MIXING FOR SMOOTH TIME-CHANGES OF GENERAL NILFLOWS

ARTUR AVILA, GIOVANNI FORNI, DAVIDE RAVOTTI, AND CORINNA ULCIGRAI

ABSTRACT. We consider irrational nilflows on any nilmanifold of step at least 2. We show that there exists a dense set of smooth time-changes such that any time-change in this class which is not measurably trivial gives rise to a mixing nilflow. This in particular reproves and generalizes to any nilflow (of step at least 2) the main result proved in [AFU] for the special class of Heisenberg (step 2) nilflows, and later generalized in [Rav2] to a class of nilflows of arbitrary step which are isomorphic to suspensions of higher-dimensional linear toral skew-shifts.

1. INTRODUCTION

Dynamical systems can roughly be divided in three categories (hyperbolic, elliptic and parabolic) according to the speed of divergence (if any) of close orbits. A (non-singular) flow is called hyperbolic if nearby orbits diverge exponentially in time. We say that the flow is parabolic if there is divergence of nearby orbits, but this divergence happens at subexponential (usually polynomial) speed, while the flow is called elliptic if there is no divergence (or perhaps it is slower than polynomial). While there is a classical and well-developed theory of hyperbolic systems and also a systematic study of elliptic ones, there is no general theory which describes the dynamics of parabolic flows and only classical and isolated examples are well-understood. This paper contributes to our understanding of typical properties of parabolic flows.

Perhaps the most studied example of a parabolic flow is given by the horocycle flow on (the unit tangent bundle of) a compact negatively curved surface. In the context of homogeneous dynamics (actions given by group multiplication on quotients of Lie groups), parabolic flows coincide with unipotent flows. Horocycle flows can be seen as the simplest example of unipotent flows on semi-simple Lie groups (given by the right action of upper triangular unipotent matrices in SL(2, R) on compact1 quotients \( \Gamma \backslash SL(2, \mathbb{R}) \)). Horocycle flows are the prototype of uniformly (homogeneous) parabolic flows, since they display uniform2 polynomial shearing of nearby trajectories at every point.

In the context of area-preserving flows on (higher genus) surfaces, another important class of parabolic flows is given by locally Hamiltonian flows, which are smooth flows which preserve a smooth area-form3. A crucial feature in this context is the presence of saddle-type singularities, which create a non-uniform (and non-homogeneous) form of parabolic shearing of nearby orbits. These flows are hence an example of non-uniformly parabolic flows.

Another fundamental class of homogeneous flows is given by nilflows, or flows on (compact) quotients of nilpotent Lie groups (nilmanifolds). The prototype example in this class are Heisenberg nilflows. Let us recall that these are given by the action by right multiplication of a 1-parameter subgroup of the Heisenberg group, which can be seen as the group of 3 \times 3 upper triangular matrices, quotiented by a lattice (for example the subgroup of matrices of the same

---

1In the following, we will always tacitly assume that the lattice is cocompact, even though some of the results also hold for finite volume surfaces.

2Uniform here means that the speed of shearing of nearby orbits is uniformly bounded (on initial points) from above and below.

3Linear flows on translation surfaces (surfaces endowed with a flat metric with conical singularities) are area-preserving but smooth, due to the presence of singularities reached in finite time. They share many features of parabolic systems, but lack others. In this case the orbit divergence is entirely produced by the splitting of trajectories near the singularities. For this reason they can be considered as elliptic flows with singularities.
form, but with integer entries). Nilflows have an elliptic factor, and they are definitely not uniformly parabolic. In fact, since they have an isometric (central) direction they are an example of partially parabolic flows.

A natural and fundamental question in parabolic dynamics is hence which ergodic and spectral properties are generic among smooth parabolic flows. There is a large and quite extensive literature on the ergodic and spectral properties of these classical parabolic examples, see for example [Fu, Mar2, Pa, FU, Bur, FF1, Str, BuFo, FoST] for the horocycle flow, and more in general [St] or [AGH], and the reference therein, for homogeneous parabolic flows; [Kal, Kea, Mas1, Ye, AF] or lecture notes such as [Mas2, Yo, FoMa] and the references therein, for translation flows; [Ul2, Ul3, Rav1, KKU, BK] among others for locally Hamiltonian flows. These results do not show an entirely coherent picture and make the identification and description of characteristic parabolic features uncertain. For example, while all smooth time-changes of horocycle flows are mixing and actually mixing of all orders (see [Mar1, Mar2]), nilflows are never weakly mixing (see below). This difference in behavior can be attributed to the lack of parabolicity in certain directions (more specifically to the existence of an elliptic factor). In this paper we prove that these obstructions can be broken by a perturbation and that in a dense set of smooth-time changes all flows which are not trivially conjugate to the nilflow itself are indeed mixing. Our result therefore supports the view that mixing is a generic property among parabolic perturbations of nilflows.

1.1. Time-changes and parabolic perturbations. Starting from the classical examples of parabolic flows mentioned above, one can build new parabolic flows by considering perturbations: the simplest perturbations are perhaps time-changes (or time-reparametrizations) of a given flow (see § 2.1 for the definition), i.e. flows that move points along the same orbits, but with different speed. This construction has the advantage that certain ergodic properties, like ergodicity and cohomological properties (which only depend on the orbit structure and hence are independent of the time-change) persist and smooth time-changes of parabolic flows are still parabolic, while finer (in particular spectral) properties can (and do, as we will discuss) emerge.

Indeed even such simple perturbations as time-changes can produce genuinely new parabolic flows, i.e. flows which are not measurably conjugated to the unperturbed flow. There is an obvious way to produce (measurably or even smoothly) conjugated flows, which corresponds to the case when the time-change is (measurably or smoothly) trivial (see § 2.1 for definitions). These trivial time-changes are described by solutions of the so-called cohomological equation. A key feature of parabolic dynamics is the existence of distributional obstructions (invariant distributions) to solve the cohomological equation, that is, obstructions which are not signed measures. The structure of the space of obstructions was described in the case of translation flows (and locally Hamiltonian flows on surfaces) in [Fo1], for nilflows in [FF3] and horocycle flows, see [FF1].

As a consequence, among smooth-time-changes, smoothly trivial time-changes are rare (i.e. form a finite or countable codimension subspace) and therefore time-changes can have essentially different dynamical properties. However, the question whether a nontrivial time-change is also non-isomorphic is in general very difficult and the answer is known only in a few cases. For example, for horocycle flows, it follows from Marina Ratner’s work in [Rat] and [FF4] that sufficiently smooth time-changes which are measurably isomorphic to the horocycle flow are actually smoothly trivial. Thus, time-changes which give isomorphic flows are rare among sufficiently smooth-time changes of horocycle flows.

---

4 The first complete study of this phenomenon is perhaps Katok’s work (which although written and circulated in the 80’s only appeared in [Ka2, Ka3] on linear skew-shifts of the 2-torus, which are closely related to Heisenberg nilflows. Let us also remark that finitely many invariant distributions for horocycle flows in the finite area, non-compact case were first constructed by P. Sarnak [Sa] by methods based on Eisenstein series.

5 The space of obstructions has a different structure in each of these cases: it is in fact finite dimensional in any finite (Sobolev) order for translation flows, and infinite dimensional for any sufficiently high order for nilflows and horocycle flows. Finally, the horocycle flow has obstructions of arbitrarily high order, which nilflows lack, and therefore subsumes features of both translation flows and nilflows.
1.2. Previous results on time-changes. Perturbations of (homogeneous) parabolic flows are much less understood than the classical homogeneous parabolic examples, even in the simplest case of time-changes. For example, while ergodic and spectral properties of horocycle flows have long been well-understood (see for example [Fu, Mar2, P4]), much less is known about the ergodic theory of time-changes of unipotent flows or nilflows. It is a classical result of B. Marcus [Mar2] that smooth time-changes of horocycle flows are mixing\(^6\) (and actually mixing of all orders) (see also a previous result by Kuschnirenko [Ku], which applies to time-changes which are sufficiently small in the \(C^1\) topology). Effective mixing and spectral results are recent. Decay of correlations and the Lebesgue spectral property were proved by two of the authors (G. F. and C. U.) in [FU], thus partially confirming the Katok-Thouvenot conjecture [KT] on the nature of the spectrum (the question on the multiplicity of the spectrum was left open). The absolute continuity of the spectrum was simultaneously proved by Tiedra de Aldecoa in [Tie] (and later extended to the case of semi-simple unipotent flows by Simonelli [Si]). The methods for proving the Lebesgue spectrum property have been refined by B. Fayad, G. F. and A. Kanigowski [FFK], who treated the case of Diophantine toral flows with a sufficiently strong power singularity (Kochergin flows).

In a recent improvement of the paper [FFK] the authors were able to prove the countable multiplicity of the spectrum for Kochergin flows, as well as for time-changes of the horocycle flows (thereby completing the proof of the Katok-Thouvenot conjecture).

Recent work of A. Kanigowski, M. Lemanczyk and C. U. [KLU] has brought the insight that although horocycle flows are uniformly parabolic, they have special, non-generic properties coming from the homogeneous nature of the dynamics. In particular, all powers of the horocycle flows are isomorphic through the action of the geodesic flow. For non-trivial smooth time changes of the horocycle flow, this phenomenon does not happen and one can prove that powers are all disjoint (as shown in [KLU] and also by L. Flaminio and G. F. in [EF4] using Ratner’s work [Rat]).

Locally Hamiltonian flows on higher genus surfaces can be seen as singular time-changes of linear flows on translation surfaces (which are well-known not to be mixing, see [Ka1]). Since, as already mentioned, locally Hamiltonian flows are parabolic, but not-uniformly parabolic, the presence (or genericity) of mixing is much more delicate (and less persistent) than in the uniformly parabolic case. It turns out furthermore that whether mixing is typical depends crucially on the type of singularities: while the presence of degenerate (multi-saddle singularities) was long known to produce mixing [Ko2] if all singularities are non-degenerate (Morse) singularities, mixing is generic (on each minimal component) if there are saddle loops homologous to zero, but absence of mixing (and weak mixing) is generic if the flow has only simple saddles. Strengthenings of the mixing property, such as quantitative mixing estimates, spectral properties or multiple mixing, and exceptional mixing examples have also been studied (see e.g. [Fa2, Rav1, FFK, CW]).

Not much is known for time-changes of nilflows. In the special case of Heisenberg nilflows, mixing for non-trivial time-changes was proved by three of the authors (A. A., G. F. and C. U.) in [AFU]; G. F. and A. Kanigowski proved effective mixing results for generic nilflows in [FK1] and the Ratner property, disjointness of powers and multiple mixing [FK2] for the (measure zero) class of Heisenberg nilflows with bounded type frequencies.

---

\(^6\)Let us recall that a measure-preserving flow \((\varphi_t)_{t \in \mathbb{R}}\) on a probability space \((X, \mathcal{A}, \mu)\) is mixing if for any two square-integrable functions \(f, g \in L^2(X, \mathcal{A}, \mu)\) the correlations \(\int_X (f \circ \varphi_t) g \, d\mu\) converge to \((\int_X f \, d\mu)(\int_X g \, d\mu)\) as \(t \to +\infty\).
The advances in the ergodic theory of parabolic flows recalled above mainly concern time-changes of renormalizable flows, a class which so far includes only translation flows, horocycle flows and Heisenberg nilflows. For these flows, renormalization provides a powerful tool to analyze the fine behavior of ergodic averages, which is crucial in several results on the ergodic theory of their time-change. Results for non-renormalizable flows, such as higher-step nilflows, are much rarer. In fact, refined quantitative estimates, especially pointwise lower bounds on sets of large measures, are not available for higher step nilflows, contrary to the step two case. None of the known results seem to provide point-wise lower bounds (on sets of large measure), although lower bounds in square mean follow from representation theory.

It is however natural to ask whether the results proved for Heisenberg nilflows also hold for other nilflows, i.e. when the step and dimension are higher. A first result in this direction was obtained by D. R., who studied mixing among non-trivial time-changes in the class of quasi-Abelian filiform nilflows (this is a special class of nilflows which constitutes a natural higher dimensional extension of Heisenberg nilflows, since they, as in the Heisenberg case, have a Poincaré section which is a skew-translation on a torus).

1.3. Main results. In this paper we consider a general nilmanifold $M = \Gamma \backslash G$ (where $G$ is a nilpotent Lie group and $\Gamma < G$ a lattice) of step at least 2 (to exclude the case when $G$ is Abelian and hence $M$ is a torus). We then consider a nilflow $\phi = \{\phi_t\}_{t \in \mathbb{R}}$ on $M$ and assume only that $\phi$ is (uniquely) ergodic; by [AGH], this equivalently means that the linear flow on the toral factor is irrational (see § 2.2). In the following we will write parabolic nilflows as a shortening for a nilflow on a nilmanifold of step at least 2 (since the step 1 or Abelian case gives rise to elliptic flows).

Our main result shows that, within a dense class of smooth-time changes of ergodic parabolic nilflows, mixing arises as soon as the time-change is measurably non-trivial (i.e. it is not cohomologous to a constant with a measurable transfer function).

Theorem 1.1. For any irrational nilflow on any nilmanifold of step at least 2 there exists a dense set of smooth time-changes which are either measurably trivial or mixing.

The dense class of time-changes in the statement of the theorem will be described in detail later, in section §3.3. The assumption that the step is at least 2 is needed to exclude the abelian case of irrational flows on tori: in that case, non-singular, smooth time-changes are typically not mixing (under a full measure Diophantine condition on the frequencies) by classical KAM-type results, see e.g. [Kol]. Thus, if the step is one, we are in the elliptic (and non-parabolic) world. We stress that we do not require any Diophantine condition on the frequencies of the toral factor, only irrationality: as soon as the nilflow is (uniquely) ergodic, non-trivial time-changes in our class are mixing.

Our result also implies that smooth time-changes of an irrational nilflow which are not measurably trivial are not measurably isomorphic to the nilflow (since they are mixing while nilflows are not). Since the converse implication is obvious, we get the following corollary.

Corollary 1.2. Given any irrational nilflow on a nilmanifold of step at least 2, a time-change within the dense set given by Theorem 1.1 is measurably conjugated (isomorphic) to the original flow if and only if it is measurably trivial.

---

7 See [FT1] for horocycle flows, [FF2] for Heisenberg nilflows, and [Fo2] for translation flows (or IET’s).

In the case of Heisenberg nilflow, bounds on ergodic integrals have a long history going back to the work of Hardy and Littlewood on bounds of quadratic Weyl sums more than a century ago until the optimal bounds of H. Fiedler, W. Jurkat and O. Körner. Recent results have refined the analysis of the behavior of ergodic integrals and derived results on their limit distributions (FF3, J. Marklof, M. F. Cellarosi and J. Marklof, J. Griffin and J. Marklof, A. Fedotov and F. Klopp, G. F. and A. Kanigowski).

8 However, quantitative (but not optimal) equidistribution estimates were proved for Lebesgue almost all points in the class of quasi-Abelian nilflows by FF3; see also the related uniform estimates proved by T. Wooley, e.g. in Wol1, Wol2 and, independently, by J. Bourgain, C. Demeter and L. Guth for exponential sums (see also the Bourbaki seminar [P], section 2.1.1).
The conclusion one might want to draw from the main result of this paper (as well as the results in \cite{KL} and \cite{FF4} on disjointness for time-changes of horocycle flows) is hence that (homogeneous) nilflows (as well as the classical horocycle flow) are indeed not generic examples of parabolic flows and display "exceptional" properties. As soon as homogeneity and extra structures (such as toral factors, or isomorphisms between time-$T$ maps) are broken by a perturbation, the expected "generic" parabolic features indeed do emerge.

1.4. Open problems. Our work leaves open some natural questions.

Both in \cite{AFU} and \cite{Rav2}, as well as in Theorem \ref{thm:main}, the dynamical dichotomy in the result (between mixing and trivial time-changes) is only claimed within a special (sub)class of time-changes. Even though the class that we consider is dense in the smooth category (in the $C^\infty$ norm), a natural question (which is already open even in the Heisenberg case), is to consider (more) general smooth time-changes.

**Problem 1**: Is any measurably non-trivial smooth time-change of a uniquely ergodic parabolic (even Heisenberg) nilflow weakly mixing? mixing?

For renormalizable flows, it is possible to derive cocycle rigidity results from results on growth of ergodic integrals of smooth function which are not smooth coboundaries, see \cite{AFU, FF4}. Hence, an effective characterization of mixing time-changes can be given in terms of vanishing of invariant distributions. For general nilflows, contrary to the horocycle or Heisenberg case, we are unfortunately not able to explicitly describe the set of measurably trivial time-changes.

We therefore pose the following problem.

**Problem 2**: Give an effective description of the class of measurably trivial, non mixing time-changes for higher step nilflows.

For Heisenberg nilflows, the results in \cite{AFU} have been recently strengthened by A. Kanigowski and G. F. in \cite{FK1} and in \cite{FK2}: under a full measure Diophantine condition, quantitative mixing estimates are given in \cite{FK1} for a larger class (in fact, residual) of time-changes, while for mixing Heisenberg nilflows of bounded type multiple mixing is shown in \cite{FK2}. It it hence natural to ask whether also these results extend to higher step nilflows.

**Problem 3**: Under a Diophantine condition on the frequencies of the toral factor of an ergodic, parabolic nilflow, prove quantitative mixing estimates.

**Problem 4**: Under similar (or stronger) assumptions, prove that mixing time-changes are mixing of all orders.

We warn the reader that the above Problems 2, 3, 4 seem hard, because of the lack of fine quantitative estimates on ergodic averages discussed above. Let us remark that multiple mixing is shown in \cite{FK1} by proving the Ratner property (which, combined with mixing, implies mixing of all orders). However, the Ratner property is not known to hold even for Heisenberg nilflows which are not of bounded type.

1.5. Strategy of the proof. Results on mixing for time-changes of homogeneous parabolic flows, as well as parabolic surface flows, are all based on a geometric mechanism known as shearing: short segments transversal to the flow, pushed by the flow, get sheared in the direction of the flow\footnote{Shearing (in the direction of the flow) was for example exploited by Marcus to prove mixing for smooth time changes of horocycle flows \cite{Mar1}, and it provides the basis for all the mixing results in the context of area-preserving flows (e.g. in \cite{Ko2, Fa2, KS, Ul1, Rav1, CW}).} (or in a direction which commutes with the flow). When the curves are sheared in the flow direction and are asymptotically approximated by flow trajectories, this allows in particular to prove (quantitative) mixing by exploiting (quantitative) equidistribution of the trajectories of the (uniquely ergodic) flow.

Our result is also based on a mixing-via-shearing argument, but the source of shearing is more subtle. Indeed, recall that nilflows are only partially parabolic, hence there are central directions.
which are \textit{not} sheared before the time-change. However, with a carefully chosen inductive procedure, we are able to either show that in these directions shearing is created by the non-trivial time-change, or the nilflow can be seen as an extension over a lower dimensional nilflow. In this case mixing can be lifted through shearing in the central direction (a phenomenon that we call \textit{wrapping in the fibers}), which is exactly the mechanism responsible for the mixing property of nilflows, relative to the elliptic toral factor.

We will now explain the main ideas of the proof and the new difficulties which arise in the general case and were not present in the case of Heisenberg nilflows [AFU] or of quasi-Abelian filiform nilflows [Rav2]. The starting point in [AFU] is the representation of a (time-change of a) Heisenberg nilflow as a special flow over a skew-translation on a two dimensional torus. For the class of time-changes considered (which essentially consists of trigonometric polynomials) one can show that the curves in the 1-dimensional central isometric direction, pushed via the flow, get sheared. This gives the geometric shearing mechanism which then allows to prove mixing. A similar strategy is also used by Ravotti in [Rav2] to prove mixing in the case of quasi-Abelian filiform nilflows.

Contrary to the quasi-Abelian case, for a general nilflow the natural sections are isomorphic to non-toral nilmanifolds and return maps are niltranslations, which are more difficult to handle explicitly. Furthermore, an additional difficulty of the general case, which is not present in the quasi-Abelian filiform class, is that the center can be \textit{higher dimensional}. The key idea is still to study (short) curves in a central direction (to choose carefully, so that it is part of a \textit{Heisenberg triple}, see §3.3) and consider their pushforward by the flow. Here two possible scenarios appear: either there is \textit{shearing}, and one can try to directly prove mixing or it is also possible that shearing does \textit{not} occur (this is what we call the \textit{coboundary case}, see section 5.2). In this case, the idea is to \textit{quotient out} central toral fibers, chosen appropriately (see section 3.3) and find a \textit{factor} which is a time-change of a nilflow on a lower dimensional nilmanifold. In this case, if the factor is mixing, one can "\textit{lift}" the mixing property by the mechanism of \textit{wrapping in the fibers} mentioned above: if a central curve in the factor is pushed by the factor flow, its lift wraps in the toral fiber \textit{faster} than the equidistribution speed in the factor (this is essentially a consequence of the nilflow filtration structure).

To implement this strategy, we use an inductive argument on the dimension $\dim [G,G]$ of the commutator subgroup of the nilpotent group $G$. A delicate point is how to choose a sequence of quotients on which to apply the induction. Here the notion of \textit{Heisenberg triples} (defined in §3.3, see Definition 3.5) plays a key role. Each time we quotient by the (toral) closure of a central flow which belongs to a Heisenberg triple. The assumption that the time-change is non-trivial guarantees that, before reaching the "base" case, one has to find a nilflow factor where shearing occurs. The construction of the \textit{tower} of nilflow factors considered is described in detail in section 3.3.

This inductive presentation of the nilflow on a nilmanifold as tower of toral extensions dictates also the class of time-changes for which we can prove mixing. As in [AFU], it is crucial for the shearing estimates that the time-change has polynomial features (in the sense that derivatives should be controllable in terms of the function itself). Hence, the class that we consider consists of time-changes which, at each level of the tower, behave in each toral fiber like trigonometric polynomials (see Definition 3.9 in §3.3 and also Definition 3.1 in §3.1). This produces a $C^\infty$ dense class of time changes in which we have a dynamical dichotomy, i.e. either the time change is trivial, or it is mixing.

Let us remark that the choice of abandoning the previous set-up from [AFU], [Rav2] based on special flow representations of nilflows was motivated, in addition to the greater elegance and simplicity of the arguments, by the significant technical difficulties which arise in the general higher step case in working with special flow representations. In fact, the relation between the discrete time of the Poincaré return map and the continuous time of the nilflow is problematic in
the higher step case because of the distributional obstructions to solving cohomological equations for non-toral nilflows and the related deviation of ergodic averages from the mean.

We underline that the present paper supervenes the previous two results proved in [AFU] and [Rav2]: not only it gives an independent proof of prevalence of mixing also among Heisenberg and a larger class of time-changes than the one considered in [Rav2] for quasi-Abelian filiform nilflows, but it also follows a much more streamlined approach. The arguments, indeed, drawing on ideas of [AFU] and [Rav2], recast them in a geometric framework derived from [FU] (already adapted to Heisenberg nilflows in [FK1]). In this more intrinsic framework (neither section nor coordinate-dependent), we work directly on the manifold, and prove shearing by analyzing push-forwards of central curves. Consequently, the proof of mixing is also less intricate (instead than analyzing the pushforwards of partitions into small pieces of curves, we can analyze the pushforwards of a single curve and directly use integral estimates).

1.6. Structure of the paper. In Section 2 we recall basic definitions, such as time-changes and coboundaries (§2.1), as well as some basic material on nilflows on nilmanifolds (§2.2). In Section 3, we define the dense class of time-changes in Theorem 1.1. This requires also building the tower of extensions which will be used for the induction (in §3.3). In Section 4 we state and prove a number of results which will be used as tools to prove mixing, in particular two lemmas which reduce mixing to a statement about shearing (Lemma 4.1 and Lemma 4.2 in §4.1) and the computation of pushforwards of curves along the flow (Corollary 4.3 in §4.2). We also prove a result on growth of ergodic sums for functions which are not coordinate-dependent, we work directly on the manifold, and prove shearing by analyzing push-forwards of central curves. Consequently, the proof of mixing is also less intricate (instead than analyzing the pushforwards of partitions into small pieces of curves, we can analyze the pushforwards of a single curve and directly use integral estimates).

2. Background

Let us recall for the convenience of the reader some basic definitions and background.

2.1. Basic definitions: mixing, time-changes, coboundaries. Let \( \phi := \{ \phi_t \}_{t \in \mathbb{R}} \) be a measurable flow on a probability space \( (M, \mu) \). We recall that \( \phi \) is said to be mixing if for each pair of measurable sets \( A, B \subset M \), one has

\[
\lim_{t \to \infty} \mu(\phi_t(A) \cap B) = \mu(A) \mu(B),
\]

and weak mixing if, for each pair of measurable sets \( A, B \subset M \),

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t |\mu(\phi_s(A) \cap B) - \mu(A) \mu(B)| ds = 0.
\]

A flow \( \{ \tilde{\phi}_t \}_{t \in \mathbb{R}} \) is called a reparametrization or a time-change of a flow \( \{ \phi_t \}_{t \in \mathbb{R}} \) on \( M \) if there exists a measurable function \( \tau: M \times \mathbb{R} \to \mathbb{R} \) such that for all \( x \in M \) and \( t \in \mathbb{R} \) we have \( \tilde{\phi}_{\tau(x,t)}(x) = \phi_t(x) \). Since \( \{ \tilde{\phi}_t \}_{t \in \mathbb{R}} \) is assumed to be a flow the function \( \tau(x, \cdot): \mathbb{R} \to \mathbb{R} \) is an additive cocycle over the flow \( \{ \phi_t \}_{t \in \mathbb{R}} \), that is, it satisfies the cocycle identity:

\[
\tau(x, s + t) = \tau(\phi_s(x), t) + \tau(x, s), \quad \text{for all } x \in M, s, t \in \mathbb{R}.
\]

If \( M \) is a manifold and \( \{ \phi_t \}_{t \in \mathbb{R}} \) is a smooth flow, we will say that \( \{ \tilde{\phi}_t \}_{t \in \mathbb{R}} \) is a smooth reparametrization if the cocycle \( \tau \) is a smooth function. By the cocycle property, a smooth cocycle is uniquely determined by its infinitesimal generator, that is the function \( \alpha_\tau: M \to \mathbb{R} \) defined by the formula:

\[
\alpha_\tau(x) := \frac{\partial \tau}{\partial t}(x, 0), \quad \text{for all } x \in M.
\]
In fact, given any positive function \( \alpha: M \to \mathbb{R}^+ \), the formula
\[
\tau_\alpha(x,t) := \int_0^t \alpha(\phi_s(x)) \, ds,
\]
for all \((x,t) \in M \times \mathbb{R}\)
defines a cocycle over the flow \( \{ \phi_t \}_{t \in \mathbb{R}} \) with infinitesimal generator \( \alpha \).

The infinitesimal generators \( V \) and \( X \) of the flows \( \{ \phi_t \}_{t \in \mathbb{R}} \) and \( \{ \tilde{\phi}_t \}_{t \in \mathbb{R}} \) respectively are related by the identity:
\[
X = \frac{d\phi_t}{dt} \bigg|_{t=0} = \alpha_t \frac{d\tilde{\phi}_t}{dt} \bigg|_{t=0} = \alpha_t V, \quad \text{i.e. } V = \frac{1}{\alpha_t} X.
\]

An additive cocycle \( \tau: M \times \mathbb{R} \to \mathbb{R} \) over the flow \( \{ \phi_t \}_{t \in \mathbb{R}} \) is called a measurable (respectively smooth) coboundary if there exists a measurable (respectively smooth) function \( u: M \to \mathbb{R} \), called the transfer function, such that
\[
\tau(x,t) = u \circ \phi_t(x) - u(x), \quad \text{for all } (x,t) \in M \times \mathbb{R}.
\]

The additive cocycle \( \tau \) is a measurable (smooth) coboundary if and only if its infinitesimal generator \( \alpha_t \) is a measurable (smooth) coboundary for the infinitesimal generator \( X \) of the flow \( \{ \phi_t \}_{t \in \mathbb{R}} \), that is, if there exists a measurable (smooth) function \( u: M \to \mathbb{R} \), also called the transfer function, such that \( Xu = \alpha_t \).

Two additive cocycles are said to be measurably (respectively smoothly) cohomologous if their difference is a measurable (respectively smooth) coboundary in the above sense. A cocycle is said to be an almost coboundary if it is cohomologous to a constant cocycle.

An elementary, but fundamental, result establishes that time-changes given by cohomologous cocycles are isomorphic (see for example [Ka3], §9). The regularity class of the isomorphisms depends on the regularity class of the transfer function. A time-change defined by a measurable (smooth) almost coboundary is called measurably (smoothly) trivial.

Given \( \phi^V := \{ \phi^V_t \}_{t \in \mathbb{R}} \) a uniquely ergodic homogeneous flow on a manifold \( M \) generated by the vector field \( X \), given any \( \mathcal{C}^m \) function \( \alpha: M \to \mathbb{R}^+ \), we can consider the time-change \( \phi^V := \phi^V_t = \{ \phi^V_t \}_{t \in \mathbb{R}} \) of \( \phi^X_t \) with generator given by the formula
\[
V = \frac{1}{\alpha} X.
\]

Throughout the paper, we will use the notation \( \phi^V \) to denote the time change of \( \phi^X \) with generator \( V = \frac{1}{\alpha} X \). Remark that if \( \phi^X \) preserves the Haar measure \( \mu \), \( \phi^V \) preserves the measure \( \alpha \mu \) (which is absolutely continuous with respect to Haar when \( \alpha \) is smooth).

### 2.2. Preliminaries on nilmanifolds.

Let \( \mathfrak{g} \) be a \( k \)-step nilpotent real Lie algebra \((k \geq 2)\) with a minimal set of generators \( \mathcal{E} := \{ E_1, \ldots, E_n \} \subset \mathfrak{g} \). For all \( j \in \{ 1, \ldots, k \} \), let \( \mathfrak{g}_j \), denote the descending central series of \( \mathfrak{g} \):
\[
\mathfrak{g}_1 = \mathfrak{g}, \quad \mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}], \ldots, \quad \mathfrak{g}_j = [\mathfrak{g}_{j-1}, \mathfrak{g}], \ldots, \quad \mathfrak{g}_k \subset Z(\mathfrak{g}),
\]
where \( Z(\mathfrak{g}) \) is the center of \( \mathfrak{g} \).

Let \( G \) be the connected and simply connected nilpotent Lie group with Lie algebra \( \mathfrak{g} \). The corresponding Lie subgroups \( G_j = \exp \mathfrak{g}_j = [G_{j-1}, G] \) form the descending central series of \( G \). Let \( \Gamma \) be a lattice in \( G \). It exists if and only if \( G \) admits rational structure constants (see, for example, [Rai][CG]).

A (compact) nilmanifold is by definition a quotient manifold \( M := \Gamma \backslash G \) with \( G \) a nilpotent Lie group and \( \Gamma \subset G \) a lattice. On a nilmanifold \( M = \Gamma \backslash G \), the group \( G \) acts on the right transitively by right multiplication. By definition, the nilflow \( \phi^X \) generated by \( X \in \mathfrak{g} \) is the flow obtained by the restriction of this action to the one-parameter subgroup \( (\exp tX)_{t \in \mathbb{R}} \) of \( G \):
\[
\phi^X_t(\Gamma x) = \Gamma \exp(tX).
\]
It is plain that nilflows on $\Gamma \backslash G$ preserve the probability measure $\mu$ on $\Gamma \backslash G$ given locally by the Haar measure. To simplify the notation, the vector field on $\Gamma \backslash G$ generating the flow $\phi^X$ will also be indicated by $X$.

Every nilmanifold is a fiber bundle over a torus. In fact, the group $\overline{G} = G/[G,G]$ is Abelian, connected and simply connected, hence isomorphic to $\mathbb{R}^n$ and $\overline{\Gamma} = \Gamma/[\Gamma,\Gamma]$ is a lattice in $\overline{G}$. Thus we have a natural projection

$$p : \Gamma \backslash G \rightarrow \overline{\Gamma} \backslash \overline{G}$$

over a torus of dimension $n$. We recall the following:

**Theorem 2.1 ([CG], [AGH]).** The following properties are equivalent.

1. The nilflow $\phi^X$ on $\Gamma \backslash G$ is ergodic.
2. The nilflow $\phi^X$ on $\Gamma \backslash G$ is uniquely ergodic.
3. The nilflow $\phi^X$ on $\Gamma \backslash G$ is minimal.
4. The projected flow $\psi^X$ on $\overline{\Gamma} \backslash \overline{G} \approx \mathbb{T}^n$ is an irrational linear flow on $\mathbb{T}^n$, hence it is (uniquely) ergodic and minimal.

By property (iv) it follows that a nilflow is never (weakly) mixing, since it has a linear toral flow, which has pure point spectrum, as a factor. However, it is possible to prove by methods of representation theory that any nilflow is relatively mixing, in the sense that the limit of correlations of functions with zero average along all fibers of the projection in formula (3) is equal to zero.

The irrationality condition in our main result, Theorem 1.1, and in (iv) in the above Theorem 2.1 refers to the rational structure determined by the lattice $\Gamma \subset G$ and its Abelianized $\overline{\Gamma} \subset \overline{G}$ defined as follows. Let us first recall the definition of Malcev basis (see [CG]).

**Definition 2.2 (Malcev basis).** A Malcev basis for $\mathfrak{g}$ through the descending central series $\mathfrak{g}_j$ and strongly based at $\Gamma$ is a basis $E_1^1, E_2^1, \ldots, E_{n_1}^1, E_1^2, \ldots, E_{n_2}^2, \ldots, E_1^k, \ldots, E_{n_k}^k$, (with $n_1 = n$) of $\mathfrak{g}$ satisfying the following properties:

1. if we drop the first $\ell$ elements of the basis we obtain a basis of a subalgebra of codimension $\ell$ of $\mathfrak{g}$.
2. if we set $\mathcal{E}^j := \{E_1^j, \ldots, E_{n_j}^j\}$ the elements of the set $\mathcal{E}^j \cup \mathcal{E}^{j+1} \cup \cdots \cup \mathcal{E}^k$ form a basis of $\mathfrak{g}_j$.
3. every element of $\Gamma$ can be written as a product

$$\exp m_1^1 E_1^1 \cdots \exp m_{n_1}^1 E_{n_1}^1 \cdots \exp m_1^k E_1^k \cdots \exp m_{n_k}^k E_{n_k}^k$$

with integral coefficients $m_i^j$.

The existence of a Malcev basis can be derived by combining the proofs of Theorems 1.1.13 and 5.1.6 of [CG]. A Malcev basis determines a rational structure on the Lie algebra $\mathfrak{g}$ of the nilpotent Lie group $G$ as follows.

**Definition 2.3 (Irrational vectors).** An element $V \in \mathfrak{g}$ is called $\Gamma$-rational if it belongs to the span over $\mathbb{Q}$ of a Malcev basis strongly based at $\Gamma$, and it is called $\overline{\Gamma}$-rational if its projection $\overline{V} \in \mathfrak{g}/[\mathfrak{g},\mathfrak{g}] \approx \mathbb{R}^n$ belongs to the span over $\mathbb{Q}$ of the projection of the Malcev basis. An element $V \in \mathfrak{g}$ is called irrational (with respect to $\Gamma$) if the coordinates of its projection $\overline{V}$ with respect to a $\Gamma$-rational basis of $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ are linearly independent over $\mathbb{Q}$.

Thus, if the generators $\{E_1, \ldots, E_n\}$ of $\mathfrak{g}$ are chosen so that the elements $\{\exp E_1, \ldots, \exp E_n\}$ project onto generators $\{\exp \overline{E}_1, \ldots, \exp \overline{E}_n\}$ of $\overline{\Gamma}$, then for every $X \in \mathfrak{g}$ there exists a vector $\Omega_X := (\omega_1(X), \ldots, \omega_n(X)) \in \mathbb{R}^n$ such that

$$X = \omega_1(X) \overline{E}_1 + \cdots + \omega_n(X) \overline{E}_n.$$  

The element $X$ is called irrational (with respect to $\Gamma$) if the numbers $\omega_1(X), \ldots, \omega_n(X)$ are linearly independent over $\mathbb{Q}$. 
3. Towers and time-changes

The goal of this section is to define the class of time-changes in which the dichotomy in Theorem 1.1 holds. In §3.1, we first define trigonometric polynomials relative to a central extension (see Definition 3.1) and state some of the properties which will be needed later (see §3.2). In §3.3 we then explain how to construct inductively the tower of nilflow extensions we will be working with and, finally, in §3.4 we define the class of mixing time-changes (which consists of trigonometric polynomials relative to the tower, see Definition 3.3).

3.1. Trigonometric polynomials on nilmanifolds. Let $h_{\Gamma} \subset Z(g)$ be any $d$-dimensional $\Gamma$-rational subspace; that is, a $d$-dimensional subspace which admits a basis of $\Gamma$-rational vectors (in the sense of Definition 2.3). The subgroup $\Lambda_{\Gamma} := \log \Gamma \cap h_{\Gamma}$ is a lattice in $h_{\Gamma}$.

In the following, we identify $h_{\Gamma}$ with $\mathbb{R}^d$ and let $(\Phi_t^{h_{\Gamma}})_{t \in \mathbb{R}^d}$ denote the action of $h_{\Gamma} \approx \mathbb{R}^d$ on $M$. Since $h_{\Gamma} \subset Z(g)$, the group $\exp h_{\Gamma} \subset G$ is Abelian and $(\Phi_t^{h_{\Gamma}})_{t \in \mathbb{R}^d}$ descends to a toral action (with closed orbits).

Let us define the quotient Lie algebra $\overline{\mathfrak{g}} := \mathfrak{g}/h_{\Gamma}$, the corresponding Lie group $G := G/\exp h_{\Gamma}$ and the nilmanifold $\overline{M} := M/\exp h_{\Gamma}$ or, in other terms, $\overline{M} := \overline{G}/\Gamma \cap \exp h_{\Gamma}$). We remark that $M$ is a fiber bundle $\pi : M \to \overline{M}$ over $\overline{M}$, with fibers $\Gamma \cap \exp h_{\Gamma}$. The action of the torus $h_{\Gamma}/\Lambda_{\Gamma}$ on $M$ restricts to a faithful action on the fibers $F_{\pi}$ over $\pi \in \overline{M}$, hence each of the latter is isomorphic to $h_{\Gamma}/\Lambda_{\Gamma}$. In particular, the conditional measures $\mu_{F_{\pi}}$ supported on $F_{\pi}$ are, up to a constant, the Lebesgue measure on $h_{\Gamma}/\Lambda_{\Gamma}$. We can disintegrate the Haar measure $\mu$ with respect to this fibration as

$$\mu = \int_{\overline{M}} \mu_{F_{\pi}} \, d\overline{\mu}(\pi),$$

where $\overline{\mu} = \pi_\ast \mu$ is the Haar measure on the quotient nilmanifold $\overline{M}$.

We have an orthogonal decomposition

$$L^2(M) = \pi^\ast(L^2(\overline{M})) \oplus [\pi^\ast(L^2(\overline{M}))]^\perp,$$

where $\pi^\ast(L^2(\overline{M}))$ is the space of pull-backs of functions over $\overline{M}$, which coincides with the space of functions over $M$ which are $(\Phi_t^{h_{\Gamma}})_{t \in \mathbb{R}^d}$-invariant (and hence constant on the toral fibers). In the following, we will therefore identify functions on $\overline{M}$ with their pull-backs, which are defined on $M$ so that they are constant on each fiber $F_{\pi}$.

More precisely, if we let $\Lambda_{\Gamma}^\perp$ denote the dual lattice of $\Lambda_{\Gamma}$, we have a Fourier decomposition

$$L^2(M) = \bigoplus_{v \in \Lambda_{\Gamma}^\perp} H_v(h_{\Gamma}),$$

where

$$H_v(h_{\Gamma}) = \left\{ f \in \mathcal{C}^0(M) : f \circ \Phi_t^{h_{\Gamma}} = \exp(2\pi i v \cdot t) f \right\}.$$

In particular, for every function $f \in \mathcal{C}^0(M)$, there exists a family $(f_v)_{v \in \Lambda_{\Gamma}^\perp}$ of continuous functions such that we have the formula

$$f \circ \Phi_t^{h_{\Gamma}}(x) = \sum_{v \in \Lambda_{\Gamma}^\perp} f_v(x) \exp(2\pi i v \cdot t), \quad \text{for all } (x, t) \in M \times h_{\Gamma}.$$

We remark that the function $f_0 \in H_0(h_{\Gamma}) = \pi^\ast(L^2(\overline{M}))$ is the pull-back of a continuous function on $\overline{M}$ and hence is constant on all fibers $F_{\pi}$, while functions in $H_0(h_{\Gamma})^\perp = \bigoplus_{v \in \Lambda_{\Gamma}^\perp \setminus \{0\}} H_v(h_{\Gamma})$ are relative trigonometric polynomials without the term which is constant on the fibers $F_{\pi}$. Thus, we can write the orthogonal decomposition given in (6) also as

$$L^2(M) = H_0(h_{\Gamma}) \oplus H_0(h_{\Gamma})^\perp,$$

where $H_0(h_{\Gamma}) := \pi^\ast(L^2(\overline{M}))$, 

$$H_0(h_{\Gamma})^\perp := [\pi^\ast(L^2(\overline{M}))]^\perp = \bigoplus_{v \in \Lambda_{\Gamma}^\perp \setminus \{0\}} H_v(h_{\Gamma}).$$
**Definition 3.1** (Trigonometric polynomial with respect to $\mathfrak{h}_\Gamma$). Let $\mathfrak{h}_\Gamma \subset \mathbb{Z}(g)$ be any $\Gamma$-rational subspace. A function $f \in C^0(M)$ is called a trigonometric polynomial with respect to $\mathfrak{h}_\Gamma$ if in the expansion in the above formula (7) the family $(f_\mathbf{v})_{\mathbf{v} \in \mathcal{A}_\Gamma^+}$ has finite support.

For all $\mathbf{v} \in \mathcal{A}_\Gamma^+$, let $d(\mathbf{v}) \geq 0$ be the nonnegative integer defined by $\langle \mathbf{v}, \mathfrak{h}_\Gamma \rangle = d(\mathbf{v}) \mathbb{Z}$. The degree of a trigonometric polynomial $f$ is

$$\deg(f) = \max \{ d(\mathbf{v}) : f_\mathbf{v} \neq 0 \}.$$ 

Trigonometric polynomials with respect to $\mathfrak{h}_\Gamma$ will also be called *relative trigonometric polynomials* (when the toral fibers are clear from the context).

### 3.2. Basic properties of relative trigonometric polynomials

The following lemma provides an estimate on the measure of sub-level sets of relative trigonometric polynomials on nilmanifolds (see also [AFU] Lemma 4 for the case of trigonometric polynomials and Heisenberg nilflows).

**Lemma 3.2** (Level sets of relative trigonometric polynomials). For each $m \geq 1$ there exist constants $D_m$ and $d_m > 0$ such that the following holds. Let $f \in \bigoplus_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} H_\mathbf{v}(\mathfrak{h}_\Gamma)$ be a trigonometric polynomial with respect to $\mathfrak{h}_\Gamma$ of degree $m$ without constant term in the fibers. For any $C \geq 0$ and $\varepsilon \geq 0$, if

$$\mu \left( x \in M : \sum_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} |f_\mathbf{v}(x)| \leq C \right) \leq \varepsilon,$$

then, for all $\delta > 0$,

$$\mu \left( x \in M : |f(x)| \leq \delta \sum_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} |f_\mathbf{v}(x)| \right) \leq D_m \delta^{d_m} + \varepsilon.$$

Furthermore,

$$\mu \left( x \in M : |f(x)| \leq \delta C \right) \leq D_m \delta^{d_m} + \varepsilon.$$

**Proof.** Let us recall that each toral fiber $F_\mathbf{v}$ can be identified with the torus $\mathfrak{h}/\Lambda_\mathbf{v}$ and the measure $\mu_{\mathbf{v}}$ coincides (up to a constant) with the Lebesgue measure on $\mathfrak{h}/\Lambda_\mathbf{v}$. By classical results on level sets of trigonometric polynomials, see, e.g., [Br] Theorem 1.9, for each $m \geq 1$, if $p$ is a trigonometric polynomial defined on $F_\mathbf{v}$ of degree $m \geq 1$, then there exist constants $D_{m,p}$ and $d_m > 0$ such that for every $\delta > 0$, we have

$$\mu_{F_\mathbf{v}}(y \in F_\mathbf{v} : |p(y)| \leq \delta) \leq D_{m,p} \delta^{d_m}.$$ 

The constant $d_m$ depends only on $m$ and the constant $D_{m,p}$ depends continuously on $p$.

Let us define

$$E_C := \left\{ x \in M : \sum_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} |f_\mathbf{v}(x)| \leq C \right\}.$$ 

Recall that by assumption $\mu(E_C) \leq \varepsilon$ and notice that, since $|f_\mathbf{v}(x)|$ is $\Phi^m_{\mathbf{v}}(x)$-invariant, the set $E_C$ is a union of full fibers. In particular, the normalized trigonometric polynomial

$$\hat{f}(x) := \frac{f(x)}{\sum_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} |f_\mathbf{v}(x)|}$$

is well-defined on each fiber $F_\mathbf{v} \subset M \setminus E_C$, and its restriction $\hat{f}|_{F_\mathbf{v}}$ to $F_\mathbf{v}$, because of the normalization, lies in a compact set in the space of trigonometric polynomials of degree $1 \leq \deg(\hat{f}|_{F_\mathbf{v}}) \leq m$. Thus, there exist uniform positive constants $D_{m,d_m}$ such that the estimate (9) holds for any $\hat{f}$ as above. By (5) and Fubini Theorem, we deduce that the set

$$M_\delta := \left\{ x \in M \setminus E_C : |\hat{f}(x)| \leq \delta \right\} = \left\{ x \in M \setminus E_C : |f(x)| \leq \delta \sum_{\mathbf{v} \in \mathcal{A}_\Gamma^+ \setminus \{0\}} |f_\mathbf{v}(x)| \right\}$$
has measure at most \(D_m \delta^{d_m}\). Therefore, for all \(x \in M \setminus (M_\delta \cup E_C)\), we have
\[
|f(x)| > \delta \sum_{v \in A_1 \setminus \{0\}} |\hat{f}_v(x)| > \delta C.
\]
Using that by assumption \(\mu(M_\delta \cup E_C) \leq D_m \delta^{d_m} + \varepsilon\), the proof is complete. \(\Box\)

The following two corollaries of Lemma \(3.2\) state that the values of a trigonometric polynomial and its derivative along some direction are comparable on a large measure subset of the nilmanifold. Their proofs are given in \(\S 6\) (Appendix A).

For any element \(Z \in Z(G)\), let \([Z]_\Gamma \subset Z(G)\) denote the smallest \(\Gamma\)-rational subspace \(V \subset Z(G)\) such that \(Z \in V\).

**Corollary 3.3.** Let \(Z \in Z(G)\). For each \(m \geq 1\) there exist constants \(c_{m,Z}, D_m\) and \(d_m > 0\) such that the following holds. Let \(f \in \bigoplus_{V \in A_1 \setminus \{0\}} H_\delta(\mathfrak{h}_\Gamma)\) be a trigonometric polynomial with respect to \([Z]_\Gamma\) of degree \(m\). For any \(C \geq 0\) and \(\varepsilon \geq 0\), if
\[
\mu(x \in M : |f(x)| \leq C) \leq \varepsilon,
\]
then, for all \(\delta > 0\),
\[
\mu \left( x \in M : \left| \frac{d}{dx} f \circ \phi^x_{\delta}(x) \right| > c_{m,Z} \delta C \right) \geq 1 - D_m \delta^{d_m} - \varepsilon.
\]
**Corollary 3.4.** For each \(m \geq 1\) there exist constants \(D_m\) and \(d_m > 0\) such that the following holds. Let \(f \in \bigoplus_{V \in A_1 \setminus \{0\}} H_\delta(\mathfrak{h}_\Gamma)\) be a trigonometric polynomial with respect to \(\mathfrak{h}_\Gamma\) of degree \(m\). If
\[
\mu(x \in M : |f(x)| = 0) \leq \varepsilon_0,
\]
then, for all \(\delta > 0\),
\[
\mu \left( x \in M : \left| \frac{d}{dx} f \circ \phi^x_{\delta}(x) \right| \leq \frac{C_{m,Z}}{\delta} \left| f \circ \phi^x_{\delta}(x) \right| \text{ for all } s \in \left[ 0, \frac{\delta}{C_{m,Z}} \right] \right) \geq 1 - D_m \delta^{d_m} - \varepsilon_0,
\]
where \(C_{m,Z} := 4\pi m ||Z||_{1,m}\).

### 3.3. Algebraic induction towers

Let \(M\) and \(\{\phi^x_t\}_{t \in \mathbb{R}}\) be a nilmanifold and a nilflow satisfying the assumptions of Theorem \(1.1\). In this section we show that we can present \(M\) as a tower of extensions of nilmanifolds, so that at each step we are quotienting out a central fiber, chosen in a way that will be convenient for us to show prevalence of mixing by induction in the later sections.

The following notion of **Heisenberg triple** is central in determining how to build the tower.

**Definition 3.5** (Heisenberg triple). Let \(\mathfrak{g}\) be a nilpotent Lie algebra of step \(k \geq 2\). A triple \((X,Y,Z)\) of elements of \(\mathfrak{g}\) is called a **Heisenberg triple** if \(Z = \mathfrak{g}_k \subset Z(\mathfrak{g})\) and \([X,Y] = Z\).

The following lemma, which will be used as one step of the inductive construction, will guarantee the existence of Heisenberg triples at each step of the induction.

**Lemma 3.6** (Existence of Heisenberg triples). Let \(\mathfrak{g}\) be a nilpotent Lie algebra of step \(k \geq 2\). For any \(\Gamma\)-irrational element \(X \in \mathfrak{g}\) which is not in \([\mathfrak{g},\mathfrak{g}]\), there exist \(Y,Z\) such that \((X,Y,Z)\) is a Heisenberg triple.

**Proof.** Let \(A\) be the abelian group of rank \(n\) given by \(A = G/[G,G]\) and let \(\overline{\Gamma} = \Gamma/\left\langle \Gamma,\Gamma \right\rangle\); then \(\overline{\Gamma}\) is a lattice in \(A\). Let \(\{\exp E^1, \ldots, \exp E^n\}\) denote a set of generators of \(\overline{\Gamma}\), with \(\overline{F}_i \in \text{Lie}(A) \cong \mathfrak{g}/[\mathfrak{g},\mathfrak{g}]\). It is plain that the \(\overline{E}_i\)'s form a basis of \(\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]\).

Let \(E_{1,1}, E_{2,1}, \ldots, E_{n,1}, E_{1,2}, \ldots, E_{n,2}, \ldots, E_{1,k}, \ldots, E_{n,k}\) with \(n_1 = n\), be a Malcev basis for \(\mathfrak{g}\) through the descending central series \(\mathfrak{g}_j\) and strongly based at \(\Gamma\), as in Definition \(2.2\). Let \(\mathcal{E}'\) be as in Definition \(2.2\).

For any \(Y \in \mathcal{E}^{k-1}\) we have that, for any \(1 \leq j \leq n\), \([E_j, Y] \in \mathfrak{g}_k \subset Z(\mathfrak{g})\), since \(\mathcal{E}^{k-1} \subset \mathfrak{g}_{k-1}\). Thus, if we set
\[
Z := \sum_{j=1}^{n_1} \alpha_j(X)[E_j,Y],
\]
(where \( \omega_j(X) \), for \( 1 \leq j \leq n_1 \), are defined by (4)) we have that \( Z \in Z(\mathfrak{g}) \). We are left to check that \([X, Y] = Z \) and \( Z \) is non-zero.

By the Baker-Campbell-Hausdorff formula, we have \([\exp E^1_1, \exp Y] = \exp [E^1_1, Y] \). It follows that \([E^1_1, Y] \in \log Z(\Gamma) \) and there exists an integral matrix \((a_{ij}(Y))\) such that

\[
[E^1_1, Y] = \sum_{i=1}^{n_1} a_{ij}(Y)E^k_i.
\]

Finally, since \( \omega_0(X), \ldots, \omega_{n_1}(X) \) are linearly independent over \( \mathbb{Q} \), it follows that

\[
[X, Y] = \sum_{j=1}^{n_1} \omega_j(X)[E^1_1, Y] = \sum_{i=1}^{n_1} \left( \sum_{j=1}^{n_1} a_{ij}(X)a_{ij}(Y) \right) E^k_i \neq 0.
\]

The argument is complete.

Iterating inductively the previous Lemma (as shown in Corollary 3.8 below), we can construct towers of extensions corresponding to Heisenberg triples, in the following sense:

**Definition 3.7.** Let \( M = M^{(0)} = \Gamma^{(0)} \backslash G^{(0)} \) be a nilmanifold and let \( X = X^{(0)} \in \mathfrak{g} \) be an element not in \([\mathfrak{g}, \mathfrak{g}] \). A tower \( \mathcal{T}_{M, X} \) of Heisenberg extensions for \( M \) based at \( X \) is a sequence of nilmanifolds \( M^{(i)} = \Gamma^{(i)} \backslash G^{(i)} \), projections

\[
M = M^{(0)} \xrightarrow{\pi^{(1)}} M^{(1)} \xrightarrow{\pi^{(2)}} \cdots \xrightarrow{\pi^{(n)}} M^{(n)},
\]

and triples \((X^{(i)}, Y^{(i)}, Z^{(i)}) \in (\mathfrak{g}^{(i)})^3\), where \( \mathfrak{g}^{(i)} \) is the Lie algebra of the nilpotent group \( G^{(i)} \), such that

1. \((X^{(i)}, Y^{(i)}, Z^{(i)}) \) is a Heisenberg triple in \( \mathfrak{g}^{(i)} \) for all \( 0 \leq i \leq n \),
2. for all \( 0 \leq i \leq n-1 \), if we set \( h^{(i)} := [Z^{(i)}]_{\Gamma^{(i)}} \) and \( \Lambda^{(i)} := \log \Gamma^{(i)} \cap [Z^{(i)}]_{\Gamma^{(i)}} \), we have
   \[
   \mathfrak{g}^{(i+1)} = \mathfrak{g}^{(i)}/h^{(i)}, \quad \Gamma^{(i+1)} = \Gamma^{(i)}/\exp \Lambda^{(i)} = \Gamma^{(i)}/\left( \Gamma^{(i)} \cap \exp h^{(i)} \right),
   \]
   and \( \pi^{(i+1)} : M^{(i)} \to M^{(i+1)} \) is the canonical projection,
3. \( X^{(i+1)} = \pi^{(i+1)}(X^{(i)}) \) for all \( 0 \leq i \leq n-1 \).

A tower \( \mathcal{T}_{M, X} \) is maximal if \( G^{(n)} \) is Abelian and \( M^{(n)} \) is a torus.

Thus, at every step \( 0 \leq i < n \), \( M^{(i+1)} \) is a bundle over \( M^{(i)} \) with toral fiber \( h^{(i)}/\Lambda^{(i)} \). Iterating Lemma 3.6, we get the following corollary, which guarantees the existence of maximal towers.

**Corollary 3.8.** If \( X \in \mathfrak{g} \) is the generator of a uniquely ergodic nilflow on \( M \), then there exists a maximal tower of Heisenberg extensions for \( M \) based at \( X \).

**Proof.** The proof follows from Lemma 3.6 by induction on \( \dim[\mathfrak{g}, \mathfrak{g}] \). For each \( i = 0, \ldots, n-1 \), Theorem 2.1 implies that \( X^{(i)} \) is \( \Gamma^{(i)} \)-irrational, thus one can apply Lemma 3.6 to get the triple \((X^{(i)}, Y^{(i)}, Z^{(i)}) \) and the nilmanifold \( M^{(i+1)} \). By construction, the dimension \( \dim[\mathfrak{g}^{(i+1)}, \mathfrak{g}^{(i+1)}] \) is strictly lower than the dimension \( \dim[\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] \), thus this concludes the proof.

3.4. The class of time-changes. We now have all the tools to define the dense class of time-changes we will consider.

**Definition 3.9** (Trigonometric polynomials with respect to a tower). Let \( \mathcal{T} = \mathcal{T}_{M, X} \) be a maximal tower of Heisenberg extensions for \( M \) based at \( X \). We define the space of trigonometric polynomials \( \mathcal{P}_\mathcal{T} \) with respect to \( \mathcal{T} \) inductively in the following way.

1. Let \( \mathcal{P}^{(n)} \) denote the space of trigonometric polynomials over the torus \( M^{(n)} \).
2. For all \( 0 \leq i \leq n-1 \), let \( \mathcal{P}^{(i)} \) be the space of trigonometric polynomials \( f : M^{(i)} \to \mathbb{C} \) with respect to \( [Z^{(i)}]_{\Gamma^{(i)}} \), where the coefficient \( f_0 \) is in \( \mathcal{P}^{(i+1)} \).
3. Define \( \mathcal{P}_\mathcal{T} := \mathcal{P}^{(0)} \).
The dichotomy in Theorem 1.1 will be proved for time-changes in $\mathcal{P}_T$, i.e., positive, real-valued trigonometric polynomials $\alpha \in \mathcal{P}_T$ with respect to $T$. Let $\mathcal{C}^k(M)$ denote the space of $k$-times differentiable functions for $k \in \mathbb{N} \cup \{+\infty\}$. We show now that $\mathcal{P}_T$ is dense in $\mathcal{C}^k(M)$ (see Corollary 3.11). The key inductive step is provided by the following Lemma.

Let $\mathcal{M} = M/\exp \mathfrak{h}$, as in the beginning of §3.1 and consider relative trigonometric polynomials with respect to $\mathfrak{h}$.

**Lemma 3.10 (Relative density).** Let $\mathcal{F} \subset \mathcal{C}^k(\mathcal{M})$ be a dense set of functions. The set of relative trigonometric polynomials $f$ with respect to $\mathfrak{h}$ such that $f_0 \in \mathcal{F}$ is dense in $\mathcal{C}^k(\mathcal{M})$.

**Proof.** Since the group $\exp \mathfrak{h}_\mathcal{F}$ acts through the torus $\exp \mathfrak{h}_\mathcal{F}/\Gamma \cap \exp \mathfrak{h}_\mathcal{F}$, which is a compact Abelian group, by the theory of unitary representations for compact Abelian groups, for any functions $f \in \mathcal{C}^k(M)$ there exists an $L^2$-orthogonal expansion,

$$f = \sum_{v \in \Lambda_T} f_v,$$

into eigenfunctions of the action of $\exp \mathfrak{h}_\mathcal{F}$ on $M$. This implies that formula (7) holds in the space $\mathcal{C}^k(M)$. Since standard trigonometric polynomials (in several variables) are dense in the space $\mathcal{C}^k(\mathbb{T}^d)$, it follows that the space of all trigonometric polynomials with respect to $\mathfrak{h}$ is dense in $\mathcal{C}^k(M)$. Finally, since the set $\mathcal{F} \subset \mathcal{C}^k(\mathcal{M})$ is dense in $\mathcal{C}^k(\mathcal{M})$, it follows immediately that the set of all trigonometric polynomials with respect to $\mathcal{F}$ such that $f_0 \in \mathcal{F}$ is dense in $\mathcal{C}^k(M)$. □

**Corollary 3.11 (Density of $\mathcal{P}_T$).** Let $\mathcal{F} = \mathcal{F}_{MX}$ be a maximal tower of Heisenberg extensions for $M$ based at $X$. The space of trigonometric polynomials $\mathcal{P}_T$ is dense in $\mathcal{C}^\infty(M)$.

## 4. Tools for Mixing

In this section we prove several preliminary results which will provide fundamental tools for proving mixing of non-trivial time-changes in our class. In §4.1 we show how mixing can be deduced from a form of shearing (in measure) of curves in a certain direction. In §4.2 we compute the pushforward of curves by the flow. Finally, in the remaining sections, we prove a result on growth of ergodic sums for time-changes which are not given by coboundaries in the central fiber (see Theorem 4.4), which will be used to produce shearing in previously isometric directions. An outline of the proof of this result is given within §4.3.

### 4.1. Mixing via shearing

Let us consider an arc $\{\phi^w_r(x), 0 \leq r \leq s\}$ of the flow $\phi^w$ generated by some $W \in \mathfrak{g}$ and let us *push* it via the flow $\phi^V$. When these pushed arcs $\{\phi^v_r \circ \phi^w_r(x), 0 \leq r \leq s\}$ shear in the direction of $V$, one can hope to show, exploiting equidistribution of trajectories of $\phi^V$, that they equidistribute in $M$, namely, for every mean zero function $f \in L^2(M)$, the functions $\int_0^s f \circ \phi^v_r \circ \phi^w_r(x) \, dr$ converge to zero. The following Lemma shows in particular that if this convergence happens in measure (as well as in a weaker sense, see Lemma 4.2), this is sufficient to prove mixing. This is an instance of the mixing via shearing mechanism, mentioned in the introduction as a key strategy for proving mixing of parabolic flows.

**Lemma 4.1 (Mixing via shearing).** Let $W \in \mathfrak{g}$ be any vector field and let $f \in L^2(M)$ be bounded and with zero average. Let us assume that there exists a $\sigma > 0$ such that for all $s \in [0, \sigma]$, the functions $\int_0^t f \circ \phi^v_r \circ \phi^w_r(x) \, dr$ converge to zero in measure, i.e., for every $\eta > 0$ and any $\delta > 0$ there exists a $T > 0$ such that for every $t \geq T$

$$\mu\left(x \in M : \left|\int_0^t f \circ \phi^v_r \circ \phi^w_r(x) \, dr\right| \geq \eta\right) \leq \delta.$$

Then, for every $g \in L^2(M)$ such that $Wg \in L^2(M)$ we have

$$\langle f \circ \phi^v_r \circ g, \alpha \mu \rangle_{L^2(M, \alpha \mu)} \to 0.$$

In particular, it also follows that $\phi^v_r$ is mixing with respect to the invariant measure $\alpha \mu$. 
Instead of assuming converge to zero in measure the functions \( \int_0^s f \circ \phi_t^V \circ \phi_r^W (x) \, dr \) (i.e. that the pushed arcs equidistribute in measure) for all \( s \in [0, \sigma] \), to prove mixing via shearing it is actually sufficient to verify the assumptions in the following Lemma 4.2 which has a more delicate order of the quantifiers. This more general formulation will be used in some parts of the proof (see in particular the arguments at the end of §5.3).

**Lemma 4.2** (Mixing via shearing 2). Let \( W \in \mathfrak{g} \) be any vector field and let \( f \in L^2(M) \) be bounded and with zero average. The conclusion of Lemma 4.1 also holds if we assume that for every \( \delta > 0 \) there exists \( 0 < \sigma < 1 \) such that for every \( \eta > 0 \) there exists \( T > 0 \) such that for every \( t \geq T \) we have that for any \( s \in [0, \sigma] \),

\[
\mu \left( x \in M : \left| \int_0^s f \circ \phi_t^V \circ \phi_r^W (x) \, dr \right| \geq \eta \right) \leq \delta.
\]

Notice that the assumption of Lemma 4.2 is satisfied in particular under the assumptions of Lemma 4.1 (i.e. if the functions \( \int_0^s f \circ \phi_t^V \circ \phi_r^W (x) \, dr \) converge to zero in measure), thus we will only prove this more general result, which also implies Lemma 4.1.

**Proof of Lemma 4.2** (and hence of Lemma 4.2). As the Haar volume form is \( W \)-invariant, for any \( \sigma > 0 \) we have

\[
\langle f \circ \phi_t^V, g \rangle_{L^2(M, dm)} = \frac{1}{\sigma} \int_0^\sigma \langle f \circ \phi_t^V \circ \phi_s^W, g \circ \phi_s^W \rangle_{L^2(M, dm)} \, ds.
\]

Integrating by parts we derive the formula (as in [FU]):

\[
\frac{1}{\sigma} \int_0^\sigma \langle f \circ \phi_t^V \circ \phi_s^W, g \circ \phi_s^W \rangle \, ds = \frac{1}{\sigma} \int_0^\sigma \langle f \circ \phi_t^V \circ \phi_s^W \, ds, g \circ \phi_s^W \rangle - \frac{1}{\sigma} \int_0^\sigma \langle \int_0^s f \circ \phi_t^V \circ \phi_r^W \, dr, W g \circ \phi_s^W \rangle \, ds.
\]

Fix \( \varepsilon > 0 \) and let \( \delta > 0 \) be such that

\[
\sqrt{\delta} < \min \left\{ \frac{\varepsilon}{4 \| f \|_\infty \| g \|_2}, \frac{\varepsilon}{4 \| f \|_\infty \| W g \|_2} \right\}.
\]

Consider the associated \( 0 < \sigma < 1 \) given by the assumption and fix \( \eta > 0 \) such that

\[
\eta < \min \left\{ \frac{\sigma \varepsilon}{4 \| g \|_2}, \frac{\varepsilon}{4 \| W g \|_2} \right\}.
\]

Consider any \( t \geq T \), where \( T > 0 \) is given by the assumption. For any \( s \in [0, \sigma] \), let us define

\[
E_t^s := \left\{ x \in M : \left| \int_0^s f \circ \phi_t^V \circ \phi_r^W (x) \, dr \right| \geq \eta \right\}.
\]

Using the triangle and the Cauchy-Schwarz inequalities, we can bound the first term on the right-hand side of (10) by

\[
\begin{align*}
\left| \frac{1}{\sigma} \int_0^\sigma \langle f \circ \phi_t^V \circ \phi_s^W \, ds, g \circ \phi_s^W \rangle \right| & \leq \frac{1}{\sigma} \left| \langle \mathbb{I}_{E_t^s} \left( \int_0^\sigma f \circ \phi_t^V \circ \phi_s^W \, ds \right), g \circ \phi_s^W \rangle \right| \\
& + \frac{1}{\sigma} \left| \langle \mathbb{I}_{M \setminus E_t^s} \left( \int_0^\sigma f \circ \phi_t^V \circ \phi_s^W \, ds \right), g \circ \phi_s^W \rangle \right| \leq \| f \|_\infty \| g \|_2 \sqrt{\mu(E_t^s)} + \frac{\eta}{\sigma} \| g \|_2 \\
& \leq \| f \|_\infty \| g \|_2 \sqrt{\delta} + \frac{\eta}{\sigma} \| g \|_2.
\end{align*}
\]

By choice of the parameters \( \delta \) and \( \eta \) (see (11) and (12)), we get

\[
\left| \frac{1}{\sigma} \int_0^\sigma f \circ \phi_t^V \circ \phi_s^W \, ds, g \circ \phi_s^W \right| < \frac{\varepsilon}{2}.
\]
Similarly, for the second term on the right-hand side of (10),
\[
\left| \int_0^s f \circ \phi_t^V \circ \phi_r^W \, dr, W g \circ \phi_{\sigma}^W \right| \leq \left| \mathbb{I}_{E_t} \left( \int_0^s f \circ \phi_t^V \circ \phi_r^W \, dr \right), \mathbb{I}_{E_t} \circ \phi_{\sigma}^W \right|
\]
\[
+ \left| \mathbb{I}_{M^* \mathbb{E}_t} \left( \int_0^s f \circ \phi_t^V \circ \phi_r^W \, dr \right), \mathbb{I}_{M^* \mathbb{E}_t} \circ \phi_{\sigma}^W \right|
\]
\[
\leq \|W g\|_2 \|f\|_\infty \|\mu(E_t)\| + \|W g\|_2 \leq \frac{\varepsilon}{2},
\]
therefore, again by the choice of parameters (11) and (12), we conclude that for any \( t \geq T \)
\[
(f \circ \phi_t^V, g)_{L^2(M, d\mu)} < \varepsilon.
\]
This proves that \((f \circ \phi_t^V, g)_{L^2(M, d\mu)}\) tends to zero. Since this holds for every \( f, g \) as in the assumptions, this is also sufficient to to conclude that \( \phi^V \) is mixing with respect to its invariant measure \( \alpha \mu \) (since we can consider an observable of the form \( \alpha g \)). \( \square \)

4.2. **Pushforward of curves.** In order to apply the mixing via shearing arguments in the previous section we need to study pushforwards of curves given by some flow \( \phi^W \). In this section we hence compute the infinitesimal pushforward of a vector \( W \) in the Lie algebra.

Let \((X, Y, Z) \in \mathfrak{g}^3\) denote any Heisenberg triple. Recall we denote by \( V = \frac{1}{\alpha} X \) the time-change of \( X \). We have the commutations
\[
[V, Y] = \left[ \frac{1}{\alpha} X, Y \right] = -\left( Y \frac{1}{\alpha} \right) X + \frac{1}{\alpha} Z = \frac{\alpha}{\alpha} V + \frac{1}{\alpha} Z,
\]
\[
[V, Z] = \left[ \frac{1}{\alpha} X, Z \right] = \frac{Z \alpha}{\alpha} V.
\]
In particular, \( \{V, Y, Z\} \) generates a finite dimensional Lie subalgebra of the Lie algebra of vector fields. We will compute the tangent vector of the pushforwards by the flow \( \phi_t^V \) of curves tangent to the foliation given by \( \{V, Y, Z\} \). Let \( W \) be any vector in the Lie subalgebra generated by \( \{V, Y, Z\} \). We write
\[
(\phi_t^V)_*(W) = a_t V + b_t Y + c_t Z.
\]
By differentiation we derive
\[
\frac{da_t}{dt} V + \frac{db_t}{dt} Y + \frac{dc_t}{dt} Z = -V a_t V V b_t Y - b_t [V, Y] - V c_t Z - c_t [V, Z]
\]
\[
= - \left( V a_t + b_t \frac{Y}{\alpha} c_t \frac{Z \alpha}{\alpha} \right) V - V b_t Y - \left( \frac{1}{\alpha} V + V c_t \right) Z.
\]
or, in other terms,
\[
\frac{da_t}{dt} = -V a_t - b_t \frac{Y}{\alpha} - c_t \frac{Z \alpha}{\alpha},
\]
\[
\frac{db_t}{dt} = -V b_t,
\]
\[
\frac{dc_t}{dt} = -V c_t - b_t \frac{1}{\alpha}.
\]
It follows that
\[
\frac{d}{dt} (a_t \circ \phi_t^V) = -(b_t \circ \phi_t^V) \frac{Y}{\alpha} \circ \phi_t^V - (c_t \circ \phi_t^V) \frac{Z \alpha}{\alpha} \circ \phi_t^V,
\]
\[
\frac{d}{dt} (b_t \circ \phi_t^V) = 0,
\]
\[
\frac{d}{dt} (c_t \circ \phi_t^V) = -(b_t \circ \phi_t^V) \left( \frac{1}{\alpha} \circ \phi_t^V \right).
\]
By solving the system of ODEs above, we get the following expressions for the pushforwards of the vector fields \( Y \) and \( Z \) by the flow \( \phi_t^V \).
Lemma 4.3 (Infinitesimal pushforwards). For all \( x \in M \) and for all \( t \in \mathbb{R} \), the pushforwards of the vector fields \( Z \) and \( Y \) via \( \phi^t_\ast \) at the point \( \phi^t(x) \) are
\[
[\phi^t_\ast(Z)(\phi^t(x))] = Z(\phi^t(x)) - \left( \int_0^t \frac{Z}{\alpha} \circ \phi^t_\ast(x) \, d\tau \right) \circ \phi^t(x),
\]
\[
[\phi^t_\ast(Y)(\phi^t(x))] = Y(\phi^t(x)) - \left( \int_0^t \frac{1}{\alpha} \circ \phi^t_\ast(x) \, d\tau \right) \circ \phi^t(x)
\]
\[- \left( \int_0^t \frac{Y}{\alpha} \circ \phi^t_\ast(x) \, d\tau - \int_0^t \left( \frac{Z}{\alpha} \circ \phi^t_\ast(x) \right) \left( \int_0^\tau \frac{1}{\alpha} \circ \phi^t_\ast(x) \, dr \right) \, d\tau \right) \circ \phi^t(x).
\]

Notice that the shear, for example of an arc in direction of \( Z \) (see the first equation of Lemma 4.3) is described in terms of ergodic integrals of a function along the flow \( \phi^t \) (for example the coefficient in the direction of \( V \) for the pushforward of an arc in direction \( Z \) is given by the ergodic integral of the function \( Z\alpha/\alpha \), see first equation of Lemma 4.3). For this reason, to understand shear, we will now study the growth of ergodic integrals.

4.3. Growth of ergodic integrals. Let \((X, Y, Z) \in g \) be a Heisenberg triple and let \([Z]_\Gamma \subset Z(g)\) be the smallest \( \Gamma \)-rational subspace containing \( Z \). We recall from §3.1 that we have a Fourier
de
de
(13)
\[
L^2(M) = \bigoplus_{\nu \in \Lambda^\Gamma} H_\nu([Z]_\Gamma),
\]
where
de
(14)
\[
H_\nu([Z]_\Gamma) = \left\{ f \in L^2(M) : f \circ \Phi^{|Z|}_t = \exp(2\pi i (\nu \cdot t)) f \right\}.
\]
Let \( F_T \) denote ergodic integral
\[
F_T(x) = \int_0^T f \circ \phi^X_t(x) \, d\tau.
\]
We remark that for every \( t, T \in \mathbb{R} \), the following standard cocycle relation holds
\[
F_t(x) + F_T(\phi^t_\ast(x)) = F_{t+T}(x).
\]

We will be interested, to prove mixing in §5, to establish the growth in measure, as \( T \) grows, of the integral function \( F_T \), where \( f \) (in the integral defining \( F_T \)) is the function \( Z\alpha^t \). Indeed, by the calculations presented in §4.2 after the change of variable \( \tau = \tau(x, t) \), these types of integrals quantify the shearing of small curves which we crucially exploit to prove mixing (see the proof of Lemma 5.4 for details).

This section is devoted to the proof of the following result.

Theorem 4.4. Let \( f \in H_0([Z]_\Gamma) \subset L^2(M) \) be a trigonometric polynomial w.r.t \([Z]_\Gamma\), with \( f_\chi \in H_\nu([Z]_\Gamma) \). Assume that \( f \) is not a measurable coboundary for \( \phi^X_t \). Then, for every \( C > 1 \), we have
\[
\lim_{T \to \infty} \mu \{|F_T| < C\} = 0.
\]

The proof of this Theorem will take the remainder of the section.

4.4. Outline of the proof of Theorem 4.4. To prove Theorem 4.4 which follows the same scheme than in [AFU], we proceed in the following way. By a standard Gottschalk-Hedlund argument (see Lemma 4.5 below) we first prove that the Cesaro averages (in \( T \)) of the measure of the sets \( \{|F_T| < C\} \) tend to zero. From this, we want to derive that the measure of the set \( \{|F_T| < C\} \) tend to zero at \( T \) grows. To do this, we use an argument that we call decoupling: we show indeed that \( F_T \) and \( F_T \circ \phi^X_t \) become sufficiently independent for large values of \( t \), so that one can guarantee that the functions \( F_T \circ \phi^X_t \) and \( F_T \) are unlikely (in measure) to be simultaneously large when \( t \) is large (see Proposition 4.7 for the precise statement). The proof of Proposition 4.7 (which is a higher dimensional version of [AFU] Lemma 5), see also [Rav2] Lemma 5.2) is given separately in §4.5. At the end of this section we show that the decoupling result given by Proposition 4.7
can be combined with the convergence in average result of Lemma 4.5 to deduce the growth in Theorem 4.4. we use more precisely that one can find an arithmetic progressions $\{i_{0}\}_{i=1}^{\ell}$ such that $\mu\{|F_{0}| < C\} < \varepsilon$ (see Corollary 4.6) and then, thanks to the decoupling argument, apply a version of the inclusion-exclusion principle to conclude.

The following is a standard Gottschalk-Hedlund type of result. For completeness, we include a proof for convenience of the reader in §7 (Appendix B).

**Lemma 4.5 (Growth in Cesaro averages).** Assume $f$ is not a measurable coboundary for $\phi_{R}^{X}$. Then, for all $C > 1$,

$$\lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} \mu\{|F_{t}| < C\} \, dt = 0.$$ 

From the convergence of Cesaro averages, we can prove convergence along an arithmetic progressions in the following sense.

**Corollary 4.6 (Growth along arithmetic progressions).** Assume $f$ is not a measurable coboundary for $\phi_{R}^{X}$ and let $C > 1$. For all $\varepsilon > 0$ and for all $\ell \geq 1$, there exists an arithmetic progressions $\{i_{0}\}_{i=1}^{\ell}$ of length $\ell$ such that $\mu\{|F_{i_{0}}| < C\} < \varepsilon$.

**Proof.** Fix $\varepsilon > 0$ and consider $\text{Bad}_{\varepsilon} := \{t \geq 0 : \mu\{|F_{t}| < C\} \geq \varepsilon\}$. Let $\ell \in \mathbb{N}$ and fix $0 < \delta < 1/\ell$. By Lemma 4.5 there exists $T_{0} > 0$ such that $\text{Leb}([0, T] \cap \text{Bad}_{\varepsilon}) \leq \delta T$ for all $T \geq T_{0}$, so that, in particular, $\text{Leb}([0, jT] \cap \text{Bad}_{\varepsilon}) \leq \delta jT$ for all $j = 1, \ldots, \ell$. We want to find $t_{0} \in (0, T)$ such that $t_{0}, 2t_{0}, \ldots, t_{0} \notin \text{Bad}_{\varepsilon}$; in other words, $t_{0} \notin (0, T) \cap \text{Bad}_{\varepsilon}$ for all $j = 1, \ldots, \ell$.

We estimate the measure

$$\text{Leb}\left([0, T] \setminus \left( \bigcup_{j=1}^{\ell} (0, T) \cap \frac{1}{j} \text{Bad}_{\varepsilon}\right) \right) \geq T - \sum_{j=1}^{\ell} \text{Leb}\left((0, T) \cap \frac{1}{j} \text{Bad}_{\varepsilon}\right)$$

$$= T \left(1 - \sum_{j=1}^{\ell} \frac{1}{jT}\text{Leb}\left((0, jT) \cap \text{Bad}_{\varepsilon}\right)\right) \geq T(1 - \delta \ell),$$

which is greater than zero since $\delta < 1/\ell$. In particular, the set $\{t_{0} \in (0, T) : t_{0} \notin \text{Bad}_{\varepsilon} \text{ for all } j = 1, \ldots, \ell\}$ is not empty and the result follows.

The following is our decoupling result, which generalizes [AFU, Lemma 5] and [Rav2, Lemma 5.2] to the case of trigonometric polynomials with respect to $|Z|_{\Gamma}$ (in the sense of Definition 3.1). Let us consider trigonometric polynomials relative to $h := |Z|_{\Gamma}$ in the sense of Definition 3.1.

Recall the decomposition (8) of $L^{2}(M) = H_{0}(h)^{\perp} H_{0}(h)^{\perp}$

**Proposition 4.7 (Decoupling).** Let $f \in H_{0}(h)^{\perp}$ be a trigonometric polynomial w.r.t $|Z|_{\Gamma}$ of degree $m$, and let $C > 1$. For every $\varepsilon > 0$ there exist $C_{0} > 1$ and $\varepsilon_{0} > 0$ such that for every $t > 0$ for which $\mu\{|F_{t}| < C_{0}\} < \varepsilon_{0}$, there exists $T_{0} > 0$ such that for all $T \geq T_{0}$ we have

$$\mu\{|F_{T} \circ \phi_{R}^{X} - F_{T}| < 2C\} < \varepsilon.$$ 

The proof of Proposition 4.7 is given in the following subsection §4.5. We now use Corollary 4.6 and Proposition 4.7 to prove Theorem 4.4.

**Proof of Theorem 4.4.** Let $C > 1$ be fixed. By Lemma 4.5, we have that $\lim_{t \to \infty} \mu\{|F_{t}| < C\} = 0$. Since we want to prove that $\lim_{t \to \infty} \mu\{|F_{t}| < C\} = 0$, it is enough to assume by contradiction that $L := \limsup_{t \to \infty} \mu\{|F_{t}| < C\} > 0$. Let us choose $\ell \in \mathbb{N}$ and $\varepsilon > 0$ such that

$$\frac{1}{\ell} + \frac{\ell + 1}{2} \varepsilon < \frac{L}{2}. \quad (16)$$
Consider now $C_0 > 1$ and $\varepsilon_0 > 0$ given by Proposition 4.7 and let $\{i_0\}_{i=1}^\ell$ be an arithmetic progression of length $\ell$ as in Corollary 4.6 such that $\mu \{|F_{i_0}| < C_0\} < \varepsilon_0$. Let $T_0(i) > 0$ be given by Proposition 4.7 for $t = i_0$ and denote by $T_0$ the maximum of all $T_0(i)$ for $i = 1, \ldots, \ell$.

For all $T \geq T_0$, the Inclusion-Exclusion Principle yields

$$
\mu \left( \bigcup_{i=1}^\ell \phi_{i_0}^X([F_T| < C] \right) \geq \sum_{i=1}^\ell \mu \left( \phi_{i_0}^X([F_T| < C] \right) - \sum_{1 \leq j < \ell} \mu \left( \phi_{i_0}^X([F_T| < C] \cap \phi_{j_0}^X([F_T| < C] \right).
$$

For all $1 \leq j < i \leq \ell$, since $\phi_{i_0}^X$ is measure-preserving, we have

$$
\mu \left( \phi_{i_0}^X([F_T| < C] \cap \phi_{j_0}^X([F_T| < C) \right) = \mu \left( \{F_T \circ \phi_{i_0}^X \cap \phi_{j_0}^X \right) \leq \mu \left( |F_T \circ \phi_{i_0}^X - F_T \circ \phi_{j_0}^X| < 2C \right) = \mu \left( |F_T \circ \phi_{(i-j)0} - F_T| < 2C \right),
$$

which by Proposition 4.7 is less than $\varepsilon$, since $0 \leq i-j \leq \ell$ and $T \geq T_0 \geq T_0(i-j)$.

Choose now $T \geq T_0$ for which $\mu \{|F_T| < C\} > L/2$. By the computations above and since the flow $\phi_{i_0}^X$ is measure-preserving, we thus obtain

$$
1 \geq \mu \left( \bigcup_{i=1}^\ell \phi_{i_0}^X([F_T| < C] \right) \geq \ell \mu \left( |F_T| < C \right) - \ell(\ell+1)/2 \geq \ell L/2 - \ell(\ell+1)/2 \varepsilon.
$$

This yields the inequality $L/2 \leq 1/\ell + (\ell+1)/2 \varepsilon$, in contradiction with the initial assumption (16). The proof is therefore complete.

4.5. Decoupling: proof of Proposition 4.7. This subsection (the final one in this Section 4) is devoted to the proof of Proposition 4.7. Let us first make some preliminary remarks. Let

$$
f(x) = \sum_{v \in \Lambda^+_1(0)} f_v(x) \in L^2(M)
$$

be a trigonomeric polynomial with respect to $[Z]_\Gamma$ of degree $m \geq 1$ according to Definition 3.1. Denote by $N \geq 1$ the cardinality of $v \in \Lambda^+_1$ such that $f_v$ is non zero. Let

$$
F^V_T(x) = \int_0^T f_v \circ \phi^X_t(x) \, dt,
$$

so that we can write $F_T(x) = \sum_{v \in \Lambda^+_1(0)} F^v_T(x)$. By the cocycle relation for ergodic integrals (15), for any $t, T > 0$, we have

$$
F_T \circ \phi^X_t - F_T = F_t \circ \phi^X_t - F_t.
$$

Let us first notice that, since $[Z]_\Gamma \subset Z(g)$, by (7), we have

$$
F^V_t \circ \Phi^Z_{i_0}(x) = F^V_t(x) \cdot \exp(2\pi i \langle v, t \rangle);
$$

in particular, the function $F_t \circ \Phi^Z_{i_0} - F_t$ is a trigonomeric polynomial with respect to $[Z]_\Gamma$ of at most the same degree $m$ (since if $f_v = 0$ then $F^v_T = \int_0^T f_v \circ \phi^X_t \, dt = 0$ and $F_t \circ \Phi^Z_{i_0} - F_t \equiv 0$). Hence, by (7), we have

$$
(F_t \circ \Phi^Z_{i_0} - F_t) \circ \Phi^Z_{i_0}(x) = \sum_{v \in \Lambda^+_1(0)} (F^v_T \circ \Phi^Z_{i_0} - F^v_T)(x) \cdot \exp(2\pi i \langle v, t \rangle).
$$

Proof of Proposition 4.7. Let us first fix the parameters that we will use in the proof.

Choice of parameters. Let $D_m, d_m$ be given by Lemma 3.2. Let $C > 1$ and $\varepsilon > 0$; fix $C_0 > N^2$ (where $N$ defined above is the number of non zero $f_0$) and $\varepsilon_0 > 0$ such that

$$
D_m \left( \frac{4C}{\sqrt{C_0}} \right)^{d_m} + \frac{2N}{\sqrt{C_0}} + \varepsilon_0 < \varepsilon.
$$

Define $\delta = 4C/\sqrt{C_0}$. 
Fix $t > 0$ as in the assumptions of Proposition 4.1, i.e. such that $\mu \{ |F_t| < C_0 \} < \varepsilon_0$. In particular, since $|F_t| \leq \sum |F_t^v|$, we have that

\begin{equation}
\mu(N_t) < \varepsilon_0, \quad \text{where} \quad N_t := \left\{ x \in M : \sum_{v \in \Lambda_t^v \setminus \{0\}} |F_t^v(x)| < C_0 \right\}.
\end{equation}

By uniform continuity of $F_t^v$, let $s_0 > 0$ be such that if $|s| \leq s_0$ and $x' := \phi_s^v(x)$ then $|F_t^v(x) - F_t^v(x')| < 1/4$ for all $v \in \Lambda_t^v \setminus \{0\}$ (we use here that $f$ being a trigonometric polynomial, the number of non-zero $f_x$ and hence of non-zero $F_v$ is finite).

We remark that, since $|Z|_T$ is the smallest $\Gamma$-rational subspace containing $Z$, then $\langle v, Z \rangle \neq 0$ for all $v \in \Lambda_T^v \setminus \{0\}$. Let $c_{m, Z} > 0$ be the minimum of $|\langle v, Z \rangle|$ for all $v \in \Lambda_T^v \setminus \{0\}$. Fix $T_0 > 2/c_{m, Z} s_0$, and let $T \geq T_0$.

Finally, choose $\theta \in (0, \pi/2)$ such that $1/\sqrt{C_0} \leq \sin \theta \leq 2/\sqrt{C_0}$. In this way, we also have

\begin{equation}
\frac{\theta}{\pi} < \sin \left( \frac{\theta}{2} \right) \leq \sin \theta \leq \frac{1}{\sqrt{C_0}}.
\end{equation}

Strategy. We want to apply Lemma 3.2 to $F_T \circ \phi_T^v - F_T = F_T \circ \phi_T^v - F_t$. Consider the set

\[ E_{C_0} := \left\{ x \in M : \sum_{v \in \Lambda_T^v \setminus \{0\}} |(F_T^v \circ \phi_T^v - F_T^v)(x)| < \frac{\sqrt{C_0}}{2} \right\}. \]

If we can prove that

\begin{equation}
\mu(E_{C_0}) < \varepsilon_0 + \frac{2N}{\sqrt{C_0}},
\end{equation}

then Lemma 3.2 (for $\delta := 4C/\sqrt{C_0}$) and the choice of parameters (see 17) give that

\[ \mu \{ |F_T \circ \phi_T^v - F_T| < 2C \} \leq D_m \left( \frac{4C}{\sqrt{C_0}} \right)^{d_m} + \mu(E_{C_0}) \leq D_m \left( \frac{4C}{\sqrt{C_0}} \right)^{d_m} + \varepsilon_0 + \frac{2N}{\sqrt{C_0}} < \varepsilon, \]

which proves the Lemma. We hence just have to prove the estimate (20) for $\mu(E_{C_0})$.

Reduction of space to time estimates. Let us remark that, since the Haar measure $\mu$ is invariant under the flow $\phi_T^v$, we have

\begin{equation}
\mu(E_{C_0}) = \frac{1}{s_0} \int_0^{s_0} \left( \int_M \mathbb{1}_{E_{C_0}} \circ \phi_s^v \, d\mu \right) \, ds = \int_M \left( \frac{1}{s_0} \int_0^{s_0} \mathbb{1}_{E_{C_0}} \circ \phi_s^v \, ds \right) \, d\mu \end{equation}

\begin{equation}
= \int_M \frac{1}{s_0} \mathbb{Leb} \left( s \in [0, s_0] : \sum_{v \in \Lambda_T^v \setminus \{0\}} |(F_T^v \circ \phi_T^v - F_T^v)(s) \circ \phi_s^v(x)| < \frac{\sqrt{C_0}}{2} \right) \, d\mu.
\end{equation}

For any fixed $v$, consider the term $(F_T^v \circ \phi_T^v - F_T^v) \circ \phi_s^v(x)$. We notice that, by the Heisenberg commutation relations and the Baker-Campbell-Hausdorff Formula, for any $s \in \mathbb{R}$ we have

\[ \phi_{-sT}^v \circ \phi_T^v \circ \phi_s^v(x) = \Gamma \exp(sY) \exp(TX) \exp(-sY) = \Gamma \exp(TX - sZ) \]

\[ = \phi_T^v \circ \phi_{-sT}^v(x). \]

Therefore, by recalling the definition of $H_v([Z]_T)$ in §3.1

\[ (F_T^v \circ \phi_T^v - F_T^v) \circ \phi_s^v = F_T^v \circ \phi_T^v \circ \phi_s^v \circ \phi_{-sT}^v - F_T^v \circ \phi_s^v = e(v(-sT))F_T^v \circ \phi_s^v - F_T^v \circ \phi_s^v, \]

where we denoted $e_v(r) = \exp(2\pi ir \langle v, Z \rangle)$. We hence have to estimate for fixed $x$ (which will be chosen in certain set)

\[ \mathbb{Leb} \left( s \in [0, s_0] : \sum_{v \in \Lambda_T^v \setminus \{0\}} |e_v(-sT)F_T^v \circ \phi_s^v - F_T^v \circ \phi_s^v(x)| < \frac{\sqrt{C_0}}{2} \right). \]
In order to use decoupling arguments to estimate the measure of the above parameters $s \in [0, s_0]$, we first want to localize the function, i.e. show that $F_t^y \circ \phi_t^x$ does not vary much in $s$, so we can eliminate the dependence on $\phi_t^x$.

**Localization.** Let us show that we can assume that $F_t^y \circ \phi_t^x$ and $F_t^y \circ \phi_t^x$ are essentially constant and equal to $F_t^y \circ \phi_t^x$ and $F_t^y \circ \phi_t^x$ (corresponding to $s = 0$). By triangle inequality, adding and subtracting $(F_t^y + e_v(-sT)F_t^y \circ \phi_t^x)(x)$, we can write

$$
\sum_{v \in \Lambda_1^r \setminus \{0\}} |(F_t^y \circ \phi_t^x - F_t^y) \circ \phi_t^x(x)| = \sum_{v \in \Lambda_1^r \setminus \{0\}} |e_v(-sT)F_t^y \circ \phi_t^x \circ \phi_t^x(x) - F_t^y \circ \phi_t^x(x)|
$$

$$
\geq \sum_{v \in \Lambda_1^r \setminus \{0\}} \left( |e_v(-sT)F_t^y \circ \phi_t^x(x) - F_t^y(x)| - |F_t^y \circ \phi_t^x(x) - F_t^y(x)| \right)
$$

$$
- |e_v(-sT)| \cdot |F_t^y \circ \phi_t^x \circ \phi_t^x(x) - F_t^y \circ \phi_t^x(x)|;
$$

Hence, for any $0 \leq s \leq s_0$, by definition of $s_0$ and of $N$ (as number of non-zero coefficients $F_t^y$) we have

$$
\sum_{v \in \Lambda_1^r \setminus \{0\}} |(F_t^y \circ \phi_t^x - F_t^y) \circ \phi_t^x(x)| \geq \sum_{v \in \Lambda_1^r \setminus \{0\}} \left( |e_v(-sT)F_t^y \circ \phi_t^x(x) - F_t^y(x)| - \frac{1}{4} - \frac{1}{4} \right)
$$

$$
\geq \left( \sum_{v \in \Lambda_1^r \setminus \{0\}} \left| e_v(-sT)F_t^y \circ \phi_t^x(x) - F_t^y(x) \right| \right) - \frac{N}{2}.
$$

**Decoupling.** We now hence want to estimate the sum in the last equation. Let $c_1 := F_t^y(x) \in \mathbb{C}$ and $c_2 := F_t^y \circ \phi_t^x(x) \in \mathbb{C}$. Since, as we already remarked, $\langle v, Z \rangle \neq 0$ for all $v \in \Lambda_1^r \setminus \{0\}$, by elementary trigonometry, any point $c' \in \mathbb{C}$ outside the cone of 1/2-angle $\theta$ about the line $\mathbb{R}c_1$ has distance $|c' - c_1|$ from $c_1$ larger than the distance of $c_1$ from the boundary of the cone (see Figure 1). Thus, for the parameter $\theta$ chosen as in the beginning, $|e_v(-sT)c_2 - c_1| > |c_1| \sin \theta$ as long as the phase $\pi(-sT)2 \mod 2\pi$ does not fall in an interval of size $2\theta$. Summing the estimates for the (at most $N$) different $v$ which are non-zero, the measure of the set of $s \in [0, s_0]$ for which $|e_v(-sT)c_2 - c_1| \leq |c_1| \sin \theta$ is bounded by $s_0 \theta / \pi$ plus a “boundary error”: more precisely,

$$
\text{Leb} \left( s \in [0, s_0] : |e_v(-sT)F_t^y \circ \phi_t^x(x) - F_t^y(x)| \leq |F_t^y(x)| \sin \theta \right) \leq s_0 \frac{\theta}{\pi} + 2 \frac{\theta}{\pi} \cdot \frac{1}{|\langle v, Z \rangle|}.
$$

**Figure 1.** Any point $c' \in \mathbb{C}$ outside the cone of 1/2-angle $\theta$ about the line $\mathbb{R}c_1$ has distance $|c' - c_1|$ from $c_1$ larger than the distance of $c_1$ from the boundary of the cone.
Recall we defined \( c_{m,Z} > 0 \) to be the minimum of \(|\langle v, Z \rangle| \) for all \( v \in \Lambda_Γ \setminus \{0\} \). Thus, outside a set of \( s \in [0, s_0] \) of measure at most \( N(s_0/\pi + 2\theta / (\pi c_{m,Z} T)) \), we have that
\[
\left( \sum_{v \in \Lambda_Γ \setminus \{0\}} |e_v (-sT) F^X_1 \circ \phi^Y_1 (x) - F^X_1 (x)| \right) - \frac{N}{2} \geq \left( \sum_{v \in \Lambda_Γ \setminus \{0\}} |F^Y_1 (x)| \sin \theta \right) - \frac{N}{2}.
\]
\[
\geq \sum_{v \in \Lambda_Γ \setminus \{0\}} \frac{|F^Y_1 (x)|}{\sqrt{C_0}} \cdot \frac{\sqrt{C_0}}{2}.
\]
When \( x \notin N_i \), by the definition in (13), the last term above is greater or equal to \( \sqrt{C_0}/2 \). Therefore, for all \( x \notin N_i \),
\[
\text{Leb} \left( s \in [0, s_0] : \sum_{v \in \Lambda_Γ \setminus \{0\}} |(F^X_1 \circ \phi^Y_1 - F^X_1) \circ \phi^Y_1 (x)| < \frac{\sqrt{C_0}}{2} \right) \leq N \left( \frac{s_0}{\pi c_{m,Z} T} + \frac{\pi c_{m,Z} T}{\sqrt{C_0}} \right).
\]
After dividing by \( s_0 \) and integrating over \( M \setminus N_i \), we get
\[
\int_{M \setminus N_i} \frac{1}{s_0} \text{Leb} \left( s \in [0, s_0] : \sum_{v \in \Lambda_Γ \setminus \{0\}} |(F^X_1 \circ \phi^Y_1 - F^X_1) \circ \phi^Y_1 (x)| < \frac{\sqrt{C_0}}{2} \right) \, d\mu \leq N \frac{\theta}{\pi} \left( 1 + \frac{2}{s_0 c_{m,Z} T} \right) < \frac{2N}{\sqrt{C_0}}.
\]
From (21) and (22), it follows that \( \mu (E_{C_0}) < \mu (N_i) + \frac{2N}{\sqrt{C_0}} \leq \epsilon_0 + \frac{2N}{\sqrt{C_0}} \), which concludes the proof of (20) and hence of the Proposition.

5. Inductive Proof of Mixing

We can now prove the main result, Theorem 1.1. We will decompose the time-change using the towers of nilflows extensions built in §3.3 (see §5.1), then explain the two mechanisms which, at each step of the tower, either allow to directly produce mixing from stretching of Birkhoff sums (this is the non-coboundary case in §3.3), or, in the coboundary case in §5.2, allow to lift mixing from the factor to the extension. The proof by induction, which combines these two steps, is then given at the end, in §5.4.

Let us recall (see §2.1) that the time-change induced by \( V = \frac{1}{\alpha} X \) is measurably trivial if and only if \( \frac{1}{\alpha} \) is a measurable almost coboundary for \( \phi^V \), or, equivalently, if and only if \( \alpha \) is a measurable almost coboundary for \( \phi^X \).

5.1. Decomposition of the time-change along the tower. Given a uniquely ergodic nilflow \( \phi^X \) on the nilmanifold \( M \), let \( \mathcal{T}_{M,X} \) be a be a tower of Heisenberg extensions for \( M \) based at \( X \), whose existence is guaranteed by Corollary 3.8. Consider time-changes which are positive, real-valued trigonometric polynomials belonging to the class \( \mathcal{P}_\gamma \) consisting of trigonometric polynomials with respect to this tower (see Definition 3.9). Density of this class follows from Corollary 3.11.

Let \( \alpha \in \mathcal{P}_\gamma \) be a trigonometric polynomials with respect to this tower (see Definition 3.9), and set \( \alpha^{(0)} := \alpha \). Recall (from §3.3) that we denote by \( \hat{h}^{(i)} := [Z]_{\Gamma^0} \) and \( \Lambda^{(i)} := \log \Gamma^{(i)} \cap [Z]_{\Gamma^0} \) for all \( 0 \leq i \leq n \).

Assume we defined \( \alpha^{(i)} \in \mathcal{P}^{(i)} \subset \mathcal{C}^{\infty} (M^{(i)}) \) (where \( \mathcal{P}^{(i)} \) is the space of trigonometric polynomials relative to \( \hat{h}^{(i)} \) defined in Definition 3.9) for \( 0 \leq i < n \). We consider the decomposition associated to the orthogonal decomposition (defined in §3.1)
\[
L^2 (M^{(i)}) = H_0 (\hat{h}^{(i)}) \oplus H_0 (\hat{h}^{(i)})^\perp, \quad \text{where} \quad H_0 (\hat{h}^{(i)}) := (\pi^{(i+1)})^* (L^2 (M^{(i+1)})),
\]
namely we write
\[
\alpha^{(i)} (x) = \overline{(\alpha^{(i)}) (x)} + (\alpha^{(i)})^\perp (x),
\]
where
\[
\overline{(\alpha^{(i)}) (x)} = \int_{\hat{h}^{(i)}/\Lambda^{(i)}} \alpha \circ \Phi^{\hat{h}^{(i)}}_t (x) \, dt \in H_0 (\hat{h}^{(i)})
\]
is a strictly positive, \( h^{(i)} \)-invariant function, and
\[
(\alpha^{(i)})_{+}(x) = \alpha^{(i)}(x) - (\alpha^{(i)})(x) \in H_{0}(h^{(i)})_{+}.
\]
We now set \( \alpha^{(i+1)} := (\alpha^{(i)})(x) \). By definition of \( \mathcal{P}_{\gamma} \) and \( \mathcal{P}^{(i)} \) (see Definition 3.9), \( \alpha^{(i+1)} \in \mathcal{P}^{(i+1)} \), thus we can continue the inductive definition until \( i + 1 = n \).

For each \( 0 \leq i \leq n \), let us denote by \( V^{(i)} \) the time-change of the vector field \( X^{(i)} \) on \( M^{(i)} \) given by
\[
V^{(i)} = \frac{1}{\alpha^{(i)}} X^{(i)}.
\]
In the final proof, for some \( 0 \leq i \leq n \), we will need to consider two cases:

(i) \( \text{the function } (\alpha^{(i)})_{+} \text{ is a measurable coboundary w.r.t. } X^{(i)} \);

In this case (treated in §5.2), we will show that if the flow generated by \( V^{(i+1)} \) on \( M^{(i+1)} \) is mixing, then also the flow generated by \( V^{(i)} \) on \( M^{(i)} \) is mixing (see Proposition 5.2 in §5.2).

(ii) \( \text{the function } (\alpha^{(i)})_{+} \text{ is not a measurable coboundary w.r.t. } X^{(i)} \);

In this case (treated in §5.3) we will show directly (using the growth of Birkhoff sums proved in §4.3) that the flow generated by \( V^{(i)} \) on \( M^{(i)} \) is mixing (see Proposition 5.3 in §5.3).

**Notation.** In order to keep the notation simpler, in the following two sections we will drop the index \( i \), hence we will simply write \( M = \Gamma \backslash G \) for a nilmanifold, \( X \) for the generator of a uniquely ergodic nilflow, \( \{X, Y, Z\} \) for a Heisenberg triple, and \( \alpha \in C^{\infty}(M) \) for a positive, real-valued trigonometric polynomial with respect to \( h := [Z]_{\Gamma} \), that is the smallest \( \Gamma \)-rational subspace containing \( Z \). The Propositions proved in the next two sections will then be applied to \( M^{(i)}, X^{(i)} \) and the time-change \( V^{(i)} = \frac{1}{\alpha^{(i)}} X^{(i)} \) given by \( \alpha^{(i)} \) defined above.

**5.2. Case (i) (coboundary case).** In this section we consider the case in which \( \alpha_{+}(x) = \alpha(x) - 77 \in H_{0}([Z]_{\Gamma})_{+} \) is a measurable coboundary w.r.t. \( X \). In this case (thanks to the following Lemma 5.1), we will show that the flow \( \phi^{V} \) projects on a flow on the quotient manifold \( \bar{M} = M / \exp[Z]_{\Gamma} = M / h_{\Gamma} \) and that if the projected flow is mixing, also the original flow was mixing (see Proposition 5.2). This is the case in which to prove mixing we will exploit the intrinsic dynamics of nilflows and especially the shearing mechanism of *wrapping in the fibers*.

We first want to show that it is possible to define a flow on the quotient nilmanifold \( \bar{M} \). By the standard theory of time-changes, in case (i) the time-change is measurably conjugate to a time-change with \([Z]_{\Gamma}\)-invariant time-change function. We can therefore assume that the function \( \alpha \) is \([Z]_{\Gamma}\)-invariant.

**Lemma 5.1.** If \( V = \frac{1}{\alpha} X \) is the generator of a time-change with \( \alpha > 0 \) a \([Z]_{\Gamma}\)-invariant function, then
\[
\phi_{s}^{V} \circ \Phi_{t}^{[Z]_{\Gamma}} = \Phi_{t}^{[Z]_{\Gamma}} \circ \phi_{s}^{V}, \quad \text{for all } s \in \mathbb{R} \text{ and } t \in [Z]_{\Gamma}.
\]

**Proof.** Let \( W \in [Z]_{\Gamma} \). Since two smooth flows commute if and only if their generators commute, it is enough to verify that the generators \( V = \frac{1}{\alpha} X \) and \( W \) commute. Since \( W \in Z(g) \) and the function \( \alpha \) is \( W \)-invariant, we have
\[
[V, W] = \left[ \frac{1}{\alpha} X, W \right] = \frac{1}{\alpha} [X, W] - \left( W \frac{1}{\alpha} \right) X = 0.
\]
The argument is therefore complete. \( \square \)

Let now \( \mathfrak{g} \cong g / [Z]_{\Gamma}, \bar{G} = G / \exp[Z]_{\Gamma} \) and let \( \bar{M} = M / \exp[Z]_{\Gamma} \).

Since the action \( \Phi_{t}^{[Z]_{\Gamma}} \in \mathcal{R}_{\mathfrak{g}} \) induces a non-singular toral action, the quotient \( \bar{M} = \Gamma \backslash \bar{G} \) is a nilmanifold, and \( M \) is a toral bundle \( \pi: M \to \bar{M} \) over \( \bar{M} \). By Lemma 5.1 the time-change \( \phi^{V} \) projects to a time-change \( \phi^{\bar{V}} \) of a nilflow \( \phi^{\bar{V}} \) on \( \bar{M} \). We remark that the invariant measure for \( \phi^{\bar{V}} \)
from the definition that

By assumption, the latter term above has the correct asymptotics, that is

In the present case we choose

hence it is equivalent to control correlations with respect to the Haar measure

As above there exists an orthogonal decomposition

Let $f = \pi^*(\tilde{f}) \in \pi^*(L^2(M))$, $g \in L^2(M)$ and let $\pi^*(\tilde{g})$ with $\tilde{g} \in L^2(M)$ denote the orthogonal component of $g$ in $\pi^*(L^2(M))$. We have

By assumption, the latter term above has the correct asymptotics, that is

It remains to prove that whenever $f \in H_0(\|Z\|_{\Gamma}) = \pi^*(L^2(M))$ we have

In the present case we choose $W = Y$ in Lemma 5.1, and, for every $\sigma > 0$, we consider the pushed curves

For all $t \in \mathbb{R}$, let $A_t$ and $B_t$ denote the functions

By Lemma 4.3 by taking into account that $Z\alpha = 0$, we have

Let $a > 0$ denote the minimum of $1/\alpha$ on $M$. Then, for all $t \in \mathbb{R}$ and for all $x \in M$, we have $B_t(x) > a\alpha$ and, by the ergodic theorem (and unique ergodicity of irrational nilflows),

uniformly in $x \in M$. We write

and after integration by parts

By Lemma 4.1 for $W = Y$, it is enough to prove that the two terms on the right-hand side of (23) converge pointwise to zero for any $f \in H_0(\|Z\|_{\Gamma})$.
First term of the RHS of (24). By (23), we can write the second integral in the RHS of (24) in terms of a path integral along $\gamma_{\sigma}$, that is,
\[
\int_{0}^{\sigma} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, ds = \int_{\gamma_{\sigma}} f \tilde{Z}.
\]

Let us notice that the Lie derivative operator $\mathcal{L}_Z$ is invertible on $H_0([\gamma])$; we can therefore write $f = Zg$. It is not restrictive to assume that $g$ is smooth. Let $\{\tilde{V}, \tilde{Y}, \tilde{Z}\}$ be a frame of 1-forms dual to $\{V, Y, Z\}$. By (23), we have
\[
\int_{\gamma_{\sigma}} Zg \tilde{Z} = \int_{\gamma_{\sigma}} dg - \int_{\gamma_{\sigma}} Yg \tilde{Y} - \int_{\gamma_{\sigma}} Vg \tilde{V};
\]

hence, we derive the formula
\[
\int_{0}^{\sigma} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, ds = \int_{\gamma_{\sigma}} dg - \int_{0}^{\sigma} Yg \circ \phi_t^Y \circ \phi_t^Y \, ds
\]
\[
- \int_{0}^{\sigma} (A_t \circ \phi_t^Y) Vg \circ \phi_t^Y \circ \phi_t^Y \, ds.
\]

By Stokes Theorem,
\[
\left| \int_{0}^{\sigma} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, ds \right| \leq 2 \|g\|_\infty + \|Yg\|_\infty + \|Vg\|_\infty \int_{0}^{\sigma} |A_t \circ \phi_t^Y| \, ds,
\]
thus, we can bound the first term on the right-hand side of (24) by
\[
\left| (B_t \circ \phi_t^Y)^{-1} \int_{0}^{\sigma} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, ds \right|
\]
\[
\leq \frac{2 \|g\|_\infty + \|Yg\|_\infty + \|Vg\|_\infty}{a} \int_{0}^{\sigma} \frac{|A_t \circ \phi_t^Y|}{t} \, ds.
\]

Since $|A_t \circ \phi_t^Y|/t$ converges to zero uniformly in $M$, we deduce that
\[
\lim_{t \to \infty} (B_t \circ \phi_t^Y)^{-1} \int_{0}^{\sigma} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, ds = 0. \tag{25}
\]

Second term of the RHS of (24). We can rewrite the second term as
\[
- \int_{0}^{\sigma} \left( \frac{d}{ds} (B_t \circ \phi_t^Y)^{-1} \right) \left( \int_{0}^{s} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, dr \right) \, ds
\]
\[
= \int_{0}^{\sigma} \frac{d}{ds} (B_t \circ \phi_t^Y)^{-1} \left( \int_{0}^{s} (B_t \circ \phi_t^Y) f \circ \phi_t^Y \circ \phi_t^Y \, dr \right) \, ds.
\]

By definition of $B_t$, we have
\[
\frac{d}{ds} B_t \circ \phi_t^Y = \int_{0}^{t} \frac{d}{ds} \left( \frac{1}{\alpha} \circ \phi_t^Y \circ \phi_t^Y \right) \, d\tau = \int_{0}^{t} \left( [\phi_t^Y \circ \alpha(Y)] \frac{1}{\alpha} \circ \phi_t^Y \circ \phi_t^Y \right) \, d\tau,
\]
and by Lemma [4.3] since $Z\alpha = 0$, we get
\[
\frac{d}{ds} B_t \circ \phi_t^Y = \int_{0}^{t} (A_t \circ \phi_t^Y)(V \frac{1}{\alpha} \circ \phi_t^Y \circ \phi_t^Y) \, d\tau + \int_{0}^{t} Y \frac{1}{\alpha} \circ \phi_t^Y \circ \phi_t^Y \, d\tau.
\]

The term
\[
\int_{0}^{t} Y \frac{1}{\alpha} \circ \phi_t^Y \circ \phi_t^Y \, d\tau
\]
clearly grows at most linearly with time. Integration by parts gives
\[
\int_0^t A_\tau \left( \frac{1}{\alpha} \circ \phi_Y^\tau \right) d\tau = A_t \int_0^t \frac{1}{\alpha} \circ \phi_Y^\tau d\tau - \int_0^t \frac{dA_\tau}{d\tau} \left( \int_0^\tau \frac{1}{\alpha} \circ \phi_Y^\tau d\tau \right) d\tau
\]
\[
= A_t \left( \frac{1}{\alpha} \circ \phi_Y^t - \frac{1}{\alpha} \right) d\tau - \int_0^t \left( Y \frac{\alpha}{\alpha^2} \circ \phi_Y^\tau \right) d\tau - \frac{1}{\alpha} \circ \phi_Y^t + \frac{1}{\alpha} \circ \phi_Y^t - \frac{1}{\alpha} \circ \phi_Y^t d\tau.
\]
Hence, it follows from the definition of \( A_t \) that there exists a constant \( C > 0 \) such that for all \( x \in M \),
\[
\left| \frac{d}{ds} B_t \circ \phi_Y^s(x) \right| \leq Ct.
\]
Therefore,
\[
\left| \frac{d}{ds} \frac{B_t \circ \phi_Y^s}{(B_t \circ \phi_Y^s)^2} \left( \int_0^s (B_t \circ \phi_Y^r) f \circ \phi_Y^r \circ \phi_Y^r dr \right) \right| \leq C \left| \frac{1}{B_t \circ \phi_Y^t} \int_0^t (B_t \circ \phi_Y^r) f \circ \phi_Y^r \circ \phi_Y^r dr \right|
\]
which converges to zero by (25). This implies that the second term on the right-hand side of (24) converges to zero as well. We conclude that
\[
\int_0^\sigma f \circ \phi_Y^r \circ \phi_Y^r(x) ds \to 0
\]
pointwise, hence the proof is complete by an application of Lemma 4.4. \( \square \)

5.3. Case (ii) (non-coboundary case). In this subsection we consider the complementary case in which \( \alpha^\perp(x) = \alpha(x) - \mathcal{F} \in H_0([Z]_\Gamma)^\perp \) is not a measurable coboundary w.r.t. \( X \). In this case we will show directly that the flow \( \phi_Y^t \) on \( M \) is mixing, by proving the Proposition 5.3 below. It is for this case that, to prove mixing, we will exploit the shearing of curves produced by the growth of Birkhoff sums for non-coboundaries proved in §4.3.

The notation in the following proposition is the one fixed at the end of §5.1.

Proposition 5.3. If \( \alpha^\perp \) is not a measurable coboundary with respect to \( X \), the flow \( \{ \phi_Y^t \}_{t \in \mathbb{R}} \) given by \( V = \frac{1}{\alpha} X \) on \( M \) is mixing.

The rest of the section is devoted to the proof of Proposition 5.3.

We are assuming that the function \( \alpha^\perp \) is not a measurable \( X \)-coboundary. Recall that \( \alpha^\perp \) is the projection of \( \alpha \) onto \( H_0([Z]_\Gamma)^\perp \); in particular it is a trigonometric polynomial of the same degree as \( \alpha \) without constant term.

For every \( \sigma > 0 \), let \( \gamma^\sigma_{IJ} \) be the path defined as

\[
\gamma^\sigma_{IJ}(s) = (\phi_Y^s \circ \phi_T^s)(x), \quad \text{for all } s \in [0, \sigma].
\]

Let \( D_t \) denote the function on \( M \) defined as

\[
D_t(x) := - \int_0^t \frac{Z \alpha}{\alpha} \circ \phi_Y^\tau(x) d\tau.
\]

By Lemma 4.3 we have

\[
\frac{d}{ds} \gamma^\sigma_{IJ}(s) = [D_t \circ \phi_T^s(x)] V(\gamma^\sigma_{IJ}(s)) + Z(\gamma^\sigma_{IJ}(s)).
\]

It follows that

\[
\int_{\gamma^\sigma_{IJ}} f \bar{V} = \int_0^\sigma (f \circ \phi_Y^s \circ \phi_T^s)(x) (D_t \circ \phi_T^s)(x) ds.
\]
Our assumption on $\alpha^\perp$ implies, by Theorem 4.4, that ergodic integrals of $\alpha^\perp$ along $\phi^X$ grow in measure, i.e. for every $C > 0$
\[ \lim_{t \to \infty} \mu \left( \left\{ x \in M : \left| \int_0^t \alpha^\perp \circ \phi^X_t \, dt \right| < C \right\} \right) = 0. \]

The following Lemma now show that, using Corollary 3.3, we can ensure from this that the function $D_i$ grows in measure as well.

**Lemma 5.4.** For every $C > 1$, we have
\[ \lim_{t \to \infty} \mu \{ |D_i| < C \} = 0. \]

**Proof.** Let $C > 1$ and $\epsilon > 0$; let $c_{m,2}, D_m$ and $d_m$ be given by Corollary 3.3 applied to the function $D_i$. Fix $\delta > 0$ such that $D_m \delta^{d_m} \leq \epsilon / 2$. By Theorem 4.4 there exists $T_0 > 0$ such that for all $t \geq T_0$ we have
\[ \mu \left\{ x \in M : \left| \int_0^t \alpha^\perp \circ \phi^X_t \, dt \right| < \frac{C}{c_{m,2}\delta} \right\} \leq \frac{\epsilon}{2}. \]

Let us define
\[ \tilde{D}_i(x) := -\int_0^t Z(\alpha^\perp) \circ \phi^X_t(x) \, dt. \]

Remark that the endpoint of integration in $\tilde{D}_i(x) = t$, while it is $\tau(x,t)$ in the definition (26) of $D_i$. We have that
\[ \tilde{D}_i(x) = -\int_0^t \left. \frac{d}{ds} \right|_{s=0} \alpha^\perp \circ \phi^X_s(x) \, dt = -\left. \frac{d}{ds} \right|_{s=0} \int_0^t \alpha^\perp \circ \phi^X_s \, ds \circ \phi^X_s(x). \]

By Corollary 3.3 we deduce that for every $t \geq T_0$
\[ \mu \left\{ x \in M : |\tilde{D}_i(x)| \geq C \right\} \geq 1 - D_m \delta^{d_m} - \frac{\epsilon}{2} \geq 1 - \epsilon. \]

Let us recall (see 2.1) that we have $\phi^V_{\tau(x,t)}(x) = \phi^X_t(x)$, where $\tau(x,t) = \int_0^t \alpha \circ \phi^X_s \, ds$. By changing variable (setting $\tau(x,t) = \int_0^t \alpha \circ \phi^X_s \, ds$ so $d\tau = \alpha \circ \phi^X_s \, ds$) and since $\tau$ is $Z$-invariant, we can rewrite $D_i$ in (26) as
\[ D_i(x) = -\int_0^t \frac{Z\alpha}{\alpha} \circ \phi^V_{\tau(x,t)}(x) \, d\tau = \int_0^{\tilde{\tau}(x,t)} \frac{Z\alpha}{\alpha} \circ \phi^X_s(x) (\alpha \circ \phi^X_s(x)) \, ds = \int_0^{\tilde{\tau}(x,t)} Z\alpha \circ \phi^X_s \, ds - \int_0^{\tilde{\tau}(x,t)} Z(\alpha^\perp) \circ \phi^X_s \, ds = \tilde{D}_{\tilde{\tau}(x,t)}(x), \]

where $\tilde{\tau}(x,t)$ is such that $t = \int_0^{\tilde{\tau}(x,t)} \alpha \circ \phi^X_s \, ds$, or, in other words
\[ \tilde{\tau}(x,t) = \int_0^t \frac{1}{\alpha} \circ \phi^V_s(x) \, ds. \]

By unique ergodicity of $V = \frac{1}{\alpha}X$, there exists $T_1 \geq 2T_0$ such that for all $t \geq T_1$, we have $\tilde{\tau}(x,t) \geq \frac{t}{2}T_1 \geq T_0$, so that we conclude that, for all $t \geq T_1$,
\[ \mu \left\{ x \in M : |D_i(x)| < \epsilon \right\} = \mu \left\{ x \in M : |\tilde{D}_{\tilde{\tau}(x,t)}(x)| < C \right\} \leq \epsilon, \]

which proves the Lemma.

By Lemma 4.1 (and Lemma 4.2), in order to prove mixing it is sufficient to study integrals of the form
\[ \int_0^\sigma f \circ \phi^Y \circ \phi^Z(x) \, ds \]
for every $0 < \sigma \leq \sigma_0$. We write

$$
\int_0^\sigma f \circ \phi^V \circ \phi^Z ds = \int_0^\sigma \frac{1 + D_t^1 \circ \phi^Z}{1 + D_t^1 \circ \phi^Z} f \circ \phi^V \circ \phi^Z ds
$$

(28)

and estimate separately the two terms on the right-hand side of the above formula.

*First term in the RHS of (28).* We will prove that for any fixed $\sigma > 0$, the first term on the right-hand side (RHS for short) of (28) converges to zero in measure, so that we can apply Lemma 4.1.

By Lemma 5.4, $1/(1 + D_t^1) \to 0$ in measure. In particular, for every $s \in [0, \sigma]$, since $\mu$ is $\phi^Z$-invariant,

$$
\frac{1}{1 + D_t^1 \circ \phi^Z} \to 0 \quad \text{in measure}.
$$

Since we can bound

$$
\left| \int_0^\sigma \frac{1}{1 + D_t^1 \circ \phi^Z} f \circ \phi^V \circ \phi^Z ds \right| \leq \|f\|_\infty \int_0^\sigma \left| \frac{1}{1 + D_t^1 \circ \phi^Z} \right| ds,
$$

in order to conclude, we apply the following type of dominated convergence result.

**Lemma 5.5.** Let $\{g_t : M \to \mathbb{R}\}_{t \in \mathbb{R}}$ be a family of smooth functions. Assume that the functions $g_t$ are uniformly bounded and that $g_t \circ \phi^Z$ converges to zero in measure for every fixed $s \in \mathbb{R}$. Then, for every $\sigma > 0$, we have

$$
\int_0^\sigma |g_t \circ \phi^Z(x)| ds \to 0 \quad \text{in measure for } t \to \infty.
$$

We give the proof of Lemma 5.5 in §8 (Appendix C) for completeness.

*Second term in the RHS of (28).* Integrating the second term on the right-hand side of (28) by parts, we obtain

$$
\int_0^\sigma \frac{D_t^2 \circ \phi^Z}{1 + D_t^1 \circ \phi^Z} f \circ \phi^V \circ \phi^Z ds = \int_0^\sigma \frac{D_t^1 \circ \phi^Z}{1 + D_t^1 \circ \phi^Z} [(f \circ \phi^V \circ \phi^Z)(D_t \circ \phi^Z)] ds
$$

$$
= \frac{D_t \circ \phi^Z}{1 + D_t^1 \circ \phi^Z} \int_0^\sigma (f \circ \phi^V \circ \phi^Z)(D_t \circ \phi^Z) ds
$$

$$
- \int_0^\sigma \frac{D_t \circ \phi^Z}{1 + D_t^1 \circ \phi^Z} \left( \int_0^\sigma (f \circ \phi^V \circ \phi^Z)(D_t \circ \phi^Z) dr \right) ds.
$$

By (27), we then have the following identity:

$$
\int_0^\sigma \frac{D_t^2 \circ \phi^Z(x)}{1 + D_t^1 \circ \phi^Z(x)} f \circ \phi^V \circ \phi^Z(x) ds = \frac{D_t \circ \phi^Z(x)}{1 + D_t^1 \circ \phi^Z(x)} \left( \int_{\gamma_{t,x}} f \hat{V} \right)
$$

$$
- \int_0^\sigma \frac{d}{ds} \left[ \frac{D_t \circ \phi^Z(x)}{1 + D_t^1 \circ \phi^Z(x)} \right] \left( \int_{\gamma_{t,x}} f \hat{V} \right) ds.
$$

The flow $\phi^V$ is uniquely ergodic, hence we can assume that $f$ is a smooth $V$-coboundary, namely $f = Vg$ for some smooth function $g$. We can bound the terms $\int_{\gamma_{t,x}} f \hat{V}$ by

$$
\left| \int_{\gamma_{t,x}} f \hat{V} \right| \leq \left| \int_{\gamma_{t,x}} dg \right| + \left| \int_{\gamma_{t,x}} Zg \hat{Z} \right| \leq 2 \|g\|_\infty + s \|Zg\|_\infty.
$$
Thus we obtain
\[
\left| \int_0^\sigma \frac{D_t^2 \circ \phi^\Sigma_t(x) f \circ \phi^\Sigma_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \, ds \right| \leq (2 \|g\|_\infty + \sigma \|Zg\|_\infty) \left( \left| \frac{D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| + \int_0^\sigma \left| \frac{D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds \right).
\]

Since by Lemma \ref{lemma:5.4},
\[
\left| \frac{D_t \circ \phi^\Sigma_t}{1 + D_t^2 \circ \phi^\Sigma_t} \right| \to 0 \quad \text{in measure},
\]
it remains to estimate the term
\[
\int_0^\sigma \left| \frac{D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds \leq \left| 1 - \frac{D_t^2 \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \int_0^\sigma \left| \frac{\partial D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds
\]
\[
\leq \int_0^\sigma \left| \frac{\partial D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds.
\]

We cannot directly apply Lemma \ref{lemma:5.5} because the integrand functions might not be uniformly bounded. By Lemma \ref{lemma:4.2} it remains to prove that, for every \(\varepsilon > 0\), there exists \(0 < \delta < 1\) such that, for every \(\eta > 0\), there exists \(T > 0\) such that, for every \(t \geq T\) we have that, for any \(s \in [0, \sigma]\),
\[
\mu \left( x \in M : \left| D_t(x) \right| = 0 \right) \leq \frac{\varepsilon}{3}
\]

and
\[
\mu \left( x \in M : \int_0^\sigma \left| \frac{\partial D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds \geq \frac{\delta}{C_{m,2} \eta} \right) \leq \frac{\varepsilon}{2}.
\]

By Corollary \ref{corollary:3.4} applied to \(D_t\), within a set of measure at least \(1 - D_m \delta^{d_m} - \varepsilon/3\), for any \(s \in [0, \sigma]\) we have
\[
\int_0^s \left| \frac{\partial D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, dr \leq \frac{C_{m,2}}{\delta} \int_0^s \left| \frac{D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, dr \leq \frac{C_{m,2}}{\delta} \int_0^\sigma \left| \frac{D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, ds.
\]

Therefore, we conclude that for every \(s \in [0, \sigma]\),
\[
\mu \left( x \in M : \int_0^s \left| \frac{\partial D_t \circ \phi^\Sigma_t(x)}{1 + D_t^2 \circ \phi^\Sigma_t(x)} \right| \, dr \geq \frac{\delta}{C_{m,2} \eta} \right) \leq \frac{2\varepsilon}{3} + D_m \delta^{d_m} \leq \varepsilon.
\]

This proves the claim and hence completes the proof.

5.4. Final arguments. We can now combine the two cases to conclude the proof of Theorem \ref{theorem:1.1}.

Proof of Theorem \ref{theorem:1.1} Let \(\mathcal{T}_{M,X}\) be the tower of Heisenberg extensions for \(M\) based at \(X\) fixed at the beginning of this Section \ref{section:5} whose existence is guaranteed by Corollary \ref{corollary:3.8}. The dense set of smooth time-changes we want to consider is the class \(\mathcal{P}_T\) of trigonometric polynomial with respect to the tower \(\mathcal{T}_{M,X}\). This class is dense in \(C^\infty(M)\) by Corollary \ref{corollary:3.11}. Let \(\alpha \in \mathcal{P}_T\) and assume that it is not measurably trivial. The theorem will be proved if we show that the corresponding time-change is mixing.

Let \(\alpha^{(0)} := \alpha\) and let \(\alpha^{(i)}\), for \(0 \leq i \leq n\), be its projections along the tower \(\mathcal{T}_{M,X}\) (defined in \(\S5.1\)). We claim that there exists a \(0 \leq i_0 \leq n - 1\) such that \((\alpha^{(i_0)})^{-1}\) is not a coboundary for \(X^{(i_0)}\).
Indeed, assume by contradiction that \((\alpha^{(i)})^\perp\) is a coboundary for \(X^{(i)}\) for every \(0 \leq i \leq n - 1\). Recall that the base of the tower \(M^{(\alpha)}\) is a torus and that the induced vector field \(X^{(\alpha)}\) generates an irrational linear toral flow. By definition of the class of time-changes, the function \(\alpha^{(i)}\) is a trigonometric polynomial on the torus \(M^{(\alpha)} = \mathbb{T}^d\) in the classical sense. It is a standard result that \(\alpha^{(i)}\) is an almost coboundary (i.e., cohomologous to a constant) for the linear flow generated by \(X^{(\alpha)}\). Hence, its pull-back on \(M^{(\alpha^{(i-1)})}\) is an almost coboundary for \(X^{(\alpha^{(i-1)})}\). Since we are assuming that \((\alpha^{(n-1)})^\perp\) is also a measurable coboundary for \(X^{(\alpha^{(n-1)})}\), we deduce that \(\alpha^{(0)} = \alpha\) is an almost coboundary for \(X\), in contradiction with the original assumption.

Let \(0 \leq i_0 < n - 1\) be minimal \(i\) such that \((\alpha^{(i)})^\perp\) is not a coboundary for \(X^{(i)}\). Applying Proposition 5.3 (the not coboundary case) to \((\alpha^{(i)})^\perp\), we have that the flow on \(M^{(\alpha)}\) generated by \(V^{(\alpha)} = \frac{1}{\alpha^{(i)}}X^{(\alpha)}\) is mixing. By definition of \(i_0\), each \((\alpha^{(i)})^\perp\), for \(0 \leq i \leq i_0 - 1\), is a coboundary w.r.t. \(X^{(i)}\). Applying now Proposition 5.2 (the coboundary case) to all \(i = i_0 - 1, \ldots, 0\), we get in \(i_0\) steps that the flow on \(M\) generated by \(\frac{1}{\alpha}V = X\) is mixing.

6. Appendix A: Proof of Corollaries 3.3 and 3.4

**Proof of Corollary 3.3** Let \(f \in [\pi^*(L^2([M]))]^\perp = \bigoplus_{v \in \Lambda_1^+ \setminus \{0\}} H_v(h_\Gamma)\) be a trigonometric polynomial w.r.t. \([Z]_\Gamma\) of degree \(m\). Then, we can write (by the Fourier decomposition property (7))

\[
f \circ \phi_s^Z(x) = \sum_{v \in \Lambda_1^+ \setminus \{0\}} f_v(x) \cdot \exp(2\pi is\langle v, Z \rangle).
\]

We remark that the derivative w.r.t. \(Z\) is a trigonometric polynomial of the same degree, namely we can write

\[
Zf(x) = \frac{d}{ds} \bigg|_{s=0} f \circ \phi_s^Z(x) = \sum_{v \in \Lambda_1^+ \setminus \{0\}} 2\pi i\langle v, Z \rangle f_v(x).
\]

Since \(|f| \leq \sum_v |f_v|\), from the assumption we deduce that

\[
\mu \left( x \in M : \sum_{v \in \Lambda_1^+ \setminus \{0\}} |f_v(x)| \leq C \right) \leq \varepsilon.
\]

Moreover, since \([Z]_\Gamma\) is the smallest \(\Gamma\)-rational subpace which contains \(Z\), there exists a constant \(c_{m,Z} > 0\) such that \(|\langle v, Z \rangle| \geq c_{m,Z}/(2\pi)\) for all \(v \in \Lambda_1^+ \setminus \{0\}\). In particular, we deduce

\[
\mu \left( x \in M : \sum_{v \in \Lambda_1^+ \setminus \{0\}} 2\pi |\langle v, Z \rangle||f_v(x)| \leq c_{m,Z}C \right) \leq \varepsilon.
\]

By Lemma 3.2 there exist constants \(D_m, d_m > 0\) such that for any \(\delta > 0\),

\[
\mu \left( x \in M : \left| \frac{d}{ds} \bigg|_{s=0} f \circ \phi_s^Z(x) \right| \leq c_{m,Z} \delta C \right) \leq D_m \delta d_m + \varepsilon,
\]

which concludes the proof.

**Proof of Corollary 3.4** Let \(f\) be a trigonometric polynomial w.r.t. \(h_\Gamma\) of degree \(m\). As in (29), we have

\[
|Zf(x)| = \left| \frac{d}{ds} \bigg|_{s=0} f \circ \phi_s^Z(x) \right| \leq 2\pi m \|Z\|_{L^2} \sum_{v \in \Lambda_1^+ \setminus \{0\}} |f_v(x)| = \frac{C_{m,Z}}{2} \sum_{v \in \Lambda_1^+ \setminus \{0\}} |f_v(x)|.
\]

Let

\[
E_0 := \left\{ x \in M : \sum_{v \in \Lambda_1^+ \setminus \{0\}} |f_v(x)| = 0 \right\}.
\]
From the assumption, it follows that \( \mu(E_0) \leq \varepsilon_0 \). By Lemma \[L.2\] there exist constants \( D_m, d_m > 0 \) such that for any \( \delta > 0 \), the set

\[ M_\delta^c := \left\{ x \in M : |f(x)| > \delta \sum_{v \in \Lambda_t \setminus \{0\}} |f_v(x)| \right\} \subset M \setminus E_0 \]

has measure at least \( 1 - D_m \delta^{d_m} - \varepsilon_0 \). In particular, for any \( x \in M_\delta^c \), we have

\[ (31) \quad \left| \frac{d}{ds} f \circ \phi_s^Z(x) \right| \leq \frac{C_m Z}{2} \sum_{v \in \Lambda_t \setminus \{0\}} |f_v(x)| \leq \frac{C_m Z}{2\delta} |f(x)|. \]

As in the proof of Lemma \[L.2\] let us define on \( M \setminus E_0 \) the normalized trigonometric polynomial \( \hat{f}(x) = f(x) / \sum_{v \in \Lambda_t \setminus \{0\}} |f_v(x)| \). We remark that, since \( E_0 \) is a union of full fibers, it is \( \phi_s^Z \)-invariant.

By \( (30) \), we have \( |Z \hat{f}| \leq C_m Z / 2 \). Thus, for any fixed \( \sigma > 0 \), for any \( x \in M \setminus E_0 \) and for any \( s \in [-\sigma, \sigma] \), we have

\[ \left| \hat{f} \circ \phi_s^Z(x) \right| \leq \left| \hat{f}(x) \right| + \int_0^s \left| \frac{d}{dr} \hat{f} \circ \phi_r^Z(x) \right| dr \leq \left| \hat{f}(x) \right| + \frac{C_m Z}{2} s. \]

It follows that

\[ \left| \hat{f} \circ \phi_s^Z(x) \right| - \left| \hat{f}(x) \right| \leq \frac{C_m Z}{2} \sigma; \]

in particular, for any \( \delta > 0 \), if we let \( \sigma = \delta / C_m Z \), for any \( s \in [0, \sigma] \), we have that, if \( x \in M_\delta^c \), then

\[ \left| \hat{f} \circ \phi_s^Z(x) \right| \geq \left| \hat{f}(x) \right| - \frac{\delta}{2} \geq \frac{\delta}{2}. \]

This shows that, if \( x \in M_\delta^c \), then \( \phi_s^Z(x) \in M_\delta^c \) for all \( s \in [0, \delta / C_m Z] \). Therefore, by \( (31) \), for all \( \delta > 0 \), for all \( x \in M_\delta^c \)

\[ \left| \frac{d}{ds} f \circ \phi_s^Z(x) \right| \leq \frac{C_m Z}{\delta} |f \circ \phi_s^Z(x)| \quad \text{for all } s \in \left[ 0, \frac{\delta}{C_m Z} \right], \]

which concludes the proof. \[ \square \]

7. APPENDIX B: PROOF OF LEMMA 4.5

**Proof of Lemma 4.5** We notice that we can rewrite

\[ \frac{1}{T} \int_0^T \mu \{ |F_t| < C \} \, dt = \frac{1}{T} \int_0^T \left( \int_M \| (-C,C) \circ F_t(x) \| \, d\mu \right) \, dt \]

\[ = \int_M \left( \frac{1}{T} \int_0^T \| (-C,C) \circ F_t(x) \| \, dt \right) \, d\mu, \]

therefore, by the Dominated Convergence Theorem, it is sufficient to show that the function \( \frac{1}{T} \int_0^T \| (-C,C) \circ F_t(x) \| \, dt \) converges pointwise to zero.

For all \( x \in M \), denote by \( \nu_{T,x} \) the probability measure on \( M \times \mathbb{R} \) supported on the parametrized curve \( t \mapsto (\phi_t^X(x), F_t(x)) \), for \( t \in [0, T] \), defined by

\[ \nu_{T,x}(A \times [a, b]) = \frac{1}{T} \text{Leb} \left( t \in [0, T] : \phi_t^X(x) \in A \text{ and } F_t(x) \in [a, b] \right), \]

for \( A \subset M \) and \( [a, b] \subset \mathbb{R} \). We will prove that, for all \( x \in M \), \( \nu_{T,x} \) converges weakly to 0.

Suppose on the contrary that there exist \( \mathbf{x} \in M \) and a strictly increasing sequence \( T_n \to \infty \) such that \( \nu_{T_n,x} \) converges weakly to a measure \( \nu \) with non-zero total mass. We claim that \( \nu \) is \( \Psi_t \)-invariant, where \( \Psi_t(x, s) = (\phi_t^X(x), s + F_t(x)) \). Indeed, for every continuous function \( g \) on \( M \times \mathbb{R} \)
we have
\[
\int_{M \times \mathbb{R}} g d(\Psi_t) \circ \nu_{\tau, t} = \int_{M \times \mathbb{R}} g \circ \Psi_t d\nu_{\tau, t} = \int_{M \times \mathbb{R}} g(\phi_t^X(\tau), s + F_t(\tau)) d\nu_{\tau, t}
\]
\[
= \frac{1}{T_n} \int_{0}^{T_n} g(\phi_t^X, s + F_t(\tau)) d\tau - \frac{1}{T_n} \int_{0}^{T_n} g(\phi_t^X(\tau), s + F_t(\tau)) d\tau
\]
By the cocycle relation \([15]\) and by definition of \(\mu_{\tau, t}\), we obtain
\[
\left| \int_{M \times \mathbb{R}} g d(\Psi_t) \circ \nu_{\tau, t} - \int_{M \times \mathbb{R}} g d\nu_{\tau, t} \right|
\]
\[
= \frac{1}{T_n} \int_{0}^{T_n} g(\phi_t^X(\tau), s + F_t(\tau)) d\tau - \frac{1}{T_n} \int_{0}^{T_n} g(\phi_t^X(\tau), s + F_t(\tau)) d\tau \leq \frac{2\|g\|_{\infty} t}{T_n}.
\]
Therefore, \(\lim_{n \to \infty} (\Psi_t)_{\nu_{\tau, t}} = \nu\), and the claim follows from the continuity of \((\Psi_t)_{\nu}\).

Let \(\hat{v}\) be an ergodic component of \(\nu\). By unique ergodicity of \(\phi^{\delta}_t\), we have that \(\pi \circ \hat{v} = \mu\), where \(\pi: M \times \mathbb{R} \to M\) is the projection onto the nilmanifold. In particular, for almost every \(x \in M\) there exists a point \((x, s) \in M \times \mathbb{R}\) which is generic for \(\hat{v}\). Assume that there exists a fiber \(\{x\} \times \mathbb{R}\) over \(M\) with more than one generic point, that is, assume that the points \((x, s)\) and \((x, s + r)\) are both generic for \(\hat{v}\). This implies that \(\hat{v}\) is \(T_r\)-invariant, where \(T_r\) denotes the vertical translation on the fibers by \(r\): for any continuous, compactly supported function \(g \in C_c(M \times \mathbb{R})\), we have
\[
\frac{1}{T} \int_{0}^{T} g \circ \Psi_t(x, s + r) dr \to \int_{M \times \mathbb{R}} g d\hat{v},
\]
but also
\[
\frac{1}{T} \int_{0}^{T} g \circ \Psi_t \circ T_r(x, s) dr = \frac{1}{T} \int_{0}^{T} g \circ T_r \circ \Psi_t(x, s) dr \to \int_{M \times \mathbb{R}} g \circ T_r d\hat{v} = \int_{M \times \mathbb{R}} g d(\Psi_t)_{\hat{v}}.
\]
Since \(g\) was arbitrary, we deduce \(\hat{v} = (T_r)_{\hat{v}}\). As \(\hat{v}\) is a finite measure, we must have \(r = 0\), namely for almost every \(x \in M\) there exists only one point \((x, u(x)) \in M \times \mathbb{R}\) which is generic for \(\hat{v}\). The function \(u: x \mapsto u(x)\) implicitly defined above is measurable, since its graph is a measurable set. Uniqueness implies that
\[
\Psi_t(x, u(x)) = (\phi_t^X(x), u(x) + F_t(x)) = (\phi_t^X(x), u(\phi_t^X(x))),
\]
from which we deduce \(u(\phi_t^X(x)) - u(x) = F_t(x)\), in contradiction with the assumption that \(f\) is not a measurable coboundary.

\section*{Appendix C: Proof of Lemma 5.5}

\textbf{Proof of Lemma 5.5.}\ Let us fix \(\sigma > 0\). Denote by \(\nu\) the product measure \(d\nu = d\mu \, dr\) on \(\hat{M} := M \times [0, \sigma]\). We first claim that \(g_t \circ \phi^Z_s\) converges to zero in measure on \(\hat{M}\). Define
\[
E_t^\delta := \{(x, s) \in \hat{M}: |g_t \circ \phi^Z_s(x)| > \delta\}.
\]
We have
\[
\nu(E_t^\delta) = \int_{\hat{M}} \mathbb{1}_{E_t^\delta}(x, s) d\nu = \int_{0}^{\sigma} \left( \int_{M} \mathbb{1}_{E_t^\delta}(x, s) d\mu(x) \right) ds.
\]
By assumption, the term in brackets converges to zero for all \(s \in [0, \sigma]\), hence, by Lebesgue Theorem, \(\nu(E_t^\delta) \to 0\).

Let \(E_t^\delta(x) := E_t^\delta \cap \{x\} \times [0, \sigma] = \{s \in [0, \sigma]: |g_t \circ \phi^Z_s(x)| > \delta\}\) and denote by \(|E_t^\delta(x)|\) its measure. Define also
\[
\text{Bad}^\delta_t := \{x \in M: |E_t^\delta(x)| > \delta^2\}.
\]
We claim that for all $\delta_1, \delta_2 > 0$, $\mu(Bad_{\delta_1}^{\delta_2}) \to 0$. If this was not the case, there would exist $\delta_1, \delta_2, \eta > 0$ and an increasing sequence $t_n \to \infty$ such that $\mu(Bad_{\delta_1}^{\delta_2}) \geq \eta$ for all $n \in \mathbb{N}$. This would imply that for all $n \in \mathbb{N}$, by Fubini Theorem,

$$\nu(E_{t_n}^{\delta_2}) = \int_M |E_{t_n}^{\delta_1}(x)| \, d\mu(x) \geq \int_{Bad_{\delta_1}^{\delta_2}} |E_{t_n}^{\delta_2}(x)| \, d\mu(x) \geq \delta_2 \cdot \eta,$$

in contradiction with $\nu(E_{t_n}^{\delta_2}) \to 0$.

Let $C > 0$ be such that $\|g_t\|_{\infty} \leq C$. Fix $\varepsilon, \eta > 0$ and choose $\delta_1, \delta_2 > 0$ such that $C\delta_1 + \sigma \delta_2 < \eta$. Let $T > 0$ be such that $\mu(Bad_{\delta_1}^{\delta_2}) < \varepsilon$ for all $t \geq T$. For any $t \geq T$ and for any $x \notin Bad_{\delta_1}^{\delta_2}$, we have $|E_t^{\delta_2}(x)| \leq \delta_2$; hence

$$\int_0^\sigma |g_t \circ \phi^s(x)| \, ds \leq \int_0^\sigma \mathbb{I}_{E_t^{\delta_2}(x)}(s) |g_t \circ \phi^s(x)| \, ds + \int_0^\sigma \mathbb{I}_{(E_t^{\delta_2}(x))^c}(s) |g_t \circ \phi^s(x)| \, ds \leq C\delta_2 + \sigma \delta_1 < \eta.$$

Therefore,

$$\mu \left( x \in M : \int_0^\sigma |g_t \circ \phi^s(x)| \, ds > \eta \right) \leq \mu \left( Bad_{\delta_1}^{\delta_2} \right) < \varepsilon,$$

which concludes the proof.

**Acknowledgement.** The authors would like to thank Livio Flaminio for useful discussions. For their hospitality at the very initial stages of this project, C.U. thanks the Department of Mathematics of the University of Maryland and both A.A. and C.U. acknowledge the Hausdorff Institute for Mathematics in Bonn for hosting them during the program *Geometry and Dynamics of Teichmüller Space*. G.F. thanks the Department of Mathematics of the University of Bristol, UK, for its hospitality during the final stages of preparation of the paper. G.F. acknowledges support of NSF grant DMS 1600687. C.U. and D.R. acknowledge the ERC Grant *ChaParDyn* and the Leverhulme Trust for the funding which made this research possible. C.U. is also supported by the Royal Society through a Wolfson Research Merit Award. The research leading to these results has received funding from the European Research Council under the European Union Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement n. 335989.

**REFERENCES**

[AF] A. Avila & G. Forni, Weak mixing for interval exchange transformations and translation flows. *Ann. of Math.* **165** (2) (2007), 637–664. MR2299743, Zbl 1136.37003.

[AGH] L. Auslander, L. Green & F. Hahn, *Flows on homogeneous spaces*, Princeton University Press, Princeton, N.J., 1963.

[AFU] A. Avila, G. Forni & C. Ulcigrai, Mixing for the time-changes of Heisenberg nilflows, *J. of Diff. Geom.* **89** (2011), 369–410.

[BK] P. Berk & A. Kanigowski, Spectral disjointness of rescalings of some surface flows, preprint, [arXiv:1901.04724v2].

[Br] A. Brudnyi, On local behavior of analytic functions, *J. of Funct. Anal.* **169** (1999), 481–493.

[BDG] J. Bourgain, C. Demeter & L. Guth. Proof of the main conjecture in Vinogradov’s mean value theorem for degrees higher than three, *Ann. of Math.*, **184** (2) (2016), 633–682.

[Bu] A. Bufetov, Limit theorems for translation flows, *Ann. of Math.* **179** (2) (2014), 431–499.

[BuFo] A. Bufetov & G. Forni, Limit Theorems for Horocycle flows, *Ann. Sci. ENS* **47** (5) (2014), 851–903.

[Bur] M. Burger, Horocycle flow on geometrically finite surfaces. *Duke Math. J.* **61** (1990), 779–803. MR1084459, Zbl 0723.58041.

[CM] F. Cellarosi & J. Marklof, Quadratic Weyl Sums, Automorphic Functions, and Invariance Principles, *Proc. London Math. Soc.* **113** (2016), 775–828.

[CW] J. Chaika & A. Wright, A smooth mixing flow on a surface with nondegenerate fixed points, *J. Amer. Math. Soc.* **32** (2019), 81–117.

[CG] L. Corwin & F. P. Greenleaf, Representations of nilpotent Lie groups and their applications. Part 1: Basic theory and examples, Cambridge studies in advanced mathematics **18**, Cambridge University Press, Cambridge, UK, 1990.
[Fa1] B. Fayad, Polynomial decay of correlations for a class of smooth flows on the torus, *Bull. Soc. Math. France* 129 (2001), 487–503.

[Fa2] B. Fayad, Analytic mixing reparametrizations of irrational flows, * Erg. Th. Dynam. Sys.* 22 (2) (2002), 437–468.

[FKK] B. Fayad, G. Forni & A. Kanigowski, Lebesgue spectrum of countable multiplicity for conservative flows on the torus, preprint. [arXiv:1609.03757]

[FK] B. Fayad & A. Kanigowski, Multiple mixing for a class of conservative surface flows, *Inv. Math.* 203 (2) (2016), 555–614.

[FeKi] A. Fedotov & F. Klopp, An exact renormalization formula for Gaussian exponential sums and applications, *Amer. J. of Math.* 134 (2012), 711–748.

[EJK] H. Fiedler, W. B. Jurkat & O. Körner, Asymptotic expansions of finite theta series, *Acta Arithmetica* XXXII (1977) 129–146. MR0563894; Zbl 0308.10021.

[FF1] L. Flaminio & G. Forni, Invariant distributions and time averages for horocycle flows, *Duke Math. J.* 119 (3) (2003), 465–526.

[FF2] , Equidistribution of nilflows and applications to theta sums, *Erg. Th. Dynam. Sys.* 26 (2) (2006), 409–433.

[FF3] , On the cohomological equation for nilflows, *J. Mod. Dynam.* 1 (2007), 37–60.

[FF4] , Orthogonal powers and Möbius conjecture for smooth time changes of horocycle flows, preprint, [arXiv:1811.04652]

[FFRH] L. Flaminio, G. Forni & F. Rodriguez Hertz, Invariant distributions for homogeneous flows and affine transformations, *J. Mod. Dynam.* 10 (2016), 33 –79.

[ Fo1] G. Forni, Solutions of the cohomological equation for area-preserving flows on compact surfaces of higher genus. *Ann. of Math.* 146 (2) (1997), 295–344. MR1477760; Zbl 0893.58037.

[ Fo2] , Deviation of ergodic averages for area-preserving flows on surfaces of higher genus, * Ann. of Math.* (2) 155 (1) (2002), 1–103. MR1888794; Zbl 1034.37003.

[ Fo3] , On the Greenfield-Wallach and Katok conjectures, in *Geometric and Probabilistic Structures in Dynamics* (K. Burns, D. Dolgopyat and Ya. Pesin Editors). Contemporary Mathematics (Proceedings), pp. 197-215. American Mathematical Society, Providence RI, 2008.

[ FoS1] , Limit Theorems for Horocycle Flows, in *Hyperbolic Dynamics, Fluctuations and Large Deviations* (edited by D. Dolgopyat, Y. Pesin, M. Pollicott and L. Stoyanov), *Proceedings of Symposia in Pure Mathematics*, Vol. 89 (2015), 121–160.

[ FoS2] , Effective Equidistribution of nilflows and bounds on Weyl sums, in “Dynamics and Analytic Number Theory”, Proceedings of the Durham Easter School 2014. Editors: D. Badziahin, A. Gorodnik, N. Peyerimhoff, T. Ward. London Mathematical Society Lecture Note Series 437. Cambridge University Press, January 2017.

[ FK1] G. Forni & A. Kanigowski, Time-Changes of Heisenberg nilflows, preprint, [arXiv:1711.05543]

[ FK2] , Multiple mixing and disjointness for time changes of bounded-type Heisenberg nilflows, preprint, [arXiv:1810.13319]

[ FoMa] G. Forni & C. Matheus, Introduction to Teichmüller theory and its applications to dynamics of interval exchange transformations, flows on surfaces and billiards, in “Lectures from the Bedlewo Summer School 2011” edited by F. Rodriguez Hertz, *J. Mod. Dynam.* 8 (3-4) (2014), 271–436.

[ FU] G. Forni & C. Ulcigrai, Time-changes of horocycle flows, *J. Mod. Dynam.* 6 (2) (2012), 251–273.

[ Fu] H. Furstenberg. The unique ergodicity of the horocycle flow. Recent advances in topological dynamics. *Lecture Notes in Math.* 318 (1973), 95–115. Springer, Berlin, MR0393339; Zbl 0256.58009.

[ Gr] L. W. Green, Spectra of nilflows, * Bull. Amer. Math. Soc.* 67 (1961), 414–415.

[ GM] J. Griffin & J. Marklof, Limit theorems for skew translations, *J. Mod. Dynam.* 8 (2014), 177–189.

[ KKU] A. Kanigowski, J. Kulaga-Przymus & C. Ulcigrai Multiple mixing and parabolic divergence in smooth area-preserving flows on higher genus surfaces, preprint, [arXiv:1606.09189].

[ KL] A. Kanigowski, M. Lemaniczky & C. Ulcigrai, On disjointness properties of some parabolic flows, preprint, [arXiv:1810.11576v1]

[Ka1] A. B. Katok. Interval exchange transformations and some special flows are not mixing. *Israel J. of Math.* 35 (4) (1980), 301–310. MR0594335; Zbl 0437.28009.

[Ka2] , Cocycles, cohomology and combinatorial constructions in ergodic theory (in collaboration with E. A. Robinson, Jr.), in *Smooth Ergodic Theory and its applications*, Proc. Symp. Pure Math., 69 (2001), 107–173.

[Ka3] , Combinatorial Constructions in Ergodic Theory and Dynamics. University Lecture Series Vol. 30. American Mathematical Society, Providence, RI, 2003. MR2008435; Zbl 1030.37001.

[KT] A. B. Katok & J.-P. Thouvenot, Spectral Properties and Combinatorial Constructions in Ergodic Theory, in *Handbook of Dynamical Systems*, Vol 1B. (A. Katok B. Hasselblatt editors), Elsevier, 2005.

[Kea] M. Keane, Interval Exchange Transformations, *Math. Zeitschrift* 141 (1) (1975), 25–31.

[KS] K. M. Khanin & Ya. G. Sinai, Mixing for some classes of special flows over rotations of the circle, *Funktsional’nyiAnaliz i Ego Prilozheniya* 26(3) (1992), 1–21. (Translated in: *Funct. Anal. and its Appl.* 26(3) (1992), 155–169).
[Ko1] A. V. Kočergin, The absence of mixing in special flows over a rotation of the circle and in flows on a two-dimensional torus, *Dokl. Akad. Nauk SSSR* 205 (1972), 512–518. (Translated in *Soviet Math. Dokl.* 13 (1972), 949–952).

[Ko2] __________, Mixing in special flows over a shifting of segments and in smooth flows on surfaces, *Mat. Sb.* 96 (138) (1975), 471–502.

[Ko3] __________, Nondegenerate saddles and the absence of mixing in flows on surfaces, *Proc. Steklov Inst. Math.* 256, no. 1, (2007), 238. MR236903, Zbl 1153.37303.

[Kol] A. N. Kolmogorov, On dynamical systems with an integral invariant on the torus, *Dokl. Akad. Nauk SSSR (N. S.*) 93 (1953), 763–766 (in Russian). MR0062892, Zbl 0052.31004.

[Ku] A. G. Kushnirenko, Spectral properties of some dynamic systems with polynomial divergence of orbits. *Moscow Univ. Math. Bull.* 29 (1974), 101–108. MR0353369, Zbl 0276.58007.

[Mar1] B. Marcus, Ergodic properties of horocycle flows for surfaces of negative curvature, *Ann. of Math.* (2) 105 (1977), 81–105.

[Mar2] __________, The horocycle flow is mixing of all degrees, *Invent. Math.* (3) 46 (1978), 201–209.

[Mr] J. Marklof, Limit theorems for theta sums. *Duke Math. J.* 97 (1): 127–153, 1999, MR1682276, Zbl 0965.11036.

[Mas1] H. Masur, Interval exchange transformations and measured foliations, *Ann. of Math.* 115 (1982), 169–200. MR0644018, Zbl 0497.28012.

[Mas2] __________, Ergodic theory of translation surfaces, *Handbook of Dynamical Systems*, Vol. 1B, pages 527–547. Elsevier B. V., Amsterdam., 2006, MR2186247, Zbl 1130.37313.

[Ra] M. S. Rath, Discrete subgroups of Lie groups, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68, Springer-Verlag, New York, 1972.

[Rat] M. Ratner, Rigidity of time changes for horocycle flows. *Acta Math.* 156 (1986), 229–309.

[Rav1] D. Ravotti, Quantitative mixing for locally Hamiltonian flows with saddle loops on compact surfaces, *Disc. Cont. Dyn. Syst.* 18 (12) (2017), 3815–3861.

[Rav2] __________, Mixing for suspension flows over skew-translations and time-changes of quasi-abelian filiform nilflows, *Erg. Th. Dynam. Sys.* published online at https://doi.org/10.1017/etds.2018.19.

[Rav3] __________, Parabolic perturbations on unipotent flows on compact quotients of SL(R). *Comm. Math. Phys.*, published online at https://link.springer.com/article/10.1007/s00220-019-03348-0.

[Pa] O. S. Parasyuk, Flows of horocycles on surfaces of constant negative curvature. Uspehi Matem. Nauk (N.S.) 8 (3) (1953), 125–126. MR0058883, Zbl 0052.34201.

[P] L. B. Pierce, The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth], Séminaire Bourbaki 69 ème année, 2016-2017, no 1134, *Astérisque* 407 (2019), 479–564.

[Sa] P. Sarnak, Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series, *Comm. Pure Appl. Math.* 34 (1981), 719–739.

[Si] L. D. Simonelli, Absolutely continuous spectrum for parabolic flows/maps, *Disc. Cont. Dyn. Syst.* (1) 38 (2018), 263–292.

[Sc] D. Scheglov, Absence of mixing for smooth flows on genus two surfaces, *J. Mod. Dynam.* 3 (1) (2009), 13–34. MR2481330, Zbl 1183.37080.

[Sti] A. N. Starkov. Dynamical Systems on Homogeneous Spaces. *Translations of the American Mathematical Society*, 190, Providence, Rhode Island 2002, MR1746847, Zbl 1143.37300.

[Str] A. Strömbergsson, On the deviation of ergodic averages for horocycle flows, *J. Mod. Dynam.* 7 (2) (2013), 291–328.

[Tie] R. Tiede de Aldecoa, Spectral analysis of time-changes of horocycle flows, *J. Mod. Dynam.* (2) 6 (2012), 275–285.

[U11] C. Ulcigrai, Mixing of asymmetric logarithmic suspension flows over interval exchange transformations, *Ergod. Th. Dynam. Sys.* 27 (3) (2007), 991–1035.

[U12] __________, Weak mixing for logarithmic flows over interval exchange transformations, *J. Mod. Dynam.* 3 (1) (2009), 35–49. MR2481331, Zbl 1183.37081.

[U13] __________, Absence of mixing in area-preserving flows on surfaces, *Ann. of Math.* 173 (3) (2011), 1743–1778.

[Ve] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* 115 (1982), 201–242. MR064019, Zbl 0486.28014.

[Wo1] T. D. Wooley, Vinogradov’s mean value theorem via efficient congruencing, *Ann. of Math.* 175 (2012),1575–1627.

[Wo2] __________, Vinogradov’s mean value theorem via efficient congruencing, II, *Duke Math. J.* 162 (2013), 673–730.

[Yo] J.-C. Yoccoz, Interval exchange maps and translation surfaces. In *Homogeneous Flows, Moduli Spaces and Arithmetic* (Einsiedler, M., et al. editors). Proceedings of the Clay Mathematics Institute Summer School, Centro di Ricerca Matematica Ennio De Giorgi, Pisa, Italy, June 11–July 6, 2007. *Clay Mathematics Proceedings*, vol. 10, pp. 1–69. AMS/Clay Mathematics Institute, Providence/Cambridge (2010)
IMPA, Estrada Dona Castorina 110, Rio de Janeiro, Brazil,
E-mail address: artur.avila@math.uzh.ch,

Department of Mathematics, Mathematics Building, University of Maryland, College Park, MD 20742-4015, USA,
E-mail address: gforni@math.umd.edu,

Monash University, School of Mathematics, Clayton Campus, 3800 Victoria, Australia
E-mail address: davide.ravotti@gmail.com,

School of Mathematics, University of Bristol, BS8 1TW Bristol, United Kingdom, and

Universität Zürich, Institut für Mathematik, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
E-mail address: corinna.ulcigrai@bristol.ac.uk,