École Polytechnique
Université Paris-Saclay

Journal de l'École polytechnique
Mathématiques

Stefano Galatolo

Quantitative statistical stability, speed of convergence to equilibrium and partially hyperbolic skew products

Tome 5 (2018), p. 377-405.

<http://jep.cedram.org/item?id=JEP_2018__5__377_0>

Journal de l'École polytechnique
Mathématiques
Tome 5, 2018, p. 377–405 DOI: 10.5802/jep.73

QUANTITATIVE STATISTICAL STABILITY, SPEED OF CONVERGENCE TO EQUILIBRIUM AND PARTIALLY HYPERBOLIC SKEW PRODUCTS

by Stefano Galatolo

Abstract. — We consider a general relation between fixed point stability of suitably perturbed transfer operators and convergence to equilibrium (a notion which is strictly related to decay of correlations). We apply this relation to deterministic perturbations of a class of (piecewise) partially hyperbolic skew products whose behavior on the preserved fibration is dominated by the expansion of the base map. In particular, we apply the results to power law mixing toral extensions. It turns out that in this case, the dependence of the physical measure on small deterministic perturbations, in a suitable anisotropic metric, is at least Hölder continuous, with an exponent which is explicitly estimated depending on the arithmetical properties of the system. We show explicit examples of toral extensions having actually Hölder stability and non differentiable dependence of the physical measure on perturbations.

Résumé (Stabilité statistique quantitative, vitesse de convergence vers l'équilibre et produits croisés partiellement hyperboliques)

Nous considérons une relation générale entre la stabilité des points fixes d'opérateurs de transfert convenablement perturbés et la convergence vers l'équilibre (une notion strictement reliée à la décroissance des corrélations). Nous appliquons cette relation aux perturbations déterministes d'une classe de produits croisés partiellement hyperboliques (par morceaux) dont le comportement sur la fibration préservée est dominé par l'expansion de l'application de la base. Nous appliquons ces résultats aux applications sur un tore fibrant sur le cercle, partiellement hyperboliques, linéaires par morceaux, avec décroissance lente des corrélations d'allure polynomiale. Il s'avère que, dans ce cas, la dépendance de la mesure physique en les petites perturbations déterministes, dans une métrique anisotrope, est au moins Hölder-continue, avec un exposant estimé explicitement en termes des propriétés arithmétiques du système. Nous donnons des exemples explicites d'applications sur un tore fibrant sur le cercle qui ont une stabilité Hölder et une dépendance non différentiable de la mesure physique en les perturbations.

Contents

2010 Mathematics Subject Classification. — 37D30, 37C30, 37A25, 37A10.
Keywords. — Statistical stability, convergence to equilibrium, decay of correlations, transfer operator, skew product, Diophantine type, Hölder response.
The work was partially supported by EU Marie-Curie IRSES Brazilian-European partnership in Dynamical Systems (FP7-PEOPLE-2012-IRSES 318999 BREUDS), and by The Leverhulme Trust through Network Grant IN-2014-021.

1. Introduction

The concept of statistical stability of a dynamical system deals with the stability of the statistical properties of its trajectories when the system is perturbed or changed in some way. Since the statistical properties of systems and their behavior are important in many fields of mathematics and in applied science, the study of this kind of stability has important applications to many fields. Many important statistical properties of dynamics are encoded in suitable probability measures which are invariant for the action of the dynamics. Because of this the mathematical approach to statistical stability is often related to the stability of such invariant measures under perturbations of the system. In this context it is important to get quantitative estimates, such as the differentiability of the statistical properties under small perturbations (called Linear Response) or other quantitative statements, such as the Lipschitz or Hölder dependence. These questions are well understood in the uniformly hyperbolic case, where the system's derivative has uniformly expanding or contracting directions. In this case quantitative estimates are available, proving the Lipschitz and even differentiable dependence of the relevant invariant measures under perturbations of the system (see e.g. [2], [7] or [26] and related references, for recent surveys where also some result beyond the uniformly hyperbolic case are discussed). For systems not having a uniformly hyperbolic behavior, and in presence of discontinuities, the situation is more complicated and much less is known. Qualitative results and some quantitative ones (providing precise information on the modulus of continuity) are known under different assumptions or in families of cases, and there is not yet a general understanding of the statistical stability in those cases (see e.g. [1, 3, 6, 10, 8, 9, 5, 15, 17, 18, 25, 30, 31]).

We approach this question from a general point of view, using a functional analytic perturbation lemma (see Theorem 5) which relates the convergence to equilibrium speed of the system to the stability of its invariant measures belonging to suitable spaces. In our case we consider a space of signed measures equipped with a suitable anisotropic norm adapted to the system, in which the relevant invariant measures are proved to exist. We show how, with some technical work, the approach can be applied to slowly mixing partially hyperbolic skew products. The functional analytic perturbation lemma we use is quite flexible and was applied in [18] to the study of the statistical stability of maps with indifferent fixed points. A similar functional analytic construction was also applied to piecewise hyperbolic skew product maps and Lorenz-like two dimensional maps in [19].

Main results. — The paper has the following structure and main contents.

(a) A general quantitative relation between speed of convergence to equilibrium of the system and its statistical stability (Section 2).

(b) The application of this relation to a general class of skew products allowing discontinuities and a sort of partial hyperbolicity, getting quantitative estimates for their statistical stability in function of their convergence to equilibrium. Here a main ingredient is the construction of suitable spaces of regular measures adapted to these systems. (Sections 3, 4, 5).

(c) The application of this construction to the stability under deterministic perturbations of a class of piecewise constant *toral extensions having slow convergence to equilibrium*, getting Hölder stability for these examples (Section 6).

(d) Finally, we show examples of mixing piecewise constant toral extensions where a perturbation of the map of size $\delta \geqslant 0$ results in a change of physical invariant measure of the size of order δ^{β}, where $\beta \leqslant 1$ depends on the Diophantine properties of the map (Section 6.3).

Let us explain in more details the content of the items listed above:

(a) This relation is a fixed point stability statement we apply to the system's transfer operators, giving nontrivial informations for systems having different kinds of speed of convergence to equilibrium, implying for instance, that *in a system with power law convergence to equilibrium speed, under quite general additional assumptions, the physical measure is Hölder stable* (see Theorem 5 and Remark 7).

(b) We apply the relation to show a general quantitative stability statement for perturbations of a class of skew products. We consider skew products in which the *base dynamics is expanding and dominates the behavior of the dynamics on the fibers.* For this purpose we introduce suitable spaces of signed measures adapted to such systems. We consider spaces of signed measures having absolutely continuous projection on the base space $[0,1]$ (corresponding to the strongly expanding direction) and equip them with suitable anisotropic norms: the weak norm $\|\cdot\|_{\text{“}1\text{”}}$ and the strong norm $\|\cdot\|_{p\text{-BV}}$, which are defined by disintegrating along the central foliation (preserved by the skew product) and considering the regularity of the disintegration. These spaces have properties that make them work quite like L^1 and p-Bounded Variation real functions spaces in the classical theory for the statistical properties of one dimensional dynamics. In Section 4 we prove a kind of Lasota-Yorke inequality in this framework. This will be used together with a kind of Helly selection principle proved in Section 3 to estimate the regularity of the invariant measures (like it is done in the classical construction for one dimensional, piecewise expanding maps). We summarize this informally in the following (see Proposition 22 for a precise and more general statement).

Theorem 1. — *Let us consider a one parameter family of skew product maps $F_\delta : [0,1] \times M \to [0,1] \times M$, where M is a compact manifold with boundary and δ is a positive real number ranging in a neighborhood of 0. Let us suppose $F_\delta = (T_\delta, G_\delta)$ is such that the following hold uniformly on δ:*

(A1) *$T_\delta : [0,1] \to [0,1]$ is piecewise expanding with C^2 onto branches.*

(A2) *The behavior of G_δ on the fibers is dominated by the expansion of T_δ.*

(A3) *G_δ satisfies a sort of BV regularity: there exists $A > 0$, such that*[1]

$$\sup_{r \leqslant A} \frac{1}{r} \int \sup_{y \in M, x_1, x_2 \in B(x,r)} |G(x_1, y) - G(x_2, y)| dx < \infty.$$

[1] See (Sk1)–(Sk3) at beginning of Section 3 for precise statements of these assumptions.

Then the maps F_δ have invariant probability measures f_δ having an absolutely continuous projection on the base space $[0,1]$ and uniformly bounded $\|\cdot\|_{p\text{-BV}}$ norms (they are uniformly regular in the strong space).

In Section 5 we consider a class of perturbations of our skew products such that the related transfer operators are near in some sense when applied to (regular) measures and state a first *general statement on the statistical stability of such skew products* (see Proposition 25).

(c) The statement is then applied to *slowly mixing piecewise constant toral extensions*: systems of the kind (X,F), where $X=[0,1]\times\mathcal{T}^d$, $\mathcal{T}^d$ is the d dimensional torus and $F:X\to X$ is defined by

$$F(\omega,t)=(T\omega,t+\tau(\omega)), \tag{1}$$

where $T:[0,1]\to[0,1]$ is expanding and $\tau:[0,1]\to\mathcal{T}^d$ is a piecewise constant function. The qualitative ergodic theory of this kind of systems was studied in several papers (see e.g. [12, 13]). Quantitative results appeared more recently ([16, 14, 27]), proving from different points of view that the speed of correlation decay is generically fast (exponential), but in some cases where τ is piecewise constant, this decay follows a power law whose exponent depend on the Diophantine properties of τ (see [23] or Section 6.1).

We apply our general result to deterministic perturbations of these maps, showing that the physical measure of those systems varies at least Hölder continuously in our anisotropic "L^1-like" distance. We state informally an example of such an application (see Proposition 32 for a more general statement and the required definitions).

Theorem 2. — *Consider a family of skew product maps $F_\delta:[0,1]\times\mathcal{T}^d\to[0,1]\times\mathcal{T}^d$ satisfying* (A1)–(A3)*, as in Theorem* 1*. Assume that F_0 is a piecewise constant toral extension as in* (1)*, with*

$$T_0(x)=2x \bmod (1) \tag{2}$$

and

$$G_0(x,t)=(Tx,t+\theta\varphi(x)), \tag{3}$$

where $\theta=(\theta_1,\dots,\theta_d)\in\mathcal{T}^d$ has linear Diophantine type[(2)] *$\gamma_\ell(\theta)$ and $\varphi=1_{[0,1/2]}$ is the characteristic function of $[0,1/2]$. Suppose F_δ is a small perturbation of F_0 in the following sense:*

(D1) *For each δ, we have $T_\delta=T_0\circ\sigma$ for some diffeomorphism σ near to the identity, satisfying $\|\sigma-\mathrm{Id}\|_\infty\leqslant\delta$ and $\|(1/\sigma')-1\|_\infty\leqslant\delta$.*

(D2) *For each δ and $x\in[0,1]$, $y\in\mathcal{T}^d$, it holds $|G_0(x,y)-G_\delta(x,y)|\leqslant\delta$.*

(2) See Definition 26 for a recall about this Diophantine type for vectors of real numbers.

Let f_0 be the Lebesgue measure, which is invariant for F_0 and let f_δ be an invariant probability measure for F_δ with finite strong norm as found in Theorem 1. Then for each $\gamma > \gamma_\ell(\theta)$ there exists $K \geqslant 0$ such that

$$\|f_\delta - f_0\|_{``1"} \leqslant K\delta^{1/(8\gamma+1)}.$$

We remark that the perturbations allowed are quite general. In particular they allow discontinuities, and the invariant measure to become singular with respect to the Lebesgue measure after perturbation. We also remark that for a class of smooth toral extensions with fast decay of correlations, a differentiable dependence statement was proved in [17].

(d) We finally show examples of piecewise constant, mixing toral extensions where the physical measure of the system actually varies in a Hölder way (and hence not in a differentiable way) with an exponent depending on the arithmetical properties of the system. We state informally the main result about this, see Propositions 33 and 34 for precise statements.

Theorem 3. — *Consider a piecewise constant toral extension map*

$$F_0 \colon [0,1] \times \mathcal{T}^1 \longrightarrow [0,1] \times \mathcal{T}^1$$

as in (2), (3), where θ is a well approximable Diophantine irrational with $\gamma_\ell(\theta) > 2$. For every $\gamma < \gamma_\ell(\theta)$ there exist a sequence of reals $\delta_j \geqslant 0$, $\delta_j \to 0$ and a sequence of maps $\widehat{F}_{\delta_j}(x,y) = (\widehat{T}_0(x), \widehat{G}_{\delta_j}(x,y))$ satisfying (A1)–(A3), (D1), (D2) such that

$$\|\mu_0 - \mu_j\|_{``1"} \geqslant \frac{1}{9}{\delta_j}^{1/(\gamma-1)}$$

holds for every j and every μ_j, invariant Borel probability measure of $\widehat{F}_{\delta_j}$ with absolutely continuous projection on $[0,1]$.

This shows that in some sense, the general statistical stability result is sharp. We remark that recently, in [31], examples of C^r families of *mostly contracting* diffeomorphisms with strictly Hölder behavior have been given (see also [15] for previous results on Hölder stability of these kinds of partially hyperbolic maps).

2. Quantitative fixed point stability and convergence to equilibrium.

Let us consider a dynamical system (X,T), where X is a metric space and the space $\mathrm{SM}(X)$ of signed Borel measures on X. The dynamics T naturally induces a function $L : \mathrm{SM}(X) \to \mathrm{SM}(X)$ which is linear and is called transfer operator. If $\nu \in \mathrm{SM}(X)$ then $L[\nu] \in \mathrm{SM}(X)$ is defined by

$$L[\nu](B) = \nu(T^{-1}(B))$$

for every measurable set B. If X is a manifold, the measure is absolutely continuous ($d\nu = f\,dm$, where m represents the Lebesgue measure) and T is nonsingular, the operator induces another operator $\widetilde{L} : L^1(m) \to L^1(m)$ acting on measure densities ($\widetilde{L}f = d(L(f\ m))/dm$). By a small abuse of notation we will still indicate by L this operator.

An invariant measure is a fixed point for the transfer operator. Let us now see a quantitative stability statement for these fixed points under suitable perturbations of the operator. Let us consider a certain system having a transfer operator L_0 for which we know the speed of convergence to equilibrium (see (4) below). Consider a "nearby" system L_1 having suitable properties: suppose there are two normed vector spaces of measures with sign $B_s \subseteq B_w \subseteq \mathrm{SM}(X)$ (the strong and weak space) with norms $\|\cdot\|_w \leqslant \|\cdot\|_s$ and suppose the operators L_0 and L_1 preserve the spaces: $L_i(B_s) \subseteq B_s$ and $L_i(B_w) \subseteq B_w$ with $i \in \{0, 1\}$. Let us consider

$$V_s := \{f \in B_s \mid f(X) = 0\}$$

the space of zero average measures in B_s. The speed of convergence to equilibrium of a system will be measured by the speed of contraction to 0 of this space by the iterations of the transfer operator.

Definition 4. — Let $\phi(n)$ be a real sequence converging to zero. We say that the system has *convergence to equilibrium* with respect to norms $\|\cdot\|_w$, $\|\cdot\|_s$ and speed ϕ if

$$\forall\, g \in V_s, \quad \|L_0^n(g)\|_w \leqslant \phi(n)\|g\|_s. \tag{4}$$

Suppose f_0, $f_1 \in B_s$ are fixed probability measures of L_0 and L_1. The following statement relates the distance between f_0 and f_1 with the distance between L_0 and L_1 and the speed of convergence to equilibrium of L_0. The proof is elementary, we include it for completeness. Similar quantitative stability statements are used in [21], [19] and [18] to support rigorous computation of invariant measures, get quantitative estimates for the statistical stability of Lorenz-like maps and intermittent systems.

Theorem 5. — *Suppose we have estimates on the following aspects of the operators L_0 and L_1:*

(1) *(Speed of convergence to equilibrium) There exists $\phi \in C^0(\mathbb{R})$, $\phi(t)$ decreasing to 0 as $t \to \infty$ such that L_0 has convergence to equilibrium with respect to norms $\|\cdot\|_w$, $\|\cdot\|_s$ and speed ϕ.*

(2) *(Control on the norms of the invariant measures) There exists $\widetilde{M} \geqslant 0$ such that*

$$\max(\|f_1\|_s, \|f_0\|_s) \leqslant \widetilde{M}.$$

(3) *(Iterates of the transfer operator are bounded for the weak norm) There exists $\widetilde{C} \geqslant 0$ such that for each n,*

$$\|L_0^n\|_{B_w \to B_w} \leqslant \widetilde{C}.$$

(4) *(Control on the size of perturbation in the strong-weak norm) We have*

$$\sup_{\|f\|_s \leqslant 1} \|(L_1 - L_0)f\|_w =: \varepsilon < \infty.$$

Then we have the following explicit estimate

$$\|f_1 - f_0\|_w \leqslant 2\widetilde{M}\widetilde{C}\varepsilon(\psi^{-1}(\varepsilon\widetilde{C}/2) + 1),$$

where ψ is the decreasing function defined as $\psi(x) = \phi(x)/x$.

Proof. — The proof is a direct computation from the assumptions. Since f_0, f_1 are fixed probability measures of L_0 and L_1, for each $N\in\mathbb{N}$ we have

$$\begin{aligned}\|f_1-f_0\|_w &\leqslant \|L_1^N f_1 - L_0^N f_0\|_w\\ &\leqslant \|L_1^N f_1 - L_0^N f_1\|_w + \|L_0^N f_1 - L_0^N f_0\|_w\\ &\leqslant \|L_0^N (f_1-f_0)\|_w + \|L_1^N f_1 - L_0^N f_1\|_w.\end{aligned}$$

Since $f_1-f_0\in V_s$, $\|f_1-f_0\|_s\leqslant 2\widetilde{M}$, we have

$$\|f_1-f_0\|_w\leqslant 2\widetilde{M}\phi(N)+\|L_1^N f_1 - L_0^N f_1\|_w,$$

but

$$(L_0^N-L_1^N)=\sum_{k=1}^{N} L_0^{N-k}(L_0-L_1)L_1^{k-1},$$

hence

$$-(L_1^N-L_0^N)f_1=\sum_{k=1}^{N} L_0^{N-k}(L_0-L_1)L_1^{k-1}f_1=\sum_{k=1}^{N} L_0^{N-k}(L_0-L_1)f_1,$$

and therefore

$$\|f_1-f_0\|_w\leqslant 2\widetilde{M}\phi(N)+\varepsilon\widetilde{M}N\widetilde{C}.$$

Now consider the function ψ defined as $\psi(x)=\phi(x)/x$, choose N such that $\psi^{-1}(\varepsilon\widetilde{C}/2)\leqslant N\leqslant \psi^{-1}(\varepsilon\widetilde{C}/2)+1$, in this way $\phi(N)/N\leqslant\varepsilon\widetilde{C}/2\leqslant\phi(N-1)/(N-1)$ and

$$\|f_1-f_0\|_w\leqslant 2\widetilde{M}\widetilde{C}\varepsilon(\psi^{-1}(\varepsilon\widetilde{C}/2)+1). \qquad \square$$

Theorem 5 implies that a system having convergence to equilibrium is statistically stable (in the weak norm).

Corollary 6. — *Let $D\geqslant 0$, let L_δ, $\delta\in[0,D)$, be a family of transfer operators under the assumptions of Theorem 5, including $\lim_{n\to\infty}\phi(n)=0$. Let $f_\delta\in B_s$ be a fixed probability measure of L_δ. Suppose there is $C\geqslant 0$ such that for every δ*

$$\sup_{\|f\|_s\leqslant 1}\|(L_\delta-L_0)f\|_w\leqslant C\delta.$$

Then f_0 is the unique fixed probability measure in B_s and it holds

$$\lim_{\delta\to 0}\|f_\delta-f_0\|_w=0.$$

Proof. — The uniqueness of f_0 is trivial from the definition of convergence to equilibrium. For the stability, suppose there was a sequence $\delta_n\to 0$ and $\ell\geqslant 0$ such that $\|f_{\delta_n} - f_0\|_w \geqslant \ell$ for all n. Applying Theorem 5 we get successively

$$2\widetilde{M}\widetilde{C}C\delta_n(\psi^{-1}(\delta_n C\widetilde{C}/2)+1) \geqslant \ell,$$
$$\psi^{-1}(\delta_n C\widetilde{C}/2) \geqslant \frac{\ell}{\delta_n 2\widetilde{M}\widetilde{C}C} - 1,$$
$$\frac{\delta_n C\widetilde{C}}{2} \leqslant \frac{\phi((\ell/\delta_n 2\widetilde{M}\widetilde{C}C)-1)}{(\ell/\delta_n 2\widetilde{M}\widetilde{C}C)-1},$$
$$\Big(\frac{\ell}{\delta_n 2\widetilde{M}\widetilde{C}C} - 1\Big)\frac{\delta_n C\widetilde{C}}{2} \leqslant \phi((\ell/\delta_n 2\widetilde{M}\widetilde{C}C)-1),$$
$$\frac{\ell}{2\widetilde{M}} - \delta_n C\widetilde{C} \leqslant 2\phi((\ell/\delta_n 2\widetilde{M}\widetilde{C}C)-1),$$

which is impossible to hold as $\delta_n \to 0$. □

Remark 7. — In Theorem 5, if $\phi(x) = Cx^{-\alpha}$ then $\psi(x) = Cx^{-\alpha-1}$, $\varepsilon(\psi^{-1}(\varepsilon)+1) \sim \varepsilon^{1-1/(\alpha+1)}$ and we have the estimate for the modulus of continuity

$$\|f_1 - f_0\|_w \leqslant K_1 \varepsilon^{1-1/(\alpha+1)},$$

where the constant K_1 depends on $\widetilde{M}, \widetilde{C}, C$ and not on the distance between the operators measured by ε.

3. Spaces we consider

Our approach is based on the study of the transfer operator restricted to a suitable space of measures with sign. We introduce a space of regular measures where we can find the invariant measure of our systems, and the ones of suitable perturbations of it. We hence consider some measure spaces adapted to skew products. The approach is taken from [19] (see also [4]) where it was used for Lorenz-like two dimensional maps. Let us consider a map $F : X \to X$, where $X = [0,1] \times M$, and M is a compact manifold with boundary, such that

$$F(x,y) = (T(x), G(x,y)).$$

Suppose F satisfies the following conditions:

(Sk1) Suppose T is $\frac{1}{\lambda}$-expanding[(3)] and it has $C^{1+\xi}$ branches[(4)] which are onto. The branches will be denoted by T_i, $i \in \{1,\dots,q\}$.

(Sk2) Consider the F-invariant foliation $\mathcal{F}^s := \{\{x\} \times M\}_{x\in[0,1]}$. We suppose that the behavior on $\mathcal{F}^s$ is dominated by λ: there exists $\alpha \in \mathbb{R}$ with $\lambda^\xi \alpha < 1$, such that for all $x \in [0,1]$ holds

$$|G(x,y_1) - G(x,y_2)| \leqslant \alpha|y_1 - y_2| \quad \forall\, y_1, y_2 \in M.$$

(3) We suppose that $\inf_{x\in[0,1]} T'(x) \geqslant 1/\lambda$ for some $\lambda < 1$.

(4) More precisely we suppose that there are $\xi, C_h \geqslant 0$ such that

$$\frac{1}{|T_i' \circ T_i^{-1}(\gamma_2))|} - \frac{1}{|T_i' \circ T_i^{-1}(\gamma_1))|} \leqslant C_h d(\gamma_1,\gamma_2)^\xi, \quad \forall\, \gamma_1, \gamma_2 \in [0,1].$$

(Sk3) For each $p \leqslant \xi$ there exists $A > 0$, such that

$$\widehat{H} := \sup_{r \leqslant A} \frac{1}{r^p} \int \sup_{\substack{y \in M \\ x_1, x_2 \in B(x,r)}} |G(x_1, y) - G(x_2, y)| dx < \infty.$$

Remark 8. — We remark that (Sk3) allows discontinuities in G, provided a kind of bounded variation regularity is respected. (Sk2) allows a dominated expansion or contraction in the fibers direction. Furthermore, by (Sk1) the transfer operator of the map T satisfies a Lasota-Yorke inequality of the kind

$$\|L_T^n(\mu)\|_{\mathrm{BV}} \leqslant A_T \lambda^n \|\mu\|_{\mathrm{BV}} + B_T \|\mu\|_1, \tag{5}$$

where $\|\mu\|_{\mathrm{BV}}$ is the generalized bounded variation norm (see [24]) for some constant A_T and B_T depending on the map.

Definition 9. — We say that a family of maps $F_\delta = (T_\delta(x), G_\delta(x, y))$ satisfies (Sk1)–(Sk3) *uniformly*, if each T_δ is piecewise expanding, with onto $C^{1+\xi}$ branches, admitting a uniform expansion rate $1/\lambda$, a uniform α, a uniform Hölder constant C_h, a uniform second coefficient of the Lasota-Yorke inequality B_{T_δ} and furthermore the family G_δ satisfies (Sk3) with a uniform bound on the constant $\widehat{H}$.

We construct now some function spaces which are suitable for the systems we consider. The idea is to consider spaces of measures with sign, with suitable norms constructed by disintegrating measures along the central foliation. In this way a measure on X will be seen as a collection (a path) of measures on the leaves. In the central direction (on the leaves) we will consider a norm which is the dual of the Lipschitz norm. In the expanding direction we will consider the L^1 norm and a suitable variation norm. These ideas will be implemented in the next paragraphs.

Let (X, d) be a compact metric space, $g : X \to \mathbb{R}$ be a Lipschitz function and let $\mathrm{Lip}(g)$ be its best Lipschitz constant, i.e.,

$$\mathrm{Lip}(g) = \sup_{x,y \in X} \left\{ \frac{|g(x) - g(y)|}{d(x, y)} \right\}.$$

Definition 10. — Given two signed Borel measures μ and ν on X, we define a *Wasserstein-Kantorovich-like* distance between μ and ν by

$$W_1(\mu, \nu) = \sup_{\substack{\mathrm{Lip}(g) \leqslant 1 \\ \|g\|_\infty \leqslant 1}} \left| \int g d\mu - \int g d\nu \right|.$$

Let us denote

$$\|\mu\|_{W_1} := W_1(0, \mu).$$

As a matter of fact, $\|\cdot\|_{W_1}$ defines a norm on the vector space of signed measures defined on a compact metric space.

Let $\mathcal{SB}(\Sigma)$ be the space of Borel signed measures on Σ. Given $\mu \in \mathcal{SB}(\Sigma)$ denote by μ^+ and μ^- the positive and the negative parts of it ($\mu = \mu^+ - \mu^-$).

Denote by $\mathcal{AB}$ the set of signed measures $\mu \in \mathcal{SB}(\Sigma)$ such that its associated marginal signed measures, $\mu_x^{\pm} = \pi_x^* \mu^{\pm}$ are absolutely continuous with respect to the Lebesgue measure m, on $[0,1]$ i.e.,

$$\mathcal{AB} = \{\mu \in \mathcal{SB}(\Sigma) \mid \pi_x^*\mu^+ \ll m \text{ and } \pi_x^*\mu^- \ll m\},$$

where $\pi_x : X \to [0,1]$ is the projection defined by $\pi_x(x,y) = x$ and π_x^* is the associated pushforward map.

Let us consider a finite positive measure $\mu \in \mathcal{AB}$ on the space X foliated by the preserved leaves $\mathcal{F}^c = \{\gamma_\ell\}_{\ell \in [0,1]}$ such that $\gamma_\ell = {\pi_x}^{-1}(\ell)$. We will also call $\mathcal{F}^c$ as the *central foliation.* Let us denote $\mu_x = \pi_x^* \mu$ and let ϕ_μ be its density ($\mu_x = \phi_\mu m$). The Rokhlin disintegration theorem describes a disintegration of μ by a family $\{\mu_\gamma\}_\gamma$ of probability measures on the central leaves[(5)] in a way that the following holds.

Remark 11. — The disintegration of a measure μ is the μ_x-unique measurable family $(\{\mu_\gamma\}_\gamma, \phi_\mu)$ such that, for every measurable set $E \subset X$ it holds

$$\mu(E) = \int_{[0,1]} \mu_\gamma(E \cap \gamma) d\mu_x(\gamma). \tag{6}$$

Definition 12. — Let $\pi_{\gamma,y} : \gamma \to M$ be the restriction $\pi_{y|\gamma}$, where $\pi_y : X \to M$ is the projection defined by $\pi_y(x,y) = y$ and $\gamma \in \mathcal{F}^c$. Given a positive measure $\mu \in \mathcal{AB}$ and its disintegration along the stable leaves $\mathcal{F}^c$, $(\{\mu_\gamma\}_\gamma, \phi_\mu)$, we define the *restriction of μ on γ* as the positive measure $\mu_{|\gamma}$ on M (not on the leaf γ) defined as

$$\mu_{|\gamma} = \pi_{\gamma,y}^*(\phi_\mu(\gamma)\mu_\gamma).$$

Definition 13. — For a given signed measure $\mu \in \mathcal{AB}$ and its decomposition $\mu = \mu^+ - \mu^-$, define the *restriction of μ on γ* by

$$\mu_{|\gamma} = \mu_{|\gamma}^+ - \mu_{|\gamma}^-.$$

Definition 14. — Let $\mathcal{L}^1 \subseteq \mathcal{AB}$ be defined as

$$\mathcal{L}^1 = \left\{\mu \in \mathcal{AB} : \int_{[0,1]} \|\mu_{|\gamma}\|_{W_1} \, dm(\gamma) < \infty\right\}$$

and the norm $\|\cdot\|_{\text{“}1\text{”}} : \mathcal{L}^1 \to \mathbb{R}$ on it as

$$\|\mu\|_{\text{“}1\text{”}} = \int_{[0,1]} \|\mu_{|\gamma}\|_{W_1} \, dm(\gamma).$$

The notation we use for this norm is similar to the usual L^1 norm. Indeed this is formally the case if we associate to μ, by disintegration, a path $G_\mu : [0,1] \to (\mathcal{SB}(M), \|\cdot\|_{W_1})$ defined by $G_\mu(\gamma) = \mu_{|\gamma}$. In this case, this will be the L^1 norm of the path.

(5) In the following to simplify notations, when no confusion is possible we will indicate the generic leaf or its coordinate with γ.

3.1. The transfer operator associated to F and basic properties of $\mathcal{L}^1$. — Let us now consider the transfer operator L_F associated with F. Being a push forward map, the same function can be also denoted by F^* we will use this notation sometime. There is a nice characterization of the transfer operator in our case, which makes it work quite like a one dimensional operator. For the proof see [19].

Proposition 15 (Perron-Frobenius-like formula). — *Consider a skew product map F satisfying* (Sk1) *and* (Sk2). *For a given leaf $\gamma\in\mathcal{F}^s$, define the map $F_\gamma : M\to M$ by*

$$F_\gamma = \pi_y \circ F_{|\gamma} \circ \pi_{\gamma,y}^{-1}.$$

For all $\mu \in \mathcal{L}^1$ and for almost all $\gamma \in [0,1]$ it holds

$$(L_F\mu)_{|\gamma} = \sum_{i=1}^{q} \frac{\mathrm{F}^*_{T_i^{-1}(\gamma)}\, \mu_{|T_i^{-1}(\gamma)}}{|T_i' \circ T_i^{-1}(\gamma))|}.$$

We recall some results showing that the transfer operator associated to a Lipschitz function is also Lipschitz with the same constant, for the "1" distance, and moreover, that the transfer operator of a map satisfying (Sk1)–(Sk3) is also Lipschitz with the same constant for the $\|\cdot\|_{“1”}$ norm. In particular, if $\alpha \leqslant 1$ the transfer operator is a weak contraction, like it happen for the L^1 norm on the one dimensional case (for the proof and more details, see [19]).

Lemma 16. — *If $G : Y \to Y$, where Y is a metric space is α-Lipschitz, then for every Borel measure with sign μ it holds*

$$\|L_G\mu\|_{W_1} \leqslant \alpha\|\mu\|_{W_1}.$$

If $\mu \in \mathcal{L}^1$ and $F : [0,1] \times M \to [0,1] \times M$ satisfies (Sk1)–(Sk3) *then*

$$\|L_F\mu\|_{“1”} \leqslant \alpha\|\mu\|_{“1”}.$$

3.2. The strong norm. — We consider a norm which is stronger than the $\mathcal{L}^1$ norm. The idea is to consider a disintegrated measure as a path of measures on the preserved foliation and define a kind of bounded variation regularity for this path, in a way similar to what was done in [24] for real functions.

For this strong space we will prove a regularization inequality, similar to the Lasota-Yorke ones. We will use this inequality to prove the regularity of the invariant measure of the family of skew products we consider.

Let us consider $\mu \in \mathcal{L}^1$. Let us define

$$\operatorname{osc}(\mu, x, r) = \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)} (W_1(\mu_{|\gamma_1}, \mu_{|\gamma_2})) \quad \text{and} \quad \operatorname{var}_p(\mu, r) = \int_{[0,1]} r^{-p} \operatorname{osc}(\mu, x, r)\, dx.$$

Now let us choose $A > 0$ and consider $\operatorname{var}_p(\mu) := \sup_{r\leqslant A} \operatorname{var}_p(\mu, r)$. Finally we define the p-BV norm as:

$$\|\mu\|_{p\text{-BV}} = \|\mu\|_{“1”} + \operatorname{var}_p(\mu).$$

Let us consider p such that $1 \geqslant p \geqslant 0$ and the following space of measures

$$p\text{-BV} = \left\{\mu \in \mathcal{L}^1 \mid \|\mu\|_{p\text{-BV}} < \infty\right\}.$$

This will play the role of the strong space in the present case.

Remark 17. — If $\mu \in p$-BV, then it follows that

$$\operatorname*{esssup}_{\gamma} \|\mu_{|\gamma}\|_{W_1} \leqslant A^{p-1}\|\mu\|_{p\text{-BV}}.$$

See [4, Lem. 2] for a proof in the case of real functions which also works in our case.

We now prove a sort of Helly selection principle for sequences of *positive* measures with bounded variation. This principle will be used, together with a regularization inequality, proved in next section, to get information on the variation of invariant measures. First we need a preliminary lemma:

Lemma 18. — *If μ_n is a sequence of positive measures such that for each n, we have $\|\mu_n\|_{“1”} \leqslant C$, $\operatorname{var}_p(\mu_n) \leqslant M$, and $\mu_{n|\gamma} \to \mu_{|\gamma}$ for a.e. γ, in the Wasserstein distance, then*

$$\|\mu\|_{“1”} \leqslant C, \operatorname{var}_p(\mu) \leqslant M.$$

Proof. — Let us consider the $\|\cdot\|_{“1”}$ norm. Since by Remark 17 it holds $\|\mu_{n|\gamma}\|_W \leqslant A^{p-1}(C+M)$ for all γ, by the dominated convergence theorem we have $\|\mu\|_{“1”} \leqslant C$. Let us now consider the oscillation. We have that $\liminf_{n\to\infty} \operatorname{osc}(\mu_n, x, r) \geqslant \operatorname{osc}(\mu, x, r)$ for all x, r. Indeed, consider a small $\varepsilon \geqslant 0$. Because of the definition of osc(), we have that for all $\ell \geqslant \operatorname{osc}(\mu, x, r) - \varepsilon$, there exist positive measure sets A_1 and A_2 such that $W_1(\mu_{|\gamma_1}, \mu_{|\gamma_2}) \geqslant \ell$ for all $\gamma_1 \in A_1, \gamma_2 \in A_2$. Since $\mu_{n|\gamma} \to \mu_{|\gamma}$ for a.e. γ, if n is big enough there must exist sets A_1^n and A_2^n of positive measure such that $W_1(\mu_{n|\gamma_1}, \mu_{n|\gamma_2}) \geqslant \ell$ for all $\gamma_1 \in A_1^n, \gamma_2 \in A_2^n$. As a consequence we obtain $\operatorname{osc}(\mu_n, x, r) \geqslant \ell$, and then for all x, r, $\liminf_{n\to\infty} \operatorname{osc}(\mu_n, x, r) \geqslant \operatorname{osc}(\mu, x, r)$. Fatou's Lemma implies $\liminf_{n\to\infty} \operatorname{var}_p(\mu_n, r) = \liminf_{n\to\infty} \int_I r^{-p} \operatorname{osc}(\mu_n, x, r)\, dx \geqslant \operatorname{var}_p(\mu, r)$, from which the statement follows directly. □

Theorem 19 (Helly-like selection theorem). — *Let μ_n be a sequence of probability measures on X such that $\|\mu_n\|_{p\text{-BV}} \leqslant M$ for some $M \geqslant 0$. Then there exist μ with $\|\mu\|_{p\text{-BV}} \leqslant M$ and subsequence μ_{n_k} such that*

$$\|\mu_{n_k} - \mu\|_{“1”} \longrightarrow 0.$$

Proof. — Let us discretize in the vertical direction. Let us consider a continuous projection $\pi_{y,\delta} : PM(M) \to U_\delta$ of the space of probability measures on M to a finite dimensional space. Suppose $\pi_{y,\delta}$ is such that $\|\pi_{y,\delta}(\nu) - \nu\|_{W_1} \leqslant C\delta$, for all $\nu \in PM(M)$ (such a projection can be constructed discretizing the space by a partition of unity made of Lipschitz functions with support on sets whose diameter is smaller than δ, see [20] for example). Let us consider the natural extension $\pi_\delta : \mathcal{L}^1(X) \to \mathcal{L}^1(X)$ of this projection to the whole $\mathcal{L}^1(X)$ space, defined by $\pi_\delta(\mu)_{|\gamma} = \pi_{y,\delta}(\mu_{|\gamma})$.

Let us consider the sequence $\pi_\delta(\mu_n)$. We have $\|\pi_\delta(\mu_n)\|_{p\text{-BV}} \leqslant K_\delta M$, where K_δ is the modulus of continuity of π_δ. Indeed

$$\operatorname*{esssup}_{\gamma_2,\gamma_1 \in B(x,r)} (W_1(\pi_\delta(\mu)_{|\gamma_1}, \pi_\delta(\mu)_{|\gamma_1})) \leqslant K_\delta \operatorname*{esssup}_{\gamma_2,\gamma_1 \in B(x,r)} (W_1(\mu_{|\gamma_1}, \mu_{|\gamma_1})).$$

Since after projecting we now are in a space of functions with values in a finite dimensional space, to the sequence $\pi_\delta(\mu_n)$ we can apply the classical Helly selection theorem and get that there exist a limit measure μ_δ and a sub sequence n_k such that $\pi_\delta(\mu_{n_k}) \to \mu_\delta$ in $\mathcal{L}^1$ and $\pi_\delta(\mu_{n_k})_{|\gamma} \to \mu_{\delta|\gamma}$ almost everywhere. Let us consider a sequence $\delta_i \to 0$ and select inductively at every step from the previous selected subsequence μ_ℓ such that $\pi_{\delta_{i-1}}(\mu_\ell) \to \mu_{\delta_{i-1}}$ a further subsequence μ_{ℓ_k}, such that $\pi_{\delta_i}(\mu_{\ell_k}) \to \mu_{\delta_i}$ in $\mathcal{L}^1$ and almost everywhere. Since for all γ and $m \leqslant i$ we have $\|\pi_{\delta_m}(\mu_{\ell_k|\gamma}) - \mu_{\ell_k|\gamma}\|_{W_1} \leqslant C\delta_m$, it holds that for different $\delta_m, \delta_j \geqslant \delta_i$,

$$\|\pi_{\delta_i}(\mu_{\ell_k|\gamma}) - \pi_{\delta_j}(\mu_{\ell_k|\gamma})\|_{W_1} \leqslant C(\delta_i + \delta_j)$$

and then

$$\forall\,\gamma, \quad \|\mu_{\delta_m|\gamma} - \mu_{\delta_j|\gamma}\|_{W_1} \leqslant C(\delta_m + \delta_j + \delta_i)$$

uniformly in γ. Since μ_{δ_n} are positive measures, this shows that there exists a μ such that $\mu_{\delta_i} \to \mu$ in $\mathcal{L}^1$ and $\mu_{\delta_i|\gamma} \to \mu_{|\gamma}$ almost everywhere. This shows that a further subsequence μ_{n_j} can be selected in a way that

$$\pi_{\delta_i}(\mu_{n_j}) \underset{j\to\infty}{\longrightarrow} \mu_{\delta_i} \text{ for all } i, \text{ and } \mu_{n_j} \underset{j\to\infty}{\longrightarrow} \mu \text{ in } \mathcal{L}^1 \text{ and almost everywhere.}$$

Applying Lemma 18 we get $\|\mu\|_{p\text{-BV}} \leqslant M$. □

4. A regularization inequality

In this section we prove an inequality, showing that iterates of a bounded variation positive measure are of uniform bounded variation. This will play the role of a Lasota-Yorke inequality. A consequence will be a bound on the variation of invariant measures in $\mathcal{L}^1$. This will be used when applying Theorem 5 to provide the estimate needed at Item (2). The following regularization inequality can be proved.

Proposition 20. — *Let F be a skew product map satisfying assumptions* (Sk1)–(Sk3) *and let us suppose that μ is a positive measure. Let $p \leqslant \xi$ (the Hölder exponent as in* (Sk1)*). It holds*

$$\operatorname{var}_p(L_F\mu) \leqslant \lambda^p \alpha \operatorname{var}_p(\mu) + (\widehat{H}\|\mu_x\|_\infty + 3q\alpha C_h A^{\xi-p}\|\mu_x\|_\infty).$$

We recall that, here, μ_x is the marginal of the disintegration of μ (see Equation (6)) and $\|\mu_x\|_\infty$ is the supremum norm for its density.

Proof. — By the Perron Frobenius-like formula (Lemma 15)

$$(L_F\mu)_{|\gamma} = \sum_{i=1}^{q} \frac{\mathrm{F}^*_{T_i^{-1}(\gamma)}\, \mu_{|T_i^{-1}(\gamma)}}{|T_i' \circ T_i^{-1}(\gamma)|} \quad \text{for almost all } \gamma \in [0,1]$$

we have to estimate

$$\begin{aligned} I &:= \sup_{r\leqslant A} \frac{1}{r^p} \int \operatorname*{esssup}_{\gamma_2,\gamma_1 \in B(x,r)} \|(L_F\mu)_{|\gamma_1} - (L_F\mu)_{|\gamma_2}\|_{W_1}\, dm(x) \\ &= \sup_{r\leqslant A} \frac{1}{r^p} \int \operatorname*{esssup}_{\gamma_2,\gamma_1 \in B(x,r)} \Big\| \sum_{i=1}^{q} \Big(\frac{\mathrm{F}^*_{T_i^{-1}(\gamma_1)}\, \mu_{|T_i^{-1}(\gamma_1)}}{|T_i' \circ T_i^{-1}(\gamma_1)|} - \frac{\mathrm{F}^*_{T_i^{-1}(\gamma_2)}\, \mu_{|T_i^{-1}(\gamma_2)}}{|T_i' \circ T_i^{-1}(\gamma_2)|} \Big) \Big\|_{W_1} dm(x). \end{aligned}$$

To make the notation concise, let us set in the next equations

$$\mathrm{F}^*(a,b) := \mathrm{F}^*_{T_i^{-1}(a)}\,\mu_{|T_i^{-1}(b)},\ \ g_i(a) := T_i^{'} \circ T_i^{-1}(a).$$

By the triangular inequality we have

$$\begin{aligned} I \leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\Big\|\frac{\mathrm{F}^*(\gamma_1,\gamma_1)-\mathrm{F}^*(\gamma_2,\gamma_2)}{|g_i(\gamma_1)|}\Big\|_{W_1} dm \\ + \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\Big\|\mathrm{F}^*(\gamma_2,\gamma_2)(\frac{1}{|g_i(\gamma_1)|}-\frac{1}{|g_i(\gamma_2)|})\Big\|_{W_1} dm. \end{aligned}$$

Recalling that $(1/|g_i(\gamma_2))|)-(1/|g_i(\gamma_1))|)\leqslant C_h d(\gamma_1,\gamma_2)^{\xi}$, we deduce

$$\begin{aligned} I \leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int (\frac{1}{|g_i(x)|}+C_h r^{\xi}) \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*(\gamma_1,\gamma_1)-\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm \\ + \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_h r^{\xi} \operatorname*{esssup}_{\gamma_2}\|\,\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm, \end{aligned}$$

and

$$\begin{aligned} I \leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \frac{1}{|g_i(x)|} \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*(\gamma_1,\gamma_1)-\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm \\ + \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_h r^{\xi} \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*(\gamma_1,\gamma_1)-\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm \\ + \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_h r^{\xi} \operatorname*{esssup}_{\gamma_2}\|\,\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm. \end{aligned}$$

Hence

$$\begin{aligned} I \leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \frac{1}{|g_i(x)|} \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*(\gamma_1,\gamma_1)-\mathrm{F}^*(\gamma_1,\gamma_2)\|_{W_1}\, dm \\ + \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \frac{1}{|g_i(x)|} \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*(\gamma_1,\gamma_2)-\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm \\ + 3\sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_h r^{\xi} \operatorname*{esssup}_{\gamma_2}\|\,\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\, dm \\ = I_a + I_b + I_c. \end{aligned}$$

Now,

$$I_a \leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int \frac{1}{|g_i(x)|} \operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\|\,\mathrm{F}^*_{T_i^{-1}(\gamma_1)}(\mu_{|T_i^{-1}(\gamma_1)}-\mu_{|T_i^{-1}(\gamma_2)})\|_{W_1}\, dm.$$

We recall that, by Lemma 16, $\|\mathrm{F}_\gamma^*\mu\|_{W_1}\leqslant\alpha\|\mu\|_{W_1}$, so that

$$\begin{aligned}
I_a &\leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int\frac{1}{|g_i(x)|}\operatorname*{esssup}_{\gamma_2,\gamma_1\in B(x,r)}\alpha\|\mu_{|T_i^{-1}(\gamma_1)}-\mu_{|T_i^{-1}(\gamma_2)}\|_{W_1}\\
&\leqslant \sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int_{I_i}\operatorname*{esssup}_{y_1,y_2\in B(x,\lambda r)}\alpha\|\mu_{|y_1}-\mu_{|y_2}\|_{W_1}dx.\\
&=\lambda^p\sup_{h\leqslant\lambda A}\sum_{i=1}^{q}\frac{1}{h^p}\int_{I_i}\operatorname*{esssup}_{y_1,y_2\in B(x,h)}\alpha\|\mu_{|y_1}-\mu_{|y_2}\|_{W_1}dx
\end{aligned}$$

and

$$\operatorname{var}_p(L_F\mu)\leqslant\lambda^p\alpha\operatorname{var}_p(\mu)+(I_b+I_c).$$

By (Sk3), we have

$$I_b=\sup_{r\leqslant A}\sum_i\frac{1}{r^p}\int\frac{1}{|g_i(x)|}\operatorname*{esssup}_{y_1,y_2\in B(x,r)}\|\mathrm{F}^*(\gamma_1,\gamma_2)-\mathrm{F}^*(\gamma_2,\gamma_2)\|W_1\,dm\leqslant\widehat{H}\|\mu_x\|_\infty.$$

Now, let us remark that, since we are working with positive measures, we have

$$\operatorname*{esssup}_{\gamma_2}\|\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\leqslant\alpha\|\mu_x\|_\infty,$$

and therefore

$$\begin{aligned}
I_c&=3\sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_hr^\xi\operatorname*{esssup}_{\gamma_2}\|\mathrm{F}^*(\gamma_2,\gamma_2)\|_{W_1}\,dm\\
&\leqslant 3\sup_{r\leqslant A}\sum_{i=1}^{q}\frac{1}{r^p}\int C_hr^\xi\alpha\|\mu_x\|_\infty\,dm\leqslant 3qC_hA^{\xi-p}\alpha\|\mu_x\|_\infty.
\end{aligned}$$

Summarizing, we have obtained

$$\operatorname{var}_p(L_F\mu)\leqslant\lambda^p\alpha\operatorname{var}_p(\mu)+(\widehat{H}\|\mu_x\|_\infty+3q\alpha C_hA^{\xi-p}\|\mu_x\|_\infty).\tag{7}$$ □

Remark 21. — By Equation (5) it holds that for each n

$$\|L^n\mu_x\|_\infty\leqslant C_\mu:=A^{p-1}(A_T\lambda\|\mu_x\|_{\mathrm{BV}}+B_T\|\mu_x\|_1+1).$$

Iterating (7) we obtain

$$\begin{aligned}
\operatorname{var}_p(L_F^n\mu)&\leqslant(\lambda^p\alpha)\operatorname{var}_p(L^{n-1}\mu)+(\widehat{H}+3q\alpha C_hA^{\xi-p})C_\mu\\
&\leqslant\dots\\
&\leqslant(\lambda^p\alpha)^n\operatorname{var}_p(\mu)+\frac{\widehat{H}+3q\alpha C_hA^{\xi-p}}{1-\lambda^p\alpha}C_\mu.
\end{aligned}\tag{8}$$

By the Helly-like selection principle (Theorem 19) we then have:

Proposition 22. — *In a system as above there is at least an invariant positive measure in* p-BV. *For every such invariant measure* μ, *we have*

$$\operatorname{var}_p(\mu)\leqslant\frac{B_T(\widehat{H}+3q\alpha C_hA^{\xi-p})}{1-\lambda^p\alpha}\|\mu\|_{“1”}.$$

Proof. — We consider the sequence of positive measures $\mu_n = \frac{1}{n}\sum L_F^n m$. By Equation (8), this sequence has uniformly bounded variation. Applying Theorem 19, we deduce the existence of an invariant measure μ in p-BV. By the Lasota-Yorke inequality relative to the map T, we have that

$$A^{p-1}B_T\|\mu\|_{``1"} \geqslant A^{p-1}B_T\|\mu_x\|_1 \geqslant A^{p-1}\|\mu_x\|_{\mathrm{BV}} \geqslant \|\mu_x\|_\infty.$$

This gives that

$$\mathrm{var}_p(\mu) = \mathrm{var}_p(L_F\mu) \leqslant \lambda^p\alpha\,\mathrm{var}_p(\mu) + B_T\|\mu\|_{``1"}A^{p-1}(\widehat{H} + 3\alpha q C_h A^{\xi-p}),$$

from which we get the statement. □

5. Distance between the operators and a general statement for skew products

Here we consider a suitable class of perturbations of a map satisfying (Sk1)–(Sk3) such that the associated transfer operators are near in the strong-weak topology, providing one of the estimates needed to apply Theorem 5 (item (4)). In this section and in the following we set $p = 1$. We now define a topology on the space of piecewise expanding maps in order to have a notion of "allowed perturbations" for these maps.

Definition 23. — Let $n\in\mathbb{N}$, let T_1 and T_2 be to piecewise expanding maps. Denote by

$$\mathrm{Int}_n = \{A \in 2^{[0,1]} \mid A = I_1\cup,\dots,\cup I_n,\ \text{where } I_i \text{ are intervals}\}$$

the set of subsets of $[0,1]$ which is the union of at most n intervals. Set

$$\mathcal{C}(n,T_1,T_2) = \left\{\varepsilon \;\middle|\; \begin{array}{c} \exists\, A_1 \in \mathrm{Int}_n \text{ and } \exists\,\sigma : I \to I \text{ a diffeomorphism s.t.} \\ m(A_1) \geqslant 1-\varepsilon,\ T_{1|A_1} = T_2\circ\sigma_{|A_1} \\ \text{and } \forall\, x \in A_1,\ |\sigma(x)-x| \leqslant \varepsilon,\ |(1/\sigma'(x))-1| \leqslant \varepsilon \end{array}\right\}$$

and define a kind of distance from T_1 to T_2 as:

$$d_{S,n}(T_1,T_2) = \inf\{\varepsilon \mid \varepsilon \in \mathcal{C}(n,T_1,T_2)\}.$$

It holds that one dimensional piecewise expanding maps which are near in the sense of $d_{S,n}$ also have transfer operators which are near as operators from BV to L^1. If we denote by d_S the classical notion of Skorokhod distance (see e.g. [11]), it is obvious that $d_{S,n} \geqslant d_S$ for all n. By [11, Lem. 11.2.1] it follows that for all n there exists $C_{Sk} \geqslant 0$ such that, for each pair of piecewise expanding maps T_1, T_2,

$$\|L_{T_0} - L_{T_\delta}\|_{\mathrm{BV}\to L^1} \leqslant C_{Sk}\, d_{n,S}(T_1,T_2). \tag{9}$$

Let us see a statement of this kind for our skew products.

Proposition 24. — *Let $F_\delta = (T_\delta, G_\delta)$, $0 \leqslant \delta \leqslant D$ be a family of maps satisfying* (Sk1)–(Sk3) *uniformly with $\xi = 1$ and:*

(1) *There exists $n \in \mathbb{N}$ such that for each $\delta \leqslant D$, $d_{n,S}(T_0,T_\delta) \leqslant \delta$ (thus for each δ there is a set $A_1 \in \mathrm{Int}_n$ as in the definition of $\mathcal{C}(n,T_1,T_2)$).*

(2) *For each $\delta \leqslant D$ there exists a set $A_2 \in \mathrm{Int}_n$ such that $m(A_2) \geqslant 1-\delta$ and for all $x \in A_2$, $y \in M$, it holds $|G_0(x,y) - G_\delta(x,y)| \leqslant \delta$.*

Let us denote by F_δ^* *the transfer operators of* F_δ *and by* f_δ *a family of probability measures with uniformly bounded variation*
$$\mathrm{var}_1(f_\delta) \leqslant M_2.$$
Then, there exists a constant C_1 *such that for* δ *small enough*
$$\|(\mathrm{F}_0^* - \mathrm{F}_\delta^*)f_\delta\|_{``1"} \leqslant C_1\delta(M_2+1).$$

Proof. — Let us set $A = A_1 \cap A_2$. Note that $m(A^c) \leqslant 2\delta$. Let us estimate
$$\begin{aligned}(10)\qquad \|(\mathrm{F}_0^* - \mathrm{F}_\delta^*)f_\delta\|_{``1"} &= \int_I \|(F_0^* f_\delta - F_\delta^* f_\delta)_{|\gamma}\|_W \, dm(\gamma)\\ &= \int_I \|F_0^*(1_A f_\delta)|_\gamma - F_\delta^*(1_A f_\delta)_{|\gamma}\|_W \, dm(\gamma)\\ &\qquad + \int_I \|F_0^*(1_{A^c} f_\delta)|_\gamma - F_\delta^*(1_{A^c} f_\delta)_{|\gamma}\|_W \, dm(\gamma).\end{aligned}$$

By the assumptions, for a.e. γ, $\|f_\delta|_\gamma\|_W \leqslant (M_2+1)$ and $\|1_{A^c} f_\delta\|_1 \leqslant (M_2+1)\delta$. Since F^* is α-Lipschitz for the $\mathcal{L}^1$norm we have
$$\int_I \|F_0^*(1_{A^c} f_\delta)|_\gamma - F_\delta^*(1_{A^c} f_\delta)_{|\gamma}\|_W \, dm(\gamma) \leqslant 2\alpha(M_2+1)\delta.$$

Let us now estimate the first summand of (10). Let us set $\mu = 1_A f_\delta$ and let us estimate
$$\|(\mathrm{F}_0^* - \mathrm{F}_\delta^*)\mu\|_{``1"} = \int \|(F_0^*\mu - F_\delta^*\mu)_{|\gamma}\|_W \, dm(\gamma).$$
Let us denote by $T_{0,i}$, with $0 \leqslant i \leqslant q$, the branches of T_0 defined in the sets P_i, partition of I, and set $T_{\delta,i} = T_{\delta|P_i\cap A}$. These functions will play the role of the branches for T_δ. Since in A, $T_0 = T_\delta \circ \sigma_\delta$ (where σ_δ is the diffeomorphism in the definition of the Skorokhod distance), then $T_{\delta,i}$ are invertible. Then for μ_x-a.e. $\gamma \in I$, we have
$$(F_0^*\mu - F_\delta^*\mu)_{|\gamma} = \sum_{i=1}^q \frac{F^*_{0,T_{0,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{T_0(P_i\cap A)}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F^*_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{\delta,i}^{-1}(\gamma)}\chi_{T_\delta(P_i\cap A)}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|}.$$
Let us now consider $T_0(P_i \cap A)$ and $T_\delta(P_i \cap A)$, and remark that $T_0(P_i \cap A) = \sigma_\delta(T_\delta(P_i \cap A))$, where σ_δ is a diffeomorphism near to the identity. Let us denote $B_i = T_0(P_i \cap A) \cap T_\delta(P_i \cap A)$, $C_i = T_0(P_i \cap A) \triangle T_\delta(P_i \cap A)$. We have
$$\begin{aligned}(11)\quad \|(\mathrm{F}_0^* - \mathrm{F}_\delta^*)\mu\|_{``1"} &= \int_I \|(F_0^*\mu - F_\delta^*\mu)_{|\gamma}\|_W \, dm(\gamma)\\ &\leqslant \int_I \Big\| \sum_{i=1}^q \frac{F^*_{0,T_{0,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F^*_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{\delta,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \Big\|_W dm\\ &\qquad + \int_I \Big\| \sum_{i=1}^q \frac{F^*_{0,T_{0,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{C_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F^*_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{\delta,i}^{-1}(\gamma)}\chi_{C_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \Big\|_W dm.\end{aligned}$$
And since there exists K_1 such that $m(C_i) \leqslant K_1\delta$, we have
$$\int_I \Big\| \sum_{i=1}^q \frac{F^*_{0,T_{0,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{C_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F^*_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{\delta,i}^{-1}(\gamma)}\chi_{C_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \Big\|_W dm \leqslant qK_1(M_2+1)\delta.$$

Now we have to consider the first summand of (11). We have

$$\begin{aligned}
\int_I \Big\| \sum_{i=1}^{q} \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{\delta,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} \Big\|_W dm \\
\leqslant \int_I \Big\| \sum_{i=1}^{q} \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} \Big\|_W dm \\
+ \int_I \Big\| \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{\delta,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} \Big\|_W dm \\
= \int_I I(\gamma)\, dm(\gamma) + \int_I II(\gamma)\, dm(\gamma).
\end{aligned}$$

The two summands will be treated separately.

$$\begin{aligned}
I(\gamma) &= \Big\| \sum_{i=1}^{q} \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} \Big\|_W \\
&\leqslant \Big\| \sum_{i=1}^{q} \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} \Big\|_W \\
&\quad + \Big\| \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T^{-1}_{\delta,i}(\gamma))|} \Big\|_W \\
&= I_a(\gamma) + I_b(\gamma).
\end{aligned}$$

Since f_δ is a probability measure it holds posing $\beta = T^{-1}_{0,i}(\gamma)$

$$\begin{aligned}
\int I_a(\gamma)\, dm &= \int \Big\| \sum_{i=1}^{q} \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \sum_{i=1}^{q} \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} \Big\|_W dm(\gamma) \\
&\leqslant \int \sum_{i=1}^{q} \Big\| \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} \Big\|_W dm \\
&\leqslant \sum_{i=1}^{q} \int \Big\| \frac{F^*_{0,T^{-1}_{0,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} - \frac{F^*_{\delta,T^{-1}_{\delta,i}(\gamma)}\mu_{|T^{-1}_{0,i}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T^{-1}_{0,i}(\gamma))|} \Big\|_W dm \\
&\leqslant \sum_{i=1}^{q} \int_{T^{-1}_{0,i}(B_i)} \Big\| F^*_{0,\beta}\mu_{|\beta} - F^*_{\delta,T^{-1}_{\delta,i}(T_{0,i}(\beta))}\mu_{|\beta} \Big\|_W dm(\beta).
\end{aligned}$$

Remark that $T^{-1}_{0,i}(B_i) \subseteq P_i \cap A$ and $T^{-1}_{\delta,i}(T_{0,i}(T^{-1}_{0,i}(B_i))) \subseteq P_i \cap A$. Since we have $|T_{\delta,i}(\beta) - T_{0,i}(\beta)| \leqslant \delta$ and $T^{-1}_{0,i}$ is a contraction, then $|T^{-1}_{0,i} \circ T_{\delta,i}(\beta) - \beta| \leqslant \delta$. Therefore we obtain

$$\begin{aligned}
\big\| F^*_{0,\beta}\mu_{|\beta} - F^*_{\delta,T^{-1}_{\delta,i}(T_{0,i}(\beta))}\mu_{|\beta} \big\|_W \leqslant \big\| F^*_{0,\beta}\mu_{|\beta} - F^*_{\delta,\beta}\mu_{|\beta} \big\|_W \\
+ \big\| F^*_{\delta,\beta}\mu_{|\beta} - F^*_{\delta,T^{-1}_{\delta,i}(T_{0,i}(\beta))}\mu_{|\beta} \big\|_W.
\end{aligned}$$

By assumption (2),

$$\left\|F_{0,\beta}^{*}\mu_{|\beta}-F_{\delta,\beta}^{*}\mu_{|\beta}\right\|_{W}\leqslant \delta(M_2+1).$$

By assumption (Sk3),

$$\left\|F_{\delta,\beta}^{*}\mu_{|\beta}-F_{\delta,T_{\delta,i}^{-1}(T_{0,i}(\beta))}^{*}\mu_{|\beta}\right\|_{W}\leqslant \sup_{\substack{y\in M\\ x_1,x_2\in B(\beta,\delta)}}|G(x_1,y)-G(x_2,y)|(M_2+1).$$

Thus,

$$I_a(\gamma)\leqslant \delta^p(\widehat{H}+1)(M_2+1)+\delta(M_2+1).$$

To estimate $I_b(\gamma)$ we have:

$$\begin{aligned}I_b(\gamma)&=\Big\|\sum_{i=1}^{q}\frac{F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|}-\sum_{i=1}^{q}\frac{F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|}\Big\|_W\\ &\leqslant \sum_{i=1}^{q}\Big|\frac{\chi_{B_i}(\gamma)}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|}-\frac{\chi_{B_i}(\gamma)}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|}\Big|\left\|F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\right\|_W\end{aligned}$$

and

$$\int I_b\,dm\leqslant |(P_{T_0}-P_{T_\delta})(1)|\,\alpha(M_2+1)+qK_1\delta.$$

From Equation (9) we obtain

$$\int_{A_1} I_b(\gamma)\,dm(\gamma)\leqslant [C_{Sk}\alpha(M_2+1)+qK_1]\delta.$$

Now, let us estimate the integral of the second summand

$$II(\gamma)=\Big\|\sum_{i=1}^{q}\frac{F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{0,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|}-\sum_{i=1}^{q}\frac{F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\mu_{|T_{\delta,i}^{-1}(\gamma)}\chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|}\Big\|_W.$$

Setting $g_i(a):=T'_{\delta,i}\circ T_{\delta,i}^{-1}(a)$ to make the notation concise

$$\begin{aligned}\int_I II(\gamma)\,dm(\gamma)&\leqslant\sum_{i=1}^{q}\int_{B_i}\frac{1}{|g_i(\gamma)|}\left\|F^{*}_{\delta,T_{\delta,i}^{-1}(\gamma)}\big(\mu_{|T_{0,i}^{-1}(\gamma)}-\mu_{|T_{\delta,i}^{-1}(\gamma)}\big)\right\|_W dm(\gamma)\\ &\leqslant\sum_{i=1}^{q}\int_{B_i}\frac{\alpha}{|g_i(\gamma)|}\left\|\mu_{|T_{0,i}^{-1}(\gamma)}-\mu_{|T_{\delta,i}^{-1}(\gamma)}\right\|_W dm(\gamma).\end{aligned}$$

Let us consider the change of variable $\gamma=T_{\delta,i}(\beta)$. Then we obtain

$$\int_I II(\gamma)\,dm(\gamma)\leqslant\alpha\sum_{i=1}^{q}\int_{T_{\delta,i}^{-1}(B_i)}\left\|\mu_{|T_{0,i}^{-1}(T_{\delta,i}(\beta))}-\mu_{|\beta}\right\|_W dm(\beta).$$

Since $|T_{\delta,i}(\beta)-T_{0,i}(\beta)|\leqslant\delta$ and $T_{0,i}^{-1}$ is a contraction, we have $|T_{0,i}^{-1}\circ T_{\delta,i}(\beta)-\beta|\leqslant\delta$ and

$$\int_I II(\gamma)\,dm(\gamma)\leqslant\alpha\int\sup_{x,y\in B(\beta,\delta)}(\|\mu_{|x}-\mu_{|y}\|_W)\,dm(\beta)\leqslant\alpha\int\operatorname{osc}(\mu,\beta,\delta)\,d\beta,$$

and then

$$\int_I II(\gamma)\,dm(\gamma)\leqslant\alpha 2\delta(M_2+1).$$

Summing all, the statement is proved. □

The last statement, together with the results of the previous sections allows to prove the following quantitative statement for skew product maps.

Proposition 25. — *Consider a family of skew product maps $F_\delta = (T_\delta, G_\delta)$, $0 \leqslant \delta \leqslant D$ satisfying* (Sk1)–(Sk3) *uniformly, with $\xi = 1$, and let $f_\delta \in \mathcal{L}^1$ invariant probability measures of F_δ. Suppose:*

(1) *There exists $\phi \in C^0(\mathbb{R})$, $\phi(t)$ decreasing to 0 as $t \to \infty$, such that L_{F_0} has convergence to equilibrium with respect to norms $\|\cdot\|_{1\text{-BV}}$, $\|\cdot\|_{“1”}$ and speed ϕ.*

(2) *There exists $\widetilde{C} \geqslant 0$ such that for each n*

$$\|L^n_{F_0}\|_{\mathcal{L}^1 \to \mathcal{L}^1} \leqslant \widetilde{C}.$$

(3) *There exists $n \in \mathbb{N}$ such that for each $\delta \leqslant D$, $d_{n,S}(T_0, T_\delta) \leqslant \delta$.*

(4) *For each $\delta \leqslant D$ there exists a set $A_2 \in \mathrm{Int}_n$ such that $m(A_2) \geqslant 1 - \delta$ and for all $x \in A_2$, $y \in M$, we have $|G_0(x, y) - G_\delta(x, y)| \leqslant \delta$.*

Let $B = (B_T(\widehat{H} + 3q\alpha C_h)/(1 - \lambda^p\alpha)) + 1$. Consider the function ψ defined as $\psi(x) = \phi(x)/x$, then

$$\|f_\delta - f_0\|_{“1”} \leqslant 2\widetilde{C}B^2C_1\delta(\psi^{-1}(\widetilde{C}BC_1\delta/2) + 1),$$

where C_1 is the constant in the statement of Proposition 24.

Proof. — The proof is a direct application of the estimates given in the previous section into Theorem 5. The quantity $\widetilde{M}$ appearing at Item (2) of Theorem 5 is estimated by Proposition 22:

$$\widetilde{M} \leqslant \frac{B_T(\widehat{H} + 3q\alpha C_h)}{1 - \lambda^p\alpha}.$$

By Proposition 24 the distance between the operators appearing at Item (4) of Theorem 5 is bounded by $\varepsilon \leqslant C_1\delta(M_2 + 1)$, where M_2 bounds the strong norm of f_δ. □

We remark that the quantitative stability is proved here in the $\|\cdot\|_{“1”}$ topology. This topology is strong enough to control the behavior of observables which are discontinuous along the preserved central foliation, see [9] for other results on quantitative stability of the statistical properties of discontinuous observables and related applications.

In the following section we show a class of nontrivial partially hyperbolic skew products having power law convergence to equilibrium and will apply this statement to these examples.

6. Application to slowly mixing toral extensions

To give an example of application of Proposition 25 to a class of nontrivial systems, we consider a class of “partially hyperbolic” skew products with some discontinuities, having slow (power law) decay of correlations and convergence to equilibrium.

We will consider a class of skew products $F : X \to X$, with $X = [0, 1] \times \mathcal{T}^d$ (where $\mathcal{T}^d$ is the d dimensional torus), of the form $F = (T, G)$ (piecewise constant toral extensions), satisfying (Te1) and (Te2):

(Te1) There exists $\ell \in \mathbb{N}$ such that T is the piecewise expanding map on $[0,1]$ defined as

$$T(x) = \ell x \bmod (1).$$

(Te2) The map $G : [0,1] \times \mathcal{T}^d \to \mathcal{T}^d$ is defined by

$$G(x,t) = t + \theta\varphi(x),$$

where $\theta = (\theta_1, \dots, \theta_d) \in \mathcal{T}^d$ and $\varphi = 1_I$ is the characteristic function of a set $I \subset [0,1]$ which is a union of the sets P_i where the branches of T are defined. In this system the second coordinate is translated by θ if the first coordinate belongs to I.

We remark that on the system (X,F) the Lebesgue measure is invariant. We will suppose that θ is of finite Diophantine type. Let us recall the definition of Diophantine type for the linear approximation. The definition tests the possibility of approximating 0 by a linear combination of its components with integer coefficients.

The notation $\|\cdot\|$ will indicate the distance in $\mathbb{R}^d$ to the nearest vector in $\mathbb{Z}^d$, and $|k| = \sup_{1\leqslant i\leqslant d} |k_i|$ indicates the supremum norm.

Definition 26. — The Diophantine type of $\theta = (\theta_1, \dots, \theta_d)$ for the linear approximation is

$$\gamma_\ell(\theta) = \inf\{\gamma \mid \exists\, c_0 > 0,\ \|k\cdot\theta\| \geqslant c_0|k|^{-\gamma},\ \forall\, k \in \mathbb{Z}^d \smallsetminus \{0\}\}.$$

6.1. The decay of correlations. — In [29], it was observed that piecewise constant toral extensions cannot have exponential decay of correlations (in [28] by the way it is shown that for some piecewise constant $\mathrm{SU}_2(\mathbb{C})$ extensions there can be exponential decay of correlations). Quantitative estimates for the speed of decay of correlations by the arithmetical properties of the angles, have been given in [23].

In this section we recall those results and see that the systems defined above have at least polynomial decay of correlations, while for some choice of the angles the speed of decay is proved to be actually polynomial.

Definition 27 (Decay of correlations). — Let ϕ, $\psi : X \to \mathbb{R}$ be observables on X belonging to the Banach spaces B, B', let ν be an invariant measure for T. Let $\Phi : \mathbb{N} \to \mathbb{R}$ such that $\Phi(n) \underset{n\to\infty}{\to} 0$. A system (X,T,ν) is said to have decay of correlations with speed Φ with respect to observables in B and B' if

$$\left| \int \phi \circ T^n \psi d\nu - \int \phi d\nu \int \psi d\nu \right| \leqslant \|\phi\|_B \|\psi\|_{B'} \Phi(n), \tag{12}$$

where $\|\cdot\|_B, \|\cdot\|_{B'}$ are the norms in B and B'.

The decay of correlations depends on the class of observables considered. On the skew products satisfying conditions (Te1) and (Te2) as above, it is possible to establish an explicit upper bound for the rate of decay of correlations which depends on the linear type of the translation angle (see [23, Lem. 11]).

Proposition 28. — *In the piecewise constant toral extensions described above, for Lipschitz observables the rate of decay is*

$$\Phi(n) = O(n^{-1/2\gamma})$$

for any $\gamma > \gamma_\ell(\theta)$. For observables C^p, C^q, the rate of decay is

$$\Phi(n) = O(n^{-(1/2\gamma)\max(p,q,p+q-d)})$$

for any $\gamma > \gamma_\ell(\theta)$.

Remark 29. — We remark that the rate is actually polynomial in some cases. In [23, §5] (using a result of [22]) it is proved that if the Diophantine type is large, then the mixing rate of the systems we consider is actually slow, with a power law speed which depends on the Diophantine type. In a system satisfying (12), let the exponent of power law decay be defined by

$$p = \liminf_{n\to\infty} \frac{-\log \Phi(n)}{\log n}.$$

Let us consider the skew product of the doubling map and a circle rotation endowed with the Lebesgue (invariant) measure. For this example the exponent p satisfies

$$\frac{1}{2\gamma(\theta)} \leqslant p \leqslant \frac{6}{\max(2,\gamma(\theta))-2}.$$

6.2. Convergence to equilibrium. — We will use the decay of correlations of the toral extensions to get a convergence to equilibrium result with respect to the strong and weak norm of our anisotropic spaces. We have from Proposition 28 that for Lipschitz observables the rate of decay is $O(n^{-1/2\gamma})$ for any $\gamma > \gamma_\ell(\alpha)$ and for any Lipschitz observables with $\int f = 0$:

$$\Big|\int g\circ F^n \; f\, dm\Big| \leqslant C\|f\|_{\mathrm{lip}}\|g\|_{\mathrm{lip}} n^{^{-1/2\gamma}}.$$

From this we will deduce that

$$\|L^n\mu\|_{“1”} \leqslant C_4 n^{^{-1/8\gamma}}\|\mu\|_{1\text{-BV}}.$$

For this purpose our strategy is to approximate a 1-BV measure μ which is meant to be iterated with a Lipschitz density and use the decay of correlations with Lipschitz observables to estimate its convergence to equilibrium. We remark that a statement of this kind extends greatly the kinds of measures which are meant to be iterated, as the space of 1-BV measures contains measures with singular behavior in the neutral direction.

The first step in the strategy is approximating the disintegration of μ with a kind of “piecewise constant one” in the next lemma.

Lemma 30. — *Let us consider a uniform grid of size $\varepsilon = 1/m$, $m \in \mathbb{N}$, on the interval $[0,1]$. Given a measure μ with $\|\mu\|_{1\text{-BV}} < \infty$. There is a measure μ_ε such that μ_ε is* *piecewise constant on the ε-grid ($\mu_{\varepsilon|x}$ is constant on each element of the grid as x varies) and*

$$\operatorname{var}_1(\mu_\varepsilon) \leqslant 2\operatorname{var}_1(\mu), \quad \|\mu_\varepsilon\|_{\text{“}1\text{”}} \leqslant \|\mu\|_{\text{“}1\text{”}}.$$

Furthermore,

$$\|\mu - \mu_\varepsilon\|_{\text{“}1\text{”}} \leqslant 2\varepsilon \operatorname{var}_1(\mu).$$

Proof. — Let us consider μ_ε defined by averaging in the following way: let $x \in [0,1]$ and let I_i be the element of the ε-grid containing x. Then, for a measurable set $A \subseteq M$, $\mu_{|x}(A)$ is defined as

$$\mu_{\varepsilon|x}(A) = \int_{I_i} \mu_{\varepsilon|\gamma}(A) d\gamma.$$

We remark that $\mu_{|x} - \mu_{\varepsilon|x} \leqslant \operatorname{osc}(\varepsilon, x_i(x), \mu)$, where $x_i(x)$ is the grid center closest to x, and $\operatorname{osc}(\varepsilon, x_i(x), \mu) \leqslant \operatorname{osc}(2\varepsilon, x, \mu)$ then

$$\int \|\mu_{|x} - \mu_{\varepsilon|x}\|_W \leqslant 2\varepsilon \frac{\int \operatorname{osc}(2\varepsilon, x, \mu)}{2\varepsilon} \leqslant 2\varepsilon \sup_{2\varepsilon \leqslant A} \Big(\frac{\int \operatorname{osc}(2\varepsilon, x, \mu)}{2\varepsilon}\Big) \leqslant 2\varepsilon \operatorname{var}_p(\mu).$$

The other inequalities are analogous. □

Proposition 31. — *The convergence to equilibrium of a system satisfying* (Te1) *and* (Te2) *can be estimated as*

$$\|L^n \nu\|_{\text{“}1\text{”}} \leqslant C_4 n^{-1/8\gamma} \|\nu\|_{\text{1-BV}}.$$

Proof. — Consider a 1-BV measure ν, without loss of generality we can suppose $\|\nu\|_{\text{1-BV}} = 1$. Let us approximate ν it with a Lipschitz measure. First let us approximate it with a piecewise constant measure ν_ε as before. We have

$$\|\nu - \nu_\varepsilon\|_{\text{“}1\text{”}} \leqslant 2\varepsilon \operatorname{var}_1(\nu) \leqslant 2\varepsilon.$$

Let ν_i be such that $\nu_i = \nu_{\varepsilon|x_i}$ with x_i center of the ε grid as before, and f_i be the convolution $\gamma * \nu_i$, where γ is a ε_2^{-1} Lipschitz mollifier supported in $[-\varepsilon_2, \varepsilon_2]^d$. Then f_i is a ε_2^{-1} Lipschitz function. Let

$$f(x,y) = \begin{cases} f_i(y) & \text{if } |x - x_i| \leqslant (1-\varepsilon_2)\varepsilon, \\ \phi_i(x) f_i(y) + (1 - \phi_i(x)) f_{i+1}(y) & \text{if } x_i + (1-\varepsilon_2)\varepsilon \leqslant x \leqslant x_{i+1} - (1-\varepsilon_2)\varepsilon, \end{cases}$$

where ϕ_i is a linear function such that

$$\phi_i(x_i + (1-\varepsilon_2)\varepsilon) = 0 \quad \text{and} \quad \phi_i(x_{i+1} - (1-\varepsilon_2)\varepsilon) = 1.$$

We remark that f is $\sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1}$ Lipschitz, $\int f\, dm = 0.$ and $\|\nu_\varepsilon - fm\|_{\text{“}1\text{”}} \leqslant 3\varepsilon_2$. Hence

$$\|\nu - fm\|_{\text{“}1\text{”}} \leqslant 2\varepsilon \operatorname{var}_1(\nu) + 3\varepsilon_2. \tag{13}$$

Since the convolution with a Lipschitz kernel is a weak contraction in the Wasserstein norm, applying Lemma 30 we get $\operatorname{var}_1(fm) \leqslant 2\operatorname{var}_1(\nu)$ and $\|f\|_{\text{“}1\text{”}} \leqslant \|\nu\|_{\text{“}1\text{”}}$. Now we apply Proposition 28 in an efficient way. The proposition concerns the behavior of the correlation of two observables. We will consider f as one of them, and the other will be constructed in a suitable way to get the desired statement.

Let f be the Lipschitz density found above. Let $\mu = L^n fm$. Let $\mu_\varepsilon$its piecewise constant approximation defined as in Lemma 30 and $\mu_i = \mu_{\varepsilon|x_i}$. Consider 1-Lipschitz functions $\ell_i : \mathcal{T}^d \to \mathbb{R}$ such that $\big| |\int \ell_i \mu_i| - \|\mu_i\|_W \big| \leqslant \xi$, consider functions $h_i : [0,1] \to \mathbb{R}$ such that $h_i = 1$ on the central third of the i interval of the ε-grid and zero elsewhere, and $\mathrm{Lip}(h_i) = 3\varepsilon^{-1}$. Consider $g_i : X \to \mathbb{R}$ defined by $g_i(x,y) = \ell_i(y) h_i(x)$ and $g = \sum_i g_i$. By what is said above, we have

$$\|\mu_\varepsilon\|_{“1”} \leqslant 3\Big(\xi + \int g\mu_\varepsilon\Big)$$

and by Lemma 30, $\|L^n fm - \mu_\varepsilon\|_{“1”} \leqslant 2\varepsilon\,\mathrm{var}_1(\mu)$. Then

$$\begin{aligned}\|L^n(fm)\|_{“1”} &\leqslant \|\mu_\varepsilon\|_{“1”} + 2\varepsilon\,\mathrm{var}_1(\mu)\\ &\leqslant 3\Big(\xi + \int g\mu_\varepsilon\Big) + 2\varepsilon\,\mathrm{var}_1(\mu).\end{aligned}$$

Now consider $\int g\ L^n fm$. Since g is 1-Lipschitz in the y direction, we have that

$$\Big|\int g\ L^n f - \int g\ \mu_\varepsilon\Big| \leqslant \|L^n fm - \mu_\varepsilon\|_{“1”} \leqslant 2\varepsilon\,\mathrm{var}_1(\mu)$$

and

$$\|L^n f\|_{“1”} \leqslant 3\Big(\xi + \int g\ dL^n f\mu_0 + 2\varepsilon\,\mathrm{var}_1(\mu)\Big) + 2\varepsilon\,\mathrm{var}_1(\mu).$$

Now, since $\int f\,dm = 0$, by Proposition 28

$$\Big|\int g\ dL^n(fm)\Big| \leqslant C\|f\|_{\mathrm{lip}}\|g\|_{\mathrm{lip}} n^{-1/2\gamma}$$

then

$$\|L^n f\|_{“1”} \leqslant 3(\xi + C\|f\|_{\mathrm{lip}} 3\varepsilon^{-1} n^{-1/2\gamma} + 2\varepsilon\,\mathrm{var}_1(L^n(fm))) + 2\varepsilon\,\mathrm{var}_1(L^n(fm)).$$

We recall that the Lebesgue measure is invariant for the system. Then if K is a constant density such that $f + K \geqslant 0$ it holds $L^n(fm + Km) = L^n(fm) + Km$. It holds $\mathrm{var}_1(L^n f) = \mathrm{var}_1(L^n(f+K))$, since it is a positive measure, to $(f+K)m$ we can apply the regularization inequality (Proposition 20). Setting $B = B_T(\lambda^p + \widehat{H} + 3qC_h)/(1-\lambda^p\alpha)$ we get

$$\mathrm{var}_1(L^n(fm)) \leqslant \lambda^n\alpha^n\,\mathrm{var}_1(f) + B(\|f\|_{“1”} + K)$$

since f is $\sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1}$-Lipschitz and $\|f\|_{\text{1-BV}} \leqslant 1$, $\|f\|_\infty \leqslant \sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} + 1$. Therefore $\mathrm{var}(L^n(fm)) \leqslant \lambda^n\alpha^n\ \mathrm{var}(f) + B(1 + \sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} + 1)$ and

$$\begin{aligned}\|L^n(fm)\|_{“1”} &\leqslant 3\xi + 3C\|f\|_{\mathrm{lip}} 3\varepsilon^{-1} n^{-1/2\gamma} + 8\varepsilon\,\mathrm{var}_1(L^n f)\\ &\leqslant 3\xi + 3C\|f\|_{\mathrm{lip}} 3\varepsilon^{-1} n^{-1/2\gamma} + 8\varepsilon[\lambda^n\alpha^n\,\mathrm{var}_1(f) + B(1 + \sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} + 1)]\\ &\leqslant 3\xi + 3C\|f\|_{\mathrm{lip}} 3\varepsilon^{-1} n^{-1/2\gamma} + 16\varepsilon[\lambda^n\alpha^n\,\mathrm{var}_1(f) + B(1 + \sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} + 1)]\\ &\leqslant 3\xi + 3C\sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} 3\varepsilon^{-1} n^{-1/2\gamma} + 16\varepsilon[\lambda^n\alpha^n\,\mathrm{var}_1(f) + B(1 + \sqrt{2}\,\varepsilon_2^{-1}\varepsilon^{-1} + 1)].\end{aligned}$$

Taking $\xi = n^{-1/2\gamma}$, $\varepsilon_2 = n^{-1/8\gamma}$, $\varepsilon = n^{-1/8\gamma}$, recalling that $\mathrm{var}_1(f) \leqslant \mathrm{var}_1(\nu) \leqslant 1$, we obtain

$$\begin{aligned}\|L^n(fm)\|_{“1”} &\leqslant 3n^{-1/2\gamma} + 9C\sqrt{2}\, n^{3/8\gamma} n^{-1/2\gamma} + 16n^{-1/8\gamma}(\alpha\lambda^n + 2B) + \sqrt{2}\, Bn^{-1/8\gamma} \\ &\leqslant C_3 n^{-1/8\gamma}.\end{aligned}$$

Finally, by Equation (13)

$$\begin{aligned}\|L^n\nu\|_{“1”} &\leqslant \|L^n(\nu - fm)\|_{“1”} + \|L^n(fm)\|_{“1”} \\ &\leqslant 2\varepsilon\|\nu\|_{1\text{-BV}} + \|L^n(fm)\|_{“1”} + 3\varepsilon_2 \leqslant C_4 n^{-1/8\gamma}.\end{aligned}$$ □

Once we have an estimate for the speed of convergence to equilibrium, by Proposition 25, and Remark 7, the following holds directly:

Proposition 32. — *Consider a family of skew product maps $F_\delta = (T_\delta, G_\delta)$, $0 \leqslant \delta \leqslant D$ satisfying uniformly* (Sk1)–(Sk3) *and let $f_\delta \in \mathcal{L}^1$ its invariant probability measures, suppose:*

(1) *F_0 is a piecewise constant toral extension as defined in Section* 6*, with linear Diophantine type $\gamma_\ell(\theta)$.*

(2) *There exists $n \in \mathbb{N}$ such that for each $\delta \leqslant D$, $d_{n,S}(T_0, T_\delta) \leqslant \delta$.*

(3) *For each $\delta \leqslant D$ there exists a set $A_2 \in \mathrm{Int}_n$ such that $m(A_2) \geqslant 1 - \delta$ and for all $x \in A_2, y \in \mathcal{T}^d$, it holds $|G_0(x,y) - G_\delta(x,y)| \leqslant \delta$.*

Then for each $\gamma > \gamma_\ell(\theta)$ there exists K_1 such that for δ small enough

$$\|f_\delta - f_0\|_{“1”} \leqslant K_1 \delta^{1/(8\gamma+1)}.$$

6.3. An example having Hölder behavior. — In this section we show a simple example of perturbation of toral extensions satisfying assumptions (Te1) and (Te2) for which the statistical behavior is actually, only Hölder stable. This shows how that Propositions 25 and 32 give a general estimate, which is quite sharp in the case of piecewise constant toral extensions.

Proposition 33. — *Consider a well approximable Diophantine irrational θ with $\gamma_\ell(\theta) > 2$. Let us consider the map $F_0 : [0,1] \times \mathcal{T}^1$ defined as a skew product $F_0(T_0(x), G_0(x,y))$, where*

$$T_0(x) = 2x \bmod (1) \quad \text{and} \quad G_0(x,y) = y + \theta\varphi(x),$$

where $\varphi = \chi_{[1/2,1]}$. Consider $\gamma' < \gamma_\ell(\theta)$. There exist a sequence of reals $\delta_j \geqslant 0$, $\delta_j \to 0$ and a sequence of perturbed of maps $\widehat{F}_{\delta_j}(x,y) = (\widehat{T}_{\delta_j}(x), \widehat{G}_{\delta_j}(x,y))$ satisfying (Sk1)–(Sk3)*, with $\widehat{T}_{\delta_j}(x) = T_0(x)$ and $\|\widehat{G}_{\delta_j}(x,y) - G_0(x,y)\|_\infty \leqslant 2\delta_j$ such that*

$$\|\mu_0 - \mu_j\|_{“1”} \geqslant \frac{1}{9}{\delta_j}^{1/(\gamma'-1)}$$

holds for every j and every μ_j, invariant measure of $\widehat{F}_{\delta_j}(x,y)$ in $\mathcal{L}^1$.

Proof. — We remark that since there is convergence to equilibrium for F_0, the Lebesgue measure μ_0 on $[0,1]\times\mathcal{T}^1$ is the unique invariant measure in $\mathcal{L}^1$ for F_0. Consider $F_\delta=(T_0(x),\ y+(\delta+\theta)\varphi(x)))$. For a sequence of values of δ converging to 0 it holds that $(\delta+\theta)$ is rational. For this sequence the map $y\mapsto y+(\delta+\theta)$ $(:\mathcal{T}^1\to\mathcal{T}^1)$ is such that, 0 has a periodic orbit. Let $y_1=0,\dots,y_k$ be this orbit. For these parameters, consider the product measure $\mu_n=\frac{1}{k}\sum_{i\leqslant k} m\otimes\delta_{y_i}$, where m is the Lebesgue measure on $[0,1]$ and δ_{y_i} is the delta measure placed on y_i. The measure μ_n is invariant for $F_\delta(x,y)$ and is in $\mathcal{L}^1$. It is easy to see that $\|\mu_0-\mu_n\|_{“1”}\geqslant(1/9)(1/k)$. Now the Diophantine type of θ will give an estimate for the relation between δ and k. Indeed, let $\gamma'<\gamma(\theta)$, by the Diophantine type of θ we know that there are infinitely many k_j and integers p_j such that $|k_j\theta-p_j|\leqslant|1/k_j|^{\gamma'}$. Then $|\theta-p_j/k_j|\leqslant|1/k_j|^{\gamma'-1}$. Let us now consider $\delta_j=-\theta+p_j/k_j$. It holds $|\delta_j|\leqslant|1/k_j|^{\gamma'-1}$ and the angle $(\delta_j+\theta)$ generates a periodic orbit of period k_j. This happens by perturbing the second coordinate of the map by a quantity which is less than $|1/k_j|^{\gamma'-1}$. Summarizing, for the map F_{δ_j}

– we have that there is no perturbation on the first coordinate,
– for the second coordinate we have $\|G_0-G_{\delta_j}\|_\infty\leqslant\delta_j$, and
– denoting by μ_j the invariant measure on the periodic orbit defined before, it holds

$$\|\mu_0-\mu_j\|_{“1”}\geqslant\frac{1}{9}\,\delta_j{}^{1/(\gamma'-1)}.$$

This example can further be improved by perturbing the map F_{δ_j} to a new map $\widehat{F}_{\delta_j}$ in a way that μ_j (a measure supported on the attractor of $\widehat{F}_{\delta_j}$) and $\mu_j+k_j/2$ [(6)](supported on the repeller of $\widehat{F}_{\delta_j}$) are the only invariant measures in $\mathcal{L}^1$ for $\widehat{F}_{\delta_j}$ and μ_j is the unique physical measure for the system. This can be done by making a small further C^∞ perturbation on G. Let us denote again by $y_1,\dots,y_{k_j}$ the periodic orbit of 0 as before. Let us consider a C^∞ function $g:[0,1]\to[0,1]$ such that:

– g is negative on the each interval $[y_i,y_i+1/2k_j]$ and positive on each interval $[y_i+1/2k_j,y_{i+1}]$ (so that $g(y_i+1/2k_j)=0$).

– g' is positive in each interval $[y_i+1/3k_j,y_{i+1}-1/3k_j]$ and negative in $[y_i,y_{i+1}]-[y_i+1/3k_j,y_{i+1}-1/3k_j]$.

Considering $D_\delta:\mathcal{T}^1\to\mathcal{T}^1$ defined by $D_\delta(x)=x+\delta g(x)\bmod(1)$, it holds that the iteration of this map send all the space but the set $\{y_i+1/2k_j\mid i\leqslant k_j\}$ (which is a repeller) to the set $\{y_i\mid i\leqslant k_j\}$ (the attractor). Then define $\widehat{F}_{\delta_j}$ as:

$$\widehat{F}_{\delta_j}(x,y)=(T_{\delta_j}(x),D_{\delta_j}(y+(\delta_j+\theta)\varphi(x))).$$

The claim directly follows from the remark that for the map $(\widehat{F}_{\delta_j})^{k_j}$ the sets

$$\Gamma_1:=[0,1]\times\{y_i\mid i\leqslant k_j\}\quad\text{and}\quad\Gamma_2:=[0,1]\times\{y_i+1/2k_j\mid i\leqslant k_j\}$$

are invariant and the set Γ_1 attracts the whole $[0,1]\times\mathcal{T}^1\smallsetminus\Gamma_2$. ☐

[(6)]Defined as $[\mu_j+1/2k_j](A)=\mu_j(A-1/2k_j)$ for each measurable set A in $\mathcal{T}^1$, where $A-1/2k_j$ is the translation of the set A by $-1/2k_j$.

The construction done in the previous proof can be extended to show Hölder behavior for the average of a given regular observable. We show an explicit example of such an observable for a skew product with a particular angle θ.

Proposition 34. — *Consider a map F_0 as above with the rotation angle $\theta = \sum_1^\infty 2^{-2^{2i}}$, with*

$$T_0(x) = 2x \bmod (1) \quad \textit{and} \quad G_0(x,y) = y + \theta\varphi(x)$$

as in Proposition 33. Let $\widehat{F}_{\delta_j}$ be its perturbations as described in the proof of the proposition and μ_j their invariant measures in $\mathcal{L}^1$. There is an observable $\psi : [0,1] \times \mathcal{T}^1 \to \mathbb{R}$ with derivative in L^2 and $C \geqslant 0$ such that

$$\left| \int \psi d\mu_0 - \int \psi d\mu_j \right| \geqslant C\sqrt{\delta_j}.$$

Proof. — We recall that that $\sum_{n+1}^\infty 2^{-2^{2i}} \leqslant 2^{-2^{2(n+1)}+1}$. From this we deduce $\|2^{2^{2n}}\theta\| \leqslant 2^{-2^{2(n+1)}+1}$ and the Diophantine type of θ is greater than 4. Following the construction above, we have that with a perturbation of size less than $2^{-2^{2(n+1)}+1}$ the angles $\theta_j = \sum_1^j 2^{-2^{2i}}$ generate on the second coordinate of the skew product orbits of period $2^{2^{2j}}$. Now let us construct a suitable observable which can "see" the change of the invariant measure under this perturbation. Let us consider

$$\psi(x,y) = \sum_1^\infty \frac{1}{(2^{2^{2i}})^2} \cos(2^{2^{2i}} 2\pi y) \quad \text{and} \quad \psi_k(x,y) = \sum_1^k \frac{1}{(2^{2^{2i}})^2} \cos(2^{2^{2i}} 2\pi y). \tag{14}$$

Since for the observable ψ, the i-th Fourier coefficient decreases like i^{-2}, then ψ has a derivative in L^2. Let $x_1 = 0, \dots, x_{2^{2^{2j}}}$ be the periodic orbit of 0 for $y \mapsto y + \theta_j$, and $\mu_j = (1/2^{2^{2i}}) \sum \delta_{x_i}$ the physical measure supported on it. Since $2^{2^{2j}}$ divides $2^{2^{2(j+1)}}$ we have $\sum_{i=1}^{2^{2^{2j}}} \psi_k(x_i) = 0$ for every $k < j$, thus $\int \psi_{j-1}\, d\mu_j = 0$. Then

$$v_j := \int \psi \ d\mu_j \geqslant \frac{1}{(2^{2^{2j}})^2} - \sum_{j+1}^\infty \frac{1}{(2^{2^{2i}})^2} \geqslant 2^{-2^{2j+1}} - 2^{-2^{2(j+1)}+1}.$$

And for j big enough,

$$2^{-2^{2j+1}} - 2^{-2^{2(j+1)}+1} \geqslant \frac{1}{2}(2^{-2^{2j}})^2.$$

Summarizing, with a perturbation of size

$$\delta_j = \sum_{j+1}^\infty 2^{-2^{2i}} \leqslant 2 * 2^{-2^{2(j+1)}} = 2^{-2^{2(j+1)}} = 2(2^{-2^{2j}})^4$$

we get a change of average for the observable ψ from $\int \psi\, dm = 0$ to $v_n \geqslant \frac{1}{2}(2^{-2^{2j}})^2$. Hence there exists a $C \geqslant 0$ such that with a perturbation of size δ_j we get a change of average for the observable ψ of size bigger than $C\sqrt{\delta_j}$. □

Remark 35. — Using $1/(2^{2^{2i}})^{\alpha}$ instead of $1/(2^{2^{2i}})^{2}$ in (14) we can obtain a smoother observable. Using rotation angles with bigger and bigger Diophantine type it is possible to obtain a dependence of the physical measure to perturbations with worse and worse Hölder exponent. Using angles with infinite Diophantine type it is possible to have a behavior whose modulus of continuity is worse than the Hölder one.

References

[1] J. F. Alves – "Strong statistical stability of non-uniformly expanding maps", *Nonlinearity* **17** (2004), no. 4, p. 1193–1215.

[2] J. F. Alves & M. Soufi – "Statistical stability in chaotic dynamics", in *Progress and challenges in dynamical systems*, Springer Proc. Math. Stat., vol. 54, Springer, Heidelberg, 2013, p. 7–24.

[3] J. F. Alves & M. Viana – "Statistical stability for robust classes of maps with non-uniform expansion", *Ergodic Theory Dynam. Systems* **22** (2002), no. 1, p. 1–32.

[4] V. Araújo, S. Galatolo & M. J. Pacifico – "Decay of correlations for maps with uniformly contracting fibers and logarithm law for singular hyperbolic attractors", *Math. Z.* **276** (2014), no. 3-4, p. 1001–1048, Erratum: *Ibid.* **282** (2016), no. 1-2, p. 615–621.

[5] W. Bahsoun & B. Saussol – "Linear response in the intermittent family: differentiation in a weighted C^0-norm", *Discrete Contin. Dynam. Systems* **36** (2016), no. 12, p. 6657–6668.

[6] W. Bahsoun & S. Vaienti – "Metastability of certain intermittent maps", *Nonlinearity* **25** (2012), no. 1, p. 107–124.

[7] V. Baladi – "Linear response, or else", in *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. III*, Kyung Moon Sa, Seoul, 2014, p. 525–545.

[8] V. Baladi, M. Benedicks & D. Schnellmann – "Whitney-Hölder continuity of the SRB measure for transversal families of smooth unimodal maps", *Invent. Math.* **201** (2015), no. 3, p. 773–844.

[9] V. Baladi, T. Kuna & V. Lucarini – "Linear and fractional response for the SRB measure of smooth hyperbolic attractors and discontinuous observables", *Nonlinearity* **30** (2017), no. 3, p. 1204–1220, Corrigendum: *Ibid.* **30** (2017), no. 8, p. C4–C6.

[10] V. Baladi & M. Todd – "Linear response for intermittent maps", *Comm. Math. Phys.* **347** (2016), no. 3, p. 857–874.

[11] A. Boyarsky & P. Góra – *Laws of chaos. Invariant measures and dynamical systems in one dimension*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1997.

[12] M. I. Brin – "The topology of group extensions of C-systems", *Mat. Zametki* **18** (1975), no. 3, p. 453–465.

[13] K. Burns & A. Wilkinson – "Stable ergodicity of skew products", *Ann. Sci. École Norm. Sup. (4)* **32** (1999), no. 6, p. 859–889.

[14] O. Butterley & P. Eslami – "Exponential mixing for skew products with discontinuities", *Trans. Amer. Math. Soc.* **369** (2017), no. 2, p. 783–803.

[15] D. Dolgopyat – "On dynamics of mostly contracting diffeomorphisms", *Comm. Math. Phys.* **213** (2000), no. 1, p. 181–201.

[16] ______, "On mixing properties of compact group extensions of hyperbolic systems", *Israel J. Math.* **130** (2002), p. 157–205.

[17] ______, "On differentiability of SRB states for partially hyperbolic systems", *Invent. Math.* **155** (2004), no. 2, p. 389–449.

[18] S. Galatolo – "Quantitative statistical stability and convergence to equilibrium. An application to maps with indifferent fixed points", *Chaos Solitons Fractals* **103** (2017), p. 596–601.

[19] S. Galatolo & R. Lucena – "Spectral gap and quantitative statistical stability for systems with contracting fibers and lorenz like maps", `arXiv:1507.08191`, 2015.

[20] S. Galatolo, M. Monge & I. Nisoli – "Rigorous approximation of stationary measures and convergence to equilibrium for iterated function systems", *J. Phys. A* **49** (2016), no. 27, p. 274001, 22.

[21] S. Galatolo & I. Nisoli – "An elementary approach to rigorous approximation of invariant measures", *SIAM J. Appl. Dyn. Syst.* **13** (2014), no. 2, p. 958–985.

[22] S. Galatolo & P. Peterlongo – "Long hitting time, slow decay of correlations and arithmetical properties", *Discrete Contin. Dynam. Systems* **27** (2010), no. 1, p. 185–204.
[23] S. Galatolo, J. Rousseau & B. Saussol – "Skew products, quantitative recurrence, shrinking targets and decay of correlations", *Ergodic Theory Dynam. Systems* **35** (2015), no. 6, p. 1814–1845.
[24] G. Keller – "Generalized bounded variation and applications to piecewise monotonic transformations", *Z. Wahrsch. Verw. Gebiete* **69** (1985), no. 3, p. 461–478.
[25] A. Korepanov – "Linear response for intermittent maps with summable and nonsummable decay of correlations", *Nonlinearity* **29** (2016), no. 6, p. 1735–1754.
[26] C. Liverani – "Invariant measures and their properties. A functional analytic point of view", in *Dynamical systems. Part II*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003, p. 185–237.
[27] F. Naud – "Stability of the spectrum for transfer operators", *J. Spectral Theory* (to appear), arXiv:1612.05011.
[28] ———, "Exponentially mixing, locally constant skew extensions of shift maps", arXiv:1701.08659, 2017.
[29] D. Ruelle – "Flots qui ne mélangent pas exponentiellement", *C. R. Acad. Sci. Paris Sér. I Math.* **296** (1983), no. 4, p. 191–193.
[30] W. Shen & S. van Strien – "On stochastic stability of expanding circle maps with neutral fixed points", *Dyn. Syst.* **28** (2013), no. 3, p. 423–452.
[31] Z. Zhang – "On the smooth dependence of srb measures for partially hyperbolic systems", arXiv:1701.05253, 2017.

Manuscript received 24th April 2017
accepted 30th March 2018

Stefano Galatolo, Dipartimento di Matematica, Università di Pisa
Via Buonarroti 1, 56127 Pisa, Italy
E-mail : galatolo@dm.unipi.it
Url : http://users.dma.unipi.it/galatolo/