CrossMark

Multiple mixing for a class of conservative surface flows

Bassam Fayad[1] · Adam Kanigowski[2]

Abstract Arnold and Kochergin mixing conservative flows on surfaces stand as the main and almost only natural class of mixing transformations for which higher order mixing has not been established, nor disproved. Under suitable arithmetic conditions on their unique rotation vector, of full Lebesgue measure in the first case and of full Hausdorff dimension in the second, we show that these flows are mixing of any order. For this, we show that they display a generalization of the so called Ratner property on slow divergence of nearby orbits, that implies strong restrictions on their joinings, which in turn yields higher order mixing. This is the first case in which the Ratner property is used to prove multiple mixing outside its original context of horocycle flows and we expect our approach will have further applications.

Contents

✉ Adam Kanigowski
  adkanigowski@gmail.com

[1] Institut de Mathématiques de Jussieu-Paris Rive Gauche, Paris, France

[2] Institute of Mathematics, Polish Academy of Sciences, 00-956 Warsaw, Poland

1 Introduction

A major open problem in ergodic theory is Rokhlin's question on whether mixing implies mixing of all orders, also called multiple mixing [26]. In most of the known examples of mixing dynamical systems, multiple mixing is now known to hold. Moreover, a positive answer to Rokhlin's question is actually known to generally hold within various classes of mixing dynamical systems. The most noteworthy are K-systems where multiple mixing always holds [4], horocycle flows [23], mixing systems with singular spectrum that display multiple mixing by a celebrated theorem of Host [13], and finite rank systems since Kalikow showed that rank one and mixing implies multiple mixing [14], a result that was extended to finite rank mixing systems by Ryzhikov [30].

In the second half of the last century, it was discovered that mixing is possible for a class of smooth area preserving flows on surfaces called multivalued Hamiltonians since they are locally given by Hamilton's equations. The Hamiltonian H in question is associated to a closed differential 1-form $\omega = dH$, the form ω being globally defined—but of course this is not necessarily the case for H. The possibility of mixing for these flows was studied in two different cases: first, by Kochergin who obtained mixing when ω has higher order zeros and thus the flow has degenerate saddles [17], and then by Arnold [1] who suggested that mixing is possible on a minimal component even in the case where ω is Morse but the saddles on the minimal component appear in asymmetric configurations, for example because of a saddle loop. That mixing was indeed possible in the latter context was proved in a particular case by Khanin and Sinai [32]. Absence of mixing in the case of Morse forms was obtained by Kochergin in some particular cases [18] (see also [22]), and later generalized to a typical one form by Ulcigrai [35].

Considering a 1-dimensional section of a multi-valued Hamiltonian flow allows to view the dynamics on its minimal components as special flows above interval exchange transformations (IET), which in particular situations can be circular rotations. The case of non-degenerate saddles corresponds to ceiling functions with logarithmic singularities, while the case of degenerate saddle points corresponds to ceiling functions with stronger singularities such as integrable power like singularities. In the case of power singularities [17] proved mixing above any ergodic IET, while [32] established mixing in the case of a single asymmetric logarithmic singularity above a circular rotation with a typical frequency.

The study of the mixing properties of surface flows has known a revival of interest since the beginning of the 2000s, with results such as the computation of the speed of mixing [6] or extensions of the Khanin-Sinai mixing result to include all irrational translation vectors [19,20] (see also [21]), or advances in the study of multi-valued Hamiltonian flows on surfaces in the general case where the Poincaré section return map is an IET and not just a circular rotation [3,35–37].

Mixing surface flows stand today as the main and almost only natural class of mixing transformations for which higher order mixing has not been established, nor disproved. Our aim here is to prove mixing of all orders for a subclass of these systems given by special flows above circular rotations, with ceiling functions having asymmetric logarithmic singularities or integrable power like singularities. For convenience, we will speak of Arnold flows in the first case, and Kochergin flows in the second. Our results will depend on the arithmetics of the frequency α of the rotation on the base, that determines the slope of the unique rotation vector of the corresponding surface flow.

Loosely speaking our main result is as follows (it will be made precise at the end of this introduction, see Corollaries 1.6 and 1.9).

Theorem *Arnold flows are mixing of all orders for a set of $\alpha \in (0, 1)$ of full Lebesgue measure and Kochergin flows are mixing of all orders for a set of $\alpha \in (0, 1)$ of full Hausdorff dimension. This is true in the case of a single singularity and the same result holds if there are many singularities, under a non resonance condition (of full Hausdorff dimension) between the positions of the singularities and the base frequency α.*

Similar mixing mechanisms due to orbit shear as in Kochergin and Arnold flows were observed relatively recently such as in [2,3,7,35] and it should be possible to apply the techniques of the current paper to the study of higher order mixing for such *parabolic* systems.

To explain our approach we need first to make a detour by Ratner's study of horocycle flows. In the 1980s, Ratner developed a rich machinery to study horocycle flows [27–29] and, in particular, singled out a special way of controlled slow divergence of orbits of nearby points which resulted in the notion of H_p-property, later called R-property by Thouvenot [33]. This property, to which we will come back with more detail in the sequel, has important dynamical consequences, mainly expressed by a restriction on the possible joining measures of a system having the R-property with other systems, and in particular with itself.

A *joining* between two dynamical systems $(T, X, \mathscr{B}, \mu)$ and $(S, Y, \mathscr{C}, \nu)$, $(X, \mathscr{B}, \mu)$ and $(Y, \mathscr{C}, \nu)$ being standard Borel probability spaces, is a measure ρ on $X \times Y$ invariant by $T \times S$ whose marginals on X and Y are μ and ν. The

definition for flows is similar. An important notion in Ratner's theory is that of finite extension joinings (FEJ).

Definition 1.1 An ergodic flow $(T_t)_{t \in \mathbb{R}}$ is said to have *FEJ-property*, an acronym for *finite extension joining*, if for every ergodic flow $(S_t)_{t \in \mathbb{R}}$ acting on $(Y, \mathscr{C}, \nu)$ and every ergodic joining ρ of $(T_t)_{t \in \mathbb{R}}$ and $(S_t)_{t \in \mathbb{R}}$ different from the product measure $\mu \times \nu$, ρ yields a flow which is a finite extension of $(S_t)_{t \in \mathbb{R}}$.

It was shown in [31] that a mixing flow with FEJ-property is mixing of all orders. Moreover, it was proved in [27] that a flow with R-property has the FEJ-property. It follows that mixing flows with the R-property are mixing of all orders. Since the R-property for horocycle flows stemmed from polynomial shear along the orbits, and since Kochergin flows displayed a similar polynomial shear along the orbits, the idea that special flows over rotations may enjoy the R-property, and thus be multiple mixing, was then suggested by J-P. Thouvenot in the 1990s (see p. 2 in [9]).

However, whether natural classes of special flows (not necessarily mixing) over irrational rotations may have the R-property remained open until Frączek and Lemańczyk [9,10] showed that a generalized R-property holds in some classes of special flows with roof functions of bounded variation (which, by [16], are not mixing). More precisely, they have introduced a weaker notion than the R-property, called *weak Ratner* or WR-property that however still implies the FEJ-property (see Definition 2.1 and the comment after it) .

Unfortunately, in the mixing examples of special flows under piecewise convex functions with singularities such as Arnold and Kochergin flows, the shear may occur very abruptly as orbits approach the singularity and this may prevent them from having the weak Ratner property. Indeed, we believe that these flows do not have the WR-property (it was observed by the first author that Arnold and Kochergin flows do not have a natural *strong Ratner property* as described in the next paragraph). This is corroborated by the following result that shows that Kochergin flows, in the context of bounded type frequency in the base (that is *a priori* favorable to controlling the shear), do not have the WR-property.

Theorem 1 *Let* $\alpha \in \mathbb{R}$ *be irrational of bounded type and* $f(x) = x^\gamma + r$, $-1 < \gamma < 0, r > 0$. *Then the special flow* $(T^t_{\alpha, f})_{t \in \mathbb{R}}$ *defined above the circle rotation* R_α *and under the ceiling function* f *does not have the WR-property.*

Here the circle is $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We recall that an irrational α is said to be of bounded type or $\alpha \in DC(0)$, if the partial quotients in the continued fraction of α are bounded, i.e. if $\alpha = [a_0; a_1, a_2, \ldots]$ and there exists $K > 0$ such that $a_i < K$ for all $i \geqslant 1$. We refer to Sect. 3 for the exact definition of special flows. Theorem 1 has another consequence. It is known that every

horocycle flow $(h_t)_{t \in \mathbb{R}}$ is loosely Bernoulli [29]; therefore, for every irrational α, there exists a positive function in $f \in L^1(\mathbb{T})$ such that $(h_t)_{t \in \mathbb{R}}$ is measurably isomorphic to the special flow $(T^t_{\alpha, f})_{t \in \mathbb{R}}$ [25]. It follows from [16] and the fact that $(h_t)_{t \in \mathbb{R}}$ is mixing that f is of unbounded variation. Moreover, by [24], f can be made C^1 except for one point. Since the R-property implies the WR-property and the R-property is an isomorphism invariant, no special flow as in Theorem 1 is isomorphic to a horocycle flow. Actually, this line of thought can be extended to show that horocycle flows are never isomorphic to special flows above an irrational rotation and under a roof function that is convex and C^2 except at one point. For the latter result, one needs to introduce the concept of *strong Ratner property*, which is also an isomorphism invariant, that specifies the occurrence of slow divergence of nearby orbits to the first time when the orbits do split apart. This is the natural property that Ratner actually obtains for horocycle flows, and it is relatively easy to show that the Kochergin flows do not have it. What is more complicated in the proof of the absence of the general R-property for special flows, is to make sure that the slow divergence does not occur much later in the future after the nearby points have split and then came back together (see the proof of Theorem 1).

To bypass Theorem 1 and still use controlled divergence of orbits to show multiple mixing, our approach will be to further weaken the WR-property. Namely, we introduce the SWR-property, which stands for *switchable weak Ratner* property, that assumes that a pair of nearby points displays the WR-Property either under forward iteration in time or under backward iteration, and this depending on the pair of points. We show that the SWR-property is sufficient to guarantee the same FEJ consequences as the Ratner or the weak Ratner property. Consequently, a mixing flow enjoying the SWR-property is mixing of all orders.

The main idea in showing that Arnold and Kocergin special flows may have the SWR-Property is the following. For these flows, the main contribution to the shear between orbits is due to the visits of the flow lines to the neighborhood of the singularities. With the representation of these flows as special flows above irrational rotations, the shear is translated into the divergence between the Birkhoff sums of the roof functions for nearby points, and this divergence is mainly due to the visits under the base rotation to the neighborhoods of the points where the roof function has its singularities. If the base rotation angle α is of bounded type two nearby points will accumulate sufficient stretch while staying sufficiently far from the singularity either when they are iterated forward or when they are iterated backward. In the case of ceiling functions with only logarithmic singularities we are also able to exploit the progressive contribution to the shear of these visits to the singularities to obtain multiple mixing for a full measure set of numbers α.

We now introduce some notations related to the ceiling functions that will be considered and their singularities, after which we will be able to state our exact results on the SWR-property and multiple mixing.

Definition 1.2 Let h be a positive function $h \in C^2(\mathbb{T} \setminus \{0\})$, decreasing on $(0, 1)$, $\lim_{x \to 0^+} h(x) = +\infty$, h' increasing on $(0, 1)$. Let $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ for some numbers $a_1, \ldots, a_k \in \mathbb{T}$. We say that f has *singularities of type* h at $\{a_1, \ldots, a_k\}$ if

$$\lim_{x \to a_i^+} \frac{f'(x)}{h'(x - a_i)} = A_i \quad \text{and} \quad \lim_{x \to a_i^-} \frac{f'(x)}{h'(a_i - x)} = -B_i, \tag{1}$$

for some numbers $A_i, B_i \geqslant 0, i = 1, \ldots, k$.

Notice that in this definition h may only reflect a domination on the singularities of f since the coefficients A_i, B_i may be equal to zero at some or at all i's.

Definition 1.2 will allow us to state our results with some flexibility on the singularities but the reader should keep in mind that the results will target functions that are essentially of the form $A \ln(a-x)$ or $A(a-x)^\gamma$, $\gamma \in (-1, 0)$, when x is in a left neighborhood of a singularity a and similar form on the right side of a singularity. In the case of Arnold flows all the singularities will be assumed to be essentially of logarithmic type (see more general definition in Sect. 1.1) while in the case of Kochergin flows we will be dealing with functions having at least one power like singularity while the other singularities are supposed to be of equal or weaker type (see Sect. 1.2).

Furthermore, our results can deal with functions having several singularities but require a non resonance condition between these singularities and the rotation frequency α in the base of the special flow, that we now state.

Our standing assumption is that $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. We then let (q_s) be the sequence of denominators of the best rational approximations of α. Namely (q_s) is the unique increasing sequence such that $q_0 = 1$ and $\|q_s \alpha\| < \|k\alpha\|$ for any $k < q_{s+1}, k \neq q_s$. We recall (see e.g. [15]) that

$$\frac{1}{2q_{s+1}} \leqslant \|q_s \alpha\| \leqslant \frac{1}{q_{s+1}} \tag{2}$$

Definition 1.3 *[Badly approximable singularities]* Given $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, we will say that $\{a_1, \ldots, a_k\}$ are *badly approximable by* α if there exists $C > 1$ such that for every $x \in \mathbb{T}$ and every $s \in \mathbb{N}$, there exists at most one $i_0 \in \{0, \ldots, q_s - 1\}$ such that

$$x + i_0 \alpha \in \bigcup_{i=1}^{k} \left[\frac{-1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s} \right]. \tag{3}$$

Remark 1.4 It was shown in [11, Lemma3] that if $a_i - a_j \in (\mathbb{Q} + \mathbb{Q}\alpha) \setminus (\mathbb{Z} + \mathbb{Z}\alpha)$ whenever $i \neq j$ then $\{a_1, \ldots, a_k\}$ are badly approximable by α.

Note that if there is only one singularity, that is $k = 1$, then by (2) it is always badly approximable by α. The following shows that for $k \geqslant 2$ the set of singularities that are badly approximable by α is a thick set in $[0, 1]^k$.

Lemma 1.5 *[34] Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. For any $k \in \mathbb{N}$, the set $E \subset [0, 1]^k$ of $k - tuples(a_1, \ldots, a_k)$ that are badly approximable by α is a product of sets of full Hausdorff dimension in $[0, 1]$.*

Proof Define

$$B(\alpha) := \left\{ b \in \mathbb{R} : \exists C > 0, \forall k \in \mathbb{Z} \setminus 0 : \|k\alpha - b\| \geqslant \frac{1}{C|k|} \right\}$$

Then if $(a_1, \ldots, a_k)$ are such that $a_i - a_j \in B(\alpha)$ for any $i, j \in \{1, \ldots, k\}, i \neq j$, then $(a_1, \ldots, a_k)$ are badly approximable by α.

But it was proven in [34] (see also [5]) that the set $B(\alpha)$ is a winning set in the sense of Schmidt (see [5,34] and references therein). A winning set is of full Hausdorff dimension. Moreover, for a winning set $B \in \mathbb{R}$ we have that for any $x_1, \ldots, x_n$ the set $\cap_{s=1}^{n}(x_s + B)$ is winning. So, if $a_1, \ldots, a_l$ are such that $a_i - a_j \in B(\alpha)$ for any $i, j \in \{1, \ldots, l\}, i \neq j$, then the set of $a \in [0, 1]$ such that $a \in \cap_{s=1}^{l}(a_s + B(\alpha))$ is winning which means that $a_1, \ldots, a_l, a_{l+1}$ are badly approximable by α for a winning set of a_{l+1}. The statement of the Lemma follows then by induction and because a single a_1 is always badly approximable by α.   □

1.1 Logarithmic like singularities

In the case of logarithmic like singularities, the following theorem holds.

Theorem 2 *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ with the singularities $\{a_1, \ldots, a_k\}$ of type h and badly approximable by α, with some associated constant $C > 1$. Assume that $\sum_{i=1}^{k} A_i \neq \sum_{i=1}^{k} B_i$ and that there exists a constant $m_0 > 0$ and a sequence (x_s) such that for every $s \in \mathbb{N}$, we have $x_s < \frac{1}{q_s}$ and*

1. $\lim_{s \to +\infty} \frac{h'(\frac{x_s}{4C})}{q_s h(\frac{1}{2q_s})} = 0$, $\lim_{s \to +\infty} x_s q_s h(\frac{1}{2q_s}) = +\infty$;
2. $\sum_{i \notin K_\alpha} q_i x_i < +\infty$, *where* $K_\alpha := \{s \in \mathbb{N} : q_{s+1} < \frac{1}{x_s}\}$;
3. $h(\frac{1}{2q_s})/h(\frac{1}{2q_{s+1}}) > m_0$.

Then the special flow $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-property.

What is x_s in Theorem 2? When we will compute the shear in the Birkhoff sums of the ceiling function f, we will naturally be able to control only those points that do not go too close to the singularities under the iteration by R_α on the base. Then, a controllable point will be a point that stays at distance x_s from the singularities during $O(q_{s+1})$ iterates in the future *or* in the past. We choose x_s so that the contribution to the shear of a single visit to a singularity is negligible with respect to the accumulated shear (this, either in the future or in the past). To fix ideas, suppose the singularities are reduced to just one one-sided singularity at the origin and observe that if f at the origin is exactly log then the choice $x_s = 1/q_s \log^{\frac{7}{8}} q_s$ would satisfy 1. Then if $q_{s+1} < q_s \log^{\frac{7}{8}} q_s$ we see that for any $x \in \mathbb{T}$ either up to $q_{s+1}/2$ in the future or up to $-q_{s+1}/2$ in the past the orbit of x by R_α does not enter the x_s neighborhood of the origin, and this will show that the progressive accumulation in the Birkohff sums of the derivative of the ceiling function above the orbit of x in the region of time between $O(q_s)$ and $O(q_{s+1})$ in the future or $O(-q_{s+1})$ and $O(-q_s)$ in the past dominates the value of the derivative at the closest entry to the neighborhood of the origin (this is the aim of condition 1). The latter is the crucial fact that we need to show SWR. In the opposite case where for example $q_{s+1} \geqslant q_s \log q_s$ we have to discard the points that enter the x_s neighborhood of the origin between time $-q_s$ and q_s and then show that the remaining points stay away from the x_s neighborhood of the origin for $O(q_{s+1})$ iterations (either in the past or in the future) and conclude as before. This is where 2. is necessary to show that the measure of the discarded points is arbitrarily small.

We refer to the beginning of Sect. 4.1 for an outline of the proof of Theorem 2 in which the different roles of conditions 1., 2. and 3. are all explained in detail.

We now restate Theorem 2 in the particular case of exactly logarithmic singularities, that is when $h(x) = -\log(x)$, for $x \in [0, 1)$. To be able to choose x_s such that 1., 2. and 3. are satisfied we need some arithmetic restrictions on α, that we now introduce.

For $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, let $K_\alpha := \{n \in \mathbb{N} \ : \ q_{n+1} < q_n \log^{\frac{7}{8}}(q_n)\}$. We then define in view of 1. and 2. of Theorem 2

$$\mathcal{E} := \left\{ \alpha \in \mathbb{T} \setminus \mathbb{Q} \ : \ \sum_{i \notin K_\alpha} \frac{1}{\log^{\frac{7}{8}} q_i} < +\infty \right\}.$$

Indeed, for $\alpha \in \mathcal{E}$, 1. and 2. are satisfied with $x_s := \dfrac{1}{q_s \log^{\frac{7}{8}} q_s}$, for $s \geqslant 1$.

Recall first that a number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is said to be Diophantine if there exists $\tau \geqslant 0$ such that for any $p, q \in \mathbb{Z} \times \mathbb{N}^*$ we have that $|\alpha - \frac{p}{q}| \geqslant \frac{C(\alpha)}{q^{2+\tau}}$ for some $C(\alpha) > 0$. We call $\mathcal{D}$ the set of Diophantine numbers.

To have 3., it suffices to assume that α is Diophantine since an equivalent definition of $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ being Diophantine is that its sequence of denominators q_n satisfies $\forall_{n \in \mathbb{N}}$, $q_{n+1} < r_\alpha q_n^{1+\tau}$ (see (2)).

Hence we have the following

Corollary 1.6 *Consider $h(x) = -\log(x)$. Let $\alpha \in \mathcal{D} \cap \mathcal{E}$. Let $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ with the singularities $\{a_1, \ldots, a_k\}$ of type h and badly approximable by α. Assume that $\sum_{i=1}^k A_i \neq \sum_{i=1}^k B_i$. Then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-Property.*

Proof We take for x_s the sequence $\dfrac{1}{q_s \log^{\frac{7}{8}} q_s}$ and easily check the hypothesis of Theorem 2. Therefore $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-property. □

Corollary 1.6 covers a set of full Lebesgue measure of rotation angles α. Indeed, it is known that the set of Diophantine numbers $\mathcal{D}$ has full Lebesgue measure, and we will prove in Appendix B the following result. Denote by λ the Haar measure on $\mathbb{T}$.

Proposition 1.7 *It holds that $\lambda(\mathcal{E}) = 1$.*

We now recall the following results on mixing of special flows with ceiling functions having logarithmic singularities.

Theorem 3 *Let f be as in Corollary 1.6. Then*

(a) ([18]) *If $\sum_j A_j = \sum_j B_j$ then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ is not mixing for any $\alpha \in \mathbb{R} - \mathbb{Q}$.*
(b) ([19,20,32]) *If $\sum_j A_j \neq \sum_j B_j$ then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ is mixing for almost every $\alpha \in \mathbb{R} - \mathbb{Q}$.*
(c) ([19,20]) *If $A_j - B_j \neq 0$ have the same sign for all j then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ is mixing for each $\alpha \in \mathbb{R} - \mathbb{Q}$.*

In Theorem 5 below, we show that the SWR property for $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ implies the FEJ-property. As an immediate consequence of this and of Theorem 3, Corollary 1.6 and the fact that mixing flows with the FEJ-property are multiple mixing [31] we get the following.

Corollary 1.8 *Consider $h(x) = -\log(x)$. For $\alpha \in (0, 1)$, assume $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ has singularities $\{a_1, \ldots, a_k\}$ of type h and badly approximable by α. If $A_j - B_j \neq 0$ have the same sign for all j, then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ is multiple mixing for each $\alpha \in \mathcal{E} \cap \mathcal{D}$. If $\sum_j A_j \neq \sum_j B_j$ then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ is multiple mixing for almost every $\alpha \in \mathcal{E} \cap \mathcal{D}$.*

1.2 Power like singularities

Now, we will deal with power like singularities. We suppose $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ with singularities $\{a_1, \ldots, a_k\}$ of type h. We assume that there exists $i \in \{1, \ldots, k\}$ such that $A_i^2 + B_i^2 > 0$ in (1).

Theorem 4 *Let α be irrational with bounded partial quotients, that is, $\alpha \in DC(0)$. Assume that $\{a_1, \ldots, a_k\}$ are badly approximable by α with some constant $C > 1$. Assume that there exist constants $D_1, D_2 > 0$ such that for every $s \in \mathbb{N}$*

$$D_2 > \frac{-h'\left(\frac{1}{C^4 q_s}\right)}{q_s h\left(\frac{1}{2q_s}\right)} > D_1 \quad and \quad \frac{h\left(\frac{1}{2q_s}\right)}{h\left(\frac{1}{2q_{s+1}}\right)} > D_1. \tag{4}$$

Then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-property.

The most interesting case – when h has power singularities – is discussed in the corollary below.

Corollary 1.9 *Let $\alpha \in DC(0)$. Let $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ with all the singularities $\{a_1, \ldots, a_k\}$ of power-like type x^{γ_i} from the left and x^{δ_i} from the right, $-1 < \gamma_i, \delta_i < 0$. Then, if the points $\{a_1, \ldots, a_k\}$ are badly approximable by α, we have that $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-Property and is mixing of all orders.*

Proof of Corollary 1.9 It is easy to check that (4) in Theorem 4 is satisfied. This gives the SWR-Property . Mixing of $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ was established in [17]. Multiple mixing then follows from Theorem 5 and the FEJ-property. □

Remark 1.10 A stronger version of Theorem 4 actually holds : if $\alpha \in DC(0)$ and $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$ with all the singularities $\{a_1, \ldots, a_k\}$ of at most power-like type (x^{γ_i} from the left and x^{δ_i} from the right, $-1 < \gamma_i, \delta_i < 0$) and if $\gamma = \min_{1 \leqslant i \leqslant k}\{\gamma_i, \delta_i\}$, then it is sufficient to have the singularities of maximal type badly approximable with α to guarantee the SWR property, namely if the points in $E = \{a_i : \min\{\gamma_i, \delta_i\} = \gamma\} \subset \{a_1, \ldots, a_k\}$ are badly approximable by α, then $(T_{\alpha, f}^t)_{t \in \mathbb{R}}$ has the SWR-Property and is mixing of all orders. The proof in this case is similar to the proof of Theorem 4 and we omit it to avoid overloading the paper with unnecessary technicalities.

1.3 Plan of the paper

In Sect. 2 we introduce the SWR-Property and describe its joinings consequences. In Sect. 3 we give a criterion involving the Birkhoff sums of the ceiling function that guarantees that a special flow above an isometry has the SWR-property. The treatment of these sections is similar to [9,10]. In Sect. 4 we study the Birkhoff sums of logarithmic like and power like functions and prove Theorems 2 and 4. Appendix A is devoted to the proof of Theorem 1 on the absence of the SWR-Property for a subclass of Kochergin flows. Finally Appendix B is devoted to the proof that the set of frequencies for which Theorem 2 holds has full Lebesgue measure.

2 The SWR-property

Let $(X, \mathscr{B}, \mu)$ be a probability standard Borel space. We additionally assume that X is a complete metric space with a metric d. Let $(T_t)_{t \in \mathbb{R}}$ be an ergodic flow acting on $(X, \mathscr{B}, \mu)$. The R-property used by Ratner in the context of horocycle flows is a property of slow divergence between the orbits of nearby points that essentially states as follows: for any $\varepsilon > 0$ there exists $\kappa > 0$ such that if y and y' are sufficiently close to each other and if they are not on the same orbit then there exists $M(y, y')$ such that $d(T^t y, T^{t+\iota} y') < \varepsilon$ (with $\iota = \pm 1$) for $t \in [M, (1 + \kappa)M]$.

It is not difficult to see that the R-property implies the FEJ property. Indeed, if a joining λ is not a finite extension joining then one has that there exists points x, y, y' such that (x, y) and (x, y') are typical for λ while y and y' are arbitrarily close and not in the same orbit, and by the R-property and the Birkhoff ergodic theorem one obtains an extra invariance of λ, namely by $\mathrm{Id} \times T^\iota$, that implies that λ is the product measure.

Actually, it is useful to relax the R-property by asking that the controlled divergence happens for x and y outside an exceptional set of points of measure less than ε (and for most of the times in $[M, (1 + \kappa)M]$). The proof of the FEJ property remains the same in nature but becomes a bit more technical involving some standard measure theoretical arguments (see for example [33]).

In [10], Definition 4, a slightly weaker version of the R-property is given, that is called WR or Weak Ratner property, that allows the drift to vary in some fixed compact set away from zero and infinity. There again, the FEJ consequence as well as its proof follow in practically the same way as for the R-property, with some extra standard measure theoretical arguments, and under a "continuity" assumption on orbits (see below). However, as shown in [10] the WR property is more versatile than the R-property and is adapted to nonlinear situations, where the R-property is unlikely to hold, such as in our context of reparametrizations of linear flows with singularities.

Observe now that if in the proof of the FEJ property, we used that $d(T^t y, T^{t+\iota} y') < \varepsilon$ during a large interval of negative times t instead of positive times, then exactly the same conclusion of extra invariance of λ by $\mathrm{Id} \times T^\iota$ would still hold. The crucial observation here is that it suffices to check one of the two slow divergences, in the future or in the past, and this possibly depending on the pair of points that is considered. This motivates the introduction of what we call the Switchable Ratner, or SWR, property that we now formally define.

Definition 2.1 (*The switchable Ratner property*) Fix $t_0 \in \mathbb{R}_+$ and a compact set $P \subset \mathbb{R} \setminus \{0\}$. One says that the flow has the *switchable $R(t_0, P)$-property* if for every $\epsilon > 0$ and $N \in \mathbb{N}$ there exist $\kappa = \kappa(\epsilon), \delta = \delta(\epsilon, N)$ and a set $Z = Z(\epsilon, N) \in \mathscr{B}$ with $\mu(Z) > 1 - \epsilon$ such that for any $x, y \in Z$ with

$d(x, y) < \delta$, x not in the orbit of y there exist $M = M(x, y)$, $L = L(x, y) \in \mathbb{N}$ with $M, L > N$ and $\frac{L}{M} \geqslant \kappa$ and $p = p(x, y) \in P$ such that

$$\frac{1}{L}\left|\{n \in [M, M + L] \cap \mathbb{Z} : d(T_{nt_0}(x), T_{nt_0+p}(y)) < \epsilon\}\right| > 1 - \epsilon \quad (5)$$

or

$$\frac{1}{L}\left|\{n \in [M, M + L] \cap \mathbb{Z} : d(T_{n(-t_0)}(x), T_{n(-t_0)+p}(y)) < \epsilon\}\right| > 1 - \epsilon. \quad (6)$$

If the set of $t_0 > 0$ such that the flow $(T_t)_{t\in\mathbb{R}}$ has the switchable $R(t_0, P)$-property is uncountable, the flow is said to have *SWR*-property.

For the sake of completeness, we can formally compare the SWR-property with the definition of the WR-property [10]. To have WR-property, we fix $P \subset \mathbb{R} \setminus \{0\}$ and $t_0 \in \mathbb{R}.(T_t^f)_{t\in\mathbb{R}}$ has $R(t_0, P)$ property if in Definition 2.1, (5) holds (the condition (6) is not taken into account) and $(T_t^f)_{t\in\mathbb{R}}$ has WR-property if the set of $t_0 \in \mathbb{R}$ such that $(T_t^f)_{t\in\mathbb{R}}$ has $R(t_0, P)$ property is uncountable. Consequently, SWR-property is weaker than WR-property (and as Theorem 1 shows, it is strictly weaker).

As we just mentioned the proof of the FEJ implication from the SWR property is a direct adaptation of the proof of the same implication in the case of the R property or the WR property. For completeness, we will present a detailed proof of this fact that is stated in Theorem 5 below and that occupies the rest of the section.

As a standing assumption in all the sequel, we will add one more natural condition on the flow $(T_t)_{t\in\mathbb{R}}$ which can be viewed as "continuity" on orbits. The flow $(T_t)_{t\in\mathbb{R}}$ is called *almost continuous* [10] if for every $\epsilon > 0$ there exists a set $X(\epsilon)$ with $\mu(X(\epsilon)) > 1 - \epsilon$ such that for every $\epsilon' > 0$ there exists $\delta' > 0$ such that for every $x \in X(\epsilon)$, we have $d(T_t(x), T_{t'}(x)) < \epsilon'$ for $t, t' \in [-\delta, \delta]$.

Theorem 5 *Let $(T_t)_{t\in\mathbb{R}}$ be a weakly mixing almost continuous flow acting on a probability standard Borel space $(X, \mathcal{B}, \mu)$. Assume that $(T_t)_{t\in\mathbb{R}}$ satisfies the SWR-property. Let $(S_t)_{t\in\mathbb{R}}$ be an ergodic flow acting on a probability standard Borel space $(Y, \mathcal{C}, \nu)$ and let $\rho \in J((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ be an ergodic joining. Then either ρ is equal to $\mu \otimes \nu$ or is a finite extension of the measure ν.*

For the definition and properties of joinings, we refer the reader to [33] or [12]. In the proof of Theorem 5, we will need some lemmas from [10]. But first we state a useful fact that is a simple consequence of the Birkhoff ergodic theorem.

Lemma 2.2 *Let $T, S : (X, \mathscr{B}, \mu) \rightarrow (X, \mathscr{B}, \mu)$ be two ergodic automorphisms and let $A \in \mathscr{B}$. For any $\epsilon, \delta, \kappa > 0$ there exist $N = N(\epsilon, \delta, \kappa)$ and a measurable set $Z = Z(\epsilon, \delta, \kappa)$ with $\mu(Z) > 1 - \delta$ such that for any $M, L \geqslant N$ with $\frac{L}{M} \geqslant \kappa$ and any $x \in Z$ we have*

$$\left| \frac{1}{L} \sum_{i=M}^{M+L} \chi_A(T^i x) - \mu(A) \right| < \epsilon$$

and

$$\left| \frac{1}{L} \sum_{i=M}^{M+L} \chi_A(S^i x) - \mu(A) \right| < \epsilon.$$

The following is a consequence of Lemma 5.4. in [10], that is itself based on the Birkhoff ergodic theorem.

Lemma 2.3 *Let $(T_t)_{t\in\mathbb{R}}$ be a weakly mixing almost continuous flow acting on $(X, \mathscr{B}, \mu)$, and $(S_t)_{t\in\mathbb{R}}$ be another ergodic flow acting on $(Y, \mathscr{C}, \nu)$. Let $\rho \in J((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ be such that ρ is ergodic for automorphisms $T_{t_0} \times S_{t_0}$ for some $t_0 > 0$ (hence, for $T_{-t_0} \times S_{-t_0}$). Let $P \subset \mathbb{R}$ be non-empty and compact. Let $A \in \mathscr{B}$ be such that $\mu(\partial A) = 0$ and $B \in \mathscr{C}$. Then, for every $\epsilon, \delta, \kappa > 0$ there exist a natural number $N = N(\epsilon, \delta, \kappa)$ and a set $Z = Z(\epsilon, \delta, \kappa) \subset \mathscr{B} \otimes \mathscr{C}$ with $\rho(Z) > 1 - \delta$ such that for any $\mathbb{N} \ni M, L \geqslant N$ with $\frac{L}{M} \geqslant \kappa$ and any $p \in P$, we have*

$$\left| \frac{1}{L} \sum_{j=M}^{M+L} \chi_{T_{-p}A \times B}(T_{t_0 j}x, S_{t_0 j}y) - \rho(T_{-p}A \times B) \right| < \epsilon$$

and

$$\left| \frac{1}{L} \sum_{j=M}^{M+L} \chi_{T_{-p}A \times B}(T_{-t_0 j}x, S_{-t_0 j}y) - \rho(T_{-p}A \times B) \right| < \epsilon$$

for every $(x, y) \in Z$.

Proof The proof is a simple consequence of the following lemma:

Lemma 2.4 (Lemma 5.4. in [10]) *Let $(T_t)_{t\in\mathbb{R}}$ be an ergodic almost continuous flow acting on $(X, \mathscr{B}, \mu)$, and $(S_t)_{t\in\mathbb{R}}$ be another ergodic flow acting on $(Y, \mathscr{C}, \nu)$. Let $\rho \in J((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ be such that ρ is ergodic for the automorphism $T_{t_0} \times S_{t_0}$ for some $t_0 > 0$. Let $P \subset \mathbb{R}$ be non-empty and compact. Let*

$A \in \mathscr{B}$ be such that $\mu(\partial A) = 0$ and $B \in \mathscr{C}$. Then, for every $\epsilon, \delta, \kappa > 0$ there exist a natural number $N = N(\epsilon, \delta, \kappa)$ and a set $Z = Z(\epsilon, \delta, \kappa) \subset \mathscr{B} \otimes \mathscr{C}$ with $\rho(Z) > 1 - \delta$ such that for any $\mathbb{N} \ni M, L \geqslant N$ with $\frac{L}{M} \geqslant \kappa$ and any $p \in P$, we have

$$\left| \frac{1}{L} \sum_{j=M}^{M+L} \chi_{T_{-p}A \times B}(T_{t_0 j}x, S_{t_0 j}y) - \rho(T_{-p}A \times B) \right| < \epsilon$$

for every $(x, y) \in Z$.

One uses the above lemma first for the flows $(T_t)_{t\in\mathbb{R}}$ and $(S_t)_{t\in\mathbb{R}}$ and ergodic joining $\rho \in J((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ to get, for $\epsilon, \frac{\delta}{2}, \kappa > 0$, a natural number $N_+ \in \mathbb{N}$ and a set $Z_+ \subset \mathscr{B} \otimes \mathscr{C}$ with $\rho(Z_+) > 1 - \frac{\delta}{2}$. Then, for flows $(T_{-t})_{t\in\mathbb{R}}$ and $(S_{-t})_{t\in\mathbb{R}}$ and the same ergodic joining ρ to get, for $\epsilon, \frac{\delta}{2}, \kappa > 0$, a natural number $N_- \in \mathbb{N}$ and a set $Z_- \subset \mathscr{B} \otimes \mathscr{C}$ with $\rho(Z_-) > 1 - \frac{\delta}{2}$. To finish the proof one takes $N := \max(N_+, N_-)$ and $Z = Z_+ \cap Z_-$. $\qquad\square$

The next lemma is used in the proof of Theorem 3 in [27].

Lemma 2.5 *Let* $(T_t)_{t\in\mathbb{R}}$ *and* $(S_t)_{t\in\mathbb{R}}$ *be two ergodic flows. Let* $\rho \in J^e((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ *be an ergodic joining. Then if there exists a set* V *with* $\rho(V) > 0$ *such that for any points* $(x, y), (x', y) \in V$ *either* x *is in the orbit of* x' *or* $d(x, x') > c_0$ *for some constant* $c_0 > 0$, *then* ρ *is a finite extension of* v.

In what follows, we consider only (X, d) be a σ-compact metric space. Let $A \in \mathscr{B}$. For $\eta > 0$ we denote by $V_\eta(A) := \{x \in X : d(x, A) < \eta\}$.

Lemma 2.6 (cf. [10]) *For any* $A \in \mathscr{B}$ *there exists* $R \subset (0, +\infty)$ *such that* $(0, +\infty) \setminus R$ *is countable and* $\mu(\partial V_\eta(A)) = 0$ *for* $\eta \in R$. *It particular, there exists a dense family* $(B_i)_{i\geqslant1}$ *in* $\mathscr{B}$ *with the property* $\mu(\partial B_i) = 0$ *for every* $i \in \mathbb{N}$.

Remark 2.7 Since (X, d) is a Polish space, by Lemma 2.6 and regularity of μ, there exists a dense family $\{B_i\}_{i\geqslant1} \in \mathscr{B}$, such that $\mu(\partial B_i) = 0$ for $i \geqslant 1$.

Proof of Theorem 5. Let $\rho \in J((T_t)_{t\in\mathbb{R}}, (S_t)_{t\in\mathbb{R}})$ be an ergodic joining and $\rho \neq \mu \times v$. Assume that $(T_t)_{t\in\mathbb{R}}$ has the switchable $R(t_0, P)$-property and ρ is ergodic for $T_{t_0} \times S_{t_0}$ (then ρ is ergodic for $T_{-t_0} \times S_{-t_0}$). Such $t_0 > 0$ always exists because an ergodic flow can have at most countably many non-ergodic time automorphisms and, by assumptions, the property $R(t_0, P)$ is satisfied for uncountably many $t_0 > 0$. For simplicity of notation, we assume $t_0 = 1$. Let $\{B_i\}_{i\geqslant1}$ and $\{C_i\}_{i\geqslant1}$ be two countable dense families in the σ-algebras $\mathscr{B}$ and $\mathscr{C}$, respectively such that for every $i \geqslant 1$, B_i is an open set and we

have $\mu(\partial B_i) = 0$ (such $\{B_i\}_{i \geqslant 1}$ always exists by Remark 2.7). Consider the following real function:

$$\mathbb{R} \ni t \to k(t) := \sum_{i,j \geqslant 1} (1/2^{i+j})|\rho(T_t(B_i) \times C_j) - \rho(B_i \times C_j)|.$$

As in Lemma 5.4. in [10], we conclude that $k : \mathbb{R} \to \mathbb{R}$ is a continuous function and for any $t \in \mathbb{R}$, $k(t) > 0$. Indeed, it follows by the fact that if for some $r \in \mathbb{R} \setminus \{0\}$ we have for any $i, j \in \mathbb{N} \rho(T_r(B_i) \times C_j) = \rho(B_i \times C_j)$ then ρ is product measure (recall that $(T_t)_{t \in \mathbb{R}}$ is assumed to be weak mixing hence every time r of the flow is ergodic).

The set $P \subset \mathbb{R} \setminus \{0\}$ is compact, therefore there exists $\epsilon > 0$ such that $k(p) > \epsilon$ for any $p \in P$. It follows by the definition of the function k that there exists a number $R := R(\epsilon)$ such that

$$\sum_{i,j \geqslant 1}^{R} (1/2^{i+j})|\rho(T_p(B_i) \times C_j) - \rho(B_i \times C_j)| > \epsilon/2$$

for every $p \in P$. Therefore, for every $p \in P$, there exist $1 \leqslant i, j \leqslant R$ such that $|\rho(T_p(B_i) \times C_j) - \rho(B_i \times C_j)| > \epsilon$.

By Lemma 2.6 and the fact that for $0 \leqslant i \leqslant R$, B_i is open, there exists $\epsilon' < \frac{\epsilon}{8}$ such that for every $1 \leqslant i \leqslant R$

$$\mu(V_{\epsilon'}(B_i) \setminus B_i) < \epsilon, \quad \text{and} \quad \mu(\partial V_{\epsilon'}(B_i)) = 0.$$

It follows by the fact that ρ is a joining that

$$|\rho(V_{\epsilon'}(B_i) \times C_j) - \rho(B_i \times C_j)| < \frac{\epsilon}{2} \quad \text{and} \quad |\rho(S_{-t}V_{\epsilon'}(B_i) \times C_j)$$
$$-\rho(S_{-t}B_i \times C_j)| < \frac{\epsilon}{2}, \tag{7}$$

for $1 \leqslant i, j \leqslant R$ and every $t \in \mathbb{R}$. By the switchable $R(1, P)$-property, let $\kappa := \kappa(\epsilon')$. By Lemma 2.2 applied to $\frac{\epsilon}{8}$, $\frac{1}{8}$, κ, the sets $V_{\epsilon'}(B_i) \times C_j$, $1 \leqslant i, j \leqslant R$, and to automorphisms $T_1 \times S_1$ and $T_{-1} \times S_{-1}$, we get $N_1 \in \mathbb{N}$ and a set $U_1 \in \mathscr{B} \otimes \mathscr{C}$ with $\rho(U_1) > \frac{7}{8}$, such that for every $L, M \geqslant N_1$ with $\frac{L}{M} \geqslant \kappa$ and every $(x, y) \in U_1$, we have

$$\left| \frac{1}{L} \sum_{k=M}^{M+L} \chi_{V_{\epsilon'}(B_i) \times C_j}(T^k x, S^k y) - \rho(V_{\epsilon'}(B_i) \times C_j) \right| < \frac{\epsilon}{8} \tag{8}$$

$$\left| \frac{1}{L} \sum_{k=M}^{M+L} \chi_{V_{\epsilon'}(B_i) \times C_j}(T^{-k} x, S^{-k} y) - \rho(V_{\epsilon'}(B_i) \times C_j) \right| < \frac{\epsilon}{8}. \tag{9}$$

Next, by Lemma 2.3 applied to $\frac{\epsilon}{8}, \frac{1}{8}, \kappa > 0$ and the sets $B_i \times C_j$, $1 \leqslant i, j \leqslant R$, there exist $N_2 \in \mathbb{N}$ and a set $U_2 \subset \mathscr{B} \otimes \mathscr{C}$ with $\rho(U_2) > \frac{7}{8}$ such that for every $L, M \geqslant N_2$ with $\frac{L}{M} \geqslant \kappa$ and any $p \in P$, we have

$$\left| \frac{1}{L} \sum_{k=M}^{M+L} \chi_{T_{-p}B_i \times C_j}(T_k x, S_k y) - \rho(T_{-p}B_i \times C_j) \right| < \frac{\epsilon}{8} \qquad (10)$$

and

$$\left| \frac{1}{L} \sum_{k=M}^{M+L} \chi_{T_{-p}B_i \times C_j}(T_{-k} x, S_{-k} y) - \rho(T_{-p}B_i \times C_j) \right| < \frac{\epsilon}{8}. \qquad (11)$$

It follows that if we set $N_0 := \max(N_1, N_2)$ and $U_0 := U_1 \cap U_2$, then $\rho(U_0) > \frac{1}{2}$ and for every $L, M \geqslant N_0$ with $\frac{L}{M} \geqslant \kappa$, any $p \in P$, the Eqs. (8), (9), (10), (11) are satisfied for every $(x, y) \in U_0$. Using the switchable $R(1, P)$-property with $\epsilon' > 0$ and $N_0 \in \mathbb{N}$, we obtain $\delta = \delta(\epsilon', N_0)$ and $Z = Z(\epsilon', N_0)$ with $\mu(Z) > 1 - \epsilon'$. Now, we will use Lemma 2.5 with the set $U := U_0 \cap (Z \times Y)$ (then of course $\rho(U) > \frac{1}{4}$) and $\delta_0 = \delta(\epsilon', N_0)$ to prove that for every $(x, y), (x', y) \in U, d(x, x') \geqslant \delta_0$. Assume on the contrary that $d(x, x') < \delta_0$. Then by the switchable R(1,P)-property, there exist $L_0, M_0 > N_0$ with $\frac{L_0}{M_0} \geqslant \kappa$ and $p \in P$ such that

$$\frac{1}{L_0} \left| \{ n \in [M_0, M_0 + L_0] : d(T_n(x), T_{n+p}(x')) < \epsilon' \} \right| > 1 - \epsilon'$$

or

$$\frac{1}{L_0} \left| \{ n \in [M_0, M_0 + L_0] : d(T_{-n}(x), T_{-n+p}(x')) < \epsilon' \} \right| > 1 - \epsilon'.$$

Assume that the first inequality is satisfied. We will use Eqs. (8) and (10) (in case the second one is satisfied, we use Eqs. (9) and (11)). Let $1 \leqslant i_p, j_p \leqslant R$ be the numbers which satisfy $|\rho(T_p(B_{i_p}) \times C_{j_p}) - \rho(B_{i_p} \times C_{j_p})| > \epsilon$. Let $K = K(x, x', p) := \{ n \in [M_0, M_0 + L_0] : d(T_n(x), T_{n+p}(x')) < \epsilon' \}$. It follows that if $k \in K$ and $T_{k+p} x' \in B_i$ then $T_k x \in V_{\epsilon'}(B_i)$. Therefore

$$\rho(T_{-p}B_{i_p} \times C_{j_p}) \leqslant \frac{1}{L_0} \sum_{k=M_0}^{M_0+L_0} \chi_{T_{-p}B_{i_p} \times C_{j_p}}(T^k x', S^k y) + \frac{\epsilon}{8}$$

$$\leqslant \frac{\epsilon' L_0}{L_0} + \frac{1}{L_0} \sum_{k=M_0}^{M_0+L_0} \chi_{V_{\epsilon'}(B_{i_p}) \times C_{j_p}}(T^k x, S^k y) + \frac{\epsilon}{8}$$

$$\leqslant \frac{\epsilon}{2} + \rho(V_{\epsilon'}(B_{i_p}) \times C_{j_p}) < \epsilon + \rho(B_{i_p} \times C_{j_p}). \quad (12)$$

A similar arguments show that $\rho(B_{i_p} \times C_{j_p}) < \epsilon + \rho(T_{-p}B_{i_p} \times C_{j_p})$ and consequently, $|\rho(B_{i_p} \times C_{j_p}) - \rho(T_{-p}B_{i_p} \times C_{j_p})| < \epsilon$. This contradicts our assumption that $|\rho(T_p(B_{i_p}) \times C_{j_p}) - \rho(B_{i_p} \times C_{j_p})| > \epsilon$ is satisfied. Therefore, for any $(x, y), (x', y) \in U$ we have $d(x, x') \geqslant \delta_0$ and an application of Lemma 2.5 completes the proof. □

3 SWR-property for special flows

In this section, we will prove a sufficient condition for SWR-property in the case of special flows over an ergodic isometry. We start by recalling the definition of special flows. Let T be an automorphism $(X, \mathscr{B}, \mu)$. Let $f \in L^1(X, \mu)$ such that $f > 0$. The special flow $(T_t^f)_{t \in \mathbb{R}}$ defined above T and under the ceiling function f is given by

$$X \times \mathbb{R}/\sim \to X \times \mathbb{R}/\sim$$
$$(x, s) \to (x, s + t),$$

where $\sim$ is the identification

$$(x, s + f(x)) \sim (T(x), s) \tag{13}$$

Equivalently the flow $(T_t^f)_{t \in \mathbb{R}}$ is defined for $t + s \geqslant 0$ (with a similar definition for negative times) by

$$T_t^f(x, s) = (T^n x, t + s - f^{(n)}(x))$$

where n is the unique integer such that

$$f^{(n)}(x) \leqslant t + s < f^{(n+1)}(x) \tag{14}$$

and

$$f^{(n)}(x) = \begin{cases} f(x) + \cdots + f(T^{n-1}x) & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -(f(T^n x) + \cdots + f(T^{-1}x)) & \text{if } n < 0. \end{cases}$$

If T preserves a unique probability measure μ then the special flow will preserve a unique probability measure that is the normalized product measure of μ on the base and the Lebesgue measure on the fibers. If X is a metric space with a metric d, so is X^f with the metric $d^f((x, s), (x', s')) := d(x, x') + |s - s'|$. Moreover, it is easy to show that if $(T_t^f)_{t \in \mathbb{R}}$ is a special flow acting on X^f, then $(T_t^f)_{t \in \mathbb{R}}$ is almost continuous (see Sect. 2) with $X(\epsilon) = \{(x, s) \in X^f : x \in X, \epsilon < s < f(x) - \epsilon\}$.

The following general lemma is a direct consequence of Birkhoff ergodic theorem.

Lemma 3.1 *Let T be an ergodic automorphism $(X, \mathcal{B}, \mu)$. Let $f \in L^1(X, \mu)$, $\int_X f \, d\mu \neq 0$. For every $\epsilon, \kappa > 0$ there exist $N = N(\epsilon, \kappa)$ and a set $A = A(\epsilon, \kappa)$ with $\mu(A) > 1 - \epsilon$ such that for every $M \geqslant N$*

$$\left| \frac{1}{M} \sum_{i=1}^{M} f(T^i x) - \int_X f \, d\mu \right| \leqslant \frac{\kappa}{3} \left| \int_X f \, d\mu \right| \qquad (15)$$

for every $x \in A$.

Remark 3.2 Assume that additionally f is positive and bounded away from zero. Fix $\epsilon, \kappa > 0 (\kappa < |\int_X f \, d\mu| < 1/2)$. It follows that there are constants $r_1, r_2 > 0$ such that if we take $x \in A$ then for any $M, L \geqslant N$ with $\frac{L}{M} \geqslant \kappa$, we have $r_1 < \frac{f^{(M)}(x)}{M} < r_2$ and

$$r_1 < \frac{(1 - \frac{\kappa}{3}) \int f \, d\mu (M + L) - (1 + \frac{\kappa}{3}) \int f \, d\mu \cdot M}{L}$$

$$\leqslant \frac{f^{(M+L)}(x) - f^{(M)}(x)}{M} < r_2.$$

We now state the main result of this section. It is similar to Lemma 6 of [9].

Proposition 3.3 *Let $T : (X, d) \to (X, d)$ be an ergodic isometry and $f \in L^1(X, \mathcal{B}, \mu)$ a positive function bounded away from zero. Let $(T_t^f)_{t \in \mathbb{R}}$ be the corresponding special flow. Let $P \subset \mathbb{R} \setminus \{0\}$ be a compact set. Assume that for every $\epsilon > 0$ and $N \in \mathbb{N}$ there exist $\kappa = \kappa(\epsilon), \delta = \delta(\epsilon, N)$ and a set $X' = X'(\epsilon, N)$ with $\mu(X') > 1 - \epsilon$, such that for any $x, y \in X'$ with $0 < d(x, y) < \delta$ there exist $\mathbb{N} \ni M = M(x, y), L = L(x, y)$ with $M, L \geqslant N, \frac{L}{M} \geqslant \kappa$ and $p = p(x, y) \in P$ such that*

$$|f^{(n)}(x) - f^{(n)}(y) - p| < \epsilon \text{ for every } n \in [M, M + L] \qquad (16)$$

or

$$|f^{(-n)}(x) - f^{(-n)}(y) - p| < \epsilon \text{ for every } n \in [M, M + L]. \qquad (17)$$

If $\gamma > 0$ is such that the automorphism T_γ^f is ergodic, then $(T_t^f)_{t \in \mathbb{R}}$ has the switchable $R(\gamma, P)$-property. Consequently, $(T_t^f)_{t \in \mathbb{R}}$ has the SWR-property.

Proof Fix $\gamma > 0$ such that T_γ^f is ergodic. Fix also $\frac{1}{\|f\|_{L^1}} > 4\epsilon > 0$. Apply Remark 3.2 with the constants $\epsilon/4, \kappa$ to f and T, T^{-1}, respectively to obtain

constants $D_1, D_2 > 0$ such that for $x \in A$, $\mu(A) > 1 - \epsilon/2$ (the set A is the intersection of two relevant sets), we have

$$D_1 < \frac{f^{(M)}(x)}{M}, \frac{f^{(M+L)}(x) - f^{(M)}(x)}{L}, \frac{f^{(-M)}(x)}{-M},$$
$$\frac{f^{(-M-L)}(x) - f^{(-M)}(x)}{-L} < D_2. \tag{18}$$

Fix $N > \frac{2}{D_2\epsilon^2}$. Let $\epsilon' := \min(\frac{D_1\epsilon}{8(\gamma+D_2)}, \frac{\epsilon}{16})$. Let $\kappa' := \frac{D_1}{D_2}\kappa(\epsilon')$. Let us consider the set $X(\epsilon)$ on which $(T_t^f)_{t\in\mathbb{R}}$ is $\frac{\epsilon}{8}$- "almost continuous", that is

$$X(\epsilon) := \left\{ (x, s) \in X^f : \frac{\epsilon}{8} < s < f(x) - \frac{\epsilon}{8} \right\}.$$

Now, we will use ergodicity of T_γ^f and $T_{-\gamma}^f$. It follows that there exist $N_0 := N(\epsilon)$ and a set $Z := Z(\epsilon)$ with $\mu^f(Z) > 1 - \frac{\epsilon}{2}$ and for every $(x, s) \in Z$ and $n \geqslant N_0$

$$\left| \frac{1}{n} \sum_{k=1}^n \chi_{X(\epsilon)} T_{ki}^f(x, s) - \left(1 - \frac{\epsilon}{4} \right) \right| < \frac{\kappa}{\kappa+1} \frac{\epsilon}{8} \tag{19}$$

for $i = \gamma, -\gamma$. Moreover, since $f \in L^1(X, \mathcal{B}, \mu)$, there exists a set $V = V(\epsilon) \subset X$ with $\mu(V) > 1 - \frac{\epsilon}{2}$ and such that for every $x \in V$, $f(x) < \frac{2}{\epsilon^2}$. Define the set $Z' := Z \cap \{(x, s) \in X^f : x \in V\} \cap \{(x, s) \in X^f : x \in A\}$, then $\mu^f(Z') > 1 - \epsilon$.

Let $\delta' := \delta(\epsilon', 2\gamma\frac{\max(N_0,N)}{D_1})$. Take two points $(x, s), (x', s') \in Z'$, such that $x \neq x'$ and $d^f((x, s), (x', s')) < \delta'$. It follows by definition of d^f that $d(x, x') < \delta'$ and therefore by our assumptions there exist $M, L \geqslant 2\gamma\frac{\max(N_0,N)}{D_1}$ with $\frac{L}{M} \geqslant \kappa$, $p \in P$ and such that for all $n \in [M, M+L]$ either $|f^{(n)}(x) - f^{(n)}(y) - p| < \epsilon'$ or for all $n \in [M, M+L]$, $|f^{(-n)}(x) - f^{(-n)}(y) - p| < \epsilon'$. We will consider the second case (the proof in the first case goes along the same lines).

Let us define

$$M' := \frac{f^{(-M)}(x) - s}{-\gamma} \quad \text{and} \quad L' := \frac{f^{(-M-L)}(x) - f^{(-M)}(x)}{-\gamma}.$$

By (18) it follows that $L' = \frac{f^{(-L-M)}(x) - f^{(-M)}(x)}{-L} \frac{-L}{-\gamma} > \frac{-LD_1}{-\gamma} > N$. Similarly, $\frac{f^{(-M)}(x) - s}{-\gamma} > \frac{f^{(-M)}(x)}{-\gamma} > \frac{MD_1}{\gamma}$, so $M' > N$. Moreover, since $(x, s) \in Z'$, $s < \frac{2}{\epsilon^2} < ND_2 \leqslant MD_1D_2/(2\gamma)$ (by the choice of N) and therefore

$$\frac{L'}{M'} \geqslant \frac{LD_1}{\gamma} \frac{-\gamma}{f^{(-M)}(x) - s} \geqslant \frac{LD_1}{MD_2} \geqslant \kappa'.$$

It follows by the properties of $M', L' \in \mathbb{N}$ that if $(x, s) \in Z' \subset Z$ we have

$$\left| \frac{1}{L'} \sum_{k=M'}^{M'+L'} \chi_{X(\epsilon)} T_{-k\gamma}^f - (1 - \epsilon) \right| < \frac{\epsilon}{2}. \tag{20}$$

Take any $k \in [M', M' + L']$ such that $T_{-k\gamma}^f \in X(\epsilon)$ it follows that there exist a number $m_k \in [M, M + L]$ such that $T_{-k\gamma}^f(x, s) = (T^{m_k}x, -k\gamma + s - f^{(-m_k)}(x))$, where, by the fact that $T_{-k\gamma}^f \in X(\epsilon)$, $f^{(-m_k-1)}(x) + \frac{\epsilon}{8} < -k\gamma + s < f^{(-m_k)}(x) - \frac{\epsilon}{8}$. Using additionally the inequality $|s - s'| < \delta'$ we hence obtain

$$f^{(-m_k-1)}(x') \leqslant f^{(-m_k-1)}(x) + p - \epsilon' \leqslant f^{(-m_k-1)}(x) + p + \frac{\epsilon}{8} - \delta'$$
$$< -k\gamma + s' + p.$$

A similar reasoning shows that

$$-k\gamma + s' + p < -k\gamma + s + p + \delta \leqslant f^{(-m_k)}(x) + p - \frac{\epsilon}{8} + \delta \leqslant f^{(-m_k)}(x').$$

Therefore, by the definition of the special flow, we have $T_{-k\gamma+p}^f(x', s') = (T^{m_k}x', -k\gamma + s' + p - f^{(-m_k)}(x'))$. Consequently,

$$d^f(T_{-k\gamma}^f(x, s), T_{-k\gamma+p}^f(x', s')) = d^f((x, s), (x', s')) + |f^{(-m_k)}(x)$$
$$- f^{(-m_k)}(x') - p| < \epsilon.$$

Now, the number of $k \in [M', M' + L']$ such that $T_{-k\gamma}^f \in X(\epsilon)$ is, by (20), at least $(1-\epsilon)L'$ and for any such k we get that $d^f(T_{-k\gamma}^f(x, s), T_{-k\gamma+p}^f(x', s')) < \epsilon$. Hence

$$\frac{1}{L'} \left| \left\{ k \in [M', M' + L'] \ : \ d^f(T_{-k\gamma}^f(x, s), T_{-k\gamma+p}^f(x', s')) < \epsilon \right\} \right| > 1 - \epsilon.$$

This gives us the switchable $R(\gamma, P)$-property. Note that since the flow $(T_t^f)_{t \in \mathbb{R}}$ is ergodic, then the set of $\eta \in \mathbb{R}$ such that T_η^f is not ergodic, is at most countable and therefore, as a direct consequence of Proposition 3.3, we get that $(T_t^f)_{t \in \mathbb{R}}$ enjoys SWR-property. $\qquad \square$

4 SWR-property for smooth special flows with singularities

In this section we will use Proposition 3.3 to prove SWR-property for special flows given by the assumptions in Theorem 2 and Theorem 4. In all the sequel we assume $\{a_1, \ldots, a_k\}$ are badly approximable by α with a constant $C > 1$ (see Definition 1.3). We start with an easy combinatorial fact about the visits of an orbit by the rotation R_α to the neighborhood of the singularities.

Lemma 4.1 *Let $s \in \mathbb{N}$ be such that $q_{s+1} > 4Cq_s$ and $x \in \mathbb{T}$. Then*

$$\{x + j\alpha\}_{j=0}^{[\frac{q_{s+1}}{4C}]} \cap \bigcup_{i=1}^{k} \left[\frac{-1}{4Cq_s} + a_i, a_i + \frac{1}{4Cq_s} \right] \subset \{x + rq_s + i_0\alpha\}_{r=0}^{[\frac{q_{s+1}}{4Cq_s}]},$$

where $i_0 \in \{0, \ldots, q_s - 1\}$ is such that $\rho(\{x + v\alpha\}_{v=0}^{q_s-1}, \{a_i\}_{i=1}^{k}) = \rho(x + i_0\alpha, \{a_i\}_{i=1}^{k})$. For finite sets $A, B \subset \mathbb{T}$, we use the notation $\rho(A, B) = \min_{a \in A, b \in B} \|a - b\|$.

Proof Observe that for every $0 < j < q_s - 1$, $j \neq i_0$ and every $r = 0, \ldots, [\frac{q_{s+1}}{4Cq_s}]$ we have

$$\|x + rq_s + j\alpha - (x + j\alpha)\| = \|rq_s\alpha\| \leqslant r\|q_s\alpha\| \leqslant \frac{1}{4Cq_s}. \qquad (21)$$

Moreover, since $\{a_i\}_{i=1}^{k}$ are badly approximable [see (3)] and by the definition of i_0 we get

$$\{x + j\alpha\}_{j=0}^{q_s-1} \cap \bigcup_{i=1}^{k} \left[\frac{-1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s} \right] \subset \{x + i_0\alpha\}. \qquad (22)$$

Therefore, for $j \neq i_0$ and $r = 0, \ldots, [\frac{q_{s+1}}{4Cq_s}]$, by (21) and (22)

$$\rho(x + j\alpha + rq_s, \{a_i\}_{i=1}^{k}) \geqslant \rho(x + j\alpha, \{a_i\}_{i=1}^{k}) - \frac{1}{4Cq_s} \geqslant \frac{1}{4Cq_s}.$$

This finishes the proof. $\square$

We will often use the Denjoy-Koksma inequality to control the growth of the Birkhoff sums. For a reference, see for example [4].

Proposition 4.2 (Denjoy-Koksma inequality) *Let $f : \mathbb{T} \to \mathbb{R}$ be a function of bounded variation. Then*

$$\left| \sum_{k=0}^{q_n-1} f(x + k\alpha) - q_n \int_{\mathbb{T}} f \, d\lambda \right| \leqslant Var f,$$

for every $x \in \mathbb{T}$ and $n \in \mathbb{N}$.

The following lemma is a simple consequence of the Denjoy-Koksma inequality. It will be very useful in separating the contribution to the shear of the visits to the neighborhood to the singularities from the rest of the orbit.

Lemma 4.3 *Let $h \in C^2(\mathbb{T} \setminus \{0\})$ be positive and decreasing on $(0, 1)$ with h' is increasing on $(0, 1)$ and $\lim_{x \to 0^+} h(x) = \lim_{x \to 0^+}(-h'(x)) = +\infty$. Denote by $c_0 := \inf_{\mathbb{T}} h$. Then for every $x \in \mathbb{T}$ and $s \in \mathbb{N}$ we have the following estimates:*

$$-q_s \left(h \left(\frac{1}{2q_s} \right) - c_0 \right) - 2h' \left(\frac{1}{2q_s} \right) > h'^{(q_s)}(x) \geqslant h'(x + j\alpha)$$

$$-q_s \left(h \left(\frac{1}{2q_s} \right) - c_0 \right) + 2h' \left(\frac{1}{2q_s} \right)$$

where $j \in \{0, \ldots, q_s - 1\}$ is such that $\min_{\ell \in \{0, \ldots, q_s - 1\}} |x + \ell\alpha| = x + j\alpha$.

Proof Fix $s \in \mathbb{N}$. Consider

$$\bar{h}(x) = \begin{cases} 0, & \text{if } x \in \left[0, \frac{1}{2q_s} \right) \\ h'(x), & \text{otherwise.} \end{cases}$$

Then $\bar{h} \in BV(\mathbb{T})$ and we use the Denjoy-Koksma inequality to obtain $|\bar{h}^{(q_s)}(x) - q_s \int_{\mathbb{T}} \bar{h}(t)d\lambda| < \text{Var } \bar{h}$. But $\int_{\mathbb{T}} \bar{h}(t)d\lambda = h(\frac{1}{2q_s}) - c_0$ and $\text{Var } \bar{h} \leqslant -2h'(\frac{1}{2q_s})$. Moreover, by (2), the set $\{x + r\alpha\}_{r=0}^{q_s-1} \cap [0, \frac{1}{2q_s}]$ is at most a singleton. We then finish since $h'^{(q_s)}(x) = \bar{h}^{(q_s)}(x) + h'(x+j\alpha)\chi_{[0, \frac{1}{2q_s}]}(x+j\alpha)$ and $h' < 0$. $\square$

Lemma 4.4 *Let $f \in C^2(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$. Assume that for $i = 1, \ldots, k$, $\lim_{x \to a_i^+} \frac{f'(x)}{r_i(x-a_i)}$ and $\lim_{x \to a_i^-} \frac{f'(x)}{r_i(a_i-x)}$ exist and are finite, where $0 \leqslant r_i \in C^2(\mathbb{T} \setminus \{0\})$ is decreasing on $(0, 1)$ with r_i' increasing on $(0, 1)$. Then there exists a constant $H > 0$ such that*

$$|f'(x)| < H \left(\sum_{i=1}^{k} -r_i'(x - a_i) - r_i'(a_i - x) \right),$$

for each $x \in \mathbb{T}$.

Proof By assumptions, there exists a constant $z_0 > 0$ such that for every $i = 1, \ldots, k$ and for every $x \in [-z_0 + a_i, a_i) |f'(x)| < -Kr'_i(a_i - x)$ and for every $x \in (a_i, a_i + z_0]$; $|f'(x)| < -Kr'_i(x - a_i)$ for some constant $K \geqslant 0$.

Moreover, since $f' \in C^1(\mathbb{T} \setminus \{a_1, \ldots, a_k\})$; it follows that there exists a constant $R > 0$ such that for every $x \in \mathbb{T} \setminus \bigcup_{i=1}^{k}[-z_0 + a_i, a_i + z_0]$, $|f'(x)| < R$. Denote by $C_0 := \min_{i=1,\ldots,k} |\sup_{\mathbb{T}} r'_i|$ Now, the constant $H := 2\max_{i=1,\ldots,k}\{K, \frac{R}{C_0}\}$ satisfies the assertion of the lemma. $\qquad\square$

4.1 Proof of Theorem 2

4.1.1 Outline of the proof.

We first give an outline of the proof in which we suppose for simplicity that the ceiling function has just one singularity that is exactly logarithmic. The SWR-property (in the same way as the original Ratner's property) consists of two parts. First, for two "nearby" points, we need to show that their orbits drift apart by a controlled amount at some time M (M may be positive or negative in the case of the SWR property). Second, we need to make sure that their orbits keep essentially the same drift during an interval of time comparable to M. For special flows over rotations (or IET's) we gave in Proposition 3.3 a characterization of the SWR property based on the divergence for nearby points between the corresponding Birkhoff sums of the ceiling function f. The latter is controlled by the Birkhoff sums of the derivative of the ceiling function that were the subject of investigation of several other works (see e.g. [8,17,19,20,32,35]).

In Proposition 4.6 Part a, we prove the first part of the SWR property. Precisely, we want to see a macroscopic yet controlled drift between the Birkhoff sums of two points $1/(q_s \ln q_s) \geqslant d(x, y) \geqslant 1/(q_{s+1} \ln q_{s+1})$ for every s sufficiently large. We explain now the proof showing also how the argument simplifies if $K_\alpha = \{n \in \mathbb{N} : q_{s+1} < q_s(\ln q_s)^{\frac{7}{8}}\}$ contains all the integers after some n_0, in particular if α is of constant type.

a. There exists $c > 0$ such that for any point $x \in \mathbb{T}$ either $R_\alpha^i x$ is disjoint from $[-c/q_s, c/q_s]$ for every $i = 0, \ldots, l < \frac{q_s}{2}$ or $R_\alpha^{-i} x$ is disjoint from $[-c/q_s, c/q_s]$ for every $i = 1, \ldots, l < \frac{q_s}{2}$.

b. By the Denjoy-Koksma inequality, if the forward (backward) orbit of length q_s a point x by R_α is disjoint with the $1/(q_s \ln q_s)$ neighborhood of the origin we have that $|f'^{(q_s)}(x)|$ (or $|f'^{(-q_s)}(x)|$) is of order $q_s \ln q_s$ (this is easy to understand if we rearrange the orbit of x almost like c/q_s, $(c + 1)/q_s$, $(c + 2)/q_s$... with $c \in (0, 1)$ far from 0 and from 1). Hence for any point x either $|f'^{(q_s)}(x)|$ or $|f'^{(-q_s)}(x)|$ is of order $q_s \ln q_s$. This is the

content of Lemma 4.8. In case $q_{s+1} \leqslant Cq_s$, this would give the macroscopic yet controlled drift between the points x, y such that $1/(q_{s+1} \ln q_{s+1}) \leqslant d(x, y) \leqslant 1/(q_s \ln q_s)$ either in the future at time q_s or in the past at time $-q_s$. So, if α is of constant type, we would be done with the proof of Proposition 4.6 Part a.

c. In the case $q_{s+1} \gg q_s$ but $s \in K_\alpha$, we may have to consider the Birkhoff sums beyond q_s to see the drift between the orbits of x and y such that $1/(q_{s+1} \ln q_{s+1}) \leqslant d(x, y) \ll 1/(q_s \ln q_s)$. Since $s \in K_\alpha$, we have that $1/(q_s \ln q_s) \ll c/q_{s+1}$ hence a. applied to $s + 1$ implies that either up to $q_{s+1}/2$ in the future or up to $-q_{s+1}/2$ in the past the orbit of x by R_α does not enter the $1/(q_s \ln q_s)$ neighborhood of the origin (see Lemma 4.7, case $m \in K_\alpha$). Using this, the estimate of b. and the Denjoy-Koksma inequality we get by the cocycle identity that $f'^{(\iota kq_s)}(x)$ behaves like $kq_s \ln q_s$ for s sufficiently large, $k \in [1, O(q_{s+1}/q_s)]$ and $\iota = 1$ or -1 (see Lemma 4.9). As a consequence there exists a time of the form $n_0 q_s$, $n_0 \in [1, O(q_{s+1}/q_s)]$ such that $f^{(n_0 q_s)}(x) - f^{(n_0 q_s)}(y)$ is in some compact set P away from 0, which finishes the proof of Proposition 4.6 Part a in the case K_α contains all sufficiently large integers.

d. In the case $s \notin K_\alpha$, we define $x_s := 1/(q_s (\ln q_s)^{7/8})$ and use our arithmetic condition $\alpha \in \mathcal{E}$ to define a set Z of almost full measure (see definition of Z in Sect. 4.1.2) such that $R_\alpha^i x$ does not enter the x_s neighborhood of the origin (see Definition (25)) for every $i \in \{-q_s, \ldots, 0, \ldots, q_s - 1\}$ for every $s \notin K_\alpha$ sufficiently large. This actually implies that either up to $\frac{q_{s+1}}{4C}$ in the future or up to $-\frac{q_{s+1}}{4C}$ (C is a constant coming from Definition 1.3; in case there is only one singularity $C = 1$) in the past the orbit of x by R_α does not enter the x_s neighborhood of the origin (see Lemma 4.7, case $m \notin K_\alpha$). From there the proof of Proposition 4.6 Part a is similar to case 3. above except that the condition 3 of Theorem 2, namely $h(\frac{1}{2q_s})/h(\frac{1}{2q_{s+1}}) > m_0$ is used to show that a stretch of order $n_0 q_s \ln q_s$ where n_0 can be taken to be as large as $O(q_{s+1}/q_s)$ is sufficient to produce a drift between the point x, y such that $d(x, y)$ is comparable to $1/(q_{s+1} \ln q_{s+1})$. This is where the Diophantine condition 3. of Theorem 2 is crucial.

Observe that d. is the only part where we used that in the definition of SWR, it is allowed to discard a small measure set of pairs (x, y) for which the property will not be checked.

Note however that in all the cases above, it is crucial to use the possibility to control the drift in the future or in the past depending on the pair of points.

The second part (keeping the drift) is proved in Proposition 4.6 Part b. We need to consider the points $R_\alpha^{n_0 q_s} x$ and $R_\alpha^{n_0 q_s} y$ and apply similar arguments as in Part a to bound the drift during time $\kappa n_0 q_s$. The main ingredient is Lemma 4.10 which is another lemma that allows us to bound the drift between the Birkhoff sums in the future (or in the past) up to a time comparable to q_{m+1}

for points that stay away from the singularities in the future (or in the past) during this time. We then have to "situate" $\kappa n_0 q_s$ relatively to the denominators of α and check that the conditions of Lemma 4.10 are satisfied by $R_\alpha^{n_0 q_s} x$ and $R_\alpha^{n_0 q_s} y$. Of course if s is such that $q_{s+1} \gg q_s$ (for example if $s \notin K_\alpha$) and if $1/(q_{s+1} \ln q_{s+1}) \ll d(x, y) \ll 1/(q_s \ln q_s)$, then the same argument of Part a would allow to keep the drift under control for additional $\kappa n_0 q_s$ time. But in the other cases where we have in particular to interpolate between the constant type and non-constant type behavior, our proof gets a bit technical and treats different cases separately.

4.1.2 Notations and standing assumptions

In all the proofs of Theorem 2, Theorem 4 and Theorem 1, we will use T for the irrational rotation R_α. We may assume WLOG that $\sum_{i=1}^{k}(A_i - B_i) > 0$. Fix $1 \gg \epsilon > 0$ and $N \in \mathbb{N}$. Let $d = \sum_{i=1}^{k}(A_i - B_i) - \min(\frac{1}{10}, \frac{\sum_{i=1}^{k}(A_i - B_i)}{2}) > 0$. Define $\kappa = \kappa(\epsilon) := \frac{\epsilon m_0 d}{64(d+1)Hk}$, where H comes from Lemma 4.4, and $m_0 > 0$ is the constant coming from 3. in Theorem 2. With $C > 0$ the constant from the Definition 1.3 of badly approximable singularities, we let

$$P := \left[-2(d + 1), \frac{-dm_0}{32C} \right] \cup \left[\frac{dm_0}{32C}, 2(d + 1) \right], \tag{23}$$

In the sequel, we will assume $s \geqslant s_0(\epsilon, N) = s_0$, where s_0 is a sufficiently large integer, in particular $\kappa q_{s_0} > N$.

We summarize now the consequences of the hypothesis 1.,2.,3. of Theorem 2 that will be useful to us in the sequel. If $s \geqslant s_0$ we have

$$\frac{|h'\left(\frac{x_s}{4C}\right)|}{q_s h\left(\frac{1}{2q_s}\right)} < \frac{\epsilon}{2}, \quad \frac{x_s}{2C} q_s h\left(\frac{1}{2q_s}\right) > \frac{1}{\epsilon},$$

$$\sum_{s \geqslant s_0, s \notin K_\alpha} x_s q_s < \frac{\epsilon}{16k}, \quad h\left(\frac{1}{2q_s}\right) \big/ h\left(\frac{1}{2q_{s+1}}\right) > m_0 \tag{24}$$

We also note that $h\left(\frac{1}{2q_s}\right) > 8C$. Set $v_s := \frac{x_s}{4C}$ and define

$$W_s := \bigg\{ x \in \mathbb{T} \ : \ x - q_s \alpha, \ldots, x, \ldots, x + (q_s - 1)\alpha$$

$$\notin \bigcup_{i=1}^{k}(-4v_s + a_i, a_i + 4v_s) \bigg\} \tag{25}$$

and $Z := \bigcap_{s \geqslant s_0, s \notin K_\alpha} W_s$.

Observe that $\lambda(Z) \geqslant 1 - \epsilon(\lambda(W_s) \geqslant 1 - 16kv_sq_s)$.

Set $\delta := \dfrac{1}{q_{s_0}h\left(\frac{1}{2q_{s_0}}\right)}$. Consider $x, y \in Z$ with $0 < \|x - y\| < \delta$. We will assume WLOG that $x < y$ (we consider the trigonometric order on $\mathbb{T}$).

4.1.3 Controlling the drift

The following proposition implies Theorem 2 due to Proposition 3.3.

Proposition 4.5 *Consider* $x, y \in Z$ *with* $0 < \|x - y\| < \delta$. *Then there exists* $p \in P, M, L \geqslant \kappa M \geqslant N$ *such that either* (16) *holds for* $n \in [M, M + L]$ *or* (17) *holds for* $n \in [M, M + L]$.

Proposition 4.5 can be deduced from the following main result on the drift of the Birkhoff sums of a function with logarithmic like singularities. Let $s := s(x, y)(s \geqslant s_0)$ be unique such that

$$ x < y, \quad \frac{1}{q_{s+1}h\left(\frac{1}{2q_{s+1}}\right)} < \|x - y\| \leqslant \frac{1}{q_sh\left(\frac{1}{2q_s}\right)}. \tag{26} $$

We will assume that $q_{s+1} > 2q_s$. If not, then in (26), $\dfrac{m_0}{2q_sh\left(\frac{1}{2q_s}\right)} < \|x - y\|$ and we repeat the considerations below in the time interval $[q_{s-1}, q_s]$. In other words, in this case we will see the drift between x and y before time q_s.

Proposition 4.6 *Consider* $x, y \in Z$ *as in* (26).
Part a *There exists* $n_0 \in \{1, \ldots, \max(\frac{q_{s+1}}{8Cq_s}, 1)\}$ *satisfying*

$$ f^{(n_0q_s)}(x) - f^{(n_0q_s)}(y) \in P \tag{27} $$

or

$$ f^{(-n_0q_s)}(x) - f^{(-n_0q_s)}(y) \in P \tag{28} $$

and such that the following holds
Part b *Let* $X = T^{n_0q_s}x$ *and* $Y = T^{n_0q_s}y$ *if* (27) *holds, and* $X = T^{-(n_0q_s+1)}x$ *and* $Y = T^{-(n_0q_s+1)}y$ *if* (28) *holds. For* $n = 1, \ldots, [\kappa n_0q_s] + 1$ *we have*

$$A. \ |f^{(n)}(X) - f^{(n)}(Y)| < \epsilon \quad or \quad B. \ |f^{(-n)}(X) - f^{(-n)}(Y)| < \epsilon. \quad (29)$$

The rest of this section is devoted to the proof of Proposition 4.6. But before this we show how it implies Proposition 4.5 and thus Theorem 2.

Proof of Proposition 4.5. Suppose (27) holds, the other case being similar. If A. from (29) holds, set $M := n_0 q_s$, $L := [\kappa M] + 1$ and $p := f^{(n_0 q_s)}(x) - f^{(n_0 q_s)}(y) \in P$. If B. holds, we set $M := [\frac{n_0 q_s}{1 + \kappa}]$, $L := [\kappa M] + 1$ and $p := f^{(n_0 q_s)}(x) - f^{(n_0 q_s)}(y) \in P$. Notice that in both cases $M, L \geqslant \frac{1}{2} \kappa n_0 q_s \geqslant \frac{1}{2} \kappa q_s \geqslant \frac{1}{2} \kappa q_{s_0} \geqslant N$. Finally, using $A.$ or $B.$ and the cocycle identity for the Birkhoff sums and the triangular inequality shows that for $n \in [M, M + L]$, $|f^{(n)}(x) - f^{(n)}(y) - p| < \varepsilon$ for some $p \in P$. □

4.1.4 Proof of Proposition 4.6 Part a

For $m \in \mathbb{N}$, we will often use the following non resonance conditions of a pair of points (x, y) with the singularities $\{a_1, \dots, a_k\}$ [compare with (25)].

$$\bigcup_{j=-q_m}^{q_m - 1} T^j[x, y] \cap \bigcup_{i=1}^{k} [-2v_m + a_i, a_i + 2v_m] = \emptyset \quad (30)$$

$$\bigcup_{j=0}^{\max([\frac{q_{m+1}}{4C}], q_m)} T^j[x, y] \cap \bigcup_{i=1}^{k} [-v_m + a_i, a_i + v_m] = \emptyset. \quad (31)$$

$$\bigcup_{j=1}^{\max([\frac{q_{m+1}}{4C}], q_m)} T^{-j}[x, y] \cap \bigcup_{i=1}^{k} [-v_m + a_i, a_i + v_m] = \emptyset. \quad (32)$$

Lemma 4.7 *Let $x, y \in \mathbb{T}$ be as in (26). Then for every m such that $s_0 \leqslant m \leqslant s$, if we have at least one of the following*

1. *if $m \notin K_\alpha$ and (30) is satisfied*
2. *if $m \in K_\alpha$ and $q_{m+1} \geqslant 2q_m$,*

then we have at least one of (31) or (32).

Proof Observe first that since by (26), $\|x - y\| \leqslant v_s \leqslant v_m$, then it suffices to prove (31) or (32) with just x instead of $[x, y]$ on the LHS and $2v_m$ instead of v_m on the RHS.

Assume $m \notin K_\alpha$. Since $m \geqslant s_0$ we may assume that $q_{m+1} \geqslant 16 C q_m$.

Let $t_1 \in [-q_m, q_m - 1] \cap \mathbb{Z}$ and $r_1 \in \{1, \dots, k\}$ be such that

$$\rho \left(\{x + j\alpha\}_{j=-q_m}^{q_m - 1}, \{a_i\}_{i=1}^{k} \right) = \|x + t_1 \alpha - a_{r_1}\|.$$

We assume now that $t_1 < 0$ and show (31). In case $t_1 \geqslant 0$, (32) would follow similarly.

By (30) we have that $\|x + t_1 - a_{r_1}\| \geqslant 2v_m$. Since $m \notin K_\alpha$, and $t_1 < 0$ it follows that for every $r = 0, 1, \ldots, [\frac{q_{m+1}}{4Cq_m}]$,

$$\|x + (t_1 + rq_m)\alpha - a_{r_1}\| \geqslant \|x + t_1 - a_{r_1}\| \geqslant 2v_m.$$

Moreover for every $i \in \{1, \ldots, k\}, i \neq r_1$, since $a_1, \ldots, a_k$ are badly approximable by α (see Definition 1.3), we have for every $r = 0, 1, \ldots, [\frac{q_{m+1}}{4Cq_m}]$

$$\|x + (t_1 + rq_m)\alpha - a_i\| \geqslant \|x + t_1 - a_i\| - r\|q_m\alpha\| \geqslant \frac{1}{4Cq_m} \geqslant 2v_m.$$

For $j \leqslant [\frac{q_{m+1}}{4C}] - 1$ and $j \notin \{t_1, t_1 + q_m, \ldots, t_1 + [\frac{q_{m+1}}{4Cq_m}]q_m - 1\}$, Lemma 4.1 (for $i_0 = t_1$) implies

$$x + j\alpha \notin \bigcup_{i=1}^{k} \left[-\frac{1}{4Cq_m} + a_i, a_i + \frac{1}{4Cq_m} \right]$$

and since $2v_m \leqslant 1/(4Cq_m)$, this finishes the proof of (31).   □

In the following lemma we control the drift between the Birkhoff sums up to q_s or $-q_s$ between nearby points that do not go too close to the singularities. Recall that we have assumed that $d = \sum_{i=1}^{k}(A_i - B_i) > 0$.

Lemma 4.8 *For every $s \geqslant s_0$ we have the following for any points $x < y \in \mathbb{T}$ if*

$$\bigcup_{j=0}^{q_s-1} T^j[x, y] \cap \bigcup_{i=1}^{k} [-2v_s + a_i, a_i + 2v_s] = \emptyset \tag{33}$$

then

$$(d + 1)q_s h\left(\frac{1}{2q_s}\right) \|x - y\| \geqslant f^{(q_s)}(x) - f^{(q_s)}(y) \geqslant dq_s h\left(\frac{1}{2q_s}\right) \|x - y\|. \tag{34}$$

If

$$\bigcup_{j=-q_m}^{-1} T^j[x, y] \cap \bigcup_{i=1}^{k} [-2v_m + a_i, a_i + 2v_m] = \emptyset \tag{35}$$

then

$$(d+1)q_s h\left(\frac{1}{2q_s}\right) \|x - y\| \geqslant f^{(-q_s)}(x) - f^{(-q_s)}(y) \geqslant dq_s h\left(\frac{1}{2q_s}\right) \|x - y\|. \tag{36}$$

Proof We show that (33) implies (34), the second part of the Lemma being similar. By (33) and (26), $f^{(q_s)}$ is differentiable on $[x, y]$. Therefore, there exists $\theta \in [x, y]$ such that

$$f^{(q_s)}(x) - f^{(q_s)}(y) = (x - y)f'^{(q_s)}(\theta).$$

It is enough to show that there exist $d > 0$ such that for $s \geqslant s_0$

$$(d + 1)q_s h\left(\frac{1}{2q_s}\right) \geqslant -f'^{(q_s)}(\theta) \geqslant dq_s h\left(\frac{1}{2q_s}\right). \tag{37}$$

For $s \in \mathbb{N}$, define

$$\bar{f}'_s(\theta) = \begin{cases} 0, & \text{if } \theta \in \bigcup_{i=1}^{k}\left[-\frac{1}{2q_s} + a_i, a_i + \frac{1}{2q_s}\right] \\ f'(\theta), & \text{otherwise.} \end{cases}$$

It follows that $\bar{f}'_s \in BV(\mathbb{T})$ and

$$f'^{(q_s)}(\theta) = \bar{f}'^{(q_s)}_s(\theta) + \sum_{i \in J_s} f'(\theta + j_i\alpha) + \sum_{i \in L_s} f'(\theta + l_i\alpha), \tag{38}$$

where $J_s = \{i \in [1, k] : \exists j_i \in \{0, \ldots, q_s - 1\} : \theta + j_i\alpha \in [-\frac{1}{2q_s} + a_i, a_i]\}$ and $L_s := \{i \in [1, k] : \exists l_i \in \{0, \ldots, q_s - 1\} : \theta + l_i\alpha \in [a_i, a_i + \frac{1}{2q_s}]\}$. Note that for every $i \in [1, k]$ there exists at most one $j_i \in \{0, \ldots, q_s - 1\} : \theta + j_i\alpha \in [-\frac{1}{2q_s} + a_i, a_i]$.

We use the Denjoy-Koksma inequality to $\bar{f}'_s$, to get

$$q_s \int_{\mathbb{T}} \bar{f}'_s d\lambda - \text{Var}(\bar{f}'_s) \leqslant |\bar{f}'^{(q_s)}_s(\theta)| \leqslant q_s \int_{\mathbb{T}} \bar{f}'_s d\lambda + \text{Var}(\bar{f}'_s). \tag{39}$$

We have

$$\int_{\mathbb{T}} \bar{f}'_s d\lambda = \sum_{i=1}^{k} f\left(a_i + \frac{1}{2q_s}\right) - f\left(a_i - \frac{1}{2q_s}\right) \text{ and } \text{Var}(\bar{f}'_s)$$

$$= 2\sum_{i=1}^{k}\left(f'\left(a_i + \frac{1}{2q_s}\right) + f'\left(a_i - \frac{1}{2q_s}\right)\right), \tag{40}$$

(if $s \in \mathbb{N}$ is sufficiently large). It follows by the assumptions on f' and h' and (30), that if $s_0(\epsilon, N)$ is sufficiently large, then for $s \geqslant s_0$, we have for every $i = 1, \ldots, k$:

$$|f'(\theta + j_i\alpha)| \leqslant (B_i + 1)|h'(\theta + j_i\alpha)| \leqslant (B_i + 1)\left|h'\left(\frac{x_s}{4C}\right)\right|$$

$$\leqslant \epsilon q_s h\left(\frac{1}{2q_s}\right) \quad \text{for } i \in J_s. \tag{41}$$

and similarly

$$|f'(\theta + l_i\alpha)| \leqslant \epsilon q_s h\left(\frac{1}{2q_s}\right) \quad \text{for } i \in L_s. \tag{42}$$

On the other hand, by l'Hospital's rule

$$((A_i + \epsilon) - (B_i - \epsilon))h\left(\frac{1}{2q_s}\right) \geqslant f\left(a_i + \frac{1}{2q_s}\right) - f\left(a_i - \frac{1}{2q_s}\right)$$

$$\geqslant ((A_i - \epsilon) - (B_i + \epsilon))h\left(\frac{1}{2q_s}\right). \tag{43}$$

$$\left|f'\left(a_i + \frac{1}{2q_s}\right)\right| + \left|f'\left(a_i - \frac{1}{2q_s}\right)\right| \leqslant ((A_i + 1) + (B_i + 1))\left|h'\left(\frac{1}{2q_s}\right)\right|$$

$$\leqslant \epsilon q_s h\left(\frac{1}{2q_s}\right), \tag{44}$$

(by $\frac{x_s}{4C} < \frac{1}{2q_s}$).

Now, using (38)–(44), we get

$$q_s h\left(\frac{1}{2q_s}\right)\left(\left(\sum_{i=1}^{k}(A_i - B_i)\right) - 6k\epsilon\right) \leqslant -f'^{(q_s)}(\theta)$$

$$\leqslant q_s h\left(\frac{1}{2q_s}\right)\left(\left(\sum_{i=1}^{k}(A_i - B_i)\right) + 6k\epsilon\right).$$

which allows us to conclude since if we assume WLOG that ε is sufficiently small (recall that $x < y$). $\square$

We now concatenate the inequalities of Lemma 4.8 if (31) or (32) are satisfied.

Lemma 4.9 *Let* $x, y \in \mathbb{T}$ *satisfy* (26) *and* (31) *with* $m = s$, *then there exists* $n_0 \in \{1, \ldots, \max(\frac{q_{s+1}}{8Cq_s}, 1)\}$ *such that* (27) *holds. Moreover,*

$$n_0 q_s h\left(\frac{1}{2q_s}\right) \leqslant \frac{2(d+1)}{d\|x - y\|}. \tag{45}$$

If x, y satisfy (26) and (32) with $m = s$ then there exists $n_0 \in \{1, \ldots, \max$ $(\frac{q_{s+1}}{8Cq_s}, 1)\}$ such that (28) holds for some $n_0 \in \{1, \ldots, \max(\frac{q_{s+1}}{8Cq_s}, 1)\}$ satisfying (45).

Proof We will assume (31) holds, the other case being similar. We will use repeatedly (34) of Lemma 4.8 with x, y replaced by $x + rq_s\alpha$, $y + rq_s\alpha$ respectively, with $r = 0, 1, \ldots, \max([\frac{q_{s+1}}{4Cq_s}] - 1, 0)$. Indeed, by (31) the points $x + rq_s$, $y + rq_s$ satisfy (33). Summing up the obtained inequalities we get for any $R \in [0, \max([\frac{q_{s+1}}{4Cq_s}], 1)]$,

$$R\|x-y\|(d+1)q_s h\left(\frac{1}{2q_s}\right) > f^{(Rq_s)}(x) - f^{(Rq_s)}(y) \geqslant R\|x-y\|dq_s h\left(\frac{1}{2q_s}\right). \tag{46}$$

Let $e_R := f^{(Rq_s)}(x) - f^{(Rq_s)}(y)$. Then for $R \leqslant [\frac{q_{s+1}}{4Cq_s}] - 1$ we have that $e_{R+1} - e_R \leqslant d + 1$.

Moreover, by (46) and (26) and the hypothesis $h(\frac{1}{2q_s})/h(\frac{1}{2q_{s+1}}) > m_0$ we get

$$\begin{aligned} e_{\max\left(\left[\frac{q_{s+1}}{8Cq_s}\right], 1\right)} &\geqslant \max\left(\frac{q_{s+1}}{8Cq_s}, 1\right) dq_s h\left(\frac{1}{2q_s}\right) \|x - y\| \\ &\geqslant dm_0 \max\left(\frac{1}{8C} - \frac{q_s}{q_{s+1}}, \frac{q_s}{q_{s+1}}\right) \geqslant \frac{dm_0}{16C}. \end{aligned}$$

Therefore, there exists $n_0 \in \{1, \ldots, \max([\frac{q_{s+1}}{8Cq_s}], 1)\}$ such that

$$f^{(n_0q_s)}(x) - f^{(n_0q_s)}(y) = e_{n_0} \in \left[\frac{dm_0}{32C}, 2(d+1)\right] \subset P$$

Moreover, (46) with $R = n_0$ implies (45)

$$n_0 q_s h\left(\frac{1}{2q_s}\right) \leqslant \frac{2(d+1)}{d\|x - y\|}. \tag{47}$$

In case (32) is satisfied instead of (31), we show (28) using repeatedly (36) of Lemma 4.8. □

We are ready now to finish the proof of Part a. of Proposition 4.6. If $s \notin K_\alpha$, then by the fact that x, $y \in Z \subset W_s$, it follows that 1. in Lemma 4.7 is satisfied with $m = s$. If $s \in K_\alpha$ then 2. in Lemma 4.7 is satisfied with $m = s$. Therefore we can use Lemma 4.7 for x, y and $m = s$. Now, by Lemma 4.9, if (31) holds we have (27), if (32) holds we have (28). Part a. of Proposition 4.6 is settled, we turn now to Part b.

4.1.5 Proof of Proposition 4.6 Part b

For $m \geqslant s_0$ and $\mathbb{N} \cup \{0\} \ni l \leqslant \max(\frac{q_{m+1}}{8Cq_m} - 1, 0)$ we will consider the following conditions on $x, y \in \mathbb{T}$ [compare with (31) and (32)]

$$\bigcup_{j=0}^{(l+1)q_m} T^j[x, y] \cap \bigcup_{i=1}^{k}[-v_m + a_i, a_i + v_m] = \emptyset. \tag{48}$$

$$\bigcup_{j=1}^{(l+1)q_m} T^{-j}[x, y] \cap \bigcup_{i=1}^{k}[-v_m + a_i, a_i + v_m] = \emptyset. \tag{49}$$

Lemma 4.10 *Let $x, y \in \mathbb{T}$ satisfy (26) and (48) for some $m \geqslant s_0$ and $\mathbb{N} \cup \{0\} \ni l \leqslant \max(\frac{q_{m+1}}{8Cq_m} - 1, 0)$. Then*

$$\text{for every } n = 0, \ldots, (l+1)q_m, \ \left| f^{(n)}(x) - f^{(n)}(y) \right|$$
$$< 8kH\|x - y\|(l+1)q_m h\left(\frac{1}{2q_m}\right) \tag{50}$$

Let $x, y \in \mathbb{T}$ satisfy (26) and (49) for some $m \geqslant s_0$ and $\mathbb{N} \cup \{0\} \ni l \leqslant \max(\frac{q_{m+1}}{8Cq_m} - 1, 0)$. Then

$$\text{for every } n = 1, \ldots, (l+1)q_m, \ |f^{(-n)}(x) - f^{(-n)}(y)|$$
$$< 8kH\|x - y\|(l+1)q_m h\left(\frac{1}{2q_m}\right). \tag{51}$$

Proof We only give the proof of the first case since the other is similar. For every $n = 0, \ldots, \max(\frac{q_{m+1}}{4C}, q_m)$, there exists $\theta_n \in [x, y]$ such that $|f^{(n)}(x) - f^{(n)}(y)| = (x - y)f'^{(n)}(\theta_n)$. Therefore, using Lemma 4.4, for every $n = 0, \ldots, (l+1)q_m$, we have

$$\left| f^{(n)}(x) - f^{(n)}(y) \right| \leqslant H\|x - y\|\left(\sum_{i=1}^{k} -h'^{(n)}(\theta_n - a_i) - h'^{(n)}(a_i - \theta_n)\right). \tag{52}$$

Moreover, by monotonicity of h', for every $i = 1, \ldots, k$,

$$-h'^{(n)}(\theta_n - a_i) \leqslant -h'^{(n)}(x - a_i) \quad \text{and} \quad -h'^{(n)}(a_i - \theta_n) \leqslant -h'^{(n)}(a_i - y).$$

Since $-h'$ is positive, we get that

$$-h'^{(n)}(\theta_n - a_i) \leqslant -h'^{((l+1)q_m)}(x - a_i) \quad \text{and} \quad -h'^{(n)}(a_i - \theta_n)$$
$$< -h'^{((l+1)q_m)}(a_i - y). \tag{53}$$

It follows by Lemma 4.3, (48) and (24) that for every $u = 0, \ldots, l$

$$\|x - y\| |h'^{(q_m)}(T^{uq_m}x - a_i)| \leqslant \|x - y\| \left(q_m h \left(\frac{1}{2q_m} \right) - h' \left(\frac{1}{2q_m} \right) - h'(v_m) \right)$$
$$\leqslant 4\|x - y\| q_m h \left(\frac{1}{2q_m} \right).$$

Hence, summing up over $u = 0, \ldots, l$, and using the cocycle identity, (52) implies (50).

This finishes the proof. $\qquad\square$

To prove Proposition 4.6 Part b., observe first that if s_0 is sufficiently large, and up to eventually changing κ to $\kappa' = \frac{\kappa}{8C}$, one of two possibilities holds : **1**. There exists $s_0 \leqslant m \leqslant s, m \in K_\alpha$, such that $\kappa n_0 q_s < q_m \leqslant 8C\kappa n_0 q_s$, or **2**. There exist $s_0 \leqslant m \leqslant s$ and $l \geqslant 1$ such that $lq_m \leqslant \kappa n_0 q_s < (l+1)q_m \leqslant \frac{q_{m+1}}{8C}$. **Case 1.** $\kappa n_0 q_s < q_m \leqslant 8C\kappa n_0 q_s$ with $s_0 \leqslant m \leqslant s, m \in K_\alpha$. Lemma 4.7 implies that either (31) or (32) holds for $T^{n_0 q_s}x$, $T^{n_0 q_s}y$, m. Therefore, (48) or (49) holds for m and $l = 0$. We then apply Lemma 4.10 to $T^{n_0 q_s}x$, $T^{n_0 q_s}y$, m with $l = 0$, and according to whether we have (50) or (51) we will get A. or B. of Proposition 4.6 Part b. Indeed, suppose (50) holds. Then, since $\kappa n_0 q_s < q_m$, for $n = 1, \ldots, [\kappa n_0 q_s] + 1$, we have due to (47)

$$|f^{(n)}(T^{n_0 q_s}x) - f^{(n)}(T^{n_0 q_s}y)| < 8kH\|x - y\| q_m h \left(\frac{1}{2q_m} \right)$$
$$< 16CkH\kappa n_0 q_s \|x - y\| h \left(\frac{1}{2q_m} \right)$$
$$< 16CkH\kappa \frac{2(d+1)}{d} < \varepsilon.$$

Case 2. There exist $s_0 \leqslant m \leqslant s$ and $l \geqslant 1$ such that $lq_m \leqslant \kappa n_0 q_s < (l+1)q_m \leqslant \frac{q_{m+1}}{8C}$. We will first prove that $T^{n_0 q_s}x$, $T^{n_0 q_s}y$, m, l satisfy the hypothesis of Lemma 4.10. If $m \in K_\alpha$, then Lemma 4.7 implies that either (31) or (32) holds for $T^{n_0 q_s}x$, $T^{n_0 q_s}y$, m. Therefore, since $l \leqslant \frac{q_{m+1}}{8Cq_m} - 1$, either (48) or (49) holds for $T^{n_0 q_s}x$, $T^{n_0 q_s}y$, m, l. If $m \notin K_\alpha$, then we consider two cases:

I. $m = s$. In this case $n_0 > \frac{1}{\kappa}$ and therefore $q_{s+1} > 32Cq_s$. Since $l \leqslant [\kappa n_0] + 1$, $n_0 \leqslant \frac{q_{s+1}}{8Cq_s}$ and $\kappa \ll 1$, we get that

$$\bigcup_{j=0}^{(l+1)q_s} T^j[T^{n_0q_s}x, T^{n_0q_s}y] \cap \bigcup_{i=1}^{k}[-v_s + a_i, a_i + v_s]$$

$$\subset \bigcup_{j=0}^{\max\left(\left[\frac{q_{s+1}}{4C}\right],q_s\right)} T^j[x, y] \cap \bigcup_{i=1}^{k}[-v_s + a_i, a_i + v_s].$$

Similarly,

$$\bigcup_{j=0}^{(l+1)q_s} T^{-j}[T^{n_0q_s}x, T^{n_0q_s}y] \cap \bigcup_{i=1}^{k}[-v_s + a_i, a_i + v_s]$$

$$\subset \bigcup_{j=0}^{\max\left(\left[\frac{q_{s+1}}{4C}\right],q_s\right)} T^{-j}[x, y] \cap \bigcup_{i=1}^{k}[-v_s + a_i, a_i + v_s].$$

Note that by Lemma 4.7 x, y satisfy (31) or (32). Therefore the assumptions of Lemma 4.10 are satisfied for $T^{n_0q_s}x$, $T^{n_0q_s}y$, s, l.

II. $m < s$. Since $l \leqslant \frac{q_{m+1}}{8Cq_m} - 1$, it is enough to show that $T^{n_0q_s}x$, $T^{n_0q_s}y$, m satisfy (31) or (32). Due to Lemma 4.7, we just have to check (30) for $T^{n_0q_s}x$, $T^{n_0q_s}y$, m:

$$\bigcup_{j=-q_m}^{q_m-1} T^j[T^{n_0q_s}x, T^{n_0q_s}y] \cap \bigcup_{i=1}^{k}[-2v_m + a_i, a_i + 2v_m] = \emptyset. \tag{54}$$

Since $m \notin K_\alpha$ and $m < s$, we have

$$\frac{1}{q_s} \leqslant \frac{1}{q_{m+1}} \ll v_m. \tag{55}$$

Moreover, $\|T^{n_0q_s}x - T^{n_0q_s}y\| \overset{(26)}{<} \frac{1}{10v_m}$. Therefore, it is enough to show that for $j \in \{0, \ldots, q_m - 1\}$ we have

$$T^{n_0q_s}x + j\alpha \notin \bigcup_{i=1}^{k}[-3v_m + a_i, a_i + 3v_m].$$

For this aim, let i_0 and r_1 be such that $\rho(\{T^{n_0 q_s} x + j\alpha\}_{j=0}^{q_m-1}, \{a_i\}_{i=1}^k) = \|T^{n_0 q_s} x + i_0 \alpha - a_{r_1}\|$. It follows by (3), that for $i_0 \neq j \in \{0, \ldots, q_m - 1\}$,

$$T^{n_0 q_s} x + j\alpha \notin \bigcup_{i=1}^k \left[-\frac{1}{2Cq_m} + a_i, a_i + \frac{1}{2Cq_m} \right]. \tag{56}$$

Next, by the fact that $m \notin K_\alpha$ and $x \in B_m (m \geqslant s_0)$, we get that $\|x + i_0 \alpha - a_{r_1}\| \geqslant 4v_m$, and therefore

$$\|x + i_0 \alpha + n_0 q_s \alpha - a_{r_1}\| \geqslant \|x + i_0 \alpha - a_{r_1}\| - \|n_0 q_s \alpha\| \geqslant 4v_m - \frac{n_0}{q_{s+1}}$$

$$\geqslant 4v_m - \frac{1}{8Cq_s} \overset{(55)}{\geqslant} 3v_m,$$

(recall that $n_0 \leqslant \frac{q_{s+1}}{8Cq_s}$) and (54) is thus proved. So in **Case 2.** at least one of (48) or (49) is satisfied for $T^{n_0 q_s} x, T^{n_0 q_s} y, m, l$.

Therefore, we can apply Lemma 4.10 to $T^{n_0 q_s} x, T^{n_0 q_s} y, m, l$ (recall that $l q_m \leqslant \kappa n_0 q_s < (l+1)q_m$). Now and as in **Case 1.**, if (50) holds we get A., if (51) holds we get B. Indeed, assume WLOG that $T^{n_0 q_s} x, T^{n_0 q_s} y, m, l$ satisfy (50) (the proof in the other case is analogous). Using (47) and the fact that $[\kappa n_0 q_s] + 1 \leqslant (l+1)q_m \leqslant 2\kappa n_0 q_s$, we get for $n = 1, \ldots, (l+1)q_m$

$$|f^{(n)}(T^{n_0 q_s} x) - f^{(n)}(T^{n_0 q_s} y)| < 8kH \|x - y\|(l+1)q_m h\left(\frac{1}{2q_m}\right)$$

$$< 16kH\kappa n_0 q_s \|x - y\| h\left(\frac{1}{2q_s}\right)$$

$$< 16kH\kappa \frac{2(d+1)}{d} < \varepsilon.$$

So A. in Proposition 4.6 Part b. holds.

The proof of Proposition 4.6 is thus completed and Theorem 2 follows. □

4.2 Proof of Theorem 4

4.2.1 Outline of the proof

The general scheme of the proof is similar to the scheme of the proof of Theorem 2 (see the outline of the proof of the latter theorem in Subsection 4.1). Assume for simplicity that f has just one right-sided power singularity at 0 of type $x^{-\gamma}$. In this outline we will actually see that the constant type condition is an if and only if condition in the proof of Theorem 4 that we give.

Indeed, the following facts are easy to check for an interval $I = [x, y]$ such that $y - x \in [1/q_{n+1}^{\gamma+1}, 1/q_n^{\gamma+1}]$.

a. If for some $c \in (0, 1)$, $R_\alpha^i I$ is disjoint from $[-c/q_n, c/q_n]$ for every $i = 0, \ldots, l \leqslant q_n$ then $f'^{(l)}(\theta) \leqslant Cq_n^{1+\gamma}$ for any $\theta \in I$ for some C that depends on c (with a similar statement for negative iterates).

b. If for some $c' \in (0, 1)$, $R_\alpha^{i_0} I$ is included in $[-c'/q_n, 0)$ for some $i_0 \geqslant 0$ then $f'^{(i_0+1)}(\theta) - f'^{(i_0)}(\theta) \geqslant C'q_n^{1+\gamma}$ for any $\theta \in I$ for some C' that depends on c' (with a similar statement for negative iterates).

c. If α is of constant type then there exists $c' > c > 0$ such that one of the following holds if n is sufficiently large : 1. there exists $i_0 \geqslant 0$ such that $R_\alpha^{i_0} I$ is included in $[-c'/q_n, 0)$ and $R_\alpha^i I$ is disjoint from $[-c/q_n, c/q_n]$ for every $i = 0, \ldots, i_0$ or 2. there exists $i_0 < 0$ such that $R_\alpha^{i_0} I$ is included in $[-c'/q_n, 0)$ and $R_\alpha^i I$ is disjoint from $[-c/q_n, c/q_n]$ for every $i = -1, \ldots, -i_0$.

d. If α is not of constant type and if $q_{n+1} \gg q_n$ and $y - x = \varepsilon_n/q_n^{1+\gamma}$ while $|x - p/q_n| \leqslant \varepsilon_n^2/q_n$ with $\varepsilon_n \to 0$, then as long as for $i \in [0, l]$ (or for $i \in [-l, -1]$) $R_\alpha^i I$ is disjoint from $[-1/(2q_n), 1/(2q_n)]$ we have that $f'^{(l)}(\theta) \leqslant Cq_n^{1+\gamma}$ for any $\theta \in I$, while if l is the first integer such that $R_\alpha^l I$ intersects $[-1/(2q_n), 1/(2q_n)]$ then $f'^{(l)}(\theta) \geqslant q_n^{1+\gamma}/\varepsilon_n$.

Now, if α is of constant type, and if we assume that c.1 holds, then a. and b. imply that either $f'^{(i_0)}(\theta) \in [C'q_n^{1+\gamma}, Cq_n^{1+\gamma}]$ for every $\theta \in I$ or $f'^{(i_0+1)}(\theta) \in [C'q_n^{1+\gamma}, Cq_n^{1+\gamma}]$ for every $\theta \in I$. Since q_{n+1}/q_n is bounded this implies a controlled macroscopic drift between the orbit of x and y (this is the content of Proposition 4.12 Part a). As in the proof of Theorem 2, we then need to use the same type of arguments to show that the drift remains almost constant during a small additional proportion of time (we do this in the future of $i_0 + 1$ or in the past of i_0, and this is the content of Proposition 4.12 Part b and relies on Sublemma 4.16). The case c.2 is treated similarly.

In Sublemma 4.14 we essentially prove a. and in Sublemma 4.15 we essentially prove b.

We now explain why the constant type condition is necessary in our proof. Indeed, if α is not of constant type, d. gives an example of pairs x, y for which the drift between the forward orbits jumps from ε_n to $1/\varepsilon_n$ and the same happens for backward orbits, which contradict the SWR property for this pair. Furthermore, if ε_n is taken to converge very slowly to 0 such pairs can be produced with $x \in Z$ for any Z having positive measure. Observe that this does not imply that the SWR property would not reappear much later in time but this is very unlikely as demonstrated for the absence of the WR property in the particular case of Theorem 1.

Observe finally that the same type of pairs (x, y) described in d. show that it is necessary to use the SWR property instead of the WR property. Indeed, only one of the alternatives c.1 or c.2 holds for such pairs and we are obliged, if we want to see a controlled drift, to iterate in the future or in the past.

4.2.2 Notations and standing assumptions

Recall that in the proof of Theorem 4 T means R_α. We may assume WLOG that $A_k^2 + B_k^2 > 0$. Let $C_k = \max(A_k, -B_k) > 0$. Recall $H > 0$ coming from Lemma 4.4, D_1, $D_2 > 0$, the constants in the hypothesis (4) in Theorem 4 and define

$$P := \left[-12Hk(D_2 + 2), -\frac{C_k D_1^2}{16c} \right] \cup \left[\frac{C_k D_1^2}{16c}, 12Hk(D_2 + 2), \right]$$

where c is such that for every $s \in \mathbb{N}$, $q_{s+1} \leqslant cq_s$.

Fix $\varepsilon \ll 1$ and $N \in \mathbb{N}$. We will assume that $\epsilon < \frac{C_k D_1^2}{8c}$. Let $\kappa := \kappa(\epsilon) = \frac{\epsilon}{2(3D_2+2)kCH}$.

Let $s_0 \in \mathbb{N}$ be such that $q_{s_0-4} \geqslant \frac{1}{\kappa}N$, and $h(\frac{1}{2q_s}) > 6C$ for $s \geqslant s_0$, and for every $i = 1, \ldots, k$

$$\left| \frac{f'(x)}{h'(x - a_i)} \right| > \frac{A_i}{2} \text{ for } x \in \left[a_i, a_i + \frac{1}{q_{s_0-4}} \right] \text{ and } \left| \frac{f'(x)}{h'(a_i - x)} \right| > \frac{B_i}{2},$$

$$\text{for } x \in \left[-\frac{1}{q_{s_0-4}} + a_i, a_i \right]. \tag{57}$$

Define $\delta := \frac{1}{q_{s_0}h(\frac{1}{2q_{s_0}})}$. We will show that SWR-property holds for all pairs of points $x, y \in \mathbb{T}$ with $\|x - y\| < \delta$.

4.2.3 Controlling the drift

The following proposition implies Theorem 4 due to Proposition 3.3.

Proposition 4.11 *Consider $x, y \in \mathbb{T}$ with $0 < \|x - y\| < \delta$. Then there exists $p \in P, M, L \geqslant \kappa M \geqslant N$ such that either (16) or (17) holds for $n \in [M, M + L]$.*

We can assume WLOG that $x < y$. Let $s := s(x, y)$ be unique such that

$$\frac{1}{q_{s+1}h\left(\frac{1}{2q_{s+1}}\right)} \leqslant \|x - y\| < \frac{1}{q_s h\left(\frac{1}{2q_s}\right)}. \tag{58}$$

As in the precedent section, Proposition 4.11 follows from

Proposition 4.12 *Consider* $x, y \in \mathbb{T}$ *as in* (58).
Part a. *There exists* $i_0 \in \{0, \ldots, q_{s-2} - 1\}$, *such that*

$$\left| f^{(i_0)}(x) - f^{(i_0)}(y) \right| \in P \tag{59}$$

or

$$\left| f^{(-i_0)}(x) - f^{(-i_0)}(y) \right| \in P. \tag{60}$$

Part b. *Let* $X = T^{i_0}x$ *and* $Y = T^{i_0}y$ *if* (59) *holds, and* $X = T^{-i_0-1}x$ *and* $Y = T^{-i_0-1}y$ *if* (60) *holds, for* $n = 1, \ldots, [\kappa i_0] + 1$ *the following holds*

$$A. \left| f^{(n)}(X) - f^{(n)}(Y) \right| < \epsilon \quad or \quad B. \left| f^{(-n)}(X) - f^{(-n)}(Y) \right| < \epsilon. \tag{61}$$

The rest of Sect. 4.2 is devoted to the proof of Proposition 4.12.
Consider the orbit $x - q_{s-2}\alpha, \ldots, x, \ldots, x + (q_{s-2} - 1)\alpha$ (the length of this orbit is smaller than q_s). It follows by (3) that there exists at most one $t_s \in [-q_{s-2}, q_{s-2} + 1]$ such that $x + t_s\alpha \in \bigcup_{i=1}^{k}[-\frac{1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s}]$. Hence at least one of the following two holds:

$$\bigcup_{j=0}^{q_{s-2}-1} T^j[x, y] \cap \bigcup_{i=1}^{k} \left[-\frac{1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s} \right] = \emptyset \tag{62}$$

or

$$\bigcup_{j=1}^{q_{s-2}} T^{-j}[x, y] \cap \bigcup_{i=1}^{k} \left[-\frac{1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s} \right] = \emptyset. \tag{63}$$

The following Lemma directly implies the proof of Proposition 4.12.

Lemma 4.13 *If* (62) *then* (59) *and* (61) *hold. If* (63) *then* (60) *and* (61) *hold.*

4.2.4 Proof of Lemma 4.13

We will suppose (62) holds, the proof of the other case being analogous. We will need some lemmas.

Sublemma 4.14 *For* $n = 0, \ldots, q_{s-2} - 1$,

$$|f^{(n)}(x) - f^{(n)}(y)| \leqslant 2kH(3D_2 + 2).$$

Proof By (58) we have for every $i = 1, \ldots, k, a_i \notin [x + j\alpha, y + j\alpha]$ with $j \in \{0, \ldots, q_{s-2} - 1\}$. It follows that for $n = 0, \ldots, q_{s-2} -$

1, $\left| f^{(n)}(x) - f^{(n)}(y) \right| = \left| f'(\theta_n) \right| \|x - y\|$, for some $\theta_n \in [x, y]$. Hence, using Lemma 4.4, for every $n = 0, \ldots, q_{s-2}$ we have

$$\left| f^{(n)}(x) - f^{(n)}(y) \right| \leqslant H \|x - y\| \left(\sum_{i=1}^{k} (-h'^{(n)}(\theta_n - a_i) - h'^{(n)}(a_i - \theta_n)) \right). \tag{64}$$

By the monotonicity of h' on $(0, 1)$ we obtain $-h'^{(n)}(\theta_n - a_i) \leqslant -h'^{(n)}(x - a_i)$, $-h'^{(n)}(a_i - \theta_n) \leqslant -h'^{(n)}(a_i - y)$. Since $-h' > 0$,

$$-h'^{(n)}(x - a_i) - h'^{(n)}(a_i - y) \leqslant -h'^{(q_{s-2})}(x - a_i) - h'^{(q_{s-2})}(a_i - y). \tag{65}$$

Using Lemma 4.3 (applied to $x - a_i$, where $j_i \in [0, q_{s-2}] - 1$ is unique such that $x + j_i \alpha \in [a_i, a_i + \frac{1}{2q_{s-2}}]$), we obtain

$$\|x - y\| \left(-h'^{(q_{s-2})}(x - a_i) \right)$$
$$\leqslant \|x - y\| \left(q_{s-2} h \left(\frac{1}{2q_{s-2}} \right) - 2h' \left(\frac{1}{2q_{s-2}} \right) - h'(x + j_i \alpha) \right).$$

By (62), it follows that for $n = 0, \ldots, q_{s-2} - 1$ we have $x + n\alpha, y + n\alpha \notin \bigcup_{i=1}^{k} [-\frac{1}{2Cq_s} + a_i, a_i + \frac{1}{2Cq_s}]$.

By monotonicity of h', $-h'(x + j_i \alpha) < -h'(\frac{1}{2Cq_s})$ and therefore by (4) and (58)

$$\|x - y\| \left(-h'^{(q_{s-2})}(x - a_i) \right) \leqslant \|x - y\|$$
$$\times \left(q_{s-2} h \left(\frac{1}{2q_{s-2}} \right) - 2h' \left(\frac{1}{\frac{1}{2} \frac{1}{q_{s-2}}} \right) - h' \left(\frac{1}{\frac{1}{2C} \frac{1}{q_s}} \right) \right)$$
$$\leqslant \frac{q_{s-2} h \left(\frac{1}{2q_{s-2}} \right) + 2D_2 q_{s-2} h \left(\frac{1}{2q_{s-2}} \right) + D_2 q_s h \left(\frac{1}{2q_s} \right)}{q_s h \left(\frac{1}{2q_s} \right)}$$
$$\leqslant \frac{3D_2 q_s h \left(\frac{1}{2q_s} \right) + q_{s-2} h \left(\frac{1}{2q_{s-2}} \right)}{q_s h \left(\frac{1}{2q_s} \right)} \leqslant 3D_2 + 1. \tag{66}$$

Similarly we obtain $\|x - y\| \left(-h'^{(q_{s-2})}(a_i - y) \right) < 3D_2 + 1$.
Therefore using (64) and the computations above, for $n = 0, \ldots, q_{s-2} - 1$,

$$|f^{(n)}(x) - f^{(n)}(y)| < 2kH(3D_2 + 2).$$

$\square$

Sublemma 4.15 *There exists $i_0 \in \{0, \ldots, q_{s-2} - 1\}$, such that $|f^{(i_0)}(x) - f^{(i_0)}(y)| \geqslant \frac{A_k D_1^2}{4c}$.*

Proof Since $q_{s-2} - q_{s-4} > q_{s-4} + 1$, there exists $i_0 \in [q_{s-4}, q_{s-2} - 2]$ such that

$$T^{i_0}x \in \left[a_k, a_k + \frac{1}{q_{s-4}}\right]. \tag{67}$$

Indeed

$$\{T^k x\}_{k=q_{s-4}}^{q_{s-2}-2} = T^{q_{s-4}}x + \{T^k 0\}_{k=0}^{q_{s-2}-q_{s-4}-2} \subset T^{q_{s-4}}x + \{T^k 0\}_{k=0}^{q_{s-4}-1},$$

and $\{T^k 0\}_{k=0}^{q_{s-4}-1}$ is at least $\frac{1}{q_{s-4}}$-dense. We have assumed that $A_k^2 + B_k^2 > 0$. Suppose additionally $A_k \geqslant -B_k$ (if $A_k \leqslant -B_k$ then we replace $[a_k, a_k + \frac{1}{q_{s-4}}]$ by $[-\frac{1}{q_{s-4}} + a_k, a_k]$). We claim that

$$\left| \left(f^{(i_0+1)}(x) - f^{(i_0+1)}(y) \right) - \left(f^{(i_0)}(x) - f^{(i_0)}(y) \right) \right| > \frac{A_k D_1^2}{2c}.$$

Indeed, the LHS of this inequality is equal to $|f(x + i_0\alpha) - f(y + i_0\alpha)| = |f'(\theta_{i_0})| \|x - y\|$, for some $\theta_{i_0} \in [x + i_0\alpha, y + i_0\alpha]$. Now, by (58), $\theta_{i_0} \in [a_k, a_k + \frac{1}{q_{s-4}} + \frac{1}{q_s h(\frac{1}{2q_s})}] \subset [a_k, a_k + \frac{2}{q_{s-4}}]$. By (57), monotonicity of h', (4) twice (for s and $s + 1$) and (58)

$$|f'(\theta_{i_0})| \geqslant \frac{A_k}{2} \left| h'\left(\theta_{i_0} - a_k\right) \right| \geqslant \frac{A_k}{2} \left| h'\left(\frac{2}{q_{s-4}}\right) \right| \geqslant \frac{A_k}{2} \left| h'\left(2c^4 \frac{1}{q_s}\right) \right|$$

$$\geqslant \frac{A_k}{2} D_1 q_s h\left(\frac{1}{2q_s}\right) \geqslant \frac{A_k}{2} D_1^2 \frac{q_s+1}{c} h\left(\frac{1}{2q_{s+1}}\right) \geqslant \frac{A_k D_1^2}{2c} \frac{1}{\|x - y\|}; \tag{68}$$

and the claim follows. Therefore, one of the numbers $|f^{(i_0+1)}(x) - f^{(i_0+1)}(y)|$ or $|f^{(i_0)}(x) - f^{(i_0)}(y)|$ is at least $\frac{A_k D_1^2}{4c}$.   □

As a consequence of the above lemmas, we obtain that at least one of the numbers $f^{(i_0+1)}(x) - f^{(i_0+1)}(y)$ or $f^{(i_0)}(x) - f^{(i_0)}(y)$ belongs to the set P, and (59) is proved. The next result will give the proof of (61).

Sublemma 4.16 *The following hold:*

$$\left| f^{(n)}(T^{i_0+1}x) - f^{(n)}(T^{i_0+1}y) \right| < \epsilon \quad \text{for all } 0 \leqslant n \leqslant \kappa(i_0 + 1), \tag{69}$$

$$\left| f^{(-n)}(T^{i_0}x) - f^{(-n)}(T^{i_0}y) \right| < \epsilon \quad \text{for all } 0 \leqslant n \leqslant \kappa(i_0 + 1). \tag{70}$$

Proof First we show (69). Select (the unique) $m \in \mathbb{N}$ such that $q_m \geqslant \kappa(i_0 + 1) \geqslant q_{m-1}$ (note that $q_m \ll q_s$). By (3) applied to $T^{i_0}(x)$, (67) and the fact that $q_m \ll q_s$ it follows that

$$\{T^{i_0}x, \ldots, T^{i_0}x + (q_m - 1)\alpha\} \cap \bigcup_{i=1}^{k} \left[-\frac{1}{2Cq_m} + a_i, a_i + \frac{1}{2Cq_m} \right] = \{T^{i_0}x\}.$$
(71)

Analogously, by (3) applied to $T^{i_0}(x) - (q_m - 1)\alpha$, we get

$$\{T^{i_0}x - (q_m - 1)\alpha, \ldots, T^{i_0}x\} \cap \bigcup_{i=1}^{k} \left[-\frac{1}{2Cq_m} + a_i, a_i + \frac{1}{2Cq_m} \right] = \{T^{i_0}x\}.$$
(72)

By (71) and using the same arguments which precedes (64) we obtain (cf. (65)) for $n = 0, \ldots, \kappa(i_0 + 1)$

$$\left| f^{(n)}(T^{i_0+1}x) - f^{(n)}(T^{i_0+1}y) \right| \leqslant H \|x - y\|$$
$$\times \left(\sum_{i=1}^{k} -h'^{(q_m)}(T^{i_0+1}x - a_i) - h'^{(q_m)}\left(a_i - T^{i_0+1}y \right) \right). \quad (73)$$

Then for $i \in \{1, \ldots, k\}$, again by repeating that lead to (66) we obtain

$$\|x - y\| \left(-h'^{(q_m)}(T^{i_0+1}x - a_i) \right) \leqslant \frac{q_m h\left(\frac{1}{2q_m}\right) + 3D_2 q_m h\left(\frac{1}{2q_m}\right)}{q_s h\left(\frac{1}{2q_s}\right)}$$
$$\leqslant \frac{(3D_2 + 1)q_m h\left(\frac{1}{2q_m}\right)}{q_s h\left(\frac{1}{2q_s}\right)}.$$

But $q_m \leqslant c\kappa(i_0 + 1) < c\kappa q_{s-2}$, thus (by the monotonicity of h) $\|x - y\| \left(-h'^{(q_m)}(T^{i_0+1}x - a_i) \right) \leqslant (3D_2 + 1)\frac{c\kappa q_{s-2}}{q_s} = \frac{\epsilon}{4Hk}$, by the definition of κ. Similarly, we obtain $\|x - y\| \left(-h'^{(q_m)}(a_i - T^{i_0+1}y) \right) < \frac{\epsilon}{4Hk}$.

Using this and (73) we get

$$|f^{(n)}(T^{i_0+1}x) - f^{(n)}(T^{i_0+1}y)| < \epsilon,$$

which yields the first case of (69). To handle the second case we use (72) and proceed as before to obtain first $|f^{(-n)}(T^{i_0}x) - f^{(-n)}(T^{i_0}y)| = \|x - y\| \left| f'^{(n)}(\theta_n) \right|$ with $\theta_n \in [T^{i_0}x - n\alpha, T^{i_0}yn - \alpha]$ and then estimating above by

$$H\|x - y\| \left(\sum_{i=1}^{k} -h'^{(q_m)}(T^{i_0}x - (q_m - 1)\alpha - a_i) \right.$$

$$\left. -h'^{(q_m)}(a_i - (T^{i_0}y - (q_m - 1)\alpha)) \right).$$

We conclude exactly in the same way as in the first case.   □

We proceed to the proof of Lemma 4.13 in the case (62) is satisfied. If $f^{(i_0+1)}(x) - f^{(i_0+1)}(y) \in P$, then (69) gives A. in (61); if $f^{(i_0)}(x) - f^{(i_0)}(y) \in P$, then (70) gives B. in (61). The proof of Lemma 4.13 is thus completed since the case where (63) is satisfied is analogous.   □

   This finishes the proof of Proposition 4.12, thus of Theorem 4.   □

Appendix A: Absence of weak Ratner's property

In this section, we will prove Theorem 1. Let f be as in Theorem 1; for simplicity we assume that $\int_{\mathbb{T}} f = 1$. Let $c > 1$ be such that for every $s \in \mathbb{N}, q_{s+1} \leqslant cq_s$. Recall that $C > 1$ is a constant from Definition 1.3 (such a constant exists, since $k = 1$ in our case); we may assume that $C > c$.

   Fix any compact $P \subset \mathbb{R} \setminus \{0\}$. We will prove that for any $t_0 \in \mathbb{R}, (T^t_{\alpha, f})_{t \in \mathbb{R}}$ does not have $R(t_0, P)$ property. For simplicity of the notations we will assume that $t_0 = 1$. Let

$$d > c^{1-\gamma} \text{ be such that } P \subset \left[-\frac{|\gamma|d}{4}, -\frac{100c}{d} \right] \cup \left[\frac{100c}{d}, \frac{|\gamma|d}{4} \right]. \quad (74)$$

Let $\epsilon, \kappa > 0$ sufficiently small, smallness that will be determined in the course of the proof. We use Lemma 3.1 for $Tx = x + \alpha$, to $\epsilon, 3\kappa^2$ to get a set $A \subset \mathbb{T}, \lambda(A) > 1 - \epsilon$ and $N_0 \in \mathbb{N}$, such that (15) holds for $x \in A$ and $n \geqslant N_0$. Let $N > \max\left(2N_0, \frac{1}{\epsilon^2 \kappa^2} \right)$.

We will hereafter assume that $(T^t_{\alpha;f})_{t\in\mathbb{R}}$ has the $R(t_0, P)$ property (see Definition 2.1) and obtain a contradiction. Thus, assume there exist a set $Z \subset X^f$ with $\lambda^f(Z) > 1 - \epsilon$ and $0 < \delta < \epsilon$ such that for every $(x, s), (y, s') \in Z$ with $d^f((x, s), (y, s')) < \delta$, there exist $M, L \geqslant N$ with $\frac{L}{M} \geqslant \kappa$ and $p \in P$ such that

$$\frac{1}{L}|\{n \in [M, M + L] \; : \; d^f(T^f_n(x, s), T^f_{n+p}(y, s')) < \epsilon\}| > 1 - \epsilon. \quad (75)$$

Consider

$$V := \left\{(x, s) \in Z \; : \; x \in A, \; 0 \leqslant s < \frac{1}{\epsilon^2}\right\}. \quad (76)$$

It follows that $\lambda^f(V) > 1 - 4\epsilon$.

The contradiction will come from the following two Propositions, the first one of which is a consequence of (75) and (76).

Proposition 5.1 *Let $(x, s), (y, s') \in V$ with $d^f((x, s), (y, s')) < \delta$. Then there exists an interval $I = [M', M' + L']$ such that $M' \geqslant \frac{N}{2}, \frac{L'}{M'} \geqslant \frac{a\kappa}{6} (a = a(t_0) > 0$ is a constant obtained in Lemma 5.6), there exists $p \in P$ and there exists $m \in \mathbb{Z}$ such that*

$$\|x - y - m\alpha\| < \epsilon \quad \text{and for every } n \in [M', M' + L'],$$
$$|f^{(n)}(x) - f^{(n+m)}(y) - p| < 2\epsilon. \quad (77)$$

Remark 5.2 For $p \in P, n \in \mathbb{N}$, two points $(x, s), (y, s') \in V$ are called p, n-close if

$$d^f(T^n_{\alpha, f}(x, s), T^{n+p}_{\alpha, f}(y, s')) < \epsilon.$$

Then $(x, s), (y, s')$ have the WR-property [see (75)], if there exists a time interval $[M, M + L]$, such that they are p, n-close for a proportion $1 - \epsilon$ of n's in $[M, M + L]$. In general, the set on which the points are p, n-close, can be any subset of $[M, M + L]$. Proposition 5.1 says that in our context the property actually holds on a full interval of integers. This is what happens also in the original case of horocycle flows, where, once the point are drifted after time R, they stay drifted for time εR.

Proposition 5.3 *There exists a set $W_0 \subset \mathbb{T}$ such that $\lambda(W_0) > c_0(d)(c_0 = c_0(d) > 0$ being a constant depending only on d), and a number $0 < \delta_0 < \delta$*

such that for every $x \in W_0$

for every $M \geqslant \dfrac{N}{2}$, *every* $k \in \mathbb{Z}$ *such that* $\|x - (x + \delta_0) - k\alpha\|$
$< \epsilon$ *and every* $p \in P$, *if* $I = [M, M + T]$ *is such that for every* $n \in I$,

$$|f^{(n)}(x) - f^{(n+k)}(x + \delta_0) - p| < 2\epsilon \text{ then } \frac{T}{M} < \frac{a\kappa}{10}. \tag{78}$$

Remark 5.4 The points $x \in W_0$ go too close to the singularity under iteration by R_α, so that points of the form (x, s), $(x + \delta_0, s)$ split far apart before they get separated by a distance in P (Lemma 5.11 below). In other words, these points do not have the 'natural' WR-property that consists of a controlled drift starting from the first time the points split. To make sure these points cannot display the WR-Property in the future δ_0 is chosen in such a way, that if for large $MT_{\alpha,f}^M(x, s)$, $T_{\alpha,f}^M(x + \delta_0, s)$ become close, then nevertheless $d^f(T_{\alpha,f}^M(x, s), T_{\alpha,f}^M(x + \delta_0, s)) \gg \frac{1}{M^{1-\gamma}}$, and Lemma 5.5 then precludes the WR-property (see Lemma 5.10 below).

Before we prove these propositions we will see how they imply Theorem 1.

Proof of Theorem 1 Take $x \in W_0$ and $s > 0$ such that (x, s), $(x + \delta_0, s) \in V \times V$, which is possible since the measure of V is arbitrarily close to 1 if ε is sufficiently small. By Proposition 5.3, (x, s), $(x + \delta_0, s)$ satisfy (78), hence they don't satisfy (77), a contradiction. $\square$

Proof of Proposition 5.1

Lemma 5.5 *Let* $x, y \in \mathbb{T}$ *and let* I *be an integer interval such that for every* $n \in I$, $|f^{(n)}(x) - f^{(n)}(y)| < \eta$ *(where* η *is a sufficiently small number). Then* $|I| < 2c\eta^{1+\gamma}\|x - y\|^{\frac{-1}{1-\gamma}}$.

Proof We assume that $x < y$. Let $s \in \mathbb{N}$ be unique such that

$$\frac{1}{q_{s+1}^{1-\gamma}} \leqslant \|x - y\| < \frac{1}{q_s^{1-\gamma}}. \tag{79}$$

Denote $I = [a, b] \cap \mathbb{Z}$ with $a, b \in \mathbb{Z}$. Then, by the cocycle identity, the fact that $a \in I$, for $n \in \mathbb{Z}$, we have

$$|f^{(n)}(x) - f^{(n)}(y)| \geqslant |f^{(n-a)}(T^a x) - f^{(n-a)}(T^a y)| - \eta. \tag{80}$$

Let $k \in \mathbb{N}$ be unique such that

$$q_{k+1} \geqslant \frac{2\eta^{1+\gamma} q_{s+1}}{c} > q_k. \tag{81}$$

We will show that there exists $n_0 \in [0, q_{k+1}]$ such that

$$|f^{((n_0+a)-a)}(T^a x) - f^{((n_0+a)-a)}(T^a y)| = |f^{(n_0)}(T^a x) - f^{(n_0)}(T^a y)| > 2\eta. \tag{82}$$

This, by (80), gives $|f^{(n_0+a)}(x) - f^{(n_0+a)}(y)| > \eta$ and therefore $n_0 + a \notin I$. It follows that $|I| \leqslant q_{k+1} \leqslant cq_k < 2\eta^{1+\gamma} q_{s+1} \overset{(79)}{\leqslant} 2c\eta^{1+\gamma} \|x - y\|^{\frac{-1}{1-\gamma}}$ which completes the proof. Now, we show (82). By (81) and η sufficiently small, we have $s \geqslant k$.

Note that there exist $n_1 \in [0, q_{k+1})$ such that $T^a x + n_1\alpha \in [0, \frac{1}{q_{k+1}}]$. By (79) and the fact that $k + 1 \leqslant s + 1$, we obtain $T^a y + n_1\alpha \in [0, \frac{2}{q_{k+1}}]$. Therefore

$$|(f^{(n_1+1)}(T^a x) - f^{(n_1+1)}(T^a y)) - (f^{(n_1)}(T^a x) - f^{(n_1)}(T^a y))$$
$$| = |f(T^a x + n_1\alpha) - f(T^a y + n_1\alpha)| = |f'(\theta)|\|x - y\|, \tag{83}$$

for some $\theta \in [T^a x + n_1\alpha, T^a y + n_1\alpha] \subset [0, \frac{2}{q_{k+1}}]$. Thus, by the monotonicity of f' and (81)

$$|f'(\theta)|\|x - y\| \geqslant |\gamma| \left(\frac{2}{q_{k+1}}\right)^{-1+\gamma} \frac{1}{q_{s+1}^{1-\gamma}} = |\gamma| \left(\frac{q_{k+1}}{2q_{s+1}}\right)^{1-\gamma}$$

$$\geqslant |\gamma| \left(\frac{\eta^{1+\gamma}}{c}\right)^{1-\gamma} \geqslant 4\eta,$$

the last inequality by the fact that η is small enough. Therefore at least one of the numbers, $|f^{(n_1+1)}(T^a x) - f^{(n_1+1)}(T^a y)|$, $|f^{(n_1)}(T^a x) - f^{(n_1)}(T^a y)|$ is bigger than 2η; we set n_0 either n_1, or $n_1 + 1$ to obtain (82). □

The following lemma translates (75) into a property on the Birkhoff sums above R_α of the ceiling function f.

Lemma 5.6 *Let* $(x, s), (y, s') \in V$ *with* $d^f((x, s), (y, s')) < \delta$. *There exist* $M_0, L_0 \geqslant \frac{N}{2}$ *with* $\frac{L_0}{M_0} \geqslant \frac{\kappa}{2}$ *such that*

$$\frac{1}{L_0}|\{r \in [M_0, M_0 + L_0] \ : \ \exists m_r \in \mathbb{Z}, \ s.t \ |x - y - m_r\alpha|$$
$$< \epsilon \ and \ |f^{(r)}(x) - f^{(r+m_r)}(y) - p| < 2\epsilon\}| > a. \tag{84}$$

Proof Assume WLOG that $x < y$. Let $n \in [M, M + L]$ and r_n be unique such that $f^{(r_n)}(x) \leqslant n + s < f^{(r_n+1)}(x)$. We will show that

$$\frac{1+\kappa^2}{1-\kappa^2}M \geqslant r_M \geqslant \frac{1}{1+\kappa^2}M - 2. \tag{85}$$

Indeed, first we show that $r_M > N_0$. Indeed, if not, using Lemma 3.1 to $\frac{N}{2}$ (we have $\frac{N}{2} \geqslant N_0$)

$$M \leqslant M + s < f^{(r_M+1)}(x) < f^{(\frac{N}{2})}(x) < (1 + \kappa^2)\frac{N}{2} < N,$$

a contradiction. Secondly, by the fact that $(x, s) \in V$ (hence $s < \frac{1}{\epsilon^2}$) and $r_M \geqslant N_0$, using Lemma 3.1 to r_M and the definition of N ($N \geqslant \frac{1}{\epsilon^2\kappa^2}$), we get ($M \geqslant N$)

$$(1 - \kappa^2)r_M < f^{(r_M)}(x) \leqslant M + s < (1 + \kappa^2)M$$

and $M \leqslant M + s < f^{(r_M+1)}(x) \leqslant (1 + \kappa^2)(r_M + 1)$. Now, 85 follows.

Analogously we prove that

$$\frac{1}{1 + \kappa^2}(M + L) - 2 \leqslant r_{M+L} \leqslant \frac{1 + \kappa^2}{1 - \kappa^2}(M + L). \tag{86}$$

Set $M_0 := r_M$, $L_0 = r_{M+L} - r_M$. It is easy to prove using (85) and (86) that $M_0, L_0 \geqslant \frac{N}{2}$ and $\frac{L_0}{M_0} \geqslant \frac{\kappa}{2}$. Moreover, since $f \geqslant c_\gamma > 0$, there exists a constant $a = a(t_0, \gamma) > 0$ such that for every $n \in [M, M + L]$, $|r_{n+1} - r_n| \leqslant \frac{1}{2a}$. It follows that the number of different $r_n \in [M_0, M_0 + L_0]$ is at least $2aL_0$.

Let $n \in [M, M + L]$ be such that $d^f(T_n^f(x, s), T_{n+p}^f(y, s')) < \epsilon$. By (75), there are at least $(1 - \epsilon)L$ of such $n \in [M, M + L]$. By the definition of d^f and T^f, there exist $r_n \in [M_0, M_0 + L_0]$ and $m_n \in \mathbb{N}$ such that

$$|(x + r_n\alpha) - (y + m_n\alpha)| < \epsilon \quad \text{and} \quad |f^{(r_n)}(x) - f^{(m_n)}(y) - p| < 2\epsilon.$$

We set $m_r = m_r(n) := m_n - r_n \in \mathbb{Z}$ to get $|x - y - m_r\alpha| < \epsilon$ and $|f^{(r_n)}(x) - f^{(r_n+m_r)}(y) - p| < 2\epsilon$. It follows that the number of different $r_n \in [M_0, M_0 + L_0]$ is at least $2a(1 - \epsilon)L_0$ and hence (84) follows.   □

Proof of Proposition 5.1 Denote by

$$U := \{r \in [M_0, M_0 + L_0] \; : \; \exists_{m_r} \; |x - y - m_r\alpha|$$
$$< \epsilon \text{ and } |f^{(r)}(x) - f^{(r+m_r)}(y) - p| < 2\epsilon\}.$$

It follows by (84) that $|U| \geqslant aL_0$. Let us choose in the integer interval $[M_0, M_0 + L_0]$ disjoint subintervals $I_1 = [a_1, b_1], \ldots, I_l = [a_l, b_l]$ such that $U = I_1 \cup \ldots \cup I_l$ and for every $i = 1, \ldots, l$ there exists $m_i \in \mathbb{Z}$ such that $|x - y - m_i\alpha| < \epsilon$ and for $r \in I_i$, $|f^{(r)}(x) - f^{(r+m_i)}(y) - p| < 2\epsilon$. Moreover we assume that for every $i = 1, \ldots, l$, I_i is maximal in the sense that $|f^{(h_i)}(x) - f^{(h_i+m_i)}(y) - p| \geqslant 2\epsilon$ for $h_i = a_i - 1, b_i + 1$.

We will show that there exists $i = 1, \ldots, l$ such that

$$|I_i| \geqslant \frac{aL_0}{3}. \tag{87}$$

This will obviously finish the proof of (77) with $M' = a_i$, $L' = |I_i|$, and $m = m_i \in \mathbb{Z}$.

Let us show (87). If $l \leqslant 2$ there is nothing to prove. Assume $l \geqslant 3$.

Notice that U is the set of n's such that $(x, s), (y, s') \in V$ are p, n-close. The next lemma implies that between any two disjoint integer intervals $I_j, I_{j+1} \subset U$, on which $(x, s), (y, s')$ are p, n-close, there will be an integer interval J_j much longer than I_j, such that for any $n \in J_j$, $(x, s), (y, s')$ are not p, n-close.

Sublemma 5.7 *Let $i \in \{2, \ldots, l-1\}$. There exist an interval $[c_i, d_i] = J_i \subset [M_0, M_0 + L_0]$ such that for any $r \in J_i$, $4C^3 > |f^{(r)}(x) - f^{(r+m_i)}(y) - p| \geqslant 2\epsilon$, $c_i > b_{i-1}$, $d_i < a_{i+1}$ and $|J_i| \geqslant \frac{|I_i|}{4C^3\epsilon^{1+\gamma}}$ [here $C > 0$ comes from (3)].*

Sublemma 5.7 will give (87). Indeed, by the definition of J_i and I_i, it follows that for $i, j = 2, \ldots, l-1$ with $j \neq i-1, i, i+1$

$$J_i \cap I_i = \emptyset \quad \text{and} \quad J_i \cap J_j = \emptyset.$$

Hence, $\sum_{i=2}^{l-1} |J_i| \leqslant 3L_0$, and

$$|I_2 \cup \ldots \cup I_{l-1}| < 12C^3\epsilon^{1+\gamma}L_0 < \frac{aL_0}{3}$$

Therefore, by the fact that $|U| > aL_0$, we have $|I_1 \cup I_l| > \frac{2aL_0}{3}$ and consequently, $|I_w| \geqslant \frac{aL_0}{3}$ for at least one of $w = 0$ or $w = 1$.

Hence to obtain (87) we just need to prove Sublemma 5.7.

Proof of Sublemma 5.7 Let $v \in \mathbb{N}$ be unique such that

$$\frac{1}{q_{v+1}^{1-\gamma}} < \|x - (y + m_i\alpha)\| \leqslant \frac{1}{q_v^{1-\gamma}}. \tag{88}$$

Consider $n \in I_i = [a_i, b_i]$. We have

$$
\begin{aligned}
2\epsilon &\geqslant |f^{(n)}(x) - f^{(n+m_i)}(y) - p| \\
&= |(f^{(a_i)}(x) - f^{(a_i+m_i)}(y) - p) + (f^{(n-a_i)}(T^{a_i}x) - f^{(n-a_i)}(T^{a_i+m_i}y))| \\
&\geqslant \|f^{(n-a_i)}(T^{a_i}x) - f^{(n-a_i)}(T^{a_i+m_i}y)| - 2\epsilon|.
\end{aligned}
\tag{89}
$$

Hence, for $n \in I_i$, $|f^{(n-a_i)}(T^{a_i}x) - f^{(n-a_i)}(T^{a_i+m_i}y)| < 4\epsilon$. It follows now by Lemma 5.5 applied to $\eta = 4\epsilon$, the points $T^{a_i}x$, $T^{a_i+m_i}y$ and $s = v$ that

$$|I_i| < 2(4\epsilon)^{1+\gamma}\|x - (y + m_i\alpha)\|^{\frac{-1}{1-\gamma}} \leqslant 2(4\epsilon)^{1+\gamma}q_{v+1}. \qquad (90)$$

Consider integer intervals $K_i = [a_i - q_{v-2}, a_i]$ and $L_i = [a_i, a_i + q_{v-2}]$. It follows by (3) with $k = 1$ and for the point $T^{a_i}x$, similarly to the proof of Theorem 4, that there exist at most one $t_0 \in K_i \cup L_i$ such that $T^{a_i}x + t_0\alpha \in [-\frac{1}{2Cq_v}, \frac{1}{2Cq_v}]$. Assume $t_0 < 0$. Then we consider L_i. Moreover, we may assume that $\epsilon^{\frac{\gamma}{1-\gamma}} > 2cC$, and therefore, using (88) we obtain $\frac{1}{2Cq_v} \leqslant \frac{1}{q_v^{1-\gamma}}$. It follows that for $n \in [0, q_{v-2}]$, $0 \notin [T^{a_i+n}x, T^{a_i+m_i+n}y]$. Hence,

$$n \to sign(f(T^{a_i+n}x) - f(T^{a_i+m_i+n}x)) \text{ is constant for } n = 0, \ldots, q_{v-2}, \qquad (91)$$

(it may happen that $T^{m_i}y < x$). It follows that $|f^{(n)}(T^{a_i}x) - f^{(n)}(T^{a_i+m_i}y)|\big|_{n=0}^{q_{v-2}}$ is increasing. Hence, for $q_{v-2} > n \geqslant b_i + 1$ we have

$$f^{(n-a_i)}(T^{a_i}x) - f^{(n-a_i)}(T^{a_i+m_i}y) > 4\epsilon.$$

Moreover, by Lemma 4.3 (the RHS of the inequality) to $h = f, x = \theta$ (where $f^{(q_{v-2})}(T^{a_i}x) - f^{(q_{v-2})}(T^{a_i+m_i}y) = f'^{(q_{v-2})}(\theta)\|x - y - m_i\alpha\|$) and $s = v - 2$, we obtain

$$|f^{(n)}(T^{a_i}x) - f^{(n)}(T^{a_i+m_i}y)| < |f^{(q_{v-2})}(T^{a_i}x) - f^{(q_{v-2})}(T^{a_i+m_i}y)|$$
$$\leqslant 9C^2 + 4 < 2C^3,$$

(if necessary, to get the last inequality, we consider a bigger C). We set $J_i = [b_i+1, a_i+q_{v-2}](b_i \leqslant a_i+2(4\epsilon)^{1+\gamma}q_{v+1}$, by (90)). It follows that for $n \in J_i$, we have by cocycle identity

$$2\epsilon = 4\epsilon - 2\epsilon < |f^{(n-a_i)}(T^{a_i}x) - f^{(n-a_i)}(T^{a_i+m_i}y)| - |f^{(a_i)}(x)$$
$$- f^{(a_i+m_i)}(y) - p| \leqslant |f^{(n)}(x) - f^{(n+m_i)}(y) - p|$$
$$\leqslant |f^{(a_i)}(x) - f^{(a_i+m_i)}(y) - p| + |f^{(n-a_i)}(T^{a_i}x)$$
$$- f^{(n-a_i)}(T^{a_i+m_i}y)| \leqslant 2\epsilon + 2C^3 < 4C^3. \qquad (92)$$

Now, by (90)

$$d_i - c_i = |J_i| \geqslant q_{v-2} - 2(4\epsilon)^{1+\gamma}q_{v+1}$$
$$= 2(4\epsilon)^{1+\gamma}q_{v+1}\left(\frac{q_{v-2}}{2(4\epsilon)^{1+\gamma}q_{v+1}} - 1\right) \geqslant |I_i|\frac{1}{4C^3\epsilon^{1+\gamma}}, \qquad (93)$$

since $\frac{q_{v-2}}{q_{v+1}} \geqslant \frac{1}{c^3} \geqslant \frac{1}{C^3}$ and ϵ is small enough.

Suppose $n \in J_i \cap I_{i+1}$. If $m_{i+1} = m_i$, then immediately we have a contradiction. If $m_{i+1} \neq m_i$ then

$$
\begin{aligned}
|f^{(n)}(x) - f^{(n+m_{i+1})}(y) - p| &= |(f^{(n)}(x) - f^{(n+m_i)}(y) - p) \\
&\quad + f^{(m_{i+1}-m_i)}(T^{n+m_i})y \\
&\geqslant |m_{i+1} - m_i| \inf_{\mathbb{T}} f - 4C^3 > 2\epsilon, \quad (94)
\end{aligned}
$$

since $|m_{i+1} - m_i|$ is of order $\frac{1}{\epsilon}$ because both $\|x - y - m_{i+1}\alpha\|$ and $\|x - y - m_i\alpha\|$ are close to zero. Hence $d_i < a_{i+1}$ and Lemma 5.7 has been proved. □

This finishes the proof of Proposition 5.1. □

Proof of Proposition 5.3

Sublemma 5.8 *Fix a number* $0 < \zeta \leqslant \frac{|\gamma|}{50}$. *For every* $v \in \mathbb{N}$, $v \geqslant v_0$ *and* v_0 *sufficiently large, there exists* $0 < \delta_0^v < \delta$, *satisfying*

$$
\delta_0^v \in \left[\frac{1}{q_v^{1-\gamma}}, \frac{2c}{q_v^{1-\gamma}} \right], \quad (95)
$$

such that for every $\mathbb{N} \ni |k| \geqslant \varepsilon^{-\frac{1}{2}}$,

$$
\|\delta_0^v - k\alpha\| \geqslant \frac{1}{|k|^{1+\zeta}}. \quad (96)
$$

Sublemma 5.9 *For every* $w \geqslant w_0$, w_0 *sufficiently large, there exists a set*

$$
W_0^w \subset A \cap (A - \delta_0), \quad (97)
$$

with $\lambda(W_0^w) \geqslant c_0(c_0(d)$ *will be specified in the proof), such that the following holds for* $x \in W_0^w$ *and* $y := x + \delta_0^w$:

$$
|f^{(n)}(x) - f^{(n)}(y)| < 100dc \quad \text{for every } n = 0, \ldots, q_{w-2} \quad (98)
$$

there exists $i_0 \in \{0, \ldots, q_{w-l} - 1\}$ *such that*

$$
x + i_0\alpha \in \left[\frac{1}{2dq_w}, \frac{1}{dq_w} \right], \quad (99)
$$

for some $\mathbb{N} \ni l \geqslant 1$ *depending on w, to be specified later,*

$$0 < f^{(n)}(x) - f^{(n)}(y) < \frac{100c}{d} \quad \text{for } n \leqslant i_0 \quad \text{and} \quad f^{(n)}(x) - f^{(n)}(y)$$

$$> \frac{|\gamma|d}{2} \quad \text{for } i_0 < n < wq_w. \tag{100}$$

Before we proof the above Lemmas, let us first show, how they imply Proposition 5.3.

Let $w \in \mathbb{N}$ be such that

$$\frac{1}{2}w \geqslant \frac{1}{\kappa^2}. \tag{101}$$

Denote $W_0 := W_0^w$ and $\delta_0 := \delta_0^w$. By definition,

$$\delta_0 \in \left[\frac{1}{q_w^{1-\gamma}}, \frac{2c}{q_w^{1-\gamma}} \right]. \tag{102}$$

Lemma 5.10 *We have that* (78) *holds for* $k \neq 0$.

Proof Fix any $M \geqslant \frac{N}{2}$, any $p \in P$ and any $k \neq 0$ such that $\|x - y - k\alpha\| < \epsilon$.

Let $I = [M, M + R]$ be such that for $n \in [M, M + R]$, we have $|f^{(n)}(x) - f^{(n+k)}(y) - p| < 2\epsilon$.

Note that since $\|x - y - k\alpha\| < \varepsilon$ and $\|x - y\| < \delta < \varepsilon$, then $|k| \geqslant \frac{1}{2c^3\epsilon}$. By the cocycle identity and the triangle inequality, this implies that for every $n \in [0, R]$

$$|f^{(n)}(x + M\alpha) - f^{(n)}(y + M\alpha + k\alpha)| \leqslant 4\epsilon.$$

Therefore, be Lemma 5.5,

$$R < 2c(4\varepsilon)^{1+\gamma} \|x - y + k\alpha\|^{\frac{-1}{1-\gamma}} \overset{(96)}{\leqslant} 2c(4\varepsilon)^{1+\gamma} |k|^{\frac{1+\zeta}{1-\gamma}} \leqslant |k|^{\frac{1+\zeta}{1-\gamma}}. \tag{103}$$

On the other hand, since $x \in W_0$ and $|f^{(n)}(x) - f^{(n+k)}(y) - p| < 2\epsilon$ we get that

$$2\kappa M > |f^{(M)}(x) - f^{(M)}(y)| \geqslant |f^{(k)}(T^M y)| - 2\epsilon - p$$

$$\geqslant |k| \inf_{\mathbb{T}} f - 2\epsilon - p > \frac{1}{2}|k| \inf_{\mathbb{T}} f \geqslant \frac{\inf_{\mathbb{T}} f}{4c^3\varepsilon}. \tag{104}$$

Hence, and using (98) (ϵ is small enough),

$$M > \max\left(q_{w-2}, \frac{1}{2}|k| \right). \tag{105}$$

Hence, by (103) and (105),

$$\frac{R}{M} \leqslant \frac{|k|^{\frac{1+\zeta}{1-\gamma}}}{\max\left(q_{w-2}, \frac{|k|}{2}\right)} \leqslant 2q_{w-2}^{\frac{\gamma-\zeta}{1-\gamma}} \overset{(101)}{<} \frac{a\kappa}{10}.$$

This proves (78) for $k \neq 0$. □

Lemma 5.11 *We have that* (78) *holds for* $k = 0$.

Proof Fix any $M \geqslant \frac{N}{2}$, $p \in P$ and let $I = [M, M + R]$ be such that for every $n \in I$, $|f^{(n)}(x) - f^{(n)}(y) - p| < 2\epsilon$. By (100) and (74), for every $n \in [0, wq_w]$ and every $p \in P$, $|f^{(n)}(x) - f^{(n)}(y) - p| > 2\epsilon$. Therefore, for $M \leqslant \frac{1}{2}wq_w$, (78) holds. It follows by Lemma 5.5 applied to x and y (we have $\|x - y\| = \delta_0 \overset{(102)}{\geqslant} \frac{1}{q_w^{1-\gamma}}$), that $R = |I| < 2(4\epsilon)^{1+\gamma}\|x - y\|^{\frac{-1}{1-\gamma}} \leqslant 2(4\epsilon)^{1+\gamma}q_w < q_w$. Therefore, by (101), for $M > \frac{1}{2}wq_w$,

$$\frac{R}{M} \leqslant \frac{q_w}{\frac{1}{2}wq_w} < \frac{a\kappa}{10}.$$

So, (78) holds for $k = 0$. □

We thus proved (78) in Proposition 5.3. Let us now complete the proof by proving Sublemmas 5.8 and 5.9.

Proof of Sublemma 5.8 Fix $v \in \mathbb{N}$. To simplify the notations, we will write δ_0 instead of δ_0^v. Given $u \in \mathbb{N}$, set

$$B_u := \left\{\eta \in \mathbb{T} \ : \ d\left(\eta, \{i\alpha\}_{i=-q_u}^{q_u-1}\right) \geqslant \frac{1}{2u^2q_u}\right\}. \tag{106}$$

Let $t \in \mathbb{N}$ be unique such that

$$\frac{1}{q_{t+2}} < \frac{1}{q_v^{1-\gamma}} \leqslant \frac{1}{q_{t+1}}. \tag{107}$$

Let $c_1 = 4c$, then $q_{t+1} \geqslant 4cq_{t-c_1}$ (since t depends on v which is sufficiently large, $t - c_1 > v - 4$ by (107)).

$$\left[\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}}\right] \cap \left(\bigcap_{i \geqslant t-c_1} B_i\right). \tag{108}$$

We will show below that this set is not empty. Now, let δ_0 be any number in this set. This, by the definition of B_i, will give (95) and (96). Indeed, (95) follows from (107) and (108). To show (96), note that for $|k| < q_{t-c_1}$,

$$\|\delta_0 - k\alpha\| \overset{(95),(107)}{\geqslant} \|k\alpha\| - \delta_0 \geqslant \frac{1}{2c|k|} - \frac{1}{q_{t+1}} \geqslant \frac{1}{2c|k|} - \frac{1}{4cq_{t-c_1}} \geqslant \frac{1}{|k|^{1+\varsigma}},$$

since $|k| \geqslant \varepsilon^{-\frac{1}{2}}$. If $|k| \geqslant q_{t-c_1}$, let ℓ be unique, such that $q_{\ell+1} > |k| \geqslant q_\ell$. By definition of $B_{\ell+1}$

$$\|\delta_0 - k\alpha\| \geqslant \sup_{|i| \leqslant q_{\ell+1}} \|\delta_0 - i\alpha\| \geqslant \frac{1}{(\ell+1)^2 q_{\ell+1}} \geqslant \frac{1}{|k|}^{1+\varsigma},$$

where the last inequality follows if v is sufficiently large (then by (107), t is large, so $|k|$ is large and therefore ℓ is large).

Sublemma 5.12

$$\left[\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}}\right] \cap \left(\bigcap_{i \geqslant t-c_1} B_i\right) \neq \emptyset. \tag{109}$$

Proof The proof goes by induction. We will show that for every $k \geqslant t - c_1$ there exists a closed interval $E_k \subset \left[\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}}\right] \cap \left(\bigcap_{i \geqslant t-c_1}^{k} B_i\right)$ and $E_{k+1} \subset E_k$. Moreover, we will show (for the induction purpose) that for every $k \geqslant t - c_1, |E_k| \geqslant \frac{1}{c^{c_1+2}q_k}$. Indeed, we have $\min_{-q_{t-c_1}+1 \leqslant i \leqslant q_{t-c_1}-1} \|i\alpha\| \geqslant \frac{1}{2q_{t-c_1}} \geqslant \frac{2}{q_{t+1}}$. Set $E_{t-c_1} := [\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}} - \frac{1}{2(t-c_1)^2 q_{t-c_1}}]$. It follows by the fact that $t - c_1 \geqslant v - 4$ and v is sufficiently large (by taking a bigger c, we may assume that $c > 2$) that

$$|E_{t-c_1}| = \frac{1}{q_{t+1}} - \frac{1}{2(t-c_1)^2 q_{t-c_1}} \geqslant \frac{1}{c^{c_1+2}q_{t-c_1}} \left(c - \frac{1}{c}\right) \geqslant \frac{1}{c^{c_1+2}q_{t-c_1}}.$$

Moreover, since $t - c_1$ is sufficiently large, $\frac{1}{q_{t+1}} > \frac{1}{2(t-c_1)^2 q_{t-c_1}}$. Moreover, $c_1 \geqslant 4$ and therefore $\frac{2}{q_{t+1}} - \frac{1}{2(t-c_1)^2 q_{t-c_1}} \leqslant \min_{i \in \{-q_{t-c_1},...,q_{t-c_1}-1\}} \|i\alpha\| - \frac{1}{2(t-c_1)^2 q_{t-c_1}}$. Therefore, by definition of E_{t-c_1} and B_{t-c_1},

$$E_{t-c_1} \subset B_{t-c_1} = \bigcap_{i=t-c_1}^{t-c_1} B_i.$$

Suppose that for some $k \geqslant t - c_1$ we have a closed interval $E_k \subset \left[\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}} \right] \cap \left(\bigcap_{i=t-c_1}^{k} B_i \right)$ such that $|E_k| \geqslant \frac{1}{c^{c_1+2}q_k}$. It follows that

$$E_k \cap \{i\alpha\}_{i=-q_k}^{q_k-1} = \emptyset.$$

Let $E_{k+1} \subset E_k$ be the longest closed subinterval (in E_k) such that

$$E_{k+1} \cap B_{k+1}^c = \emptyset. \tag{110}$$

It follows that $E_{k+1} \subset E_k \subset \left[\frac{1}{q_{t+1}}, \frac{2}{q_{t+1}} \right]$, and by (110), $E_{k+1} \subset \left(\bigcap_{i=t-c_1}^{k+1} B_i \right) (E_{k+1} \subset E_k)$. It remains to prove that $|E_{k+1}| \geqslant \frac{1}{c^{c_1+2}q_{k+1}}$. To do this note that

$$|E_{k+1}| \geqslant \frac{|E_k|}{\left[4|E_k|q_{k+1} + 1 \right]} - \frac{1}{(k+1)^2 q_{k+1}}.$$

Indeed, $|E_k \cap \{i\alpha\}_{-q_{k+1}}^{q_{k+1}-1}| \leqslant 4|E_k|q_{k+1}$ and around each point of the form $i\alpha, i = -q_{k+1}, \ldots, q_{k+1} - 1$, we discard an interval of length $\frac{1}{(k+1)^2 q_{k+1}}$ (see (106), for $u = k + 1$). We use the induction assumption, the fact that $k + 1 \geqslant t - c_1 \geqslant v - 4$ (and v is sufficiently large) to obtain

$$\frac{|E_k|}{\left[4|E_k|q_{k+1} + 1 \right]} - \frac{1}{(k+1)^2 q_{k+1}} \geqslant \frac{1}{c^{c_1+2}q_{k+1}}.$$

Hence (109) is proved. □

The proof of Sublemma 5.8 is thus finished. □

Proof of Sublemma 5.9 We will determine the set W_0^w. To simplify notation, we will write W_0 instead of W_0^w, the dependence on w will be clear from the context. Let $l \in \mathbb{N}$ be such that

$$\left(\frac{q_{w-l}}{q_w} \right)^{-\gamma+1} \leqslant \frac{1}{d} < \left(\frac{q_{w-l+1}}{q_w} \right)^{-\gamma+1}, \tag{111}$$

since w is large, $l \geqslant 1$. Set

$$W_{0,1} := \left\{ x \in \mathbb{T} : \{x, x+\alpha, \ldots, x+(wq_w - 1)\alpha\} \cap \left[-\frac{2c}{q_w^{1-\gamma}}, \frac{2c}{q_w^{1-\gamma}} \right] = \emptyset \right\},$$

$$W_{0,2} := \left\{ x \in \mathbb{T} : \{x, x+\alpha, \ldots, x+(q_{w-l} - 1)\alpha\} \cap \left[\frac{1}{2dq_w}, \frac{1}{dq_w} \right] \neq \emptyset \right\}.$$

We have $\lambda(W_{0,1}) \geqslant 1 - wq_w \frac{4c}{q_w^{1-\gamma}} = 1 - \frac{4cw}{q_w^{-\gamma}}$.

As $l \geqslant 1$, for $i = 0, \ldots, q_{w-l} - 1$, the sets $T^i[\frac{1}{2dq_w}, \frac{1}{dq_w}]$ are pairwise disjoint. Therefore, by (111)

$$\lambda(W_{0,2}) = q_{w-l}\frac{1}{2dq_w} \geqslant \frac{1}{c}\frac{q_{w-l+1}}{q_w}\frac{1}{2d} \geqslant \frac{1}{2c}\left(\frac{1}{d}\right)^{\frac{1}{1-\gamma}}\frac{1}{d} = \frac{1}{2c}\left(\frac{1}{d}\right)^{\frac{2-\gamma}{1-\gamma}}.$$

Now we set

$$W_0 := W_{0,1} \cap W_{0,2}. \tag{112}$$

Since $w \geqslant w_0$ is sufficiently large, $\lambda(W_0) \geqslant c_0$, where $c_0 = c_0(d) > 0$. We may assume that

$$W_0 \subset A \cap (A - \delta_0),$$

if not we take $W_0 := W_0 \cap A \cap (A - \delta_0)$ and use the fact that $\lambda(A) > 1 - \epsilon$, and $\delta_0 < \delta$ is small. This gives (97). Note that by (112) and the definition of $W_{0,2}$, (99) follows. Let us show (98).

By (99), we have

$$\{x, x + \alpha, \ldots, x + (q_{w-2} - 1)\alpha\} \cap \left[0, \frac{1}{2dq_w}\right] = \emptyset. \tag{113}$$

Therefore, using (102) and $\|x - y\| = \delta_0$, we have for every $i = 0, \ldots, q_{w-2} - 1$

$$T^i([x, y]) \cap \left[0, \frac{1}{6dq_w}\right] = \emptyset. \tag{114}$$

By (114), for $i = 0, \ldots, q_{w-2} - 1, 0 \notin [x + i\alpha, y + i\alpha]$, and therefore $|f^{(i)}(x) - f^{(i)}(y)| \leqslant |f^{(q_{w-2})}(x) - f^{(q_{w-2})}(y)|$. Therefore and by (114), (102), monotonicity of f', (114), Lemma 4.3 (to $h = f'$, and some $\theta \in [x, y]$), it follows that for $n \leqslant q_{w-2}$

$$|f^{(n)}(x) - f^{(n)}(y)| \leqslant |f^{(q_{w-2})}(x) - f^{(q_{w-2})}(y)| \leqslant \|x - y\||f'^{(q_{w-2})}(\theta)|$$
$$\leqslant \frac{2c}{q_w^{1-\gamma}}26dq_w^{1-\gamma} \leqslant 100dc, \tag{115}$$

and (98) follows.

Now we will show (100).

Moreover, for $x \in W_0 \subset W_{0,1}$, and $y := x + \delta_0, 0 \notin [x + i\alpha, y + i\alpha]$ for $i = 0, \ldots, wq_w - 1$

$$\left(f^{(n)}(x) - f^{(n)}(y)\right)_{n=0}^{wq_w - 1} \text{ is an increasing sequence.} \tag{116}$$

By the fact that $i_0 < q_{w-l}$ and (99) it follows that $\|x + j\alpha - 0\| > \frac{1}{2q_{w-l}}$ for $j \neq i_0, 0 \leqslant j \leqslant q_{w-l} - 1$. Moreover, since $w \geqslant w_0$ is large enough, $\|x - y\| = \delta_0 \leqslant \frac{2c}{q_w^{1-\gamma}} < \frac{1}{6q_w}$, and we obtain

$$\|x + j\alpha\|, \|y + j\alpha\| > \frac{1}{6q_{w-l}} \quad \text{for } j = 0, \ldots, q_{w-l} - 1, j \neq i_0. \quad (117)$$

Moreover, by (116), for $n \leqslant i_0$

$$0 < f^{(n)}(x) - f^{(n)}(y) \leqslant \left(f^{(q_{w-l})}(x) - f^{(q_{w-l})}(y) \right) \\ - (f(x + i_0\alpha) - f(y + i_0\alpha)). \quad (118)$$

Let us consider

$$\bar{f}(x) = \begin{cases} f(x) & \text{if } x > \frac{1}{6q_{w-l}}; \\ 0 & \text{if otherwise.} \end{cases}$$

Hence, by (117)

$$f^{(q_{w-l})}(x) - f(x + i_0\alpha) = \bar{f}^{(q_{w-l})}(x) \text{ and } f^{(q_{w-l})}(y) - f(y + i_0\alpha) \\ = \bar{f}^{(q_{w-l})}(y). \quad (119)$$

By $0 \notin [x + i\alpha, y + i\alpha]$ for $i = 0, \ldots, wq_w - 1$ there exists $\theta \in [x, y]$ such that $\bar{f}^{(q_{w-l})}(x) - \bar{f}^{(q_{w-l})}(y) = \|x - y\| |\bar{f}'^{(q_{w-l})}(\theta)|$. But as in Lemma 4.3 we get the following:

$$|\bar{f}'^{(q_{w-l})}(\theta)| \leqslant q_{w-l}(2q_{w-l})^{-\gamma} + 2|\gamma|(2q_{w-l})^{1-\gamma} + |\gamma|(6q_{w-l})^{1-\gamma} \\ \leqslant 46q_{w-l}^{1-\gamma}.$$

Therefore, using (118) and (119), (102) and (111), for every $n \leqslant i_0$, we have

$$0 < f^{(n)}(x) - f^{(n)}(y) \leqslant \|x - y\| |\bar{f}'^{(q_{w-l})}(\theta)| \leqslant 46q_{w-l}^{1-\gamma} \frac{2c}{q_w^{1-\gamma}} < \frac{100c}{d}.$$

For $wq_w > n > i_0$, by (116) and monotonicity of f', we have for some $\theta_0 \in [x + i_0\alpha, y + i_0\alpha] \overset{(99),(102)}{\subset} [0, \frac{2}{dq_w}]$, and

$$f^{(n)}(x) - f^{(n)}(y) \geqslant f(x + i_0\alpha) - f(y + i_0\alpha) = \|x - y\| |f'(\theta_0)| \\ \geqslant \frac{1}{q_w^{1-\gamma}} |\gamma| \left(\frac{dq_w}{2} \right)^{1-\gamma} > \frac{|\gamma|d}{2}.$$

This finishes the proof of (100).

The proof of Sublemma 5.9 is complete. □

This finishes the proof of Theorem 1.

Appendix B: Proof of Proposition 1.7

Consider the Gauss map $T : [0, 1) \to [0, 1)$, $Tx := \left\{\frac{1}{x}\right\}$, $T(0) = 0$, and let μ be its invariant probability measure given by its density with respect to the Lebesgue measure $\frac{1}{\log 2} \frac{1}{1+x} dx$.

Lemma 6.1 *There exists a constant $C > 0$ such that for every $a \in \mathbb{T}$ and for every $k \neq l \in \mathbb{N}$*

$$\mu(T^{-k}((0, a)) \cap T^{-l}((0, a))) \leqslant C\mu((0, a))^2. \qquad (120)$$

Proof Assume that $l > k$. Then (μ is T-invariant)

$$\mu(T^{-k}((0, a)) \cap T^{-l}((0, a))) = \mu((0, a) \cap T^{k-l}(0, a)).$$

Note that $T^{k-l+1}(0, a) = \bigcup_{i=1}^{+\infty}(c_i, d_i)$ for some disjoint intervals $(c_i, d_i)i = 1, \ldots, +\infty$. We will prove that for every $i \in \mathbb{N}$

$$\mu(T^{-1}(c_i, d_i) \cap (0, a)) \leqslant C\mu((c_i, d_i))\mu((0, a)) \qquad (121)$$

which implies (120) since

$$\mu((0, a) \cap T^{k-l}(0, a)) \leqslant C\mu((0, a))\mu\left(\bigcup_{i=1}^{+\infty}(c_i, d_i)\right)$$
$$= C\mu(T^{k-l+1}(0, a))\mu((0, a)) = C\mu((0, a))^2.$$

To prove (121), note that $T^{-1}(c_i, d_i) = \bigcup_{j=1}^{+\infty}(\frac{1}{d_i+j}, \frac{1}{c_i+j})$. It follows that

$$\sum_{j=1}^{+\infty}\left(\frac{1}{d_i + j}, \frac{1}{c_i + j}\right) \cap (0, a) \subset \bigcup_{j \geqslant \frac{1}{a}-d_i}\left(\frac{1}{d_i + j}, \frac{1}{c_i + j}\right).$$

Therefore,

$$\mu(T^{-1}(c_i, d_i) \cap (0, a)) \leqslant \sum_{j \geqslant \frac{1}{a} - d_i} \mu\left(\left(\frac{1}{d_i + j}, \frac{1}{c_i + j}\right)\right)$$

$$\leqslant C \sum_{j \geqslant \frac{1}{a} - d_i} \frac{c_i - d_i}{(c_i + j)(d_i + j)} \leqslant C\mu((0, a))\mu((c_i, d_i))$$

for some constant $C > 0$ (since the density function $f(x) = \frac{1}{1+x}$ is bounded from above and below on $[0, 1]$). This completes the proof. $\qquad\square$

Proposition 6.2 *Let $d > 0$ and set*

$$\mathcal{A} := \left\{x = [0; a_1, ...] : \exists_{N_0 = N_0(x)} \forall_{n \geqslant N_0} \left|\left\{k \in [n^2, (n+1)^2] : a_k \geqslant dk^{\frac{7}{8}}\right\}\right| < 2\right\}.$$

Then $\lambda(\mathcal{A}) = 1$.

Proof We will prove that $\lambda(\mathcal{A}^c) = 0$. To do this we will prove that $\mu(\mathcal{A}^c) = 0$ (λ and μ are equivalent). Note that for $k \in \mathbb{N}$ if $x = [0; a_1, \ldots,]$ is the continued fraction of x, then $T^k(x) = \cfrac{1}{a_k + \cfrac{1}{a_{k+1} + \cdots}}$. Therefore $\mathcal{B} \subset \mathcal{A}$, where

$$\mathcal{B} := \left\{x \in \mathbb{T} : \exists_{N_0 = N_0(x)} \forall_{n \geqslant N_0} \left|\left\{k \in [n^2, (n+1)^2] : T^k x \leqslant \frac{1}{dk^{\frac{7}{8}}}\right\}\right| < 2\right\}.$$

We will prove that $\mu(\mathcal{B}^c) = 0$. To do this note that

$$\mathcal{B}^c = \bigcap_{N_0 = 1}^{+\infty} \left(\bigcup_{n \geqslant N_0} B_n\right), \tag{122}$$

where $B_n := \left\{x \in \mathbb{T} : \left|\{k \in [n^2, (n+1)^2] : T^k x \leqslant \frac{1}{dk^{\frac{7}{8}}}\}\right| \geqslant 2\right\}$. Moreover,

$$B_n \subset \bigcup_{i_1 \neq i_2 \in [n^2, (n+1)^2]} B_{i_1, i_2}^n, \tag{123}$$

where $B_{i_1, i_2}^n := \{x \in \mathbb{T} : T^{i_1} x, T^{i_2} x \in (0, \frac{1}{dn^{\frac{7}{4}}}]\}$. Let us note that

$$B_{i_1, i_2}^n = T^{-i_1}\left(\left(0, \frac{1}{dn^{\frac{7}{4}}}\right)\right) \cap T^{-i_2}\left(\left(0, \frac{1}{dn^{\frac{7}{4}}}\right)\right).$$

By (120) from Lemma 6.1, we get that $\mu(B^n_{i_1,i_2}) \leqslant Cn^{-\frac{7}{2}}$. Therefore, using (123) and summing up over all $i_1 \neq i_2 \in [n^2, (n+1)^2]$, we get that $\mu(B_n) \leqslant Cn^{-\frac{3}{2}}$. This and (122) yield

$$\mu(\mathcal{B}^c) = \lim_{N_0 \to +\infty} \mu\left(\bigcup_{n \geqslant N_0} B_n\right) = 0.$$

This finishes the proof. □

Lemma 6.3 *Let $\alpha \in \mathcal{A}$. Then $\sum_{s \notin K_\alpha} \frac{1}{\log^{\frac{7}{8}} q_s} < +\infty$.*

Proof Let $N_0 := N_0(\alpha)$ be the number resulting from the fact that $\alpha \in \mathcal{A}$. We will prove that $\sum_{s \notin K_\alpha, s \geqslant N_0} \frac{1}{\log^{\frac{7}{8}}(q_s)} < +\infty$. There exists a constant $d > 0$ such that for any $s \in \mathbb{N}$

$$\log(q_s) \geqslant (2d)^{\frac{8}{7}} s \tag{124}$$

(indeed, the sequence $(q_s)_{s=1}^{+\infty}$ grows exponentially fast). Let $s \notin K_\alpha, s \geqslant N_0$. Then

$$a_{s+1} q_s + q_{s-1} = q_{s+1} \geqslant q_s \log^{\frac{7}{8}} q_s,$$

and therefore, for $s \notin K_\alpha$, by (124)

$$a_{s+1} \geqslant (\ln(q_s))^{\frac{7}{8}} - 1 \geqslant ds^{\frac{7}{8}}. \tag{125}$$

Since $\alpha \in \mathcal{A}$, for every $k \geqslant 1$ in every interval of the form $[(N_0 + k)^2, (N_0 + k + 1)^2]$ there is at most one s such that $a_{s+1} \geqslant ds^{\frac{7}{8}} \geqslant (d(N_0 + k)^{\frac{7}{8}})$.

Therefore

$$\sum_{s \notin K_\alpha, s \geqslant N_0} \frac{1}{\log^{\frac{7}{8}} q_s} \leqslant 2d \sum_{s \notin K_\alpha, s \geqslant N_0} \frac{1}{s^{\frac{7}{8}}} \leqslant 2d \sum_{k \geqslant 1} \frac{1}{(N_0 + k)^{2\frac{7}{8}}}$$

$$= 2d \sum_{k \geqslant 1} \frac{1}{(N_0 + k)^{\frac{7}{4}}} < +\infty.$$

This finishes the proof. □

Hence we proved that $\mathcal{A} \subset \mathcal{E}$ and $\lambda(\mathcal{A}) = 1$, therefore $\lambda(\mathcal{E}) = 1$. □

References

1. Arnold, V.: Topological and ergodic properties of closed 1-forms with incommensurable periods. Funktsionalnyi Analiz i Ego Prilozheniya **25**(2), 1–12 (1991). [Translated In: Functional Analyses and its Applications **25**(2), 81–90 (1991)]
2. Avìla, A., Forni, G., Ulcigrai, C.: Mixing times changes for Heisenberg nilflows, preprint. arXiv:1003.4636
3. Chaika, J., Wright, A.: A smooth mixing flow on a surface with non-degenerate fixed points. arXiv:1501.02881v1
4. Cornfeld, I.P., Fomin, S.V., Sinai, Ya.G.: Ergodic theory. In: Grundlehren der Math. Wissenschaften, vol. 245. Springer, New York (1982)
5. Einsiedler, M., Tseng, J.: Badly approximable systems of affine forms, fractals, and Schmidt games. J. Reine Angew. Math. **660**, 83–97 (2011)
6. Fayad, B.: Polynomial decay of correlations for a class of smooth flows on the two torus. Bull. SMF **129**, 487–503 (2001)
7. Fayad, B.: Analytic mixing reparametrizations of irrational flows. Erg. Theory Dynam. Syst. **22**, 437–468 (2002)
8. Fayad, B., Lemańczyk, M.: On the ergodicity of cylindrical transformations given by the logarithm. Mosc. Math. J. **6**, 657–672 (2006)
9. Frączek, K., Lemańczyk, M.: On mild mixing of special flows over irrational rotations under piecewise smooth functions. Ergod Theory Dynam. Syst. **26**, 1–21 (2006)
10. Frączek, K., Lemáczyk, M.: Ratner's property and mild mixing for special flows over two-dimensional rotations. J. Mod. Dyn. **4**, 609–635 (2010)
11. Frączek, K., Lemańczyk, M., Lesigne, E.: Mild mixing property for special flows under piecewise constant functions. Discrete Contin. Dyn. Syst. **19**, 691–710 (2007)
12. Glasner, E.: Ergodic Theory via Joinings. In: Mathematical Surveys and Monographs, vol. 101. American Mathematical Society, Providence (2003)
13. Host, B.: Mixing of all orders and pairwise independent joinings of system with singular spectrum. Israel J. Math. **76**, 289–298 (1991)
14. Kalikow, S.: Twofold mixing implies threefold mixing for rank one transformations. Ergod. Theory Dynam. Syst. **7**, 531–557 (1987)
15. Khinchin, A.Y.: Continued fractions. Dover Publications Inc., Mineola, NY, russian edn (1997). ISBN 0-486-69630-8 (with a preface by B. V. Gnedenko, Reprint of the 1964 translation)
16. Kochergin, A.V.: On the absence of mixing in special flows over the rotation of a circle and in flows on a two-dimensional torus. Dokl. Akad. Nauk SSSR **205**, 949–952 (1972)
17. Kochergin, A.V.: Mixing in special flows over a shifting of segments and in smooth flows on surfaces. Mat. Sb. **96**(138), 471–502 (1975)
18. Kochergin, A.V.: Nonsingular saddle points and the absence of mixing. Mat. Zametki 19 **3**, 453–468 (1976) (Translated. In: Math. Notes **19**(3), 277–286)
19. Kochergin, A.V.: Nondegenerate fixed points and mixing in flows on a two-dimensional torus I. Sb. Math. **194**, 1195–1224 (2003)
20. Kochergin, A.V.: Nondegenerate fixed points and mixing in flows on a two-dimensional torus II. Sb. Math. **195**, 317–346 (2004)
21. Kochergin, A.V.: Well-approximable angles and mixing for flows on $\mathbb{T}^2$ with nonsingular fixed points. Electron. Res. Announc. Am. Math. Soc. **10**, 113–121 (2004)
22. Lemánczyk, M.: Sur l'absence de melange pour des flots speciaux au-dessus d'une rotation irrationnelle. Colloq. Math. **84/85**(part 1), 29–41 (2000)
23. Marcus, B.: The horocycle flow is mixing of all degrees. Invent. Math. **46**(3), 201–209 (1978)
24. Ornstein, D.S., Smorodinsky, M.: Continuous speed changes for flows. Israel J. Math. **31**(2), 161–168 (1978)

25. Ornstein, D., Rudolph, D., Weiss, B.: Equivalence of measure preserving transformations. Memors AMS **37**(262) (1982)
26. Rokhlin, V.: On endomorphisms of compact commutative groups. Izv. Akad. Nauk SSSR **13**, 329–340 (1949)
27. Ratner, M.: Horocycle flows, joinings and rigidity of products. Ann. Math. **118**, 277–313 (1983)
28. Ratner, M.: Rigidity of horocycle flows. Ann. Math. **115**, 597–614 (1982)
29. Ratner, M.: Horocycle flows are loosely Bernoulli. Israel J. Math. **31**, 122–132 (1978)
30. Ryzhikov, V.V.: Joinings and multiple mixing of finite rank actions. Funct. Anal. Appl. **27**(2), 128–140 (1993)
31. Ryzhikov, V.V., Thouvenot, J.-P.: Disjointness, divisibility, and quasi-simplicity of measure-preserving actions. Funkts. Anal. Prilozh. **40**(3), 85–89 (2006)
32. Sinai, Ya.G., Khanin, K.M.: Mixing for some classes of special flows over rotations of the circle. Funktsionalnyi Analiz i Ego Prilozheniya 26 **3**, 1–21 (1992) [(Translated In: Functional Analysis and its Applications **26**(3), 155–169 (1992)]
33. Thouvenot, J.-P.: Some properties and applications of joinings in ergodic theory. In: Ergodic theory and its connections with harmonic analysis (Aleksandria, 1993), London Math. Soc. Lecture Note Series, vol. 205. Cambridge University Press, Cambridge, pp. 207–235 (1995)
34. Tseng, J.: Badly approximable affine forms and Schmidt games. J. Number Theory **129**(12), 3020–3025 (2009)
35. Ulcigrai, C.: Mixing of asymmetric logarithmic suspension flows over interval exchange transformations. Ergod. Theory Dyn. Syst. **27**, 991–1035 (2007)
36. Ulcigrai, C.: Weak mixing for logarithmic flows over interval exchange transformations. J. Mod. Dynam. **3**, 35–49 (2009)
37. Ulcigrai, C.: Absence of mixing in area-preserving flows on surfaces. Ann. Math. **173**, 1743–1778 (2011)