J. DIFFERENTIAL GEOMETRY
89 (2011) 369-410

MIXING FOR TIME-CHANGES OF HEISENBERG NILFLOWS

Artur Avila, Giovanni Forni & Corinna Ulcigrai

Abstract

We consider reparametrizations of Heisenberg nilflows. We show that if a Heisenberg nilflow is uniquely ergodic, all non-trivial time-changes within a dense subspace of smooth time-changes are mixing. Equivalently, in the language of special flows, we consider special flows over linear skew-shift extensions of irrational rotations of the circle. Without assuming any Diophantine condition on the frequency, we define a dense class of smooth roof functions for which the corresponding special flows are mixing whenever the roof function is not a coboundary. Mixing is produced by a mechanism known as stretching of ergodic sums. The complement of the set of mixing time-changes (or, equivalently, of mixing roof functions) has countable codimension and can be explicitly described in terms of the invariant distributions for the nilflow (or, equivalently, for the skew-shift), producing concrete examples of mixing time-changes.

1. Introduction

In this paper we give a contribution to the smooth ergodic theory of parabolic flows. We prove that for any uniquely ergodic Heisenberg nilflow, all non-trivial time-changes, within a dense subspace of time-changes, are mixing. The set of trivial time-changes has countable codimension and can be explicitly described in terms of invariant distributions for the nilflow.

A non-singular flow is called *parabolic* if nearby orbits diverge polynomially in time. If nearby orbits diverge exponentially, the flow is called *hyperbolic*; if there is no divergence (or perhaps it is slower than polynomial), the flow is called *elliptic*. In contrast with the hyperbolic case, and to a lesser extent with the elliptic case, there is no general theory that describes the dynamics of parabolic flows. The main (typical) ergodic properties often associated with parabolic dynamics are unique ergodicity, mixing, polynomial speed of convergence of ergodic averages, polynomial decay of correlations for smooth functions and, of course,

Received 5/03/2010.

zero entropy. Another important feature of parabolic flows is the presence of infinitely many independent distributional obstructions to the solution of the so-called *cohomological equation* (which are not signed measures as in the hyperbolic case). This important property allows for the existence of non-trivial time-changes which are not given by the existence of fast periodic approximations (Liouvillean phenomenon) as in the classical, better understood, elliptic case.

A fundamental example of a parabolic flow is given by *horocycle flows* on compact negatively curved surfaces. It is well known that horocycle flows are uniquely ergodic [**15**], mixing of all orders [**28**], and have countable Lebesgue spectrum [**34**]. Kuschnirenko [**27**] has proved that all time-changes are mixing under an explicit condition which holds if the time-change is sufficiently small (in the C^1 topology). It follows by a result by Marcus [**29**] that this result extends to all smooth time-changes. Nothing is known about the spectral properties of time-changes. A. Katok has conjectured that countable Lebesgue spectrum persists at least under Kuschnirenko's condition.

Other important examples of flows which are sometimes considered parabolic are given by *area-preserving flows on surfaces* of higher genus (genus greater than two) with saddle-like singularities. In this case, the orbit divergence is entirely produced by the splitting of trajectories near the singularities. In particular, directional flows on translation surfaces, often called *translation flows*, which appear in the study of the geodesic flow on a surface endowed with a flat metric with conical singularities (we refer for example to the survey [**33**] for definitions), have been studied in depth in the past thirty years. The unique ergodicity of almost any translation flow is a fundamental result of H. Masur [**32**] and W. Veech [**44**], while the first two authors proved that typical translation flows are weak mixing [**2**]. Translation flows are never mixing, as known since the work of Katok [**19**]. This leads to the question of mixing in reparametrizations of translation flows.

Time-changes of translation flows can be represented as special flows over interval exchange transformations (IET's), which are one-dimensional piecewise isometries. A reparametrization of translation flows which appears naturally in physical problems is the locally Hamiltonian parametrization, which was studied since Novikov and his school in the nineties. The corresponding flows on surfaces are known as flows given by a *multi-valued Hamiltonian* and can be represented as special flows over IET's with a roof function which has singularities. If the zeros of the flow are degenerate, i.e. they are multi-saddles, they give rise to power-like singularities of the roof function; if they are non-degenerate (Morse) saddles, they give rise to logarithmic singularities. If the flow has saddle loops, logarithmic singularities are typically asymmetric; otherwise they are symmetric.

The mixing properties of special flows have been studied in depth by many authors. The situation can perhaps be summarized as follows. On one hand, weak mixing is typical and it does not require any assumptions on the singularities of the roof functions: in the work by the first two authors [**2**] already mentioned above, it is proven that for any piecewise constant roof function, weak mixing holds for typical IET's. The third author proved in [**43**] that a simple mechanism shows weak mixing in the case of roof functions with logarithmic singularities over typical IET's.

On the other hand, mixing relies crucially on the presence of singularities. Indeed, for roof functions of bounded variations (thus in particular for smooth roofs), A. Katok [**19**] proved absence of mixing. Kočergin proved in [**23**] that a flow given by a roof function with *power-like singularities* over a typical IET (with minimal combinatorics) is mixing, and mixing is produced as an effect of the shear at the singularities. When the singularities are *logarithmic*, mixing depends on whether the singularities satisfy a certain symmetry condition. Even though each side of a saddle produces shearing, the symmetry condition in fact leads to the mutual cancellation of the mixing effect. In the *asymmetric* case, typical mixing was proved by Khanin and Sinai [**21**] for flows over circle rotations and by the third author for flows over IET's on any number of intervals [**42**]. In the *symmetric* case, Kočergin proved the absence of mixing for flows over circle rotations [**22**, **24**]. This result was extended to typical IET's first by Scheglov [**37**], who treated the case of IET's of four and five intervals, and finally to typical IET's on any number of intervals by the third author [**41**].

Another important class of (homogeneous) parabolic flows is given by *nilflows*. By classical results of homogeneous dynamics (see [**1**]), minimal nilflows are uniquely ergodic. However, in contrast with horocycle flows, they are never mixing, not even weak mixing. However, there is a clear geometric obstruction to the (weak) mixing property, that is, every nilflow is only partially parabolic, in the sense that it has an elliptic factor given by a linear flow on a torus. For observables in the orthogonal complement of the span of the pull-back to the nilmanifold of the toral characters, any nilflow has countable Lebesgue spectrum [**16**, **1**]; hence it is mixing. Thus, nilflows have the properties of *relative* Lebesgue spectrum and mixing.

Our results confirm some heuristic principles on the dynamics of parabolic flows. In particular, for time-changes of any uniquely ergodic Heisenberg nilflow (without Diophantine conditions), mixing is prevalent, and it occurs unless the flow is only partially parabolic (because of the presence of a measurable elliptic factor), which in this case means that the time-change is trivial. As a consequence, weak and strong mixing are equivalent. This picture is in agreement with the result of Marcus

[**29**] on mixing for smooth time-changes of the horocycle flow. It shows that Heisenberg nilflows, as horocycle flows, differ significantly from translation flows or area-preserving flows on higher genus surfaces. As outlined above, in the latter case, the typical (non-trivial) time-change is weak mixing, but not mixing, and mixing can only be produced by shear at the singularities. In other words, area-preserving flows on surfaces are better classified as elliptic flows with singularities than as parabolic flows.

Our approach to mixing for nilflows has the advantage of not requiring Diophantine conditions. However, it does not seem to be possible to derive quantitative information on the decay of correlations. A natural conjecture is that if the elliptic toral factor is a Diophantine linear flow, then the decay of correlations of smooth functions is *polynomial* in time. This conjecture is consistent with Ratner's result [**35**] on the decay of correlations for horocycle flows and with the rate of relative mixing for Heisenberg nilflows (which can be estimated by Fourier analysis). In fact, several results on parabolic flows suggest the following heuristic principle: *a uniquely ergodic smooth flow with polynomial speed of convergence of ergodic averages is a smooth time-change of a smooth flow with polynomial decay of correlations (for smooth functions)*. For horocycle flows, the rate of mixing [**35**] as well as the speed of convergence of ergodic averages [**45**, **36**, **3**, **17**, **10**, **40**] are polynomial. To the best of our knowledge, no quantitative mixing estimates are available for smooth time-changes of the horocycle flow [**29**]. For minimal ergodic area-preserving flows and translation flows, the polynomial decay of ergodic averages (for smooth functions vanishing at sufficiently high order at the singularities) was conjectured by A. Zorich [**46**] and M. Kontsevich [**26**] and proved by the second author in [**14**]. According to the above-mentioned heuristic principle, the decay of correlations for time-changes with a degenerate saddle should also be polynomial. While mixing is known after Kočergin's result [**23**], to the authors' best knowledge polynomial decay of correlations (under a Diophantine condition) has been proved only for the particular case of flows on the 2-torus with a single degenerate saddle of restricted type [**6**]. For Heisenberg nilflows, the speed of convergence of ergodic averages of smooth functions is polynomial and the optimal exponents (which depend on the Diophantine properties of the toral factor) are known [**11**]. This result is related to optimal bounds for Weyl sums of quadratic polynomials; see [**9**, **30**]. According to the heuristics proposed above, there should be mixing time-changes with polynomial decay of correlations.

On the *spectral properties* of our mixing time-changes of Heisenberg nilflows, it is reasonable to conjecture that they have countable Lebesgue spectrum. However, this seems to be a difficult problem, beyond the reach of the methods of this paper.

The mechanism that we use to produce mixing is sometimes known as *stretching of ergodic sums*. The stretching of ergodic sums for Heisenberg nilflows is derived from a theorem on the growth of ergodic sums of functions which are not coboundaries with a *measurable* transfer function. This result is quite general and can be proved for all nilflows. In fact, it is essentially based on a measurable Gottschalk-Hedlund theorem, which holds for any volume preserving uniquely ergodic dynamical system, and on the parabolic divergence of orbits (although in a quite explicit form). Finally, we prove a theorem on cocycle effectiveness for the Heisenberg case, which states that if a smooth function is a coboundary with a measurable transfer function, then the transfer function is in fact smooth. This result is based on sharp bounds for ergodic sums which are only available in the Heisenberg case [**9**, **30**, **11**]. The cocycle effectiveness allows a concrete description of mixing time-changes in terms of the non-vanishing of any of the distributional obstructions to the existence of smooth solutions of the cohomological equation.

A similar mixing mechanism was used by Marcus to prove that horocycle flows on surfaces with variable negative curvature and smooth time-changes of the horocycle flow are mixing [**29**], and by Fayad in [**8**] to produce smooth (analytic) mixing time-changes of some elliptic flows, i.e. linear flows on tori $\mathbb{T}^n$, with $n \geq 3$ and Liouvillean frequencies. It is also worth recalling that if $n = 2$, smooth time-changes of a linear flow on $\mathbb{T}^2$ are never mixing (for example, as a consequence of the result of A. Katok [**19**] quoted above). Moreover, for Diophantine linear flows on $\mathbb{T}^n$, all smooth time-changes are trivial (since all smooth functions of zero average are smooth coboundaries) by the generalization to all dimensions [**18**] of a well-known theorem of Kolmogorov [**25**]. In dimension $n = 2$, the Denjoy-Koksma inequality explains the absence of mixing time-changes even for Liouvillean frequencies, but does not prevent the existence of weak mixing examples, which are in fact topologically generic, as proved in [**7**]. In higher dimensions, the failure of the Denjoy-Koksma inequality opens the way for mixing examples with Liouvillean frequency [**8**]. Thus, in this elliptic realm, the phenomenon of stretching of ergodic sums and mixing time-changes is not generic and can occur only for Liouvillean frequencies, in contrast to our result for nilflows, where mixing time-changes are generic for any uniquely ergodic nilflow, or equivalently, as long as the frequency of the elliptic factor is irrational.

Outline. In Section 2 we give the definitions of Heisenberg nilflows (§ 2.1), special flows (§ 2.4), and time-changes (§ 2.3), and recall how to represent a Heisenberg nilflow as a special flow (§ 2.2). We then state our main results for time-changes of nilflows in § 2.3 (Theorem 3) and in § 2.4 in the language of special flows (Theorem 4). The class of mixing time-changes is defined in § 2.5 (Definition 2) and, as explained in § 2.6, it can be explicitly characterized in terms of invariant distributions for the nilflow (Theorem 7). Sections 3, 4, 5, and 6 are devoted to proofs: in Section 3 we prove that non-triviality of the time-change guarantees that there is stretch of ergodic sums (Theorem 6). Using this stretch, in Section 4 we implement the mixing mechanism and prove mixing (Theorem 5). Section 5 contains the proof of the effective characterization of non-trivial time-changes (Theorem 7) which allows us to exhibit explicit examples of mixing time-changes. The proofs of Theorem 3 and Theorem 4 then follow easily in Section 6.

Acknowledgements. A. Avila and C. Ulcigrai would like to thank the University of Mary- land for the hospitality during the visit when part of this work was completed. A. Avila is partially supported by the Balzan Research Project of J. Palis. G. Forni acknowledges support of the NSF grant DM 0800673. C. Ulcigrai is currently supported by an RCUK Academic Fellowship and the EPSRC First Grant EP/I019030/1, whose support is fully acknowledged.

2. Definitions and main results

2.1. Heisenberg nilflows. The 3-dimensional *Heisenberg group* N is the unique connected, simply connected Lie group with 3-dimensional Lie algebra $\mathfrak{n}$ on two generators X, Y satisfying the Heisenberg commutation relations

$$[X,Y] = Z\,, \quad [X,Z] = [Y,Z] = 0\,.$$

Up to isomorphisms, N is the group of upper triangular unipotent matrices

$$[x,y,z] := \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \qquad x,y,z \in \mathbb{R}. \tag{1}$$

A basis of the Lie algebra $\mathfrak{n}$ satisfying the Heisenberg commutations relations is given by the matrices

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2}$$

The abelianized Lie algebra $\mathfrak{n}/[\mathfrak{n},\mathfrak{n}]$ of the Heisenberg Lie algebra is isomorphic to $\mathbb{R}^2$ (as a Lie algebra); hence the abelianized Lie group $N/[N,N]$ of the Heisenberg group is also isomorphic to $\mathbb{R}^2$ (as a Lie group). In fact, both the center $Z(N)$ and the commutator subgroup $[N,N]$ of the Heisenberg group N are equal to the one-parameter subgroup $\{[0,0,z] : r \in \mathbb{R}\}$ and the maps

$$z \mapsto [0,0,z] \qquad \text{and} \qquad [x,y,z] \mapsto (x,y) \tag{3}$$

define a (non-split) exact sequence

$$0 \to \mathbb{R} \to N \to \mathbb{R}^2 \to 0 \,, \tag{4}$$

which exhibits N as a line bundle over $\mathbb{R}^2$.

A compact *Heisenberg nilmanifold* is the quotient $M := \Gamma\backslash N$ of the Heisenberg group over a co-compact lattice $\Gamma < N$. It is well known that there exists a positive integer $E \in \mathbb{N}$ such that, up to an automorphism of N, the lattice Γ coincides with the lattice

$$\Gamma := \left\{ \begin{pmatrix} 1 & x & z/E \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \;:\; x, y, z \in \mathbb{Z} \right\}.$$

Let $\overline{\Gamma} := \Gamma/[\Gamma,\Gamma] < \mathbb{R}^2$ denote the abelianized lattice. The canonical projection homomorphism $N \to N/[N,N] \approx \mathbb{R}^2$ defined in (3) induces a Seifert fibration $\pi : M \to \mathbb{T}^2 = \overline{\Gamma}\backslash\mathbb{R}^2$, that is, M is a circle bundle over the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ with fibers given by the orbits of the flow by right translation of the central one-parameter subgroup $Z(N) = \{\exp(zZ)\}_{z\in\mathbb{R}}$. The left invariant fields X, Y on M define a connection whose total curvature (the Euler characteristic of the fibration) is exactly E. Any Heisenberg nilmanifold M has a natural probability measure μ locally given by the Haar measure of N.

The group N acts on the right transitively on M by right multiplication:

$$R_g(x) := x\,g, \quad x \in M,\, g \in N.$$

By definition, *Heisenberg nilflows* are the flows obtained by the restriction of this right action to the one-parameter subgroups on N. The measure μ defined above, which is invariant for the right action of N on M, is, in particular, invariant for all nilflows on M.

Thus each $W := w_x X + w_y Y + w_z Z \in \mathfrak{n}$ defines a measure preserving flow (ϕ_W, μ) on M where $\phi_W := \{\phi_W^t\}_{t\in\mathbb{R}}$ is given by the formula

$$\phi_W^t(x) = x \exp(tW), \quad x \in M \;, t \in \mathbb{R}\,.$$

The projection $\overline{W}$ of W into $\mathbb{R}^2$ is the generator of a linear flow $\psi_{\overline{W}} := \{\psi_{\overline{W}}^t\}_{t\in\mathbb{R}}$ on $\mathbb{T}^2 \approx \mathbb{R}^2\backslash\overline{\Gamma}$ defined by

$$\psi_{\overline{W}}^t(x,y) = (x + tw_x, y + tw_y).$$

The canonical projection $\pi : M \to \mathbb{T}^2$ intertwines the flows ϕ_W and $\psi_{\overline{W}}$. We recall the following basic result:

Theorem 1. [**16**, **1**] *The following conditions are equivalent:*

1) *The nilflow (ϕ_W, μ) is ergodic.*
2) *The nilflow ϕ_W is uniquely ergodic.*
3) *The nilflow ϕ_W is minimal.*
4) *The projected flow $\psi_{\overline{W}}$ is an irrational linear flow on $\mathbb{T}^2$ and hence it is minimal and uniquely ergodic.*

Results on the speed of equidistribution of Heisenberg nilflows for smooth functions were proved in [**11**] by L. Flaminio and the second author. Similar results can be proved by bounds on Weyl sums for quadratic polynomials; see [**9**, **30**].

Nilflows are clearly not weak mixing, and hence not mixing. In fact, all eigenfunctions of linear toral flows (that is, all characters of the group $\mathbb{T}^2$) pull-back to eigenfunctions of all nilflows on $M = \Gamma\backslash N$. However, all nilflows are *relatively mixing* in the following sense. Let $H := \pi^* L^2(\mathbb{T}^2) \subset L^2(M)$ be the subspace obtained by pull-back of the square-integrable functions on the torus $\mathbb{T}^2$ and let $H^\perp \subset L^2(\mathbb{T})$ be its orthogonal complement. The following result holds.

Theorem 2. [**16**, **1**] *The restriction of any nilflow (ϕ_W, μ) to the N-invariant subspace $H^\perp \subset L^2(\mathbb{T})$ has countable Lebesgue spectrum; hence it is mixing.*

In fact, it is possible to prove by the theory of unitary representations of the Heisenberg group (the Stone-Von Neumann theorem; see for example [**5**], §2.2) that for all sufficiently smooth functions in $H^\perp$ the decay of correlations is polynomial (it is faster than any polynomial for infinitely differentiable functions in $H^\perp$).

2.2. Return maps of Heisenberg nilflows. Any uniquely ergodic Heisenberg nilflow has a smooth compact transversal surface, isomorphic to a 2-dimensional torus. One can compute the return map and the return time function (see [**39**], §3). It turns out that the return time is constant and the return map is a linear *skew-shift* over an irrational rotation of the circle. We recall this well-known construction for the convenience of the reader.

Let $\Sigma \subset M$ be the smooth surface defined as follows:

$$\Sigma := \{\Gamma \exp(xX + zZ) : \ (x,z) \in \mathbb{R}^2\}\,.$$

Since the subspace $< X, Z >$ generated in $\mathfrak{n}$ by $X, Z \in \mathfrak{n}$ is an abelian ideal, the surface Σ is isomorphic to a 2-dimensional torus. The isomorphism is given by the map

$$j(x,z) = \Gamma \exp(xX + zZ)\,, \quad \text{for all } (x,z) \in \mathbb{T}^2_E := \mathbb{R}^2/(\mathbb{Z} \times \mathbb{Z}/E)\,.$$

Let $W := w_x X + w_y Y + w_z Z$ be the generator of a uniquely ergodic nilflow and let $\phi^W = \{\phi^W_t\}_{t\in\mathbb{R}}$ denote the corresponding Heisenberg nilflow.

Lemma 1. *The first return time function of the flow ϕ^W to the transverse section Σ is constant equal to $1/w_y$ and the first return (Poincaré) map $P_W : \Sigma \to \Sigma$ is given by the following formula:*

$$(5)\quad P_W \circ j(x,z) = j(x + \frac{w_x}{w_y}, z + x + \frac{w_z}{w_y} + \frac{w_x}{2w_y})\,, \quad \textit{for all } (x,z) \in \mathbb{T}^2_E\,.$$

Proof. Since the nilflow is uniquely ergodic, we have $w_y \neq 0$, which implies that the surface Σ is transverse to the nilflow. The set of all return times of the nilflow to Σ is a subset of the set of all return times of the projected linear flow $\psi_{\overline{W}}$ on the torus $\mathbb{T}_\Gamma$, which is equal to the subgroup $\mathbb{Z}/w_y \subset \mathbb{R}$. Finally, by the Baker-Campbell-Hausdorff formula, since $[\mathfrak{n}, [\mathfrak{n}, \mathfrak{n}]] = 0$, we have

$$\begin{aligned}&\exp(-Y)\exp(xX + zZ)\exp\left(W/w_y\right)\\ &\qquad = \exp\left[(x + \frac{w_x}{w_y})X + (z + x + \frac{w_z}{w_y} + \frac{w_x}{2w_y})Z\right] .\end{aligned}$$

Since by definition $\exp(-Y) \in \Gamma$, it follows from the above formula that the forward first return time is equal to $1/w_y$ for all $(x, z) \in \mathbb{T}^2_E$ and that the forward first return time map is given by formula (5) as claimed. q.e.d.

Lemma 1 implies that any (uniquely ergodic) Heisenberg nilflow is smoothly isomorphic to a *special flow* over a linear skew-shift of the form (5) with constant *roof function*. The notion of a special flow is recalled below in Section 2.4.

2.3. Mixing time-changes. We recall below basic notions about time-changes of flows and state our main theorem on mixing of time-changes of Heisenberg nilflows.

A flow $\{\widetilde{h}_t\}_{t\in\mathbb{R}}$ is called a *reparametrization* or a *time-change* of a flow $\{h_t\}_{t\in\mathbb{R}}$ on X if there exists a measurable function $\tau : X \times \mathbb{R} \to \mathbb{R}$ such that for all $x \in X$ and $t \in \mathbb{R}$ we have $\widetilde{h}_t(x) = h_{\tau(x,t)}(x)$. Since $\{\widetilde{h}_t\}_{t\in\mathbb{R}}$ is assumed to be a flow (a one-parameter group), the function $\tau(x, \cdot) : \mathbb{R} \to \mathbb{R}$ is an *additive cocycle* over the flow $\{\widetilde{h}_t\}_{t\in\mathbb{R}}$; that is, it satisfies the cocycle identity:

$$\tau(x, s + t) = \tau(\widetilde{h}_s(x), t) + \tau(x, s)\,, \qquad \text{for all } x \in X\,,\ s, t \in \mathbb{R}\,.$$

If X is a manifold and $\{h_t\}_{t\in\mathbb{R}}$ is a smooth flow, we will say that $\{\widetilde{h}_t\}_{t\in\mathbb{R}}$ is a smooth reparametrization if the cocycle τ is a smooth function. By the cocycle property a smooth cocycle is uniquely determined by its infinitesimal generator, that is, the function $\alpha_\tau : X \to \mathbb{R}$ defined by the formula:

$$\alpha_\tau(x) := \frac{\partial \tau}{\partial t}(x, 0)\,, \qquad \text{for all } x \in X\,.$$

In fact, given any positive function $\alpha : X \to \mathbb{R}^+$, the formula

$$\tau_\alpha(x, t) := \int_0^t \alpha(\tilde{h}_s(x))ds\,, \qquad \text{for all}\,(x, t) \in X \times \mathbb{R}$$

defines a cocycle over the flow $\{\tilde{h}_t\}_{t\in\mathbb{R}}$ with infinitesimal generator α.

The infinitesimal generators $\widetilde V$ and V of the flows $\{\widetilde h_t\}_{t\in\mathbb{R}}$ and $\{h_t\}_{t\in\mathbb{R}}$ respectively are related by the identity:

$$\widetilde V := \frac{d\widetilde h_t}{dt}\Big|_{t=0} = \alpha_\tau \frac{dh_t}{dt}\Big|_{t=0} := \alpha_\tau V\,.$$

An additive cocycle $\tau : X\times\mathbb{R}\to\mathbb{R}$ over the flow $\{\widetilde h_t\}_{t\in\mathbb{R}}$ is called a measurable (respectively smooth) *coboundary* if there exists a measurable (respectively smooth) function $u: X\to\mathbb{R}$, called the *transfer function*, such that

$$\tau(x,t) = u\circ \widetilde h_t(x) - u(x)\,, \qquad \text{for all } (x,t)\in X\times\mathbb{R}\,.$$

The additive cocycle τ is a measurable (smooth) coboundary if and only if its infinitesimal generator α_τ is a measurable (smooth) *coboundary* for the infinitesimal generator V of the flow $\{\widetilde h_t\}_{t\in\mathbb{R}}$, that is, if there exists a measurable (smooth) function $u : X\to\mathbb{R}$, also called the transfer function, such that $\widetilde V u = \alpha_\tau$. 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 (see [**20**], def. 9.4).

An elementary, but fundamental, result establishes that time-changes given by measurably (smoothly) cohomologous coycles are measurably (smoothly) isomorphic (see for example [**20**], §9). The regularity of the isomorphisms depends on the regularity of the transfer function. A time-change defined by a measurable (smooth) almost coboundary is called *measurably (smoothly) trivial*.

Let $\{h_t\}_{t\in\mathbb{R}}$ be a uniquely ergodic homogeneous flow on the Heisenberg nilmanifold M. For any function $\alpha : C^\infty(M)\to\mathbb{R}^+$ let $h^\alpha := \{h^\alpha_t\}_{t\in\mathbb{R}}$ be the time-change with generator given by the formula

$$\frac{dh^\alpha_t}{dt}\Big|_{t=0} = \alpha\,\frac{dh_t}{dt}\Big|_{t=0}\,.$$

We recall that a measure preserving flow $\varphi := \{\varphi_t\}_{t\in\mathbb{R}}$ on a probability space (X,μ) is said to be *weak mixing* if, for each pair of measurable sets A, $B\subset X$,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t |\mu(\varphi_s(A)\cap B) - \mu(A)\mu(B)|\,\mathrm{d}s = 0\,,$$

and *mixing* if for each pair of measurable sets A, $B\subset X$, one has

$$\lim_{t\to\infty}\mu(\varphi_t(A)\cap B) = \mu(A)\mu(B)\,.$$

Theorem 3 (Mixing time-changes for Heisenberg nilflows)**.** *There exists a subspace $\mathcal{T}_h \subset \mathcal{A} \subset C^\infty(M)$ of countable dimension and countable codimension in a dense subspace $\mathcal{A} \subset C^\infty(M)$ such that, for any positive function $\alpha \in \mathcal{A}$, the following properties are equivalent:*

1) *the function $\alpha \in \mathcal{M}_h := \mathcal{A} \setminus \mathcal{T}_h$;*
2) *the time-change h^α is not smoothly trivial;*
3) *the time-change h^α is weak mixing;*
4) *the time-change h^α is mixing.*

Thus, for time-changes within the dense subspace $\mathcal{A} \subset C^\infty(M)$, mixing is equivalent to weak mixing and any non smoothly trivial time-change is automatically mixing. Theorem 3 is proved in § 6. Our results leave open several natural questions on possible generalizations of Theorem 3 and on the dynamics of the mixing flows constructed.

Questions.

a) Does Theorem 3 hold within the class of all smooth time-changes?

b) Does it extend to nilflows on 2-step nilmanifolds on several generators?

c) Does it extend to nilflows on s-step nilmanifolds for any $s \geq 3$?

d) Is the correlation decay polynomial in time for sufficiently smooth functions (under a Diophantine condition on the frequency)?

e) Is the spectrum of mixing time-changes singular continuous or absolutely continuous? Is it Lebesgue with countable multiplicity?

2.4. Mixing special flows over skew shifts on $\mathbb{T}^2$. In this section we recall the notion of a special flow and the representation of time-changes in terms of special flows. We then state our main theorem for special flows over uniquely ergodic skew-shifts on $\mathbb{T}^2$.

Let $f : \Sigma \to \Sigma$ be a Poincaré return map of the flow $\{h_t\}_{t\in\mathbb{R}}$ on X to a measurable transverse section $\Sigma \subset X$, and let $\Phi : \Sigma \to \mathbb{R}^+$ be the *return time function* (in general defined only almost everywhere). The flow $\{h_t\}_{t\in\mathbb{R}}$ is isomorphic to a *special flow* over the map $f : \Sigma \to \Sigma$ with *roof function* $\Phi > 0$, defined as we now recall. Given any function Φ on Σ, let Φ_n denote the n^{th} ergodic sums along the orbits of the map $f : \Sigma \to \Sigma$, that is, the function

$$\Phi_n := \sum_{k=0}^{n-1} \Phi \circ f^k \,. \tag{6}$$

If $\Phi > 0$ is a continuous positive function, we let $f^\Phi = \{f^\Phi_t\}_{t\in\mathbb{R}}$ be the special flow over f with roof function Φ, which is defined as the quotient of the unit speed vertical flow $\dot z = 1$ on the phase space $\{(x, z) \in \Sigma \times \mathbb{R}\}$ with respect to the equivalence relation $\sim_\Phi$ defined by $(x, \Phi(x) + z) \sim_\Phi (f(x), z)$, for all $x \in \Sigma$, $z \in \mathbb{R}$. The flow f^Φ can thus be seen as defined on the fundamental domain $\{(x, z) : x \in \Sigma,\ 0 \leq z < \Phi(x)\}$ and explicitly given by the formula

$$f_t^{\Phi}(x,z) = \left(f^{n_t(x,z)}(x), z + t - \Phi_{n_t(x,z)}(x)\right), \tag{7}$$

where $n_t(x,z)$ is the maximum $n \in \mathbb{N}$ such that $\Phi_n(x) \le t + z$. For any f-invariant measure ν on Σ, the finite measure obtained by the restriction of the product measure $\nu \times \mathrm{Leb}$ (where Leb is the Lebesgue measure in the z-fiber) to the domain of f^{Φ} is invariant by the special flow f^{Φ}.

A function $\Phi : \Sigma \to \mathbb{R}$ is called a measurable (smooth) *coboundary* for the map $f : \Sigma \to \Sigma$ if and only if there exists a measurable (smooth) function $u : \Sigma \to \mathbb{R}$, also called the *transfer function*, such that $\Phi = u \circ f - u$. Two functions are called measurably (smoothly) *cohomologous* if their difference is a measurable (smooth) coboundary. As time-changes defined by measurably (smoothly) cohomologous cocycles are measurably (smoothly) isomorphic, similarly special flows over the same map under measurably (smoothly) cohomologous roof functions are measurably (smoothly) isomorphic (we refer for example to [**20**]).

Any time-change $\{h_t^{\alpha}\}_{t\in\mathbb{R}}$ of $\{h_t\}_{t\in\mathbb{R}}$ determines the same return map $f : \Sigma \to \Sigma$, but a different return time function $\Phi^{\alpha} : \Sigma \to \mathbb{R}^+$. In fact, the following elementary result holds.

Lemma 2. *The return time function $\Phi^{\alpha} : \Sigma \to \mathbb{R}^+$ is given by the formula:*

$$\Phi^{\alpha}(x) = \int_0^{\Phi(x)} (\alpha \circ h_t)(x)dt\,, \qquad \textit{for all } x \in \Sigma\,.$$

It follows in particular from Lemma 2 that the return time functions Φ^{α} and Φ are *cohomologous* with respect to the return map $f : \Sigma \to \Sigma$ if and only if the function $\alpha : X \to \mathbb{R}$ is cohomologous to the constant function equal to 1 for the infinitesimal generator of the flow $\{h_t(x)\}_{t\in\mathbb{R}}$.

In the rest of this section we will consider the case when $\Sigma = \mathbb{T}^2$ and $f : \mathbb{T}^2 \to \mathbb{T}^2$ is a *linear skew-shift* over a circle rotation, defined as

$$f(x,y) := (x+\alpha, y+x+\beta), \quad \text{where } (x,y) \in \mathbb{T}^2,\ \alpha,\beta \in \mathbb{R}. \tag{8}$$

We will also assume that f is uniquely ergodic, which is equivalent to $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ (see [**4**]). As we saw in § 2.2 (see Lemma 1), any uniquely ergodic Heisenberg nilflow ϕ^W has a global cross section on which the first return (Poincaré) map has the form (8). We will denote by Leb (respectively Leb_2) the one-dimensional (respectively the two-dimensional) Lebesgue measure and by μ the probability measure obtained by normalization of the restriction of the measure $\mathrm{Leb}_2 \times \mathrm{Leb}$ on $\mathbb{T}^2 \times \mathbb{R}$ to the domain of f^{Φ} (the normalizing factor is equal to $1/\int_{\mathbb{T}^2} \Phi(x,y)\,\mathrm{d}x\,\mathrm{d}y$). By construction μ is invariant under the special flow f^{Φ} on $\Sigma \times \mathbb{R}/\sim_{\Phi}$.

It is well known (see Lemma 16, § 6) that if the roof function $\Phi > 0$ is a *measurable (smooth) almost coboundary*, that is, if there exists a measurable (smooth) function $u : X \to \mathbb{R}$ such that

$$u \circ f - u = \Phi - \int_{\mathbb{T}^2} \Phi \, \mathrm{dLeb}\,,$$

then the special flow f^Φ is measurably (smoothly) isomorphic to a special flow with constant roof function over the skew-shift. In this case, we will call the special flow f^Φ *measurably (smoothly) trivial*. Any measurably trivial special flow is not weak-mixing, and hence not mixing (see again Lemma 16, § 6).

We will show that there is a class $\mathcal{M}_f$ of smooth roof functions which correspond to smooth *mixing* special flows over a uniquely ergodic skew-shift and that $\mathcal{M}_f$ is *generic* in a precise sense. In fact, we prove the following.

Theorem 4 (Mixing special flows)**.** *There exists a subspace $\mathcal{T}_f \subset \mathcal{R} \subset C^\infty(\mathbb{T}^2)$ of countable dimension and countable codimension in a dense subspace $\mathcal{R} \subset C^\infty(\mathbb{T}^2)$ such that, for any positive function $\Phi \in \mathcal{R}$, the following properties are equivalent:*

1) *the roof function $\Phi \in \mathcal{M}_f := \mathcal{R} \setminus \mathcal{T}_f$;*
2) *the special flow f^Φ is not smoothly trivial;*
3) *the special flow f^Φ is weak mixing;*
4) *the special flow f^Φ is mixing.*

It is natural to ask whether Theorem 4 generalizes to linear skew-shift on $\mathbb{T}^n$ with $n > 2$. A statement of the form 1) $\Rightarrow$ 4) could be proved for higher dimensional skew shifts; see Remark 9 in § 3. On the other side, the implication 2) $\Rightarrow$ 1) requires the analogue of the cocycle effectiveness (Theorem 7 below) which relies on estimates currently known only for $n = 2$ (see Remark 8 below and Remark 12 in § 5).

Let us remark that the generic subset $\mathcal{M}_f$ in Theorem 4 is concretely described in terms of invariant distributions (see § 5). Thus, it is possible to check *explicitly* if a given smooth roof function given in terms of a Fourier expansion belongs to $\mathcal{M}_f$ and to give concrete examples of mixing reparametrizations.

Examples.

The following roof functions all give examples of mixing special flows.

(E1) $\Phi(x, y) = \sin(2\pi y) + 2$;
(E2) $\Phi(x, y) = \cos(2\pi(kx + y)) + \sin(2\pi l x) + 3$, $k, l \in \mathbb{Z}$;
(E3) $\Phi(x, y) = \mathrm{Re} \sum_{j \in \mathbb{Z}} a_j e^{2\pi i (jx+y)} + c$, if $\sum_{j\in\mathbb{Z}} a_j e^{-2\pi i(\beta j + \alpha\binom{j}{2})} \neq 0$ and c is such that $\Phi > 0$.

Example (E1) shows that it is enough to have oscillations in the y-variable to produce mixing. We show that the roofs in the examples above belong to the class $\mathcal{M}_f$ at the end of § 5, after Corollary 2.

2.5. Mixing roof functions. Here we define the class of roof functions considered to obtain mixing special flows. Let $\pi : \mathbb{T}^2 \to \mathbb{T}$ be the projection defined as $\pi(x, y) = x$ for all $(x, y) \in \mathbb{T}^2$. The space $\pi^* L^2(\mathbb{T}) := \{\Phi \circ \pi : \ \Phi \in L^2(\mathbb{T})\}$ is a closed subspace of $L^2(\mathbb{T}^2)$, hence there is an orthogonal decomposition

$$L^2(\mathbb{T}^2) = \pi^* L^2(\mathbb{T}) \oplus \pi^* L^2(\mathbb{T})^{\perp}\,.$$

We introduce the following notation for the orthogonal projections of a function $\Phi \in L^2(\mathbb{T}^2)$ onto the components of the above splitting:

$$\phi(x, y) := \Phi(x, y) - \int \Phi(x, y) dy \in \pi^* L^2(\mathbb{T})^{\perp}, \tag{9}$$

$$\phi^{\perp}(x) := \int \Phi(x, y) dy \in \pi^* L^2(\mathbb{T}) \equiv L^2(\mathbb{T}). \tag{10}$$

Definition 1 (Roofs class $\mathcal{R}$)**.** *For an integer $d \geq 1$, let $\mathcal{P}_d$ be the space of all continuous Φ such that for each $x \in \mathbb{T}$, $\Phi(x, \cdot)$ is a trigonometric polynomial of degree at most d on $\mathbb{T}$. Let $\mathcal{P} := \bigcup_{d \geq 1} \mathcal{P}_d$.*

The function $\Phi \in \mathcal{R}$ if and only if $\Phi \in \mathcal{P}$, Φ is real-valued and positive, and its projection $\phi^{\perp}$, defined in (10)*, is a trigonometric polynomial on $\mathbb{T}$.*

We remark that if $\Phi \in \mathcal{R}$, we can write $\Phi(x, y) = \sum_{k=-d}^{d} c_k(x) e^{2\pi i k y}$, where $c_{-k}(x) = \overline{c_k}(x)$ since $\Phi \in \mathcal{P}_d$ and is real, and $c_0(x)$ is a trigonometric polynomial. By definition the set $\mathcal{R}$ is a dense subspace of $C^{\infty}(\mathbb{T}^2)$.

Definition 2 (Trivial roofs $\mathcal{T}_f$ and mixing roofs $\mathcal{M}_f$)**.** *A function Φ belongs to $\mathcal{T}_f$ if and only if $\Phi \in \mathcal{R}$ and its projection ϕ defined in* (9) *is a measurable coboundary for the map $f : \mathbb{T}^2 \to \mathbb{T}^2$. Let us set $\mathcal{M}_f := \mathcal{R} \setminus \mathcal{T}_f$, so that Φ belongs to $\mathcal{M}_f$ if and only if $\Phi \in \mathcal{R}$ and ϕ is* not *a measurable coboundary.*

One of the two main steps in the proof of Theorem 4 is given by the following theorem.

Theorem 5 (Mixing)**.** *For any positive roof function Φ belonging to the class $\mathcal{M}_f$ in Definition 2, the special flow f^{Φ} is mixing.*

The crucial ingredient in the proof of Theorem 5 is given by the following result on the growth of ergodic sums of the skew-shift.

Theorem 6 (Stretch of ergodic sums)**.** *Assume that $\Phi \in \mathcal{M}_f$; thus ϕ is not a measurable coboundary. Then for each $C > 1$,*

$$\lim_{n \to \infty} \mathrm{Leb}_2(|\phi_n| < C) = 0\,.$$

Let us remark that a similar result was proved by K. Schmidt in [**38**] in much greater generality under the assumption that the base transformation f is mixing.

The proof of Theorem 6 is given in § 3, while the proof of Theorem 5 is in § 4.

2.6. Cocycle effectiveness. The following effectiveness result for coboundaries (in the sense of [**20**], def. 11.4) leads to a complete explicit description of the set $\mathcal{M}_f$ in terms of Fourier series, and it constitutes another main step in the proof of Theorem 4 (see § 6).

We recall that, as found by A. Katok [**20**] (see §11.6.1), there are countably many independent obstructions (which are not signed measures) to the existence of smooth solutions of the cohomological equation $u \circ f - u = \phi$. Such obstructions are invariant distributions for the skew-shift (see Theorem 10, § 5). If $\phi \in \pi^* L^2(\mathbb{T})^\perp$ is smooth and belongs to the kernel of all f-invariant distributions, then it is a coboundary and the *transfer function*, that is, the unique zero average solution of the cohomological equation, is smooth.

We will show that, if a sufficiently smooth function Φ such that $\phi^\perp = 0$ is a coboundary for a skew-shift f on $\mathbb{T}^2$ with a measurable transfer function, then Φ belongs to the kernel of the (infinite dimensional) space of all f-invariant distributions and the transfer function is smooth. More precisely, let $W^s(\mathbb{T}^2)$ denote the standard Sobolev space on $\mathbb{T}^2$, that is, the space of all functions $\Phi = \sum_{(m,n)\in\mathbb{Z}^2} \Phi_{m,n} \exp(2\pi i(mx+ny))$ such that

$$\|\Phi\|_s := \left(\sum_{(m,n)\in\mathbb{Z}^2} (1+m^2+n^2)^s |\Phi_{m,n}|^2 \right)^{1/2} < +\infty\,.$$

Theorem 7 (Cocycle Effectiveness)**.** *Let f be any uniquely ergodic skew-shift on $\mathbb{T}^2$ as in* (8)*. For any function $\phi \in \pi^* L^2(\mathbb{T})^\perp \cap W^s(\mathbb{T}^2)$ for $s > 3$, the following holds. If ϕ is a measurable coboundary, then it belongs to the kernel of all f-invariant distributions and the transfer function $u \in W^t(\mathbb{T}^2)$ for all $t < s-1$.*

The proof of Theorem 7, given in § 5, is based on the quantitative estimates on equidistribution of nilflows by L. Flaminio and the second author in [**11**].

Remark 8. Theorem 7 above answers a question posed by A. Katok in [**20**], §11.6.1, p. 88. For higher dimensional skew-shifts (or for any other higher dimensional nilpotent linear map), the analogue of Theorem 7 is not known.

3. Stretch of ergodic sums

In this section we prove Theorem 6. Let Φ be any continuous function on $\mathbb{T}^2$ such that its projection ϕ (see (9)) is *not a measurable coboundary*. Since the map f is uniquely ergodic, we can derive the following result by a standard Gottschalk-Hedlund technique.

Lemma 3. *For each constant $C > 1$ and for all $(x, y) \in \mathbb{T}^2$,*

$$\frac{1}{N}\#\{0 \leq n \leq N-1 \, s.t. \; |\phi_n(x,y)| < C\} \xrightarrow{N\to\infty} 0.$$

Proof. Let $\mu_{N,x,y}$ be a probability measure on $\mathbb{T}^2 \times \mathbb{R}$ with atoms of equal mass along $(f^k(x,y), \phi_k(x,y))$, $0 \leq k \leq N-1$. It is enough to prove that $\mu_{N,x,y} \to 0$ in the weak-$*$ topology, as $N \to \infty$, independently of $(x,y) \in \mathbb{T}^2$. If this did not happen, we would be able to take a non-trivial limit, which would be a measure μ with non-zero mass, such that $F_*\mu = \mu$, where $F(x,y,z) = (f(x,y), z + \phi(x,y))$.

By unique ergodicity of f, $\pi_*\mu$ is a multiple of Leb_2, where $\pi(x,y,z) = (x,y)$, and the conditional measures $\mu_{x,y}$ coincide up to translation: for almost every x, y, x', y', $\mu_{x,y} = T_*\mu_{x',y'}$ where $T(z) = z + t$, with $t = t(x,y,x',y')$. By invariance, we have $t(x,y,f(x',y')) = t(x,y,x',y') + \phi(x',y')$. Choosing (x_0,y_0) in a full measure set, and defining $u(x,y) = t(x_0,y_0,x,y)$, we get $\phi = u \circ f - u$, which contradicts the assumption that ϕ is not a measurable coboundary. q.e.d.

Corollary 1. *For each constant $C > 1$,*

$$\frac{1}{N}\sum_{n=0}^{N-1} \mathrm{Leb}_2(|\phi_n| < C) \xrightarrow{N\to\infty} 0.$$

Proof. The functions $\frac{1}{N}\sum_{n=0}^{N-1} \chi_{(-C,C)} \circ \phi_n$, where $\chi_{(-C,C)}$ is the characteristic function of the interval $(-C,C) \subset \mathbb{R}$, converge pointwise to zero by Lemma 3. Thus, the corollary follows immediately by integration over $\mathbb{T}^2$ and by the Lebesgue dominated convergence theorem. q.e.d.

Lemma 4. *For each $d \geq 1$ and for any norm $\|\cdot\|_d$ on $\mathbb{C}^{2d}$, there exist constants B_d and $b_d > 0$ such that if $\mathbf{c} = (c_{-d}, \ldots, c_{-1}, c_1, \ldots, c_d) \in \mathbb{C}^{2d}$ is a vector of unit norm (that is, $\|\mathbf{c}\|_d = 1$), then for every $\delta > 0$*

$$\mathrm{Leb}\left(\left|\sum_{0<|k|\leq d} c_k e^{2\pi i k x}\right| < \delta\right) < B_d \delta^{b_d}. \tag{11}$$

Proof. Since $\sum_{|k|\leq d} c_k e^{2\pi i k x}$ is a function of class C^{2d} with critical points of order at most $2d$, an estimate of the form (11) holds for any $\delta > 0$, where the exponent b_d depends only on d and the constant involved depends continuously on the coefficients c_k. Since the set of trigonometric polynomials of the form $\sum_{|k|\leq d} c_k e^{2\pi i k x}$ with $\|\mathbf{c}\|_d = 1$ forms a compact set of the space of functions of class C^{2d} on $\mathbb{R}$ with critical points of order at most $2d$, the estimate follows. q.e.d.

From now on we assume that $\Phi > 0$ and $\Phi \in \mathcal{R}$.

Lemma 5. *Let $C > 1$. For any $\epsilon > 0$, there exist $C' > 1$ and $\epsilon' > 0$ such that for all $n \geq 1$ such that $\mathrm{Leb}_2(|\phi_n| < C') < \epsilon'$, there exists $N_0 := N_0(C, \epsilon, n) \in \mathbb{N}$ such that for all $N \geq N_0$, we have*

$$\mathrm{Leb}_2(|\phi_N \circ f^n - \phi_N| < 2C) < \epsilon\,.$$

Proof. Indeed, let us write

$$\phi_n(x, y) = 2\mathrm{Re} \sum_{0<k\leq d} c_{k,n}(x) e^{2\pi iky} = \sum_{0<|k|\leq d} c_{k,n}(x) e^{2\pi iky}\,,$$

with $c_{-k,n}(x) = \overline{c_{k,n}}(x)$ for all $0 < k \leq d$, $x \in \mathbb{T}$. Then

$$\begin{aligned} \phi_N \circ f^n(x, y) - \phi_N(x, y) &= \phi_n \circ f^N(x, y) - \phi_n(x, y) \\ &= \sum_{0<|k|\leq d} c_{k,N,n}(x) e^{2\pi iky}, \end{aligned} \tag{12}$$

where we have denoted

$$c_{k,N,n}(x) := e^{2\pi ik[\binom{N}{2}\alpha + N\beta]} c_{k,n}(x + N\alpha) e^{2\pi ikNx} - c_{k,n}(x).$$

Given any two complex numbers $c_i = \rho_i e^{\theta_i}$, $i = 1, 2$, for each $0 \leq \theta < \pi/2$, if $\theta_2 + 2\pi kNx \notin (\theta_1 + \pi - \theta, \theta_1 + \pi + \theta) + 2\pi\mathbb{Z}$, then by elementary trigonometry $|c_2 e^{2\pi ikNx} - c_1| \geq |c_1| \sin\theta$. Thus, for any interval $I \subset \mathbb{T}$ of length at most $\delta > 0$ and for $0 \leq \theta < \pi/2$ we have

$$\mathrm{Leb}\{x \in I \;:\; |c_2 e^{2\pi ikNx} - c_1| \leq |c_1| \sin\theta\} \leq \delta\,\frac{\theta}{\pi} + \frac{2\theta}{\pi kN}. \tag{13}$$

By uniform continuity of $c_{k,n}$, it is possible to choose $\delta > 0$ so that if $|x - x'| \leq \delta$, then $|c_{k,n}(x) - c_{k,n}(x')| \leq 1/4$. Let us decompose $\mathbb{T}$ into intervals of size at most δ. If $[x_1, x_2)$ is one of these intervals and $x \in [x_1, x_2]$, if we set

$$c_1 := c_{k,n}(x_1)\,, \quad c_2 := e^{2\pi ik[\binom{N}{2}\alpha + N\beta]} c_{k,n}(x_1 + N\alpha),$$

we have

$$|c_{k,N,n}(x)| \geq |c_2 e^{2\pi iNx} - c_1| - 1/2.$$

Using the estimate (13) on each of these intervals we get that, for every $0 < k \leq d$ and for $0 < \theta < \frac{\pi}{2}$, the following bound holds:

$$\limsup_{N\to\infty} \mathrm{Leb}\left(|c_{k,N,n}(x)| < |c_{k,n}(x)| \sin\theta - 3/4\right) \leq \frac{\theta}{\pi}.$$

Choose $\epsilon' > 0$ and $C' > 1$ to be such that

$$C' \geq 9d^2, \qquad B_d(4C/\sqrt{C'})^{b_d} + d/2\sqrt{C'} + 2\epsilon' < \epsilon\,. \tag{14}$$

Let $n \geq 1$ be such that $|\sum_{0<|k|\leq d} c_{k,n}(x) e^{2\pi iky}| \geq C'$ except for a set of $x \in \mathbb{T}$ of Lebesgue measure $\epsilon' > 0$. Choosing $\theta \in [0, \pi/2)$ such that $\sin\theta < 1/\sqrt{C'}$ and, using the fact that $\sin\theta > \frac{2}{\pi}\theta$ for all $0 < \theta < \frac{\pi}{2}$ and the choice (14) of C', for N sufficiently large, outside a set of measure $2d(\theta/\pi) + 2\epsilon' \le d/\sqrt{C'} + 2\epsilon'$, we have

$$\begin{aligned}\sum_{0<|k|\le d} |c_{k,N,n}(x)| &\ge \sum_{0<|k|\le d} |c_{k,n}(x)| \sin\theta - 3d/2 \\ &\ge \left| \sum_{0<|k|\le d} c_{k,n}(x) e^{2\pi i k y} \right| \frac{1}{\sqrt{C'}} - \frac{\sqrt{C'}}{2} \ge \frac{\sqrt{C'}}{2}.\end{aligned} \tag{15}$$

By Lemma 4 (using as norm $\|\mathbf{c}\|_d = \sum_{0\le|k|\le d} |c_k|$), outside of a set of $y \in \mathbb{T}$ of Lebesgue measure at most $B_d \left(4C/\sqrt{C'}\right)^{b_d}$, whenever $x \in \mathbb{T}$ is such that (15) holds, we have that

$$\left| \sum_{0<|k|\le d} c_{k,N,n}(x) e^{2\pi i k y} \right| \ge \frac{4C}{\sqrt{C'}} \sum_{0<|k|\le d} |c_{k,N,n}(x)| \ge 2C. \tag{16}$$

Thus, except for a set of measure at most $B_d(4C/\sqrt{C'})^{b_d} + d/\sqrt{C'} + 2\epsilon'$, which is less than ϵ by the choice (14) of C' and ϵ', combining (12) and (16) we find the desired estimate $|\phi_N \circ f^n - \phi_N| \ge 2C$. q.e.d.

Proof of Theorem 6. Let $C > 1$ and $\epsilon > 0$ be fixed. We prove below that for every N sufficiently large, $\mathrm{Leb}_2(|\phi_N| < C) < \epsilon$. Let us fix an integer $A \ge 1$ and $\epsilon_0 > 0$ such that

$$1/(A+1) + A(A+1)\epsilon_0/2 < \epsilon. \tag{17}$$

By Lemma 5 there exist $C' > 0$ and $\epsilon'' > 0$ such that if $\mathrm{Leb}_2(|\phi_n| < C') < \epsilon'$ then $\mathrm{Leb}_2(|\phi_N \circ f^n - \phi_N| < 2C) < \epsilon_0$ for all $N \ge N_0(C, \epsilon, n)$. By Corollary 1, we can find $l \ge 1$ such that for each $n = jl$ with $1 \le j \le A$ we have $\mathrm{Leb}_2(|\phi_n| < C') < \epsilon'$. Let $N_1 := \max\{N_0(C, \epsilon, jl) : 1 \le j \le A\}$. We claim that, for every $N \ge N_1$,

$$\mathrm{Leb}_2 \left(\bigcup_{0\le j<j'\le A} \left\{ \left| \phi_N \circ f^{jl} - \phi_N \circ f^{j'l} \right| < 2C \right\} \right) \le \frac{A(A+1)}{2} \epsilon_0 \,. \tag{18}$$

In fact, for every $0 \le j < j' \le A$, for $N \ge N_1 \ge N_0(C, \epsilon, (j'-j)l)$, we have $\mathrm{Leb}_2(|\phi_N \circ f^{(j'-j)l} - \phi_N| < 2C) < \epsilon_0$, but since f is measure preserving,

$$\mathrm{Leb}_2(|\phi_N \circ f^{jl} - \phi_N \circ f^{j'l}| < 2C) = \mathrm{Leb}_2(|\phi_N \circ f^{(j'-j)l} - \phi_N| < 2C) < \epsilon_0 \,.$$

The claim follows as $\#\{(j, j') : 0 \le j < j' \le A\}$ is equal to $\frac{A(A+1)}{2}$.

By construction, for all $N \ge N_1$, the sets $f^{-jl}\{|\phi_N| < C\}$ are pairwise disjoint for $j = 0, \dots, A$ outside a set of measure at most $\frac{A(A+1)}{2}\epsilon_0$ (the set in formula (18)); hence again by the measure preserving property of the map and choice (17) of A and ϵ_0,

$$\mathrm{Leb}_2(|\phi_N| < C) \le \frac{1}{A+1} + \frac{A(A+1)}{2}\epsilon_0 < \epsilon\,.$$

The proof is complete. q.e.d.

Remark 9. While Lemma 3 holds for any uniquely ergodic transformation, Lemma 5 exploits the parabolic divergence of orbits of the skew product in the neutral (isometric) direction. A similar result could be proved more in general for higher dimensional skew-product maps of $\mathbb{T}^k$. In fact, it holds most likely for return maps of arbitrary uniquely ergodic nilflows. However, in this latter case, a more indirect argument is needed since exact formulas are not available in general.

4. Mixing

In this section we give the proof of the main mixing result, Theorem 5. We are going to use the following mixing criterion.

4.1. Mixing criterion. Let $f^\Phi := \{f_t^\Phi\}_{t\in\mathbb{R}}$ be the special flow over a uniquely ergodic skew-shift $f : \mathbb{T}^2 \to \mathbb{T}^2$ of the form (8) under a roof function $\Phi : \mathbb{T}^2 \to \mathbb{R}^+$ belonging to the set $\mathcal{R}$ in Definition 2. Recall that for given $t \in \mathbb{R}$ and $(x,y) \in \mathbb{T}^2$ we denote by

$$n_t(x,y) := \max\{n \in \mathbb{N} : \Phi_n(x,y) \le t\}\,,$$

and that by the definition (7) of a special flow on a fundamental domain of $\mathbb{T}^2 \times \mathbb{R}/\sim_\Phi$, for any $(x,y) \in \mathbb{T}^2$ we have

$$f_t^\Phi((x,y),0) = (f^{n_t(x,y)}, t - \Phi_{n_t(x,y)})\,.$$

In order to prove that f^Φ is mixing, it is enough to show the following. Let us call *cube* any set of the form $[x_1,x_2] \times [y_1,y_2] \times [0,h]$ where $[x_1,x_2], [y_1,y_2] \subset \mathbb{T}$ and $0 < h < \min\Phi$. Let us call a *partial partition into intervals* of $\{x\}\times\mathbb{T}$ any collection of intervals I of the form $I = \{x\} \times [y',y'']$, $[y',y''] \subset \mathbb{T}$ with pairwise disjoint interiors.

Lemma 6 (Mixing Criterion)**.** *The flow f^Φ is mixing if, for any cube Q, any $\epsilon > 0$ and any $\delta > 0$, there exists $t_0 > 0$ and for all $t \ge t_0$ there exists a measurable set $X(t) \subset \mathbb{T}$ and for each $x \in X(t)$ there exists a partial partition $\xi(x,t)$ into intervals of $\{x\}\times\mathbb{T}$ such that*

$$\mathrm{Leb}_2\left(\mathbb{T}^2 \setminus \left(\cup_{x\in X(t)} \cup_{I\in\xi(x,t)} I\right)\right) \le \delta \tag{19}$$

and for all $x \in X(t)$ and all $[y',y''] \in \xi(x,t)$,

$$\mathrm{Leb}_x(\{x\}\times[y',y''] \cap f_{-t}^\Phi(Q)) \ge (1-\epsilon)(y''-y')\mu(Q), \tag{20}$$

where Leb_x denotes here the Lebesgue measure on the fiber $\{x\}\times\mathbb{T}$.

The proof of the lemma is an application of Fubini theorem. Details can be found in [**8**, **42**]. We are going to prove that f^Φ is mixing by constructing sets $X(t)$ and partial partitions $\xi(x,t)$ of the fibers $\{x\}\times\mathbb{T}$ with $x \in X(t)$ which satisfy the mixing estimate (20).

4.2. Mixing mechanism outline. The main mechanism that we use to prove (20) is a phenomenon of *stretching of ergodic sums* in the z-direction. Let us first give a heuristic explanation of this mechanism and an outline of the proof. We recall that this type of mechanism was used to prove mixing of horocycles flow in variable curvature by Marcus [**29**], to produce mixing reparametrizations of flows over Liouvillean rotations on $\mathbb{T}^2$ by Fayad in [**8**], and to prove mixing in a class of area-preserving flows on the torus (by Sinai and Khanin in [**21**]) and on higher genus surfaces (by the last author in [**42**]).

Fix $x_0 \in \mathbb{T}$ and let $I = \{x_0\} \times [a,b]$ be a subinterval of the y-fiber $\{x_0\}\times\mathbb{T}$. The *stretch* of Φ_n on I is by definition the following quantity:

$$\Delta\Phi_n(I) := \max_{a\le y\le b} \Phi_n(x_0,y) - \min_{a\le y\le b} \Phi_n(x_0,y).$$

We will show using Theorem 6 that one can find, for all sufficiently large t, a set of intervals $I = \{x\}\times[y',y'']$ whose union has large measure in $\mathbb{T}^2$ and which have large stretch $\Delta\Phi_n(I)$ for all times n of the form $n_t(x,y)$ for some $(x,y)\in I$. As shown in the next section (§ 4.3), large stretch implies that the variation of the number of discrete iterations $n_t(x,y)$ with $(x,y)\in I$ is large. Moreover, our construction will be such that the function $y \mapsto n_t(x,y)$ is monotone on each of the constructed intervals $[y',y'']$. If we subdivide I into intervals I_i on which $n_t(x,y)$ is constant, the image under f_t^Φ of each I_i in the interior of I is a 1-dimensional curve $\gamma_i = f_t^\Phi(I_i)$ which goes from the base (i.e. the set $\mathbb{T}^2\times\{0\}$) to the roof (i.e. the set $\{(x,y,\Phi(x,y)) : (x,y)\in\mathbb{T}^2\}$). Since f sends y-fibers to y-fibers and preserves distances within y-fibers, the projection of each curve γ_i under the map $(x,y,z)\mapsto(x,y)$ is an interval in another y-fiber of the same length as I_i. If the intervals I are chosen sufficiently small, the projections of the curves γ_i shadow with good approximation an orbit of f. Moreover, one can estimate the distortion of the curves γ_i and show that they are close to segments in the z-direction. Using that the skew-product f is uniquely ergodic, together with estimates on the distortion, we can hence show that $f_i^\Phi(I)$, which is the union of the curves γ_i, becomes equidistributed and hence prove the mixing estimate (20).

4.3. Stretching and discrete number of iterations. In the following sections we will denote by $\overline{\Phi}$ and $\underline{\Phi}$ respectively the maximum and the minimum of Φ on $\mathbb{T}^2$. By assumption $\underline{\Phi} > 0$. We will need later the following simple estimate on the variation of the discrete number of iterations $n_t(x,y)$ on a fiber interval $I = \{x\}\times[a,b]$ in terms of the stretch on I.

Lemma 7. *Let $I = \{x\} \times [a, b]$. Let us denote by*

$$\underline{n}_t(I) := \min_{a \le y \le b} n_t(x, y), \qquad \overline{n}_t(I) := \max_{a \le y \le b} n_t(x, y).$$

We have

$$\frac{\Delta\Phi_{\underline{n}_t(I)}(I)}{\overline{\Phi}} - \frac{\overline{\Phi}}{\underline{\Phi}} \le \overline{n}_t(I) - \underline{n}_t(I) \le \frac{\Delta\Phi_{\underline{n}_t(I)}(I)}{\underline{\Phi}} + \frac{\overline{\Phi}}{\underline{\Phi}}. \tag{21}$$

Clearly (21) is meaningful when the stretch $\Delta\Phi_{\underline{n}_t(x,b)}(I)$ is large and hence shows that in this case also the variation of $n_t(x, y)$ on I is large.

Proof. Let us write for brevity of notation $\overline{n}_t := \overline{n}_t(I)$ and $\underline{n}_t := \underline{n}_t(I)$. Let $\underline{y}, \overline{y} \in [a, b]$ be such that respectively $n_t(x, \underline{y}) = \underline{n}_t$ and $n_t(x, \overline{y}) = \overline{n}_t$. Remark that by definition of special flow we then have

$$t - \overline{\Phi} < \Phi_{\underline{n}_t}(x, \underline{y}), \Phi_{\overline{n}_t}(x, \overline{y}) \le t. \tag{22}$$

Writing $\Phi_{\overline{n}_t}(x, \overline{y}) = \Phi_{\underline{n}_t}(x, \overline{y}) + \Phi_{\overline{n}_t - \underline{n}_t}(f^{\underline{n}_t}(x, \overline{y}))$ and using (22) and the trivial estimate $\Phi_n \ge n\underline{\Phi}$, we have

$$\begin{aligned} (\overline{n}_t - \underline{n}_t)\underline{\Phi} &\le \Phi_{\overline{n}_t - \underline{n}_t}(f^{\underline{n}_t}(x, \overline{y})) \\ &= \Phi_{\overline{n}_t}(x, \overline{y}) \pm \Phi_{\underline{n}_t}(x, \underline{y}) - \Phi_{\underline{n}_t}(x, \overline{y}) \\ &\le t - (t - \overline{\Phi}) + \Delta\Phi_{\underline{n}_t}(I), \end{aligned} \tag{23}$$

where the latter inequality holds since by definition

$$\Delta\Phi_{\underline{n}_t}(I) \ge \Phi_{\underline{n}_t}(x, \underline{y}) - \Phi_{\underline{n}_t}(x, \overline{y}).$$

This proves the upper bound in (21). To prove the lower bound, let $y_n, y_M \in [a, b]$ such that

$$\Phi_{\underline{n}_t}(x, y_M) = \max_{a \le y \le b} \Phi_{\underline{n}_t}(x, y) \quad \text{and} \quad \Phi_{\underline{n}_t}(x, y_m) = \min_{a \le y \le b} \Phi_{\underline{n}_t}(x, y).$$

Reasoning as in (23), since $\Phi_{\underline{n}_t}(x, y_M) \le \Phi_{n_t(x, y_M)}(x, y_M) \le t$, we get

$$\begin{aligned} (\overline{n}_t - \underline{n}_t)\overline{\Phi} \ge (n_t(x, y_m) - \underline{n}_t)\overline{\Phi} &\ge \Phi_{n_t(x, y_m) - \underline{n}_t}(f^{\underline{n}_t}(x, y_m)) \\ &= \Phi_{n_t(x, y_m)}(x, y_m) \pm \Phi_{\underline{n}_t}(x, y_M) - \Phi_{\underline{n}_t}(x, y_m) \\ &\ge (t - \overline{\Phi}) - t + \Delta\Phi_{\underline{n}_t}(I). \end{aligned} \tag{24}$$

This concludes the proof of (21). q.e.d.

4.4. From discrete to continuous time stretching. Theorem 6 shows that $|\phi_n|$ (and hence the stretch on y-fibers) grows (in measure) as $n \to \infty$. To prove that f^Φ is mixing, we need to show that the stretch grows as t tends to infinity. The following lemma is used to make this connection between discrete and continuous time. In its proof we use the fact that the roof function belongs to the class $\mathcal{R}$ introduced in Definition 1.

We recall that, for a given $t > 0$, $n_t(x, y)$ is the maximum $n \in \mathbb{N}$ such that $\Phi_n(x, y) \le t$. For each $x \in \mathbb{T}$, let $\underline{n}_t(x) = \min_{y\in\mathbb{T}} n_t(x, y)$. For each $C > 0$, let

$$X(t, C) := \{x \mid \quad \exists\, y_x \text{ s.t. } |\phi_{\underline{n}_t(x)}(x, y_x)| > C\ \}. \tag{25}$$

Lemma 8. *Let $\Phi \in \mathcal{R}$. For each $C > 1$,*

$$\lim_{t\to\infty} \mathrm{Leb}_x\{\mathbb{T}\backslash X(t, C)\} = 0. \tag{26}$$

Proof. We will prove the lemma by contradiction. Remark that if $x \notin X(t, C)$, then for all $y \in \mathbb{T}$, $|\phi_{\underline{n}_t(x)}(x, y)| \le C$. Thus, if (26) does not hold, there exists $C > 0$, $\delta > 0$ and a subsequence $t_k \to \infty$ as $k \to \infty$ such that for all $k \in \mathbb{N}$

$$\mathrm{Leb}_2\{(x, y) \in \mathbb{T}^2 : |\phi_{\underline{n}_{t_k}(x)}(x, y)| \le C\} \ge \mathrm{Leb}\{\mathbb{T}\backslash X(t, C)\} \ge \delta. \tag{27}$$

Let us show that in this case $n_{t_k}(x)$ as $x \in \mathbb{T}\backslash X(t_k, C)$ assumes a finite number of values uniformly bounded in k. Since $\phi^\perp$ is a trigonometric polynomial by Definition 1, one can easily see using Fourier analysis that $\phi^\perp - \int \Phi(x, y)\,\mathrm{d}x\,\mathrm{d}y$ is a coboundary, i.e. there exists a g such that $\phi^\perp(x) = g(x + \alpha) - g(x) + \int \Phi(x, y)\,\mathrm{d}x\,\mathrm{d}y$, and moreover g is also a trigonometric polynomial.

For a fixed $t > 0$, let $\underline{y}(x)$ be such that $\underline{n}_t(x) = n_t(x, \underline{y}(x))$. From the definition of special flow, we have that $t - \Phi(f^{\underline{n}_t(x)-1}(x, \underline{y}(x))) < \Phi_{\underline{n}_t(x)}(x, \underline{y}(x)) \le t$. Moreover, by the decomposition $\Phi = \phi + \phi^\perp$, using that $\phi^\perp - \int \Phi(x, y)\,\mathrm{d}x\,\mathrm{d}y$ is a coboundary and $\int \Phi\,\mathrm{d}x\,\mathrm{d}y = 1$, we have

$$\Phi_{\underline{n}_t(x)}(x, \underline{y}(x)) = \phi_{\underline{n}_t(x)}(x, \underline{y}(x)) + \underline{n}_t(x) + g(x + \underline{n}_t(x)\alpha) - g(x, y).$$

So, denoting by $\overline{g} = \max g$ (well defined since here g is a trigonometric polynomial and hence continuous) and by $\overline{\Phi} = \max \Phi(x, y)$, if $t = t_k$ for some k and $x \notin X(t_k, C)$, we have $t_k - C - 2\overline{g} - \overline{\Phi} \le \underline{n}_{t_k}(x) \le t_k + C + 2\overline{g}$. This shows that there exists $N > 0$ independent on k, such that $\left|\underline{n}_{t_k}(x) - t_k\right| \le N$ for all $x \notin X(t_k, C)$. From this, recalling (27), we can find for each k some $n_k \in \mathbb{N}$ such that $n_k = \underline{n}_{t_k}(x)$ for some $x \notin X(t_k, C)$ and $\mathrm{Leb}_2\{(x, y) \in \mathbb{T}^2 : |\phi_{n_k}(x, y)| < C\} \ge \delta/N$. Since $\min_{x\in\mathbb{T}} n_{t_k}(x) \ge t_k/\overline{\Phi}$, $n_k \to \infty$ as $k \to \infty$ and this shows that $\mathrm{Leb}_2\{(x, y) \in \mathbb{T}^2 : |\phi_n(x, y)| < C\}$ does not converge to zero as $n \to \infty$. On the other side, ϕ is not a coboundary by Definition 2, so we reached a contradiction with Theorem 6. We conclude that (26) holds. q.e.d.

4.5. Choice of parameters. Let $Q = [x_1, x_2] \times [y_1, y_2] \times [0, h]$ be a given cube. Given $\epsilon, \delta > 0$, let us define the sets $X(t)$ and the partial partitions $\xi(x, t)$ which satisfy the conclusion in Lemma 6. Let us first fix parameters $\delta_0, \epsilon_0, N_0, C_0, t_0$ as follows. The reader can skip these definitions at first (their use will become clear during the proofs).

Let us denote by

$$\Phi'(x,y) := \frac{\partial \Phi(x,y)}{\partial y}, \qquad \Phi''(x,y) := \frac{\partial^2 \Phi(x,y)}{\partial y^2}.$$

(P1) Choose $0 < \delta_0 < 1$ such that $(2d+1)\delta_0 + B_d \delta_0^{b_d} \leq \delta$, where d is the degree of $y \mapsto \Phi(x,y)$ and B_d, b_d are as in Lemma 4.

(P2) Choose $\epsilon_0 > 0$ such that $\epsilon_0 < \min\{\frac{\delta_0}{4d}, \underline{\Phi}, \frac{1}{2}\}$ and $(1-\epsilon_0)^5 \leq (1-\epsilon)$.

(P3) Let χ be a continuous function such that

$$\chi \equiv \begin{cases} 1 & \text{on } [x_1,x_2] \times [y_1, y_2 - \epsilon_0(y_2-y_1)]; \\ 0 & \text{outside } [x_1,x_2] \times \left[y_1, y_2 - \frac{\epsilon_0}{2}(y_2-y_1)\right]. \end{cases}$$

Let us remark that Φ' and Φ'' have zero average on $\mathbb{T}^2$ while Φ has integral equal to 1. Since f is uniquely ergodic and ergodic sums of continuous functions over a uniquely ergodic transformation converge uniformly (see [**4**]), there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ and for all $(x,y) \in \mathbb{T}^2$ all the following bounds hold simultaneously:

(P3a) $|\Phi'_n(x,y)| \leq \epsilon_0 n$;

(P3b) $|\Phi''_n(x,y)| \leq \epsilon_0 n$;

(P3c) $|\frac{\Phi_n(x,y)}{n} - 1| \leq \frac{\epsilon_0}{1+\epsilon_0}$;

(P3d) $\left|\frac{\chi_n(x,y)}{n} - \int_{\mathbb{T}^2} \chi\right| < \epsilon_0$.

(P4) Let $\overline{|\Phi'|}$ and $\overline{|\Phi''|}$ denote respectively the maximum of $|\Phi'|$ and $|\Phi''|$ on $\mathbb{T}^2$ and choose $C_0 > 0$ such that

$$C_0 > \max\left\{ \frac{16d\overline{\Phi}}{\delta_0^2 \underline{\Phi}}, \frac{dN_0\overline{|\Phi'|}}{\delta_0}, \frac{dN_0\overline{|\Phi''|}}{\delta_0}, \frac{d(N_0+1)^2}{\delta_0^2}, \frac{d^2\pi^2 \max\{\overline{|\Phi'|}, 1\}^2}{\epsilon_0^{10}\underline{\Phi}^2(y_2-y_1)^2} \right\}.$$

(P5) Let $X(t, C_0)$ be as in (25) in § 4.4. By Lemma 8, we can choose t_0 such that for each $t \geq t_0$ we have $\mathrm{Leb}_x(\mathbb{T} \setminus X(t, C_0)) < \delta_0$.

4.6. Definition of $X(t)$ and preliminary y-fibers partitions. Fix any $t \geq t_0$ where t_0 is as in (P5) in § 4.5. For brevity we will denote by $X(t) := X(t, C_0)$, where C_0 is the constant defined in (P4) in § 4.5. For any $x \in X(t)$, let us define a preliminary partition $\xi_1(x,t)$ of the fiber $\{x\} \times \mathbb{T}$ into intervals. We will later refine $\xi_1(x,t)$ to obtain a partition $\xi(x,t)$ with the properties in Lemma 6.

In what follows we will abuse the notation and denote by $\xi(x,t)$ both a partial partition of $\mathbb{T}$ into intervals and the corresponding support, that is, the set of $y \in \mathbb{T}$ which belongs to some interval of $\xi_1(x,t)$. Let us write

$$\phi_{\underline{n}_t(x)}(x,y) = \mathrm{Re} \sum_{k=1}^{d} c_k(x) e^{2\pi iky}, \quad \frac{\partial}{\partial y}\phi_{\underline{n}_t(x)}(x,y) = \mathrm{Re} \sum_{k=1}^{d} c'_k(x) e^{2\pi iky},$$

where $c_k'(x) = 2\pi i k c_k(x)$. Let δ_0 be as in (P1) in § 4.5 and let us define

$$\xi_0(x,t) := \left\{ y \in \mathbb{T} : |\phi'_{\underline{n}_t(x)}(x,y)| \geq \delta_0 \max_{1\leq k\leq d} |c_k'(x)| \right\}. \tag{28}$$

Clearly $\xi_0(x,t)$ is a union of intervals. Let $\xi_1(x,t)$ be the partial partition of $\{x\} \times \mathbb{T}$ obtained by discarding from $\xi_0(x,t)$ all intervals which have length less than δ_0.

For brevity, we will denote in the following sections

$$\phi'(x,y) := \partial\phi(x,y)/\partial y \quad \text{and} \quad \phi''(x,y) := \partial^2\phi(x,y)/\partial y^2.$$

Lemma 9. *The following identities hold:*

$$\begin{aligned} \Phi_n'(x,y) &:= \frac{\partial}{\partial y}\Phi_n(x,y) = \phi_n'(x,y), \\ \Phi_n''(x,y) &:= \frac{\partial^2}{\partial y^2}\Phi_n(x,y) = \phi_n''(x,y). \end{aligned}$$

Proof. The identity $\frac{\partial}{\partial y}(\Phi_n) = (\frac{\partial}{\partial y}\Phi)_n$ holds since f is a skew-product of the form (5). In fact, it is clear from (5) that f commutes with all translations in the y coordinate, and hence it commutes with the derivative $\partial/\partial y$. By definition, the function

$$\Phi_n(x,y) - \phi_n(x,y) = \sum_{k=0}^{n-1} \int \Phi(x + k\alpha, y)\,\mathrm{d}y$$

depends only on $x \in \mathbb{T}$, hence $\phi_n(x,y_1) - \phi_n(x,y_2) = \Phi_n(x,y_1) - \Phi_n(x,y_2)$, for any $n \in \mathbb{N}$ and for any $x, y_1, y_2 \in \mathbb{T}$. The lemma follows. q.e.d.

The following lemma shows that for any interval $[a,b] \in \xi_1(x,t)$ both the derivative $\Phi'_{\underline{n}_t(x)}$ at any point $y \in [a,b]$ and the stretch are large and are of the same order.

Lemma 10. *The partial partitions $\xi_1(x,t)$, $x \in X(t)$, are such that*

$$\mathrm{Leb}_2\left(\mathbb{T}^2 \backslash \left(\cup_{x\in X(t)} \cup_{I\in\xi_1(x,t)} I\right)\right) \leq \delta \tag{29}$$

and for all $x \in X(t)$,

$$|\Phi'_{\underline{n}_t(x)}(x,y)| \geq \frac{2\pi\delta_0}{d} C_0, \quad \text{for all } y \in \xi_1(x,t); \tag{30}$$

$$|\Phi'_{\underline{n}_t(x)}(x,y_1)| \geq \frac{\delta_0}{d} |\Phi'_{\underline{n}_t(x)}(x,y_2)|, \quad \text{for all } y_1, y_2 \in \xi_1(x,t). \tag{31}$$

Moreover, for each $I = \{x\} \times [a,b]$ with $[a,b] \in \xi_1(x,t)$, we have

$$\Delta\Phi_{\underline{n}_t(x)}(I) \geq \frac{2\pi\delta_0^2}{d} C_0, \tag{32}$$

$$\frac{\delta_0^2}{d} \max_{y\in\xi_1(x,t)} |\Phi'_{\underline{n}_t(x)}(x,y)| \leq \Delta\Phi_{\underline{n}_t(x)}(I) \leq \frac{d}{\delta_0} \min_{y\in\xi_1(x,t)} |\Phi'_{\underline{n}_t(x)}(x,y)|, \tag{33}$$

$$\max_{a\leq y\leq b} |\Phi''_{\underline{n}_t(x)}(x,y)| \leq \frac{2\pi d^2}{\delta_0^2} \Delta\Phi_{\underline{n}_t(x)}(I). \tag{34}$$

Proof. Let us first prove the estimate on the total measure (29). By applying Lemma 4 (with the norm $\|\mathbf{c}\|_d = \max_{0\leq|k|\leq d} |c_k|$) to the trigonometric polynomial

$$\frac{\phi'_{\underline{n}_t(x)}}{\max_{1\leq k\leq d} |c'_k(x)|} = \frac{\sum_{0<|k|\leq d} c'_k(x) e^{2\pi i k y}}{\max_{1\leq |k|\leq d} |c'_k(x)|}, \quad \text{where } c_{-k,n}(x) := \overline{c_{k,n}}(x),$$

we get

$$\mathrm{Leb}_x((\{x\}\times\mathbb{T}) \setminus \xi_0(x,t)) \leq B_d \delta_0^{b_d}.$$

Since the solutions of $\phi'_{\underline{n}_t(x)}/\max_k |c'_k(x)| = \pm\delta_0$ are a subset of the zeros of a polynomial of degree at most $2d$, there are at most $2d$ solutions for each level set. Thus, $\xi_1(x,t)$ is obtained by removing at most $2d$ intervals of length smaller than δ_0 from $\xi_0(x,t)$ and for each $x \in X(t)$ we have $\mathrm{Leb}_x(\xi_1(x,t)) \geq 1 - B_d\delta_0^{b_d} - 2d\delta_0$. Applying Fubini and recalling that $\mathrm{Leb}(X(t)) \geq 1-\delta_0$ by (P5) in § 4.5, we get (29) by the choice (P1) of δ_0 in § 4.5.

Clearly, for all (x,y) we have $|\phi'_{\underline{n}_t(x)}(x,y)| \leq d\max_k |c'_k(x)|$. Thus, for $x \in X(t)$, from the definition (28) of $\xi_0(x,t) \supset \xi_1(x,t)$, we immediately have

$$\min_{y\in\xi_1(x,t)} |\phi'_{\underline{n}_t(x)}(x,y)| \geq \delta_0 \max_{y\in\xi_1(x,t)} |\phi'_{\underline{n}_t(x)}(x,y)|/d\,,$$

and hence (31) by Lemma 9. Moreover, since by definition of $X(t)$, there exists $y(x)$ such that $|\phi_{\underline{n}_t(x)}(x,y(x))| \geq C_0$, we also have $\max_k |c_k(x)| \geq C_0/d$ and since $\max_k |c'_k(x)| \geq 2\pi \max_k |c_k| \geq 2\pi C_0/d$, again from the definition of $\xi_0(x,t)$ and Lemma 9 we get $|\Phi'_{\underline{n}_t(x)}(x,y)| \geq 2\pi\delta_0 C_0/d$, concluding the proof of (30).

The estimates (32, 33) on the stretch follow simply by using mean value from (30) and (31) respectively, and from the lower estimate on the size of intervals in $\xi_1(x,t)$. The last estimate (34) is obtained by combining the trivial upper estimate $|\Phi''(x,y)| \leq 2\pi d^2 \max |c'_k(x)|$ with $\Delta\Phi_{\underline{n}_t(x)}(I) \geq \delta_0^2 \max_k |c'_k(x)|$, which follows from mean value and definition of $\xi_0(x,t)$ and $\xi_1(x,t)$. q.e.d.

Let us denote by $\overline{n}_t(x) = \max_{y\in\mathbb{T}} n_t(x,y)$. The choices of parameters in § 4.5 and Lemma 10 guarantee that derivatives and stretch are not only large for $n = \underline{n}_t(x)$, but also remain large for all further iterates up to $\overline{n}_t(x)$, as stated below.

Lemma 11. *For all $x \in X(t)$ and $I = \{x\} \times [a,b]$ with $[a,b] \in \xi_1(x,t)$, the sign of Φ'_n on I for all $\underline{n}_t(x) \le n \le \overline{n}_t(x)$ is the same, and for all $\underline{n}_t(x) \le n \le \overline{n}_t(x)$ we have:*

$$\frac{|\Phi'_{\underline{n}_t(x)}(x,y)|}{2} \le |\Phi'_n(x,y)| \le \frac{3|\Phi'_{\underline{n}_t(x)}(x,y)|}{2}, \quad \text{for all } y \in [a,b]; \tag{35}$$

$$\frac{1}{2}\frac{\delta_0}{d}\,\Delta\Phi_{\underline{n}_t(x)}(I) \le \Delta\Phi_n(I) \le \frac{3}{2}\frac{d}{\delta_0}\Delta\Phi_{\underline{n}_t(x)}(I); \tag{36}$$

$$|\Phi''_n(x,y)| \le \frac{4\pi d^2}{\delta_0^2}\,\Delta\Phi_{\underline{n}_t(x)}(I), \quad \text{for all } y \in [a,b]\ . \tag{37}$$

Proof. Fix $x \in X(t)$ and n with $\underline{n}_t(x) \le n \le \overline{n}_t(x)$. We have $\Phi'_n(x,y) = \Phi'_{\underline{n}_t(x)}(x,y) + \Phi'_{n-\underline{n}_t(x)}(x_t,y_t)$ where $(x_t,y_t) := f^{\underline{n}_t(x)}(x,y)$. Let us show that $|\Phi'_{n-\underline{n}_t(x)}(x_t,y_t)| \le |\Phi'_{\underline{n}_t(x)}(x,y)|/2$, from which it follows that (35) holds and also that $\Phi'_n(x,y)$ has constant sign for all $\underline{n}_t(x) \le n \le \overline{n}_t(x)$ and $y \in [a,b]$ (recall that $\Phi'_{\underline{n}_t(x)}$ has no zeros on I by construction).

Consider N_0 defined in (P3) in § 4.5. On one hand, if $n - \underline{n}_t(x) \le N_0$, we have

$$|\Phi'_{n-\underline{n}_t(x)}\,(x_t,y_t)\,| \le N_0|\overline{\Phi'}| \le \delta_0 C_0/d\,,$$

by choice of C_0 in (P4), § 4.5. Thus, $|\Phi'_{n-\underline{n}_t(x)}(x_t,y_t)| \le |\Phi'_{\underline{n}_t(x)}(x,y)|/2$ by (30). On the other hand, if $n - \underline{n}_t(x) \ge N_0$, by (4.5) in § 4.5, then by Lemma 7 and then by mean value, (31), and $\epsilon_0 \le 1$, we get

$$\begin{aligned} |\Phi'_{n-\underline{n}_t(x)}(x_t,y_t)| &\le \epsilon_0(n - \underline{n}_t(x)) \\ &\le \epsilon_0 \frac{\Delta\Phi_{\underline{n}_t(x)}(\{x\}\times\mathbb{T}) + \overline{\Phi}}{\underline{\Phi}} \\ &\le \epsilon_0 \frac{d}{\delta_0}|\Phi'_{\underline{n}_t(x)}(x,y)| + \frac{\overline{\Phi}}{\underline{\Phi}}. \end{aligned} \tag{38}$$

The two terms in the last expression are both less than $|\Phi'_{\underline{n}_t(x)}(x,y)|/4$, the first by choice of ϵ_0 in (P2) in § 4.5 and the second using (30) and $\pi\delta_0 C_0/2d \ge \overline{\Phi}/\underline{\Phi}$, which follows by choice of C_0 in (P4) in § 4.5 remarking that $2\delta_0 < \pi$ since by choice $\delta_0 < 1$. This concludes the proof of (35).

The estimate (36) follows from (35) by mean value and (31). To get (37), separating as before the cases $n - \underline{n}_t(x) \le N_0$ and $n - \underline{n}_t(x) \ge N_0$ and, in the second case, using (4.5) in § 4.5 and reasoning as in the proof of (38), we have

$$|\Phi''_n(x,y)| \le |\Phi''_{\underline{n}_t}(x,y)| + \max\left\{N_0|\overline{\Phi''}|, \frac{\epsilon_0}{\underline{\Phi}}\Delta\Phi_{\underline{n}_t(x)}(\{x\}\times\mathbb{T}) + \frac{\overline{\Phi}}{\underline{\Phi}}\right\}.$$

The final estimate (37) follows from here estimating $|\Phi''_{\underline{n}_t}(x,y)|$ by (34) and controlling the first term in the maximum by using mean value, (33) and $\epsilon_0 \le \underline{\Phi} \le \pi d \underline{\Phi}$ (recall the choice of ϵ_0 in (P2) in § 4.5) and estimating the second term in the maximum by (32) and the choice of C_0 in (P4) in § 4.5, which guarantees that $2\pi d^2 \Delta\Phi_{\underline{n}_t(x)}/\delta_0^2 \geq 4\pi^2 dC_0 \geq C_0 \geq \overline{\Phi}/\underline{\Phi}$. q.e.d.

4.7. Fibers partitions. Let us refine the partial partitions $\xi_1(x,t)$ so that we can prove that intervals of the refined partition become equidistributed. Let us fix $x \in X(t)$ and $I = \{x\}\times[a,b]$ with $[a,b] \in \xi_1(x,t)$. Let us recall that we denote by $\underline{n}_t(I) = \min_{a\leq y\leq b} n_t(x,y)$ and by $\overline{n}_t(I) = \max_{a\leq y\leq b} n_t(x,y)$ and let $\Delta n_t(I) := \overline{n}_t(I) - \underline{n}_t(I) + 1$. The previous construction guarantees the following properties.

Lemma 12. *For each $x \in X(t)$ and each $[a,b] \in \xi_1(x,t)$, the function $y \mapsto n_t(x,y)$ is monotone on $[a,b]$ and $\Delta n_t(I) \geq \pi\delta_0^2 C_0/d\underline{\Phi}$.*

Proof. By Lemma 11, the sign of Φ_n' on $I := \{x\} \times [a,b]$ is the same for all $\underline{n}_t(I) \leq n \leq \overline{n}_t(I)$. Let us assume that $\Phi'_{\underline{n}_t(I)} < 0$ so that Φ_n is monotonically decreasing on I for all $\underline{n}_t(I) \leq n \leq \overline{n}_t(I)$ (the construction is analogous in the other case), and let us show that this implies that $y \mapsto n_t(x,y)$ is increasing on $[a,b]$. If $a \leq y_1 < y_2 \leq b$, we have $\Phi_{n_t(x,y_1)}(x,y_2) < \Phi_{n_t(x,y_1)}(x,y_1) \leq t$ by monotonicity and definition of $n_t(x,y_1)$. Thus, by definition of $n_t(x,y_2)$ this shows $n_t(x,y_2) \geq n_t(x,y_1)$. From Lemma 7, $\overline{n}_t(I) - \underline{n}_t(I) \geq \Delta\Phi_{\underline{n}_t(I)}(I)/\underline{\Phi} - \overline{\Phi}/\underline{\Phi}$ and by (32) we have $\Delta\Phi_{\underline{n}_t(I)}(I)/\underline{\Phi} \geq 2\pi\delta_0^2 C_0/d\underline{\Phi}$. This gives the desired estimate for $\Delta n_t(I)$ since $\overline{\Phi}/\underline{\Phi} \leq \pi\delta_0^2 C_0/d\underline{\Phi}$ by choice of C_0 in (P4) § 4.5. q.e.d.

From Lemma 12, we know that I can be subdivided into exactly $\Delta n_t(I)$ maximal intervals on which $y \mapsto n_t(x,y)$ is locally constant. Let us assume without loss of generality that $\Phi'_{\underline{n}_t(I)} < 0$. In this case, more precisely, for each $1 \leq j \leq \overline{n}_t(I) - \underline{n}_t(I)$ there is a unique $y_j \in [a,b]$ such that $\Phi_{\underline{n}_t(I)+j}(x,y_j) = t$ and moreover $y_j < y_{j+1}$ for all $1 \leq j \leq \overline{n}_t(I) - \underline{n}_t(I)$. Thus, setting $y_0 := a$ and $y_{\overline{n}_t(I)-\underline{n}_t(I)+1} := b$, for each $0 \leq j \leq \overline{n}_t(I) - \underline{n}_t(I)$ the interval $[y_j, y_{j+1})$ is a maximal interval on which $n_t(x,y)$ is equal to $\underline{n}_t(I) + j$. (Notice that if $\Phi'_{\underline{n}_t(I)} > 0$, the continuity intervals have the form $(y_j, y_{j+1}]$.)

Let $N_t(I) := [\sqrt{\Delta n_t(I)}]$, where $[z]$ denotes the integer part of z. Let us group the intervals (y_j, y_{j+1}) into $N_t(I)+1$ groups, each of the first $N_t(I)$ made by exactly $N_t(I)$ consecutive intervals, the last by the remaining ones, which are at most $2\sqrt{\Delta n_t(I)}$. In this way we obtain a subdivision of the interval I of the partition $\xi_1(x,t)$ into intervals of the form $(y_{kN_t(I)}, y_{(k+1)N_t(I)-1})$ for $k = 0, \dots, N_t(I) - 2$ or, in the case of the last interval, of the form $(y_{N_t(I)(N_t(I)-1)}, b)$.

Let $\xi(x,t)$ be the partition obtained refining $\xi_1(x,t)$ by repeating the above subdivision for each interval $I \in \xi_1(x,t)$. The elements $J \in \xi(x,t)$ have the following properties, used in the following section (§ 4.8) to prove the equidistribution estimate (20).

Lemma 13 (Properties of fiber partitions)**.** *For each interval $J = \{x\}\times(y', y'')$ with $x \in X(t)$ and $(y', y'') \in \xi(x, t)$, the following properties hold.*

$$\left|\frac{\Delta n_t(J)}{\Delta\Phi_{\underline{n}_t(J)}(J)} - 1\right| \leq \epsilon_0, \quad \text{where}\, \Delta n_t(J) = \overline{n}_y(J) - \underline{n}_y(J) + 1; \tag{39}$$

$$\frac{1}{\Delta n_t(J)} \sum_{n=\underline{n}_y(J)}^{\overline{n}_y(J)} \chi\left(f^n(x, y)\right) \geq (1-\epsilon_0)^2 (x_2 - x_1)(y_2 - y_1); \tag{40}$$

$$\mathrm{Leb}_x(J) = |y'' - y'| \leq \min\left\{\frac{\epsilon_0(y_2 - y_1)}{2}, \frac{\epsilon_0}{2|\overline{\Phi'}|}, \frac{\delta_0^3}{4\pi d^3}\epsilon_0\right\}; \tag{41}$$

$$\left|\frac{\Delta\Phi_{\underline{n}_t(J)}(J)}{\Delta\Phi_n(J)} - 1\right| \leq \epsilon_0, \quad n = \underline{n}_t(J), \ldots, \overline{n}_t(J). \tag{42}$$

Moreover, if for $h > 0$ and $\underline{n}_t(J) \leq n \leq \overline{n}_t(J)$ we denote by

$$J_n^h := \{y \in (y', y'') : t - h < \Phi_n(x, y) \leq t\}, \tag{43}$$

we also have

$$\left|\frac{\Delta\Phi_n(J)\mathrm{Leb}_x(J_j^h)}{(y'' - y')h} - 1\right| \leq \epsilon_0, \qquad n = \underline{n}_t(J), \ldots, \overline{n}_t(J). \tag{44}$$

Proof. Let $J = \{x\} \times (y', y'')$ with $(y', y'') \in \xi(x, t)$ and let $I = \{x\} \times (a, b)$ with $(a, b) \in \xi_1(x, t)$ be such that $J \subset I$. Without loss of generality, let us assume that $y \mapsto n_t(x, y)$ is increasing on (y', y''). We will assume that J does not contain either of the endpoints of I. The proof in the latter case requires minor adjustments to take care of the intervals where $n_t(x, y) = \underline{n}_t(I)$ or $\overline{n}_t(I)$, which we leave to the reader. Let us remark that in the case considered, the values assumed by $n_t(x, y)$ on J, which by definition are $\Delta n_t(J)$, are by construction exactly equal to $N_t(I)$.

Since $y \mapsto n_t(x, y)$ is increasing, we have that $\Phi_{\underline{n}_t(J)}$ is decreasing and that $\underline{n}_t(J) = n_t(x, y')$ and $\overline{n}_t(J) = n_t(x, y'') - 1$, so that $\Phi_{\underline{n}_t(J)}(x, y') = t = \Phi_{\overline{n}_t(J)+1}(x, y'')$. Thus,

$$\begin{aligned}
\Delta\Phi_{\underline{n}_t(J)}(J) &= \Phi_{\underline{n}_t(J)}(x, y') - \Phi_{\underline{n}_t(J)}(x, y'') \\
&= \Phi_{\overline{n}_t(J)+1}(x, y'') - \Phi_{\underline{n}_t(J)}(x, y'') \\
&= \Phi_{\Delta n_t(J)}(f^{\underline{n}_t(J)}(x, y'')).
\end{aligned} \tag{45}$$

This shows that (39) follows from (4.5) in § 4.5, which can be applied since $\Delta n_t(J) = N_t(I) \geq \sqrt{\Delta n_t(I)} - 1 \geq N_0$ by Lemma 12 and by the inequality $\pi\delta_0^2 C_0/d\underline{\Phi} \geq N_0 + 1$, which holds by choice of C_0 in condition (P4) in § 4.5. For the same reason, condition (4.5) in § 4.5

also holds and gives (40) by remarking that, by definition of χ (see (P3), § 4.5),

$$\int \chi - \epsilon_0 \geq \int \chi(1-\epsilon_0) \geq (x_2-x_1)(y_2-y_1)(1-\epsilon_0)^2\,.$$

To estimate the size of J, let us remark that by mean value

$$\Delta\Phi_{\underline{n}_t(J)}(J) = |y''-y'||\Phi'_{\underline{n}_t(J)}(x,\tilde{y})|\,, \qquad \text{for some } \tilde{y} \in (y',y'')\,.$$

By (39) (since $\Delta n_t(J) = N_t(I)$ by construction and $\epsilon_0 < 1/2$), we have $\Delta\Phi_{\underline{n}_t(J)}(J) \leq 2N_t(I)$. Since moreover $|\Phi'_{\underline{n}_t(J)}(x,\tilde{y})| \geq \delta_0 \Delta\Phi_{\underline{n}_t(I)}(I)/2d$ by (35) and (33), we get $|y''-y'| \leq \frac{4d}{\delta_0}\frac{N_t(I)}{\Delta\Phi_{\underline{n}_t(I)}(I)}$. From Lemma 7 and the definition of $N_t(I)$, we have $\Delta\Phi_{\underline{n}_t(I)} \geq N_t(I)^2\underline{\Phi} - \overline{\Phi} - \underline{\Phi}$. Thus, since using Lemma 12, $\underline{\Phi} \leq 1$, and the definition of C_0 (see (P4) in § 4.5) we have $N_t(I) \geq \frac{\delta_0\sqrt{C_0}}{\sqrt{d}} - 1 \geq \frac{\delta_0\sqrt{C_0}}{2\sqrt{d}}$ and $N_t(I)^2\underline{\Phi} \geq 4\overline{\Phi}$, we get the estimate

$$|y'-y''| \leq \frac{4d\ N_t(I)}{\delta_0(N_t(I)^2\underline{\Phi} - 2\overline{\Phi})} \leq \frac{8d}{\delta_0 N_t(I)\underline{\Phi}} \leq \frac{16d\sqrt{d}}{\delta_0^2\sqrt{C_0}\underline{\Phi}}.$$

From here, one can check that by choice of C_0 and ϵ_0 in (P2), (P3c) in § 4.5, we have $|y''-y'| \leq \epsilon_0^4(y_2-y_1)/\pi\max\{\overline{|\Phi'|},1\}$ and thus that $|y''-y'|$ satisfies (41).

To prove the estimate (42), let us remark that, using the definition of stretch, the cocycle properties of ergodic sums, and then mean value, we can write $|\Delta\Phi_{\underline{n}_t(J)}(J) - \Delta\Phi_n(J)| \ \leq |\Phi'_{n-\underline{n}_t(J)}(x,\tilde{y})|\ |y''-y'|$ for some $\tilde{y} \in (y',y'')$. Thus, since $n - \underline{n}_t(J) \leq N_t(I)$, we get

$$\frac{|\Delta\Phi_{\underline{n}_t(J)}(J) - \Delta\Phi_n(J)|}{\Delta\Phi_n(J)} \leq \frac{N_t(I)\overline{|\Phi'|}|y''-y'|}{\Delta\Phi_n(J)},$$

which, using that $\frac{N_t(I)}{\Delta\Phi_n(J)} \leq 2$ by (39), is less than ϵ_0 by (41).

Let us finally prove (44). Remark that J_n^h is an interval since Φ_n is monotone by Lemma 11. Since by mean value theorem there exists $\eta_1,\eta_2 \in (y',y'')$ such that $h = |\Phi'_n(x,\eta_1)|\mathrm{Leb}_x(J_n^h)$ and $\Delta\Phi_n(J) = |\Phi'_n(x,\eta_2)||y''-y'|$, (44) follows if we prove that $\left|\frac{|\Phi'_n(x,\eta_2)|}{|\Phi'_n(x,\eta_1)|} - 1\right| \leq \epsilon_0$. Let us show that this holds by showing that

$$\max_{y'\leq y\leq y''}|\Phi''_n(x,y)||y''-y'| \leq \epsilon_0 \min_{y'\leq y\leq y''}|\Phi'_n(x,y)|.$$

This follows from (41), since by (37) we have that

$$\max_{y'\leq y\leq y''}|\Phi''_n(x,y)| \leq \frac{4\pi d^2}{\delta_0^2}\,\Delta\Phi_{\underline{n}_t(x)}(I)$$

and by (35) and (33) we have that

$$\min_{y'\leq y\leq y''}|\Phi'_n(x,y)| \geq \frac{\delta_0}{2d}\Delta\Phi_{\underline{n}_t(x)}(I)\,.$$

q.e.d.

4.8. Final equidistribution estimates. Let us use the properties in Lemma 13 to show that, for each $t \geq t_0$, each $J = \{x\} \times (y', y'')$ with $x \in X(t)$ and $(y', y'') \in \xi(x,t)$ verifies the equidistribution estimate (20) in Lemma 6.

Let us first prove that, if $Q = [x_1, x_2] \times [y_1, y_2] \times [0, h]$ is the cube fixed at the beginning of § 4.5 and Leb_x in the LHS denotes as before the 1-dimensional Lebesgue on the fiber $\{x\} \times \mathbb{T}$, we have

$$\mathrm{Leb}_x(\{x\} \times [y', y''] \cap f^{\Phi}_{-t} Q) \geq \sum_{n=\underline{n}_t(J)}^{\overline{n}_t(J)} \chi\left(f^n(x, y')\right) \mathrm{Leb}_x\left(J^h_n\right), \tag{46}$$

where J^h_n was defined in (43) and χ is the smoothened characteristic function of the base of Q defined in (P3), § 4.5. Let us remark that by definition of J^h_n, if $y \in J^h_n$, then $\Phi_n(x, y) \leq t$ but $\Phi_{n+1}(x, y) < t$ since $h < \underline{\Phi}$, so $n_t(x, y) = n$ and $\{x\} \times J^h_n$ is contained in $n_t(x, y) = n$. Thus, the intervals J^h_n, $\underline{n}_t(J) \leq n \leq \overline{n}_t(J)$, are all disjoint. Hence, to prove (46), it is enough to show that if $\chi\left(f^n(x, y')\right) > 0$, then $\{x\} \times J^h_n \subset \{x\} \times [y', y''] \cap f^{\Phi}_{-t} Q$. If $y \in J^h_n$, since as we remarked $n_t(x, y) = n$, by definition of special flow (7) we have $f^{\Phi}_t\left((x, y), 0\right) = (f_n(x, y), t - \Phi_n(x, y))$ with $0 \leq t - \Phi_n(x, y) < h$ by definition of J^h_n. If $\chi(f^n(x, y')) > 0$, by definition of χ (see (P3) in § 4.5), $f^n(x, y') \in [x_1, x_2] \times [y_1, y_2 - \frac{\epsilon_0}{2}(y_2 - y_1)]$. Since $|y' - y| \leq \epsilon_0(y_2 - y_1)/2$ by (41) and f^n preserves y-fibers and distances between points in y-fibers, we also have $f^n(x, y) \in [x_1, x_2] \times [y_1, y_2]$. This shows that $f^{\Phi}_t(\{x\} \times J^h_n) \subset Q$ and concludes the proof of (46).

Let us now estimate the RHS of (46). For $t \geq t_0$, using (39, 42, 44) and then (40), we get

$$\begin{aligned}
&\sum_{n=\underline{n}_t(J)}^{\overline{n}_t(J)} \chi\left(f^n(x, y')\right) \mathrm{Leb}_x\left(J^h_n\right) \\
&= \sum_{\underline{n}_t(J)}^{\overline{n}_t(J)} \frac{\chi\left(f^n(x, y')\right) \frac{\Delta n_t(J)}{\Delta \Phi_{\underline{n}_t(J)}(J)} \frac{\Delta \Phi_{\underline{n}_t(J)}(J)}{\Delta \Phi_n(J)} \frac{\Delta \Phi_n(J) \mathrm{Leb}_x\left(J^h_n\right)}{h|y''-y'|} h|y'' - y'|}{\Delta n_t(J)} \\
&\geq (1 - \epsilon_0)^5 h|y'' - y'|(x_2 - x_1)(y_2 - y_1) = (1 - \epsilon_0)^5 (y'' - y')\mu(Q).
\end{aligned}$$

This, together with (46), concludes the proof of (20) by choice of ϵ_0 in (P2), § 4.5. Since the partitions $\xi(x,t)$ are by construction a subdivision of the partition $\xi_1(x,t)$, we already verified in Lemma 10 that the partitions satisfy also the first assumption (19) of Lemma 6. Thus, mixing of f^{Φ} follows from Lemma 6, concluding the proof of Theorem 5.

5. Cocycle effectiveness

In this section we prove Theorem 7. We begin by recalling basic results on the cohomological equation for the skew-shift essentially due to A. Katok at the beginning of the 80's (published in [**20**], §11.6.1).

Let us consider the *cohomological equation* for a linear skew-shift of the form (8), that is, the linear difference equation

$$u \circ f - u = \Phi \,, \tag{47}$$

for a given function Φ on $\mathbb{T}^2$. By the decomposition $\Phi := \phi + \phi^\perp$ of any function $\Phi \in L^2(\mathbb{T}^2)$ into a sum of a function $\phi \in \pi^* L^2(\mathbb{T})^\perp \equiv L^2(\mathbb{T})$ and a function $\phi^\perp \in \pi^* L^2(\mathbb{T})$ (see definitions (9) and (10)), the equation can be decomposed into the cohomological equation

$$u^\perp \circ R_\alpha - u^\perp = \phi^\perp$$

for the rotation R_α of the circle of angle $\alpha \in \mathbb{R}$ and the cohomological equation (47) with a right hand side satisfying the property

$$\int_{\mathbb{T}} \Phi(x,y) dy = 0 \,, \text{ for all } x \in \mathbb{T} \,. \tag{48}$$

The space $L^2(\mathbb{T}^2)$ further decomposes into orthogonal irreducible components for the action of the skew-shift. Let $A \in GL(2,\mathbb{R})$ be the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} .$$

Let $\{e_{m,n} : (m,n) \in \mathbb{Z}^2\}$ be the standard Fourier basis of $L^2(\mathbb{T}^2)$, that is,

$$e_{m,n}(x,y) := \exp[2\pi i (mx + ny)] \,, \qquad \text{for all } (x,y) \in \mathbb{T}^2 \,.$$

Let $\mathcal{O}_A$ be the set of orbits of the action of the matrix A on $\mathbb{Z}^2$. For any $\omega \in \mathcal{O}_A$, let $H_\omega \subset L^2(\mathbb{T}^2)$ be the subspace defined as follows:

$$H_\omega = \bigoplus_{(m,n)\in\omega} \mathbb{C} e_{m,n} \,.$$

The following result is well known and easy to verify:

Lemma 14. *The space $L^2(\mathbb{T}^2)$ admits an orthogonal splitting*

$$L^2(\mathbb{T}^2) = \bigoplus_{\omega \in \mathcal{O}_A} H_\omega \,;$$

all the components H_ω are invariant under the skew-shift $f : \mathbb{T}^2 \to \mathbb{T}^2$, that is,

$$f^*(H_\omega) = H_\omega \,, \qquad \text{for all } \omega \in \mathcal{O}_A \,.$$

The existence of solutions of the cohomological equation can therefore be investigated in each irreducible component H_ω. We describe below the space $\mathcal{O}_A$ and the irreducible components H_ω, $\omega \in \mathcal{O}_A$ in more detail.

Let $(m,n) \in \mathbb{Z}^2$. If $n = 0$, the A-orbit $[(m,0)] \subset \mathbb{Z}^2$ of $(m,0)$ is reduced to a single element. The space

$$H_0 := \bigoplus_{m\in\mathbb{Z}} H_{[(m,0)]} = \pi^* L^2(\mathbb{T})$$

is the space of functions which factor through a square-integrable function on the circle. For such functions the cohomological equation is reduced to the cohomological equation for circle rotations. We are especially interested in functions in the orthogonal complement of H_0, that is, functions of zero average along the fibers of the projection $\pi : \mathbb{T}^2 \to \mathbb{T}$ (see (48)).

If $n \neq 0$, then the A-orbit $[(m,n)] \subset \mathbb{Z}^2$ of (m,n) can be described as follows:

$$[(m,n)] = \{(m+jn, n) : j \in \mathbb{Z}\}\,.$$

It follows that every A-orbit can be labeled uniquely by a pair $(m,n) \in \mathbb{Z}_{|n|} \times \mathbb{Z} \setminus \{0\}$. Let $H_{(m,n)}$ denote the corresponding factor. Let us denote by $C^\infty(H_{(m,n)})$ the subspace of smooth functions in $H_{(m,n)}$. By definition every function $\Phi \in C^\infty(H_{(m,n)})$ has a Fourier expansion of the form

$$\Phi = \sum_{j\in\mathbb{Z}} \Phi_j e_{m+jn,n}\,.$$

For every $s > 0$, let $W^s(H_{(m,n)})$ be the standard Sobolev space, that is, the completion of $C^\infty(H_{(m,n)})$ with respect to the norm:

$$\|\Phi\|_s := \left(\sum_{j\in\mathbb{Z}} (1 + (m+jn)^2 + n^2)^s |\Phi_j|^2 \right)^{1/2}\,.$$

Theorem 10. *(**[20]**, th. 11.25) There exists a unique distributional obstruction to the existence of a smooth solution $u \in C^\infty(H_{(m,n)})$ of the cohomological equation (47) with right hand side $\Phi \in C^\infty(H_{(m,n)})$. Such an obstruction is the invariant distribution $D_{(m,n)} \in W^{-s}(\mathbb{T}^2)$ for all $s > 1/2$ defined as follows:*

$$D_{(m,n)}(e_{a,b}) := \begin{cases} e^{-2\pi i[(\alpha m+\beta n)j + \alpha n\binom{j}{2}]} & \text{if } (a,b) = (m+jn, n)\,; \\ 0 & \text{otherwise.} \end{cases}$$

The solution of the cohomological equation for any $\Phi \in C^\infty(H_{(m,n)})$ such that $D_{(m,n)}(\Phi) = 0$ is given by the following formula. If $\Phi =$

$\sum_{j\in\mathbb{Z}} \Phi_j e_{m+jn,n}$, the solution $u = \sum_{j\in\mathbb{Z}} u_j e_{m+jn,n}$ is:

$$
\begin{aligned}
u_j &= -e^{2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]} \sum_{k=-\infty}^{j} \Phi_k e^{-2\pi i[(\alpha m+\beta n)k+\alpha n\binom{k}{2}]} \\
&= e^{2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]} \sum_{k=j+1}^{\infty} \Phi_k e^{-2\pi i[(\alpha m+\beta n)k+\alpha n\binom{k}{2}]} .
\end{aligned} \tag{49}
$$

If $\Phi \in W^s(H_{(m,n)})$ for any $s > 1$ and $D_{(m,n)}(\Phi) = 0$, then the above solution $u \in W^t(H_{(m,n)})$ for all $t < s - 1$ and there exists a constant $C_{s,t} > 0$ such that

$$\|u\|_t \le C_{s,t} \|\Phi\|_s .$$

The results below establish the quantitative behavior of ergodic averages for smooth functions under the skew-shift.

Lemma 15. *Let $(m, n) \in \mathbb{Z}_{|n|} \times \mathbb{Z} \setminus \{0\}$ and let $s > 1/2$. There exists a constant $C_s > 0$ such that, for any $\Phi \in W^s(H_{(m,n)})$,*

$$
\begin{aligned}
C_s^{-1}|D_{(m,n)}(\Phi)| &\le \liminf_{N\to+\infty} \frac{1}{N^{1/2}} \left\| \sum_{k=0}^{N-1} \Phi \circ f^k \right\|_{L^2(\mathbb{T}^2)} \\
&\le \limsup_{N\to+\infty} \frac{1}{N^{1/2}} \left\| \sum_{k=0}^{N-1} \Phi \circ f^k \right\|_{L^2(\mathbb{T}^2)} \le C_s |D_{(m,n)}(\Phi)| .
\end{aligned} \tag{50}
$$

Proof. Let us write the Fourier expansion of a function $\Phi \in H_{(m,n)}$ and directly compute the ergodic sums. We obtain the formula

$$\left\| \sum_{k=0}^{N-1} \Phi \circ f^k \right\|^2_{L^2(\mathbb{T}^2)} = \sum_{\ell\in\mathbb{Z}} \left| \sum_{j=\ell-N+1}^{\ell} \Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]} \right|^2$$

from which the result follows. Let us first prove the lower bound, which is the relevant one for our paper. Since $\Phi \in W^s(H_{(m,n)})$, by Hölder inequality, for all $M \in \mathbb{N} \setminus \{0\}$,

$$\left| \sum_{|j|\ge M} \Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]} \right| \le K_s \|\Phi\|_s M^{-(s-1/2)} . \tag{51}$$

It follows that there exists a constant $K'_s > 0$ such that

$$
\begin{aligned}
&\frac{1}{N} \sum_{\ell=N/4}^{N/2} \left| \sum_{j=\ell-N+1}^{\ell} \Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]} \right|^2 \\
&\qquad \ge \frac{|D_{(m,n)}(\Phi)|^2}{8} - K'_s \|\Phi\|_s^2 N^{-2(s-\frac{1}{2})} ,
\end{aligned} \tag{52}
$$

which implies the lower bound on the lower limit claimed in the statement.

As for the upper bound, it can be proved as follows. For any $\eta \in (0,1)$, the following bound can be derived from the estimate in formula (51):

$$\frac{1}{N}\sum_{\ell=N^\eta}^{N-N^\eta}\left|\sum_{j=\ell-N+1}^{\ell}\Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]}\right|^2 \leq 2|D_{(m,n)}(\Phi)|^2 + K'_s\|\Phi\|_s^2 N^{-2\eta(s-\frac{1}{2})}\,. \tag{53}$$

By applying again formula (51), we can derive the following bounds:

$$\frac{1}{N}\sum_{\ell\geq N+N^\eta}\left|\sum_{j=\ell-N+1}^{\ell}\Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]}\right|^2 \leq K'_s\|\Phi\|_s^2 N^{-2\eta(s-\frac{1}{2})}\,;$$

$$\frac{1}{N}\sum_{\ell\leq -N^\eta}\left|\sum_{j=\ell-N+1}^{\ell}\Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]}\right|^2 \leq K'_s\|\Phi\|_s^2 N^{-2\eta(s-\frac{1}{2})}\,.$$

Finally, the following estimates hold:

$$\frac{1}{N}\sum_{\ell=N-N^\eta}^{N+N^\eta}\left|\sum_{j=\ell-N+1}^{\ell}\Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]}\right|^2 \leq 2K'_s\|\Phi\|_s^2 N^{-(1-\eta)}\,;$$

$$\frac{1}{N}\sum_{\ell=-N^\eta}^{N^\eta}\left|\sum_{j=\ell-N+1}^{\ell}\Phi_j e^{-2\pi i[(\alpha m+\beta n)j+\alpha n\binom{j}{2}]}\right|^2 \leq 2K'_s\|\Phi\|_s^2 N^{-(1-\eta)}\,.$$

The upper bound on the upper limit claimed in the statement follows immediately by combining the former two estimates. q.e.d.

The uniform norm of the ergodic averages of sufficiently smooth functions can be controlled sharply along a subsequence of times. This result can be derived from classical (sharp) number theory results on Weyl sums of quadratic polynomials (see [**9**] and references therein or [**30**]) or from the results of [**11**] on the quantitative equidistribution of Heisenberg nilflows.

Theorem 11. *Let $\alpha \in \mathbb{R}\setminus\mathbb{Q}$ be any irrational number and let $s>3$. There exist a constant $M_s>0$ and a (positively) diverging sequence $\{N_\ell\}_{\ell\in\mathbb{N}}$ (depending on α) such that, for all $\Phi \in W^s(\mathbb{T}^2)\cap H_0^\perp$ and for all $(x,y)\in\mathbb{T}^2$,*

$$\frac{1}{N_\ell^{1/2}}\left|\sum_{k=0}^{N_\ell-1}\Phi\circ f^k(x,y)\right| \leq M_s\,\|\Phi\|_s\,. \tag{54}$$

Proof. Since the special flow with a constant roof function $r > 0$ of a uniquely ergodic linear skew-shift is smoothly equivalent to a uniquely ergodic Heisenberg nilflow, it is sufficient to prove the result for Heisenberg nilflow. In fact, let $\{f_t^r\}$ denote the special flow over f with roof function $r > 0$. Let $\chi \in C_0^\infty(0, r)$ be a compactly supported function of integral equal to 1 on $\mathbb{R}$. For any function $\Phi \in \mathbb{T}^2$, let $\hat{\Phi}_\chi : \mathbb{T}^2 \times [0, r] \to \mathbb{R}$ be the smooth function defined as follows:

$$\hat{\Phi}_\chi(x, y, z) = \Phi(x, y)\chi(z)\,, \qquad \text{for all } (x, y) \in \mathbb{T}^2,\ z \in [0, r]\,.$$

Since $\mathbb{T}^2 \times [0, r]$ is a fundamental domain for the quotient $\mathbb{T}^2 \times \mathbb{R}/\sim_r$, the function $\hat{\Phi}_\chi$, which vanishes at the boundary with all its derivatives, projects to a well-defined function Φ_χ on $M \approx \mathbb{T}^2 \times \mathbb{R}/\sim_r$. The function $\Phi_\chi \in W^s(M) \cap \pi^* L^2(\mathbb{T}^2)^\perp$ if and only if $\Phi \in W^s(\mathbb{T}^2) \cap H_0^\perp$. By construction, since the function $\chi \in C_0^\infty(0, r)$ has integral equal to 1 on $(0, r)$, for all $N \in \mathbb{N}$ and for all $(x, y) \in \mathbb{T}^2$, we have

$$\sum_{k=0}^{N-1} \Phi \circ f^k(x, y) = \int_0^N \Phi_\chi \circ f_t^r(x, y, 0)dt\,. \tag{55}$$

Thus the statement of the theorem can be derived from the following claim. For every $s > 3$ and for every uniquely ergodic Heisenberg nilflow $\{\phi_t^W\}$ on $M = \Gamma\backslash N$, there exist a constant $C_s > 0$ and a (positively) diverging sequence $\{T_\ell\} \subset \mathbb{R}$ such that, for all $\Psi \in W^s(M) \cap \pi^* L^2(\mathbb{T}^2)^\perp$ and for all $x \in M$,

$$\frac{1}{T_\ell^{1/2}} \left|\int_0^{T_\ell} \Psi \circ \phi_t^W(x)dt\right| \le C_s \|\Psi\|_s\,. \tag{56}$$

The above claim follows from lemma 5.5 and lemma 5.8 in [**11**]. Let $\overline{W} = (1, \alpha) \in \mathbb{R}^2$ be the projection of the generator $W \in \mathfrak{n}$ onto the abelianized Lie algebra $\mathfrak{n}/[\mathfrak{n}, \mathfrak{n}] \approx \mathbb{R}^2$. For any compact set $K \subset PSL(2, \mathbb{Z})\backslash PSL(2, \mathbb{R})$, there exists a constant $C_s := C_s(K)$ such that the bound (56) holds under the condition that $\log T_\ell \in \mathbb{R}^+$ is a return time to K of the trajectory of the point

$$PSL(2, \mathbb{Z}) \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \in PSL(2, \mathbb{Z})\backslash PSL(2, \mathbb{R})$$

under the geodesic flow on $PSL(2, \mathbb{Z})\backslash PSL(2, \mathbb{R})$ (a space isomorphic to the unit tangent bundle of the modular surface). Thus the above claim follows from the recurrence to a compact subset of all irrational points of the homogeneous space $PSL(2, \mathbb{Z})\backslash PSL(2, \mathbb{R})$ under the geodesic flow.

Let $\{T_\ell\} \subset \mathbb{R}^+$ be any diverging sequence such that the bound (56) holds. By the identity (55), the bound (54) holds for the diverging sequence $\{[T_\ell]\} \subset \mathbb{N}$. The proof of the theorem is completed. q.e.d.

Remark 12. The theory on the existence of smooth solutions of the cohomological equation outlined above generalizes to skew-shifts in any dimensions (in fact, to nilflows on any nilpotent manifold [**12**]). However, as far as we know, Theorem 11 is not established for higher dimensional skew-shifts, not even for typical rotation numbers. Bounds on ergodic averages of higher dimensional skew-shifts are closely related to bounds on Weyl sums for polynomials of degree greater than 2.

We conclude by proving Theorem 7, which states that any sufficiently smooth function $\Phi \in H_0^\perp$ is a smooth coboundary for a uniquely ergodic (irrational) skew-shift if and only if it is a measurable coboundary.

Proof of Theorem 7. Consider the sequence ergodic sums of the function $\Phi \in W^s(\mathbb{T}^2) \cap H_0^\perp$ along the sequence $\{N_l\}_{l\in\mathbb{N}}$ constructed in Theorem 11, that is,

$$\Phi_N(x,y) = \sum_{k=0}^{N-1} \Phi \circ f^k(x,y), \qquad \text{where } N \in \{N_\ell\}_{l\in\mathbb{N}}.$$

For any $\ell \in \mathbb{N}$ let $S_\epsilon^\ell \subset \mathbb{T}^2$ be the set defined as follows:

$$S_\epsilon^\ell := \{(x,y) \in \mathbb{T}^2 : |\Phi_{N_\ell}(x,y)| \geq \epsilon N_\ell^{1/2}\}\,. \tag{57}$$

Theorem 11 implies by an elementary estimate that for any $\ell \in \mathbb{N}$

$$\|\Phi_{N_\ell}\|^2_{L^2(\mathbb{T}^2)} \leq M_s^2 \|\Phi\|_s^2 \mathrm{Leb}_2(S_\epsilon^\ell) N_\ell + \epsilon^2 N_\ell (1 - \mathrm{Leb}_2(S_\epsilon^\ell))\,.$$

It follows that, if the function Φ does not belong to the kernel of all invariant distributions, by Lemma 15 there exists a constant $c_\Phi > 0$ such that

$$c_\Phi N_\ell \leq M_s^2 \|\Phi\|_s^2 \mathrm{Leb}_2(S_\epsilon^\ell) N_\ell + \epsilon^2 (1 - \mathrm{Leb}_2(S_\epsilon^\ell)) N_\ell\,,$$

hence

$$(c_\Phi - \epsilon^2) \leq (M_s^2 \|\Phi\|_s^2 - \epsilon^2) \mathrm{Leb}_2(S_\epsilon^\ell).$$

Thus, there exist $\epsilon > 0$ and $\eta(\epsilon) > 0$ such that

$$\mathrm{Leb}_2(S_\epsilon^\ell) \geq \eta_\epsilon\,, \quad \text{for all } \ell \in \mathbb{N}\,. \tag{58}$$

We conclude the argument by proving that if the lower bound (58) holds, the function Φ is not a measurable coboundary. In fact, let us assume it is and derive a contradiction. Let u be a measurable transfer function on $\mathbb{T}^2$. Since u is almost everywhere finite, there exists a constant $M_\epsilon > 0$ such that

$$\mathrm{Leb}_2\{(x,y) \in \mathbb{T}^2 : |u(x,y)| \leq M_\epsilon/2\} \geq 1 - \eta(\epsilon)/4\,.$$

Thus, by the identity $\Phi_N(x,y) = u \circ f^N(x,y) - u(x,y)$, it follows that for any $\ell \in N_\ell$ we have

$$\mathrm{Leb}_2\{(x,y) \in \mathbb{T}^2 : |\Phi_{N_\ell}(x,y)| \leq M_\epsilon\} \geq 1 - \eta(\epsilon)/2\,. \tag{59}$$

However, by definition (57), the set $S_\ell(\epsilon)$ and the set $\{(x,y) \in \mathbb{T}^2 : |\Phi_{N_\ell}(x,y)| \leq M_\epsilon\}$ are disjoint for all $N_\ell > \epsilon^{-2} M_\epsilon^2$, hence

$$1 + \eta(\epsilon)/2 \leq \mathrm{Leb}_2(S_\ell(\epsilon)) + \mathrm{Leb}_2\{(x,y) \in \mathbb{T}^2 : |\Phi_{N_\ell}(x,y)| \leq M_\epsilon\} \leq 1\,,$$

which is the desired contradiction. q.e.d.

Let us show that the class $\mathcal{M}_f$ of mixing roof functions in Definition 2 contains the complement of a countable codimension subspace of a dense subspace of the space of smooth functions which can be described explicitly.

Corollary 2. *The class $\mathcal{M}_f$ contains the set*

$$\mathcal{P}^+_{\mathbb{T}^2} \setminus \left(\cap_{n\in\mathbb{Z}\setminus\{0\}} \cap_{m\in\mathbb{Z}_{|n|}} \ker D_{(m,n)} \right) \cap \ker D_{(0,0)} \,, \tag{60}$$

where $\mathcal{P}^+_{\mathbb{T}^2}$ denotes the space of all real-valued positive functions on $\mathbb{T}^2$ which are trigonometric polynomials in both variables, and $D_{(m,n)}$ are the invariant distributions for the linear skew-shift described in Theorem 10.

Proof. Let Φ be in the set (60) and let us show that $\Phi \in \mathcal{M}_f$. Since $\Phi \in \mathcal{P}^+_{\mathbb{T}^2}$ is real-valued, positive, and is a trigonometric polynomial in all variables, the inclusion $\mathcal{P}^+_{\mathbb{T}^2} \subset \mathcal{R}$ holds trivially (see Definition 1 of $\mathcal{R}$). Thus, by Definition 2 of $\mathcal{M}_f$, to show that $\Phi \in \mathcal{M}_f$, it is enough to show that Φ is not a measurable coboundary. By Theorem 7 and Theorem 10, Φ is a measurable coboundary if and only if it belongs to the kernel of all invariant distributions in Theorem 10. Since by definition the set (60) is not contained in the intersection of the kernels of all invariant distributions, $\Phi \in \mathcal{R}$. q.e.d.

Let us now prove that the roofs functions of the examples at the end of § 2.4 belong to the class $\mathcal{R}$. We will prove only that roofs in (E3) are in $\mathcal{R}$, since one can check that (E1) is a special case of (E3) and (E2) differs by a function in (E3) by a coboundary. Let Φ be as in (E3). By Corollary 2, it is enough to find a distribution $D_{(m,n)}$ as in Theorem 10 which is not zero. One can check that the roof function Φ belongs to the subspace $H_{[(0,1)]} + H_{[(0,-1)]}$ and, by Theorem 10 and by assumption,

$$D_{(0,1)} \left(\sum_{j\in\mathbb{Z}} a_j e^{2\pi i (jx+y)} \right) = \sum_{j\in\mathbb{Z}} a_j e^{-2\pi i (\beta j + \alpha \binom{j}{2})} \neq 0 \,.$$

6. Non-triviality, weak mixing and mixing equivalences

In this section we give the proofs of the equivalences in Theorem 3 and Theorem 4. We first recall for the convenience of the reader the following well-known elementary result about special flows that relates non-triviality of time-changes and weak mixing (see for instance [**20**], §9.3.4).

Lemma 16 (Non-triviality and weak mixing)**.** *Let f be a measure preserving transformation on a probability space (Σ, ν). For any measurable almost coboundary $\Phi : \Sigma \to \mathbb{R}^+$, the special flow f^Φ over f with roof function Φ is measurably trivial; hence it is not weak mixing.*

Proof. Since Φ is an almost coboundary, there exist a constant $C_\Phi > 0$ and a measurable function $u : X \to \mathbb{R}$ such that

$$\Phi - C_\Phi = u \circ f - u. \tag{61}$$

Let $I : X \times \mathbb{R} \to X \times \mathbb{R}$ be the map

$$I(x, y) = (x, y + u(x)) , \quad \text{for all } (x, y) \in X \times \mathbb{R} .$$

One can easily check that the map I is a measurable isomorphism of $X \times \mathbb{R}$ which conjugates the vertical flow to itself. Since the phase space of the special flow under Φ is defined as the quotient $X/ \sim_\Phi$ with respect to the equivalence relation $(x, \Phi(x) + y) \sim_\Phi (f(x), y)$, for all $x \in X, y \in \mathbb{R}$, it is sufficient to prove that the map I has a well-defined projection on the quotient spaces $X/ \sim_\Phi$ and $X/ \sim_{C_\Phi}$. Since u is a solution of the cohomological equation (61), the following identities hold:

$$\begin{aligned} I(x, \Phi(x) + y) &= (x, \Phi(x) + u(x) + y) = \\ &= (x, C_\Phi + u \circ f(x) + y) \sim_{C_\Phi} (f(x), u \circ f(x) + y) \\ &= I(f(x),) , \end{aligned}$$

hence the map $I : X \times \mathbb{R} \to X \times \mathbb{R}$ passes to the quotient as claimed. It is well known and immediate to verify that no special flow with constant roof function is weak mixing. q.e.d.

Proof of Theorem 4. Let $\mathcal{T}_f$ and $\mathcal{M}_f = \mathcal{R} \setminus \mathcal{T}_f$ be as in Definitions 1 and 2. As a consequence of the cocycle effectiveness (Theorem 7), $\mathcal{T}_f$ is in fact the intersection of the dense space $\mathcal{R}$ with the kernel of countable many linear functionals, as stated in Theorem 4 (see Corollary 2). Let us prove the equivalences of $1) - 4)$. The implication $1) \Rightarrow 4)$ is exactly the content of Theorem 5. The implication $4) \Rightarrow 3)$ is obvious. If Φ is smoothly trivial, and hence in particular measurably trivial, f^Φ is *not* weak mixing (see Lemma 16). Thus, taking counterpositives, $3) \Rightarrow 2)$. We are left to prove $2) \Rightarrow 1)$. Let us again prove the counterpositive implication and, since $\Phi \in \mathcal{R}$ by the assumptions in Theorem 4, if 1) does not hold, we know that $\Phi \in \mathcal{R} \setminus \mathcal{M}_f$. This means, by Definition 2 of $\mathcal{M}_f$, that the projection ϕ defined in (9) is a measurable coboundary for f. Since clearly $\mathcal{R} \subset W^s(\mathbb{T}^2)$, $s > 3$, by Theorem 7 we then know that ϕ belongs to the kernel of all f-invariant distributions and it is a *smooth* almost coboundary. It is easy to check solving the cohomological equation in Fourier coefficients that any trigonometric polynomial on $\mathbb{T}$ is a smooth almost coboundary for any irrational circle rotation. Thus, $\Phi = \phi + \phi^\perp$ is a smooth almost coboundary for the skew-shift and f^Φ is smoothly trivial, or equivalently, 2) does not hold. This concludes the proof of the equivalences. q.e.d.

Proof of Theorem 3. We will deduce Theorem 3 from Theorem 4. As summarized in § 2.2, any uniquely ergodic Heisenberg nilflow ϕ^W has a global transverse smooth transverse surface $\Sigma \approx \mathbb{T}^2_E = \mathbb{R}^2/(\mathbb{Z} \times \mathbb{Z}/E)$ and the Poincaré map $P_W : \mathbb{T}^2_E \to \mathbb{T}^2_E$ is a uniquely ergodic skew-shift (over a circle rotation). Let $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. It follows from Lemma 1 that there exist $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $\beta \in \mathbb{R}$ such that the *linear skew-shift* over a circle rotation, defined in (8), is a covering map of finite order $E \in \mathbb{N} \setminus \{0\}$ of the Poincaré map, in the sense that the canonical projection $\pi_E : \mathbb{T}^2 \to \mathbb{T}^2_E$ yields a semi-conjugacy between the skew-shift f on $\mathbb{T}^2$ and the Poincaré map P_W on $\mathbb{T}^2_E$. It is sufficient to prove the theorem in the particular case $E = 1$, when the Poincaré map is isomorphic to a uniquely ergodic standard skew-shift of the form (8). In fact, all other cases can be treated similarly or reduced to this one by considering the appropriate covering map on $\mathbb{T}^2$.

Let us say that a positive function α belongs to the class $\mathcal{A}$ (respectively to the class $\mathcal{M}_h$) iff the return time function Φ^α given by Lemma 2 where $\Phi \equiv 1$ belongs to $\mathcal{R}$ (respectively to $\mathcal{M}_f$). The proof of Theorem 3 now reduces simply in a rephrasing $1) - 4)$ in Theorem 3 using the dictionary between time-changes of flows and special flows recalled in § 2.4 and checking that they correspond to $1) - 4)$ in Theorem 4. q.e.d.

References

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

[2] A. Avila & G. Forni, *Weak mixing for interval exchange transformations and translation flows,* Annals of Mathematics, **165** (2): 637–664, 2007, MR 2299743, Zbl 1136.37003.

[3] M. Burger, *Horocycle flow on geometrically finite surfaces,* Duke Math. J. 61: 779–803, 1990, MR 1084459, Zbl 0723.58041.

[4] I.P. Cornfeld, S.V. Fomin & Y.G. Sinai, *Ergodic Theory*, Springer-Verlag, 1980, MR 0832433, Zbl 0493.28007.

[5] L.J. Corwin & F.P. Greenleaf, *Representations of nilpotent Lie groups and their applications. Part I: Basic theory and examples,* Cambridge studies in advanced mathematics **18**, Cambridge University Press, Cambridge, UK, 1990, MR 1070979, Zbl 0704.22007.

[6] B.R. Fayad, *Polynomial decay of correlations for a class of smooth flows on the two torus,* Bull. Soc. Math. France **129**: 487–503, 2001, MR 1894147, Zbl 1187.37009.

[7] ———. *Weak mixing for reparameterized linear flows on the torus,* Ergodic Theory Dynam. Systems **22** (1): 187–201, 2002, MR 1889570, Zbl 1001.37006.

[8] ———. *Analytic mixing reparametrizations of irrational flows,* Ergodic Theory and Dynamical Systems, **22** (2): 437–468, 2002, MR 1898799, Zbl 1136.37307.

[9] H. Fiedler, W.B. Jurkat & O. Körner, *Asymptotic expansions of finite theta series,* Acta Arithmetica, XXXII, 129–146, 1977, MR 0563894, Zbl 0308.10021.

[10] L. Flaminio & G. Forni. *Invariant distributions and time averages for horocycle flows,* Duke Math. J. **119** (3): 465–526, 2003, MR 2003124, Zbl 1044.37017.

[11] ———. *Equidistribution of nilflows and applications to theta sums,* Ergodic Theory and Dynamical Systems, **26**:2:409–433, 2006, MR 2218767, Zbl 1087.37026.

[12] ———. *On the cohomological equation for nilflows,* Journal of Modern Dynamics **1** (1): 37–60, 2007, MR 2261071, Zbl 1114.37004.

[13] G. Forni, *Solutions of the cohomological equation for area-preserving flows on compact surfaces of higher genus,* Annals of Mathematics, **146** (2): 295–344, 1997, MR 1477760, Zbl 0893.58037.

[14] ———. *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus,* Annals of Mathematics (2), **155** (1): 1–103, 2002, MR 1888794, Zbl 1034.37003.

[15] H. Furstenberg, *The unique ergodicity of the horocycle flow. Recent advances in topological dynamics,* Lecture Notes in Math., **318**: 95–115, 1973, Springer, Berlin, MR 0393339, Zbl 0256.58009.

[16] L.W. Green, *Spectra of nilflows,* Bull. Amer. Math. Soc. **67**: 414–415, 1961, MR 0126504, Zbl 0099.39104.

[17] D.A. Hejhal, *On the uniform equidistribution of long closed horocycles,* Loo-Keng Hua: a great mathematician of the twentieth century, Asian J. Math. **4** (4): 839–853, 2000, MR 1870662, Zbl 1014.11038.

[18] M.R. Herman, *Examples de flots hamiltoniens dont aucune perturbation en topologie C^∞ n'a d'orbites périodiques sur un ouvert de surfaces d'énergies,* C. R. Acad. Sci. Paris **312**: 989–994, 1991, MR 1113091, Zbl 0759.58016.

[19] A. B. Katok, *Interval exchange transformations and some special flows are not mixing,* Israel Journal of Mathematics, **35** (4): 301–310, 1980, MR 0594335 , Zbl 0437.28009.

[20] ———. *Combinatorial Constructions in Ergodic Theory and Dynamics,* University Lecture Series Vol. 30. American Mathematical Society, Providence, RI, 2003, MR 2008435, Zbl 1030.37001.

[21] K.M. Khanin & Y.G. Sinai, *Mixing for some classes of special flows over rotations of the circle,* Funktsional'nyi Analiz i Ego Prilozheniya, **26**(3):1–21, 1992, (translated in: *Functional Analysis and Its Applications*, **26**:3:155–169, 1992) MR 1189019, Zbl 0797.58045.

[22] 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**: 512–518, 1972 (translated in Soviet Math. Dokl., **13**: 949-952, 1972), MR 0306629, Zbl 0262.28015.

[23] ———. *Mixing in special flows over a rearrangement of segments and in smooth flows on surfaces,* Mat. Sb., **96**(138): 471–502, 1975, MR 0516507, Zbl 0326.28030.

[24] A.V. Kochergin, *Nondegenerate saddles and the absence of mixing in flows on surfaces,* Proc. Steklov Inst. Math., **256** (2007), no. 1, 238–252, MR 2336903, Zbl 1153.37303.

[25] A.N. Kolmogorov, *On dynamical systems with an integral invariant on the torus,* Dokl. Akad. Nauk SSSR (N. S.) **93**: 763–766, 1953 (in Russian), MR 0062892, Zbl 0052.31904.

[26] M. Kontsevich, *Lyapunov exponents and Hodge theory, in "The mathematical beauty of physics",* Saclay, 1996, Adv. Ser. Math. Phys. **24**, World Scientific, River Edge, NJ, 318–332, 1997, MR 1490861, Zbl 1058.37508.

[27] A.G. Kushnirenko, *Spectral properties of some dynamic systems with polynomial divergence of orbits,* Moscow Univ. Math. Bull **29**: 101–108, 1974, MR 0353369, Zbl 0276.58007.

[28] B. Marcus, *The horocycle flow is mixing of all degrees,* Invent. Math., 46, (3): 201–209, 1978, MR 0488168, Zbl 0395.28012.

[29] B. Marcus. *Ergodic properties of horocycle flows on surfaces of negative curvature,* Ann. of Math., **105** (2): 81–105, 1977, MR 0458496, Zbl 0322.28012.

[30] J. Marklof, *Limit theorems for theta sums,* Duke Math. J. **97** (1): 127–153, 1999, MR 1682276, Zbl 0965.11036.

[31] S. Marmi, P. Moussa & J.-C. Yoccoz, *The Cohomological Equation for Roth-Type Interval Exchange Maps,* Journal of the American Mathematical Society, **18** (4): 823–872, 2005, MR 2163864, Zbl 1112.37002.

[32] H. Masur, *Interval exchange transformations and measured foliations,* Annals of Mathematics, **115**: 169–200, 1982, MR 0644018, Zbl 0497.28012.

[33] ———. *Ergodic theory of translation surfaces*, pages 527–547. Handbook of Dynamical Systems, Vol. 1B. Elsevier B. V., Amsterdam., 2006, MR 2186247, Zbl 1130.37313.

[34] O.S. Parasyuk, *Flows of horocycles on surfaces of constant negative curvature,* Uspehi Matem. Nauk (N.S.), **8** (3): 125–126, 1953, MR 0058883, Zbl 0052.34201.

[35] M. Ratner, *The rate of mixing for geodesic and horocycle flows,* Ergodic Theory Dynam. Systems, **7** (2): 267–288, 1987, MR 0896798, Zbl 0623.22008.

[36] P. Sarnak, *Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series,* Comm. Pure Appl. Math. **34**: 719–739, 1981, MR 0634284, Zbl 0501.58027.

[37] D. Scheglov, *Absence of mixing for smooth flows on genus two surfaces,* Journal of Modern Dynamics., **3** (1): 13–34, 2009, MR 2481330, Zbl 1183.37080.

[38] K. Schmidt, *Dispersing cocycles and mixing flows under functions,* Fund. Math. **173**(2): 191–199, 2002, MR 1924814, Zbl 1032.37004.

[39] A.N. Starkov, *Dynamical Systems on Homogeneous Spaces,* Translations of the American Mathematical Society, **190**, Providence, Rhode Island 2002, MR 1746847, Zbl 1143.37300.

[40] A. Strömbergsson., *On the uniform equidistribution of long closed horocycles,* Duke Math. J. **123** (3): 507–547, 2004, MR 2068968, Zbl 1060.37023.

[41] C. Ulcigrai, *Absence of mixing in area-preserving flows on surfaces,* Annals of Mathematics, **173**(3): 1743–1778, 2011.

[42] ———. *Mixing of asymmetric logarithmic suspension flows over interval exchange transformations,* Ergodic Theory Dynam. Systems **27**(3), 991–1035, 2007, MR 2322189, Zbl 1135.37004.

[43] ———. *Weak mixing for logarithmic flows over interval exchange transformations,* Journal of Modern Dynamics, **3** (1): 35–49, 2009, MR 2481331, Zbl 1183.37081.

[44] W.A. Veech, *Gauss measures for transformations on the space of interval exchange maps,* Annals of Mathematics, **115**: 201–242, 1982, MR 0644019, Zbl 0486.28014.

[45] D. Zagier, *Eisenstein series and the Riemann zeta function, in Automorphic Forms, Representation Theory and Arithmetic (Bombay, 1979),* Tata Inst. Fund. Res. Studies in Math. 10, Tata Inst. Fund. Res., Bombay, 275–301, 1981, MR 0633666, Zbl 0484.10019.

[46] A. Zorich., *Deviation for interval exchange transformations,* Ergodic Theory Dynam. Systems **17**: 1477–1499, 1997, MR 1488330, Zbl 0958.37002

CNRS UMR 7586
Institut de Mathématiques de Jussieu
175 rue du Chevaleret
75013-Paris, France
and
IMPA
Estrada Dona Castorina, 110
22460-320, Rio de Janeiro, Brazil
E-mail address: artur@math.jussieu.fr

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

School of Mathematics
University of Bristol
University Walk
BS8 1TW Bristol, United Kingdom
E-mail address: corinna.ulcigrai@bristol.ac.uk