepsilon + O(K^n \varepsilon^{-d} \gamma_N). \quad \square \tag{57}$$

Remark. The above argument comes from [8] (cf. also [24]).

Corollary 6.5. *If $\alpha \in C_\theta(\Sigma, X)$ then*

$$| \int_{W^u_{loc}} B(T^N(\omega, \alpha(\omega))) - \int B(\omega, x) d\mu(\omega) dx| \leq \text{Const} \|B\|_{k,\theta} N^{-\tilde{\tilde{\beta}}(k)}.$$

PROOF: Use Corollary 6.3 and equations (56) and (57) . $\square$

6.3. **Random walks on homogeneous spaces.** Let $X = G/H$. Take a finite set $W = \{g_1, g_2 \dots g_d\} \subset G$ and let $\vec{p} = \{p_1, p_2 \dots p_d\}$ be a probability distribution on W. Consider a Markov chain with the initial distribution dx and $x_n = g_j x_{n-1}$ with the probability p_j. Denote by g_j^a and g_j^s the projections of g_j on $G/[G,G]$ and $G/\text{Center}(G)$ respectively. We say that x_n **satisfies CLT** if there is $r > 0$ such that for any function $A \in C^r(X)$ with zero mean $\frac{\sum_{j=0}^{n-1} A(x_j)}{\sqrt{n}}$ converges in distribution to a Gaussian random variable with zero expectation.

Proposition 6.6. *Our Markov chain satisfies Central Limit Theorem if and only if*

$$\langle W^s \rangle = G/\text{Center}(G) \tag{58}$$

and

$$\{g_j^a - g_k^a\} \quad \text{is} \quad \text{Diophantine} \tag{59}$$

PROOF: (1) Suppose that (58) and (59) are satisfied. Consider the subshift of finite type with alphabet W, transition matrix $\mathbf{A}_{jk} \equiv 1$ and measure $\mu(\mathcal{C}_{w_1 \dots w_n}) = \prod_{j=1}^n p_{w_j}$. Consider the skew extension with $\tau(\omega) = g_{\omega_0}$. Since T is reduced, Γ_e is generated by $\{g_j\}$ and Γ_t is generated by $\{g_j g_k^{-1}\}$. Thus T is ergodic, and by Corollary 4.9 and Theorem 4.1 it is rapidly mixing. Thus, by Corollary 6.1 x_n satisfies CLT.

(2) Let x_n satisfy CLT. Then $< W >$ is ergodic since otherwise there would exist a non-constant W invariant function A of zero mean and so $\frac{\sum_{j=0}^{n-1} A(x_j)}{n}$ would not converge to 0 in distribution. Thus (58) holds. If (59) would fail there would exist $m_l \to \infty$ such that

$$|\exp[2\pi i(m_l, (g_j^a - g_k^a))] - 1| \leq \frac{1}{m_l^l}.$$

By passing to a subsequence, we can assume that $m_{l+1} \geq m_l^{2r+8}$. Let $A(x) = \sum_l \varepsilon_l \frac{1}{m_l^{r+2}} \exp[2\pi i(m_l, x^a)]$, where $\varepsilon_l \in \{0, 1\}$ and x^a denotes the projection of x to $X/[G,G]$. Then

$$\sum_{j=1}^{m_l^{2r+6}} A(x_j) = \gamma_l(x_0) + \varepsilon_l m_l^{r+4} \exp[2\pi i(m_l, x_0^a)] + O(m_l^{-\zeta})$$

where $\zeta = (r+2)(2r+8) - (r+6) = 2r^2 + 11r + 10$ and γ_l depends only on $\varepsilon_1, \varepsilon_2 \dots \varepsilon_{l-1}$. Thus we can choose ε_l in such a way that

$$\text{Prob}\left(\frac{\sum_{j=1}^{m_l^{2r+6}} A(x_j)}{m_l^{r+3}} > \frac{m_l}{4}\right) \geq \frac{1}{2}$$

and so A does not satisfy CLT. □

7. Conclusions.

Here we describe how are results fit into the general theory of weakly hyperbolic dynamical systems and present some open questions related to this subject.

7.1. **Mixing rates of skew extensions of Axiom A diffeomorphisms.**

Question. *Is stable ergodicity (stable rapid mixing and so on) generic in the space of compact group extensions of Axiom A diffeomorphisms?*

This question is easier when the set of non–wandering points of the base transformation is large, for example, when it is connected (see [26].) On the other hand nothing seems to be known if the base is a horseshoe, especially in higher dimensions.

There are also some questions on the optimality of the bounds we have obtained.

Question. *Is exponential mixing generic among compact group extensions of (say, volume preserving) Anosov diffeomorphisms?*

Question. *If G is semisimple and T is mixing, is it also exponentially mixing?*

More generally, in the case when non-wandering set of F is large there are not so many situations where we can get an asymptotics of the mixing rate.

Problem. *Construct some examples where correlation functions could be computed explicitly.*

7.2. **Partially hyperbolic systems.** It is interesting to see how much of the theory presented here can be extended to general partially hyperbolic systems where the central bundle is generated by the orbits of some symmetry group. Examples of such systems include frame flows on compact negatively curved manifolds (or products of such manifolds) systems obtained by applying compact extension construction several times, e.g. nilpotent extensions, etc (see [15] for more discussion).

Problem. *Generalize the results of Section 2 to compact group extensions of partially hyperbolic systems with accessibility property.*

See [34] for some results along these lines.

More generally, here as well as in [20, 22], we showed how to derive mixing properties of transversely hyperbolic systems with symmetries from the the property of the holonomy maps along short loops (the first result in this direction appeared in [18]). One can ask what can be said about more general partially hyperbolic systems. The best result in this direction so far is a theorem of Pugh and Shub [41] saying that partially hyperbolic volume preserving, dynamically coherent and center-bunched systems with accessibility property are K-systems. In particular, they enjoy mixing of all orders.

Question. *Let $f : M \to M$ be a map satisfying conditions of Pugh and Shub and moreover is locally transitive. What can be said about its rate of mixing?*

In the context of skew extensions of Axiom A some results are given by Theorem 1.1 and Corollaries 1.2 and 1.3. Similarly one should compare results of [16] and [22].

In full generality this question seems to be very hard but an advance in this direction would drastically increase our understanding of partially hyperbolic systems.

Question. *Does there exist a stably ergodic diffeomorphism which is not mixing? Could the mixing rate of a stably ergodic diffeomorphism be arbitrary slow? Could one get a uniform bound on the mixing rate of a stably ergodic diffeomorphism?*

Another question along the same lines is

Question. *Must a stably ergodic diffeomorphism be Bernoulli? stably Bernoulli?*

At present not much is known about stably Bernoulli systems apart from some examples constructed in [2, 5, 45].

7.3. **Skew extensions of non-uniformly hyperbolic systems.** In recent papers [48, 49] Young introduced a class of non-uniformly hyperbolic systems which have statistical properties similar to that of Axiom A attractors.

Problem. *Generalize the results of this paper and* [22] *to the compact group extensions of non-uniformly hyperbolic systems satisfying the conditions of Young.*

Examples of systems one would like to understand along these lines are billiard flows, frame flows on manifolds without conjugated points.

7.4. **Random walks on homogeneous spaces.** The next question deals with improving estimates of the Appendix.

Question. *In the case G is semisimple give more information about the spectrum of the operator J defined in (61).*

So far in all the examples where estimates could be obtained J has a spectral gap. See [29] for the survey of known cases as well as some numerical simulations.

Other questions deal with the situation of Subsection 6.3 without the assumption that G is compact.

Problem. *Give necessary and sufficient conditions for x_n to satisfy CLT.*

This question appears to be hard especially if $< W >$ is nilpotent, but still it is possible that a nice characterization could be obtained for large class of pairs (G, X).

Question. *Is it true that generically x_n satisfies CLT?*

Some important special cases are studied in [6, 31].

7.5. **Non rapidly mixing extensions.** This subsection deals with the classification of non rapidly mixing extensions. For example, if Y is an infranilmanifold and $F : Y \to Y$ is Anosov and $X = G$, Corollary 1.2 tells us that stably ergodic maps are rapidly mixing. Now by [17] non stably ergodic extensions can be characterized by the fact that by a coordinate change T can be reduced to a subextension with skewing function $\tau(x)$ belonging to a coset of a proper subgroup of G. Our Theorem 3.3 has a similar conclusion.

Question. *Is the same conclusion valid without the assumption that $X = G$?*

Some special cases are analyzed in [12].

Another question deals with a geometric characterization of exponential mixing similar to our Theorem 1.1.

Question. *Does exponential mixing depend on which Gibbs potential we consider?*

We plan to address some of these problems elsewhere.

Appendix A. Diophantine approximations.

Here we study some problems related to Diophantine approximations. Let a compact group G act transitively on a manifold X. Let $\mathcal{H}_s$ denote the $s-th$ Sobolev space: if $f = \sum_\lambda f_\lambda$, $f_\lambda \in \mathbb{H}_\lambda$ then $||f||_s^2 = \sum_\lambda ||f_\lambda||_{L^2}^2 \lambda^{2s}$. $||\cdot||$ will denote L^2-norm. Recall that $\pi(g)f$ denotes

$$[\pi(g)f](x) = f(g^{-1}x).$$

Definition. A subset $W \subset G$ is called **Diophantine** if $\exists \alpha_1, C_1$ so that $\forall f \in \mathbb{H}_\lambda$, $\lambda \neq 0$, $\exists g \in W$ such that

$$\|(1-\pi(g))f\| \geq C_1 \lambda^{-\alpha_1} \|f\|. \tag{60}$$

We shall say that W is Diophantine on X if it is not clear which action of G we are considering.

Recall that $\langle W \rangle$ denotes the smallest Lie subgroup of G containing W.

Proposition A.1. *(a)If W is Diophantine then $\langle W \rangle$ acts transitively on G.*

(b) W is Diophantine if and only if $W \bigcup W^{-1}$ is Diophantine.

Proof: (a) is clear, since otherwise there would be an $\langle W \rangle$–invariant function.

(*b*) $\|(1-\pi(g))f\| \geq C_1 \lambda^{-\alpha_1} \|f\| \Leftrightarrow \|(1-\pi(g^{-1}))f\| \geq C_1 \lambda^{-\alpha_1} \|f\|$. □

Now we consider the case when W is a finite set: $W = \{g_1, g_2 \dots g_d\}$. Let S_n denote the set of all words in W and W^{-1} of length at most n : $S_n = \{g_{i_1}^{\pm 1} g_{i_2}^{\pm 1} \dots g_{i_k}^{\pm 1}\}_{k \leq n}$. Define

$$J(f) = \frac{1}{d} \sum_{l=1}^{d} \pi(g_l) f. \tag{61}$$

Let $C_0^\infty(X)$ denote the space of C^∞-functions on X with zero mean.

Theorem A.2. *The following conditions are equivalent:*
(1) W is Diophantine;
(2) $1-J$ is invertible on $C_0^\infty(X)$ and there is a constant α_2 such that for all s there is a constant C_s such that

$$\|(1-J)^{-1} f\|_s \leq C_s \|f\|_{s+\alpha_2};$$

(3) $\exists C_3, \alpha_3$, x_0 such that $\forall \varepsilon$, $S_{[C_3(\frac{1}{\varepsilon})^{\alpha_3}]} x_0$ is an ε–net in X;
(4)$\exists C_4, \alpha_4$, such that $\forall x_0$ $\forall \varepsilon$, $S_{[C_4(\frac{1}{\varepsilon})^{\alpha_4}]} x_0$ is an ε–net in X;
(5) $\exists C_5, \alpha_5$ such that $\forall f : \int f(x)dx = 0$, $\int |f(x)|^2 dx = 1$, $|\Delta f(x)| \leq \lambda^2$ there are $g \in W$ and $x_0 \in X$ such that $|f(x_0) - f(gx_0)| \geq C_5 \lambda^{-\alpha_5}$.

Proof: (1)$\Rightarrow$(2) : For $f \in \mathbb{H}_\lambda$

$$([1-J]f, f) = \frac{1}{d} \sum_l ([1-\pi(g_l)]f, f) \tag{62}$$

Let $j(f)$ be an index such that

$$\|(1-\pi(g_{j(f)}))f\| \geq C_1 \lambda^{-\alpha_1} \|f\|$$

Since each term in (62) is positive

$$([1-J]f,f) \geq \frac{1}{d}([1-\pi(g_{j(f)})]f,f) = \frac{1}{d}(\|f\|^2 - (\pi(g_{j(f)})f,f)) =$$

$$\frac{1}{2d}\|[1-\pi(g_{j(f)})]f\|^2 \geq \frac{C_1^2}{2d}\lambda^{-2\alpha_1}\|f\|^2;$$

(2)⇒(4) : We have to find N such that for all x, y there is $g \in S_N$ such that $g\mathcal{B}(x, \frac{\varepsilon}{2}) \bigcap \mathcal{B}(y, \frac{\varepsilon}{2}) \neq \emptyset$. If $f, h \in \mathbb{H}_\lambda$ then

$$\frac{1}{N}\left|\left((J - J^{N+1})(1-J)^{-1}f, h\right)\right| \leq \text{Const}\|f\|_0\|h\|_0\frac{\lambda^{\alpha_2}}{N}.$$

Take $f, h \in C^\infty(X)$ such that $\operatorname{supp} f \subset \mathcal{B}(x, \frac{\varepsilon}{2})$, $\operatorname{supp} h \subset \mathcal{B}(y, \frac{\varepsilon}{2})$, $f, g \geq 0$, $\int f(x)\, dx = \int g(x) dx = 1$ and $\|f\|_0, \|h\|_0 \leq \text{Const}\varepsilon^{-m}$, $\|f\|_2, \|h\|_2 \leq \text{Const}\varepsilon^{-m-2}$ where $m = \frac{\dim X}{2}$. Decompose $f = \sum_\lambda f_\lambda$, $h = \sum_\lambda h_\lambda$, where $f_\lambda, h_\lambda \in \mathbb{H}_\lambda$. We have

$$\left|\left(\frac{1}{N}\sum_{k=1}^N J^k f, h\right) - 1\right| = \left|\left(\frac{1}{N}\sum_{k=1}^N J^k f, h\right) - \int f(x)dx \int h(x)dx\right| \leq$$

$$\sum_{\lambda\neq 0}\left|\left(\frac{1}{N}\sum_{k=1}^N J^k f_\lambda, h_\lambda\right)\right| = \sum_{\lambda\neq 0}\left|\left(\frac{1}{N}(1-J^{N+1})(1-J)^{-1})f_\lambda, h_\lambda\right)\right| =$$

$$(I) + (II)$$

where in (I) sum is taken over $\lambda \leq \lambda_0$ and in (II) $\lambda > \lambda_0$.

$$(I) \leq \text{Const}\frac{\lambda^{\alpha_2}}{N}\sum \|f_\lambda\|_0\|h_\lambda\|_0 \leq \text{Const}\frac{\lambda^{\alpha_2}}{N}\varepsilon^{-2m}.$$

$$(II) \leq \sum_{\lambda>\lambda_0}\|f_\lambda\|_0\|h_\lambda\|_0 \leq \lambda_0^{-4}\sum_{\lambda>\lambda_0}\|f_\lambda\|_2\|h_\lambda\|_2 \leq \lambda_0^{-4}\varepsilon^{-2m+4}.$$

Take $\lambda_0 = \varepsilon^{-\frac{2m+5}{4}}$, $N = \lambda^{\alpha^2}\varepsilon^{-2m-1}$. Then

$$\left|\left(\frac{1}{N}\sum_{k=1}^N J^k f, h\right) - 1\right|$$

$$\leq \text{Const}\varepsilon,$$

so for some k $(J^k f, h) > 0$ and hence $\exists g \in S_N$ such that

$$g\mathcal{B}(x, \frac{\varepsilon}{2}) \bigcap \mathcal{B}(y, \frac{\varepsilon}{2}) \neq \emptyset.$$

Therefore $S_N x$ is an ε–net in X;

Clearly (4)⇒(3);

(3)⇒(5) : Let $N = [C_3(\frac{1}{\varepsilon})^{\alpha_3}]$. Take $g \in S_N$ such that $|f(gx_0) - f(x_0)| \geq \frac{1}{2}$. Let $g = g_{i_1}^{\epsilon_1} g_{i_2}^{\epsilon_2} \dots g_{i_N}^{\epsilon_N}$, where $\epsilon_j \in \{-1, +1\}$ then

$$\frac{1}{2} \leq \sum_k \left| f(g_{i_k}^{\epsilon_k} \dots g_{i_N}^{\epsilon_N} x) - f(g_{i_{k+1}}^{\epsilon_{k+1}} \dots g_{i_N}^{\epsilon_N} x) \right|.$$

So at least one of the terms is greater than $\frac{1}{2N}$;

(5)⇒(1) : Let $f \in \mathbb{H}_\lambda$. By the Sobolev embedding Theorem

$$\|\nabla f\|_{C(X)} \leq C\lambda^\beta.$$

Thus if $|f(gx_0) - f(x_0)| \geq C_5 \lambda^{-\alpha_5}$ then for $x \in B(x_0, \frac{C_5}{4C}\lambda^{-(\alpha_5+\beta)})$

$$|f(gx_0) - f(x_0)| \geq \frac{C_5}{2}\lambda^{-\alpha_5}$$

and so $\int |f(gx) - f(x)|^2 \, dx \geq \mathrm{Const}\lambda^{-\gamma}$, $\gamma = (\alpha_5 + \beta)\mathrm{dim}X + 2\alpha_5$. □

We now turn to the case when G is semisimple $X = G$ and the action is left translation.

Theorem A.3. *W is Diophantine if and only if $\langle W \rangle = G$, moreover there is a constant $\epsilon_0 = \epsilon_0(G)$ such that any ϵ_0–net in G is Diophantine.*

Proof: In view of Proposition A.1 we only have to prove the "moreover" part. We proceed by induction. Start with some small ε_0 and consider n_0 so that S_{n_0} is an ε_0–net in G. Define

$$\varepsilon_{j+1} = (\tilde{C}\varepsilon_j)^{\frac{4}{3}}, \quad n_{j+1} = \left[\bar{C} n_j \varepsilon_j^{-\frac{1}{3}}\right].$$

Assuming that $\tilde{C}, \bar{C}$ are large enough we show that $S_{n_{j+1}}$ is an ε_{j+1}–net if S_{n_j} is ε_j–net. Consider $\mathfrak{g}$ with an invariant scalar product. We say that a basis $\{X_k\}$ is aligned if

$(a)\frac{1}{2} \leq \|X_k\| \leq 2$;

$(b)\angle(X_k, X_l) \geq \frac{\pi}{4}$.

If ε_j is small enough, there is an aligned basis $\{X_k^{(j)}\}$ such that $x_k^{(j)} = \exp(\varepsilon_j^{\frac{2}{3}} X_k^{(j)}) \in S_{n_j}$. Thus

$$x_{kl}^{(j)} = [x_k^{(j)}, x_l^{(j)}] = \exp\left(\varepsilon_j^{\frac{4}{3}}[X_k^{(j)}, X_l^{(j)}] + O(\varepsilon_j^2)\right) \in S_{2n_j}.$$

Now the space of all aligned bases is compact and $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, so we can extract from $\{[X_k^{(j)}, X_l^{(j)}]\}$ a basis $\{Y_m^{(j)}\}$ so that

$$c_1 \leq \|X_k\| \leq c_2 \quad \angle(X_k, X_l) \geq c_3 \tag{63}$$

By the same argument the set of the bases satisfying (63) is compact and so given δ we can find $C_1, C_2, C_3 > 1$ such that $\forall Y : C_1 \leq \|Y\| \leq C_2$

$\exists a_k \in \mathbb{Z}$, $|a_k| \leq C_3$ such that

$$\|Y - \sum_k a_k Y_k^{(j)}\| \leq \delta \|Y\|.$$

So if ε_j is small enough then $\exists C_1', C_2', C_3'$ such that $\forall y$: $C_1' \varepsilon_j^{\frac{4}{3}} \leq \text{dist}(y, id) \leq C_2' \varepsilon_j^{\frac{4}{3}}$ $\exists a_{kl} \in \mathbb{Z}$, $|a_{kl}| \leq C_3'$ such that $\text{dist}(y, \prod_{kl} (x_{kl}^{(j)})^{a_{kl}}) \leq \delta \varepsilon_j^{\frac{4}{3}}$. We are now ready to establish our claim. Choose a neighborhood U of the identity in G and introduce a coordinate system on U. Partition U into coordinate cubes $U = \bigcup_t \mathcal{C}_t$ of diameter $\bar{C} \varepsilon_j^{\frac{4}{3}}$. By assumption $\forall t$ $\exists t'$: $\mathcal{C}_{t'} \bigcap S_{n_j} \neq \emptyset$ and $\text{dist}(\mathcal{C}_t, \mathcal{C}_{t'}) \leq \text{Const} \varepsilon_j$. Thus we can join $\mathcal{C}_{t'}$ and $\mathcal{C}_t$ by a chain $\mathcal{C}_{t'} = \mathcal{C}^{(0)}, \mathcal{C}^{(1)} \dots \mathcal{C}^{(N)} = \mathcal{C}_t$ of at most $\text{Const} \varepsilon_j^{-\frac{1}{3}}$ elements. Now if $\bar{C}$ is large enough the considerations above imply that $\exists C$ such that if $S_n \bigcap \mathcal{C}^{(i)} \neq \emptyset$ then $S_{n+Cn_j} \bigcap \mathcal{C}^{(i+1)} \neq \emptyset$. Let $N_j = C n_j \varepsilon_j^{-1/3}$. Then S_{N_j} intersects all $\mathcal{C}_t$'s and so is ε_{j+1}–net in U. But if ε_0 is small enough then $S_{n_0} U = G$. $\square$

To pass to the general case we need a slight generalization of this result.

Proposition A.4. *Let G be a compact group and $W \subset G$ be a finite subset. Then*

(a) W is Diophantine on $\langle W \rangle / \text{Center}(\langle W \rangle)$;

(b) $\dim[\langle W \rangle, \langle W \rangle]$ is a lower semicontinuous function on $G^{|W|}$.

PROOF: (a) is a direct consequence of Theorem A.3 since the group $\langle W \rangle / \text{Center}(\langle W \rangle)$ is semisimple;

(b) We establish the following statement

Let $H \subset G$ be a semisimple Lie subgroup, $\mathfrak{h} = L(H)$, then $\exists \varepsilon_0$ such that if $X_1 \dots X_k$ is a basis in $\mathfrak{h}$ such that

$$\text{(Al)} \quad \max(||X_j||) \leq 2 \min(||X_j||), \quad \angle(X_i, X_j) \geq \frac{\pi}{4},$$

$||X_j|| \leq \varepsilon_0$ *then* $\forall Y_1 \dots Y_k$ *such that* $||Y_j - X_j|| \leq \varepsilon_0 ||X_j||$, *the group* $F =< \{\exp(Y_j)\} >$ *satisfies* $\dim([F, F]) \geq \dim(H)$.

The proof is by induction on $\text{codim}(H)$. If $H = G$ this follows from Theorem A.3. For inductive step let $y_j^{(1)} = \exp(Y_j)$. We proceed as in the proof of Theorem A.3 constructing $y_j^{(m)} = \exp(Y_j^{(m)}) \in [F, F]$ where $Y_j^{(m)}$ satisfy (Al) and their norms decreases with m. Let δ be sufficiently small constant. There are two cases:

(1) $\forall m, j$ $\angle(Y_j^{(m)}, \mathfrak{h}) < \delta$. The proof is completed as in Theorem A.3

(2) $\exists m, j$ such that $\angle(Y_j^{(m)}, \mathfrak{h}) \geq \delta$. Consider the one with minimal m. Let $L(G) = \mathfrak{h}+\mathfrak{h}_1+\mathfrak{h}_2$, where $\mathrm{ad}(H) = 0$ on $\mathfrak{h}_2$ and is non-degenerate on $\mathfrak{h}_1$. It is easy to see that $||\pi_{\mathfrak{h}_1} Y_j^{(m)}|| > \delta ||Y_j^{(m)}||/2$. Then the statement follows from the inductive assumption applied to $L(Y_1, \dots Y_k, Y_j^{(m)})$. ∎

Now we can dispose of the assumption that W is finite.

Corollary A.5. *Let G, X be as above. For infinite W the following conditions are equivalent*

(1) W is Diophantine;

(2) W contains a finite Diophantine subset;

(3) $\langle W \rangle = G$.

Proof: We already know that (2)⇒(1)⇒(3). On the other hand the proof of Theorem A.3 shows that if ε_0 is small enough then any ε_0–net is Diophantine. So if (3) holds then S_{n_0} is an ε_0–net for some n_0 and we can extract a finite subnet $V \subset S_{n_0}$. Now if W' is a finite subset of W such that $\langle W' \rangle \supset V$ then W' is Diophantine by Theorem A.3. □

Remark. The above statements fail for the torus. For example the set of all elements of finite order is Diophantine but it obviously does not have finite Diophantine subset. On the other hand there are plenty of t's such that $\langle \{t\} \rangle = \mathbb{T}$ but $\{t\}$ is not Diophantine.

Corollary A.6. *Let X be an arbitrary transitive G-space (G semisimple). Then*

(a) W is Diophantine iff $\langle W \rangle$ acts transitively on X;

(b) If W is Diophantine and $\tilde{W}$ is close to W in the Hausdorff sense, then $\tilde{W}$ is Diophantine.

Proof: (a) follows from Proposition 2.8(b) and Proposition A.4(a); (b) follows from (a), Corollary A.5 and Proposition A.4(b) . □

Corollary A.7. *Let G be any compact group and X be a transitive G-space. Let $p : \langle W \rangle \to \langle W \rangle / [\langle W \rangle, \langle W \rangle]$ be the natural projection. A set W is Diophantine iff $\langle W \rangle$ acts transitively on X and $p(W)$ is Diophantine on $[\langle W \rangle, \langle W \rangle] \backslash X$*

Proof: Let W be a set satisfying the conditions of the corollary. Suppose that $\forall N$ there is a sequence $f_N \in \mathbb{H}_{\lambda_N}$ such that $\|f_N - \pi(g) f_N\| \leq C \lambda_N^{-N}$ for $g \in W$. We want to get a contradiction if N is large enough. By Corollary A.5 and the proof of Theorem A.3 there is a finite set $\tilde{W} \subset S_m(W)$ such that $\tilde{W} \subset [W, W]$ and $\tilde{W}$ generates $[G, G]$. Let $\bar{f}(x)$ denote $\bar{f} = \int_{[G,G]} f_N(gx) dg$. Then $\forall g \in W$

$\|\bar{f}_N - \pi(g)\bar{f}_N\| \leq C\lambda_N^{-N}$, for

$$\bar{f}_N(gx) = \int_{[G,G]} f_N(hgx)dh = \int_{[G,G]} f_N(g(g^{-1}hgx)dh = \overline{(f_N \circ g)}(x)$$

By Corollary A.5 $\tilde{W}$ is Diophantine on $[G, G]$ so there are constants $C(\tilde{W}), \beta(\tilde{W})$ such that $||f_N - \bar{f}_N|| \leq C\lambda_N^{N-\beta}$. But $\bar{f}_N$ can be regarded as a function on $[\langle W\rangle, \langle W\rangle]\backslash X$ so $\exists g \in W$ such that $\|\pi(g)\bar{f} - \bar{f}\| \geq$ Const$^{-\alpha(W)}$, a contradiction if $N > \alpha + \beta \quad \square$.

Remark. This corollary yields quite comprehensive criterion for an action to be Diophantine. Indeed Diophantine subsets of tori are well studied. Also topological generators of semisimple groups are well understood. For example [4] proves that for a semisimlple G, pairs generating G form open dense set. Hence we have

Corollary A.8. *Let G be a compact group acting transitively on X. Then almost all two point sets are Diophantine.*

References

[1] R. Adler, B. Kitchens & M. Shub *Stably ergodic skew products,* Discrete & Continuous Dyn. Sys. **2** (1996) 249-250, Correction **5** (1999) 456.

[2] J. Alves, C. Bonatti & M. Viana *SRB measures for partially hyperbolic systems whose central direction is mostly expanding,* Inv. Math. **140** (2000) 351–398.

[3] D. Anosov *Geodesic flows on closed Riemannian manifolds of negative curvature* Proc. Steklov Inst. Math. **90** (1967).

[4] H. Auerbach *Sur les groupes lineares bornes (III)* Studia Math **5** (1934) 43-49.

[5] C. Bonatti & M. Viana *SRB measures for partially hyperbolic systems whose central direction is mostly contracting,* Israel J. Math. **115** (2000) 157–193.

[6] P. Bougerol & J. Lacroix *Products of random matrices with application to Schrodinger operators,* Progress in Prob. & Stat. **8** (1985) Birkhauser, Boston.

[7] R. Bowen *Equilibrium states and the ergodic theory of Anosov diffeomorphisms* Lect. Notes Math. **470** (1975) Springer-Verlag, Berlin–New-York.

[8] R. Bowen *Weak mixing and unique ergodicity on homogeneous spaces* Isr. J. Math. **23** (1976) 267-273.

[9] R. Bowen *On Axiom A diffeomorphisms* AMS Regional Conf. Proc. **35** (1978).

[10] M. Brin *Topological transitivity of one class of dynamical systems and flows of frames on manifolds of negative curvature* Func. An. & Appl. **9** (1975) 8-16.

[11] M. Brin *Topology of group extensions of Anosov systems* Russ. Math. Notes **18** (1975) 858-864.

[12] M. Brin & M. Gromov *On the ergodicity of frame flows* Inv. Math. **60** (1980) 1-7.

[13] M. Brin & Ya. Pesin *Partially hyperbolic dynamical systems* Math. USSR-Izvestiya **8** (1974) 177-218.

[14] L. Bunimovich *On one class of special flows* Izvestya **38** (1974) 213-227.

[15] K. Burns, C. Pugh, M. Shub & A. Wilkinson *Recent results about stable ergodicity,* to appear in Proc. AMS Summer Research Institute, Seattle-1999.

[16] K. Burns., C Pugh & A. Wilkinson *Stable ergodicity and Anosov flows,* Topology **39** 2000 149–159.
[17] K. Burns & A. Wilkinson *Stable ergodicity of skew products*, Ann. Sci. Ecole Norm. Sup (4) **32** (1999) 859–889.
[18] N. Chernov *Markov approximations and decay of correlations for Anosov flows,* Ann. Math. **147** (1998) 269–324.
[19] I. Cornfeld, S. Fomin & Ya. Sinai *Ergodic Theory* Grundlehren der Mathematischen Wissenschaften **245** (1982) Springer, New York–Berlin.
[20] D. Dolgopyat *Prevalence of rapid mixing-I,* Erg. Th.& Dyn. Sys. **18** (1998) 1097-1114.
[21] D. Dolgopyat *Prevalence of rapid mixing-II,* Erg. Th.& Dyn. Sys. **20** (2000)
[22] D. Dolgopyat *On decay of correlations in Anosov flows* Ann. Math. **147** (1998) 357-390.
[23] D. Dolgopyat *Limit Theorems for partially hyperbolic systems,* preprint.
[24] R. Ellis & W. Perrizo *Unique ergodicity of flows on homogeneous spaces* Isr. J. Math. **29** (1978) 276-284.
[25] M. Field *Generating sets for compact semisimple Lie groups*, Proc. AMS **127** (1999) 3361–3365.
[26] M. Field & M. Nicol *Ergodic Theory of equivariant diffeomorphisms II: stable ergodicity,* preprint.
[27] M. Field & W. Parry *Stable ergodicity of skew extensions by compact Lie groups,* Topology **38** (1999) 167–187.
[28] M. Grayson, C. Pugh & M. Shub *Stably ergodic diffeomorphisms* Ann. Math. **140** (1994) 295-329.
[29] A. Gamburd, D. Jakobson & P. Sarnak *Spectra of elements in the group ring of SU(2),* JEMS **1** (1999) 51–85.
[30] M. Gromov *Groups of polynomial growth and expanding maps* Publ. IHES **53** (1981) 53–73.
[31] Y. Guivarch, M. Keane & B. Roynette *Marches aleatoires sur les groups de Lie,* Lect. Notes in Math. **624** (1977) Springer, Berlin–New-York.
[32] N. Haydn, *Canonical product structure of equilibrium states* Random Comput. Dynam. **2** (1994), 79–96.
[33] M. Hirsch, C. Pugh & M. Shub *Invariant manifolds,* Lect. Notes in Math **583** (1977) Springer, Berlin–New-York.
[34] A. Katok & A. Kononenko *Cocycle stability for partially hyperbolic systems* Math. Res. Lett. **3** (1996) 191-210.
[35] C. Liverani *Central limit theorem for deterministic systems* in International conference on dynamical systems (Eds. F. Ledrappier, J. Lewowics & S. Newhouse), Pitman Res. Notes **363** (1996) 56-75.
[36] D. Ornstein & B. Weiss *Geodesic flows are Bernoulli* Isr. J. Math **14** (1973) 184-198.
[37] W. Parry & M. Pollicott *Zeta functions and periodic orbit structure of hyperbolic dynamics* Asterisque **187-188** (1990).
[38] W. Parry & M. Pollicott *Livsic cocycle equation for compact Lie group extensions of hyperbolic systems,* J. Lond. Math. Soc. **56** (1997) 405–416.
[39] W. Parry & M. Pollicott *Stability of mixing for toral extensions of hyperbolic systems*, Proc. of Steklov Inst. **216** (1997) 354–363.

[40] C. Pugh & M. Shub *Stably ergodic dynamical systems and partial hyperbolicity*, J. of Complexity **13** (1997) 125-179.

[41] C. Pugh & M. Shub *Stable ergodicity and juliene quasiconformality*, JEMS **2** (2000) 1–52.

[42] M. Ratner *Anosov flows with Gibbs measures are also Bernoullian* Isr. J. Math **17** (1974) 380-391.

[43] D. Rudolph *Classifying the isometric extensions of Bernoulli shifts* J. d'Analyse Math. **34** (1978) 36-60.

[44] M. Shub *Endomorphisms of compact differentiable manifolds* Am. J. Math. **91** (1969) 175-199.

[45] Shub M. & Wilkinson A. *Pathological foliations and removable zero exponents* Inv. Math. **139** (2000) 495–508.

[46] Ya. Sinai *Dynamical systems with countably multiple Lesbegue spectrum-II* Izvestiya **30** (1966) 15-68.

[47] Ya. Sinai *Gibbs measures in ergodic theory* Russ. Math. Surv. **27** no. 4 (1972) 21-70.

[48] L.–S. Young *Statistical properties of dynamical systems with some hyperbolicity,* Ann. Math. **147** (1998) 585–650.

[49] L.–S. Young *Recurrence times and rates of mixing,* Israel J. Math. **110** (1999) 153–188.

Department of Mathematics
Penn State University
University Park, PA 16801, USA
email: dolgop@math.psu.edu
website: http://www.math.psu.edu/dolgop/