Polynomial Decay for Non-uniformly Expanding Maps

Huyi Hu [*] Sandro Vaienti [†]

July 1, 2013

Abstract

We give conditions under which nonuniformly expanding maps have polynomial decay of correlations. We show that if the Lasota-Yorke type inequalities for the transfer operator of a first return map are satisfied in a Banach space $\mathcal{B}$, and the absolutely continuous invariant measure obtained is weak mixing, in terms of aperiodicity, then under some renewal condition, the maps has polynomial decay of correlations for observables in $\mathcal{B}$. We also provide some general conditions that give aperiodicity for expanding maps in higher dimensional spaces. As applications, we obtain polynomial decay, including lower bound in some cases, for piecewise expanding maps with an indifferent fixed point and for which we also allow non-markov structure and unbounded distortion. The observables are functions that have bounded variation or satisfy quasi-Hölder conditions respectively.

Contents

[*]Mathematics Department, Michigan State University, East Lansing, MI 48824, USA. e-mail: <hu@math.msu.edu>.

[†]UMR-7332 Centre de Physique Théorique, CNRS, Université d'Aix-Marseille, Université du Sud Toulon-Var and FRUMAM, Fédération de Recherche des Unités des Mathématiques de Marseille, CPT Luminy, Case 907, F-13288 Marseille CEDEX 9. e-mail: <vaienti@cpt.univ-mrs.fr>. HH was supported by the University of Marseille and the University of Toulon during his visits at the Center of Theoretical Physics in Luminy. SV was supported by the ANR-Project *Perturbations*, by the CNRS-PEPS *Mathematical Methods of Climate Theory* and by the PICS (Projet International de Coopération Scientifique), *Propriétés statistiques des systèmes dynamiques detérministes et aléatoires*, with the University of Houston, n. PICS05968. Part of this work was done while he was visiting the *Centro de Modelamiento Matemático, UMI2807*, in Santiago de Chile with a CNRS support (délégation).

0 Introduction

The purpose of this paper is to study polynomial decay of correlations for invariant measures which are absolutely continuous with respect to some reference measures. Typically the maps T which we consider are non uniformly expanding and they may neither have a Markov partition nor exhibit bounded distortion. The main tool we use is the transfer operator on induced subsystems endowed with the first return map. Let us call $||R_n||$ a suitable norm (see below) of the n-th power of the transfer operator restricted to the level sets with first return time $\tau = n$. We will show that if Lasota-Yorke inequalities can be verified for the transfer operator of the first return maps, and if $||R_n||$ converges to 0 at a speed $1/n^{\beta+1}$, $\beta > 1$, then the decay rates are given by the measure of the sets $\{\tau = n\}$. In the second part of the paper we apply the results to piecewise expanding maps with an indifferent fixed point in one dimensional and higher dimensional spaces to get polynomial decay of correlations. The results for maps in higher dimensional spaces with $Df_p = \mathrm{id}$ at the indifferent fixed point p is new, and in all the cases, the observables are more general than Hölder functions.

We now explain in more details the content of this paper. Let us consider a non uniformly expanding map T defined on a compact subset $X \subset \mathbb{R}^n$, with or without discontinuities. Since we do not have necessarily bounded distortion or Markov partitions, Hölder continuous functions are not preserved under the transfer operator. Therefore we will work on Banach spaces $\mathcal{B}$ consisting of some L^1 functions, and endow a norm $\|\cdot\|_{\mathcal{B}}$ stronger than the L^1 norm $\|\cdot\|_{L^1}$. We give some conditions on $\mathcal{B}$ under which the results apply (see Assumption B); notice that the norm of R_n will be taken in $\mathcal{B}$.

Let us now take a subset $\widetilde{X} \subset X$ and define the first return map $\widehat{T}$. The *first ingredient* of our theorem is the *Lasota-Yorke inequality* for the transfer operator $\widehat{\mathscr{P}}$ of $\widehat{T}$ with respect to the norm $\|\cdot\|_{\mathcal{B}}$ and $\|\cdot\|_{L^1}$. Hence, $\widehat{\mathscr{P}}$ has a fixed point $\hat{h}$ that defines an absolutely continuous measure $\hat{\mu}$ invariant under $\widehat{T}$. The measure $\hat{\mu}$ can be extended to a measure μ on X invariant under T. We may assume ergodicity for $\hat{\mu}$, otherwise we can take an ergodic component. Then the ergodicity of $\hat{\mu}$ gives ergodicity of μ. However, we also need some mixing property for μ. Therefore our *second ingredient* is to require that the function τ given by the first return time is *aperiodic*, which is equivalent to the weak mixing of μ for T. The *third ingredient* is the *renewal condition*, which could be stated by asking that $||R_n||$ decays at least as $n^{-(\beta+1)}$, with $\beta > 1$. Such a decay gives also an estimate of the error term, which should be faster than the decay rates of $\mu(\tau > n)$ in order to get an optimal rate for the decay of correlations. With these conditions our theorem (Theorem A) states that the decay of correlations $\mathrm{Cov}(f, g \circ T^n) := |\int f\ g \circ T^n\ d\mu - \int f d\mu\ \int g\ d\mu|$, is polynomial for functions $f \in \mathcal{B}$ and $g \in L^\infty(X, \nu)$ with $\operatorname{supp} f, \operatorname{supp} g \subset \widehat{X}$.

The assumption on aperiodicity is usually difficult to check. We provide some general conditions in Theorem B for the maps T under which aperiodicity follows automatically. The conditions include piecewise smoothness, finite image and uniform expansion for an induced maps and topological mixing for the original maps.

As applications we studied piecewise smooth expanding maps with an indifferent fixed point in one and higher dimensional spaces. In the one-dimensional case we use the set of bounded variation functions for the Banach space $\mathcal{B}$, and we found that the decay rates are of the order $n^{\beta-1}$ if near the fixed point the maps has the form $T(x) \approx x + x^{1+\gamma}$, $\gamma \in (0,1)$ and $\beta = 1/\gamma$.

One of our main goal in the paper is to obtain polynomial decay of correlations for piecewise smooth expanding maps with an indifferent fixed point in *higher dimensional space*.

For a large class of those maps, we constructed, in a previous paper ([HV]), absolutely continuous invariant measures by using the Lasota-Yorke inequality. Our maps could be written in the form of (5.4) near the indifferent fixed point p, where the local behavior is precisely given by an isometry plus homogeneous terms and higher order terms. In the present paper we show that such maps have polynomial decay of correlations for observables in $\mathcal{B}$. As we said above, in the estimates we should compare the decay of $\|R_n\|$ with the measure of the level sets with the first return time larger than n. The former could be determined by the norms $\|DT^{-n}\|$ or the determinants $|\det DT^{-n}|^*$, where the latter, denoted by $\mu(\tau > n)$, is often of order $n^{-m/\gamma}$, with $m = \dim X$ and γ is given in (5.4). If $\|R_n\|$ decreases as $|\det DT^{-n}|$, then it usually approaches to 0 faster than $\mu(\tau > n)$ does, and therefore both upper and lower estimates

*Notice that T^{-n} denotes the inverse of T^n restricted to the domain of injectivity containing p.

for decay rates of correlations are of the same order as $\sum_{k\geq n}\mu(\tau > k)$. This is the case we have *optimal* rates of decay of correlations. We obtain optimal rates under the assumption that all preimages of some neighborhood of p do not intersect discontinuities, (see Theorem E and examples in Section 7 for more details). This is satisfied whenever T has a Markov partition or a finite range structure (see Remark 6.1). Whenever $\|R_n\|$ decreases as $||DT^{-n}||$, we get polynomial upper bound (Theorem D).

Our first theorem, Theorem A, is based on works of Sarig ([Sr], see also [Go]), but contrarily to them, we do not assume existence of absolutely continuous invariant measures. Our conditions are given by the Lasota-Yorke type inequalities, which imply existence of absolutely continuous invariant measures and the conditions on spectral gap as they required. Since such conditions are easier to verify for maps without Markov partitions, it makes it possible to verify those inequalities for observables beyond Hölder continuous functions.

The proof of aperiodicity in Theorem B is particularly technical. We use some results in the theory developed in the paper [ADSZ], where aperiodicity is proved for a large class of interval maps, and some methods in [AD] for skew product rigidity. We extended aperiodicity result to the multidimensional setting without Markov partition. The authors in [ADSZ] mentioned that aperiodicity for non-Markov case was not so well understood. Our results indicate that under some general conditions, if we could find a suitable Banach space for which a Lasota-Yorke type of inequality (see (1.6)) can be verified, then aperiodicity follows.

For piecewise expanding interval maps with indifferent fixed points, it is relatively easy to get the desired spectral properties on the space of bounded variation functions and to estimate decreasing rates for $\|R_n\|$: our theorem allows then to get optimal polynomial decay rates of correlation.

The higher dimensional case is much more complicated. Part of reason is due to unbounded distortion of the systems caused by different expansion rates in different directions as a point move away from the indifferent fixed point. Moreover it is not easy to estimate the decreasing rates of the norm $\|R_n\|$ for quasi-Hölder spaces: Theorems D and E deals with these situations, by assuming certain hypothesis. One surely needs more work to weaken those assumptions and achieve optimal decay for a much larger class of maps.

To study statistic properties for non uniformly hyperbolic or expanding systems, it is common to find some "good" part on which we can get bounded distortion, like Pesin's blocks ([Ps]), elements in Young's tower ([Yo1, Yo2]), or some neighborhood near points that have hyperbolic time ([ABV]). Another approach is to work directly on some Banach spaces, like bounded variation functions ([LY]) or quasi-Hölder functions ([Ss]), that are preserved by the transfer operator of the dynamical system. Our paper follows the latter way and we give some conditions on Banach spaces through which one can obtain some statistical properties such as existence of physical measure and decay of correlations.

We would like to remark at this point that the functional space is abstract, as long as certain general assumptions (Assumption B(d) to (f)) are satisfied. In the applications, we present two type of Banach spaces. It seems interesting to find more different spaces to deal with different kinds of dynamical systems.

Part I: Conditions for Polynomial Decay Rates

1 Assumptions and statements of results

Let $X \subset \mathbb{R}^m$ be a subset with positive Lebesgue measure ν. We assume $\nu X = 1$. Let d be the (euclidean) metric induced from $\mathbb{R}^m$.

The transfer (Perron-Frobenius) operator $\mathscr{P} = \mathscr{P}_\nu : L^1(X,\nu) \to L^1(X,\nu)$ is defined by $\int \psi \circ T\phi d\nu = \int \psi \mathscr{P}\phi d\nu$ $\forall \phi \in L^1(X,\nu)$, $\psi \in L^\infty(X,\nu)$.

Let $\widehat{X} \subset X$ be a measurable subset of X with positive Lebesgue measure.

Recall that the first return map of T with respect to $\widehat{X} \subset X$ is defined by $\widehat{T}(x) = T^{\tau(x)}(x)$, where $\tau(x) = \min\{i \geq 1 : T^i x \in \widehat{X}\}$ is the return time. We put $\hat{\nu}$ the normalized Lebesgue measure on $\widehat{X}$. Then we let $\widehat{\mathscr{P}} = \widehat{\mathscr{P}}_{\hat{\nu}}$ be the transfer operator of $\widehat{T}$.

Moreover we define

$$R_n f = 1_{\widehat{X}} \cdot \mathscr{P}^n(f 1_{\{\tau = n\}}) \quad \text{and} \quad T_n f = 1_{\widehat{X}} \cdot \mathscr{P}^n(f 1_{\widehat{X}}) \tag{1.1}$$

for any function f on $\widehat{X}$. For any $z \in \mathbb{C}$, denote $R(z) = \sum_{n=1}^{\infty} z^n R_n$. It is clear that $\widehat{\mathscr{P}} = R(1) = \sum_{n=1}^\infty R_n$.

For simplicity of notation, we regard the space $L^1(\widehat{X},\hat{\nu})$ as a subspace $L^1(X,\nu)$ consisting of functions supported on $\widehat{X}$, and we denote it with $L^1(\hat{\nu})$ or L^1 sometimes and when no ambiguity arises. We point out that in the following we will mainly work on the induced space and its "objects" will be indicated with an "$\widehat{\ }$".

Suppose that there is a seminorm $|\cdot|_{\mathcal{B}}$ for functions in $L^1(\widehat{X},\hat{\nu})$. Consider the set $\mathcal{B} = \mathcal{B}(\widehat{X}) = \{f \in L^1(\widehat{X},\hat{\nu}) : |f|_{\mathcal{B}} < \infty\}$. Define a norm on $\mathcal{B}$ by

$$\|f\|_{\mathcal{B}} = |f|_{\mathcal{B}} + \|f\|_1$$

for $f \in \mathcal{B}$, where $\|f\|_1$ is the L^1 norm. We assume that $\mathcal{B}$ satisfies the following.

Assumption B. (a) (Compactness) *$\mathcal{B}$ is a Banach space and the inclusion $\mathcal{B} \hookrightarrow L^1(\hat{\nu})$ is compact; that is, any bounded closed set in $\mathcal{B}$ is compact in $L^1(\hat{\nu})$.*

(b) (Boundness) *The inclusion $\mathcal{B} \hookrightarrow L^\infty(\hat{\nu})$ is bounded; that is, $\exists C_b > 0$ such that $\|f\|_\infty \leq C_b\|f\|_{\mathcal{B}}$ for any $f \in \mathcal{B}$.*

(c) (Algebra) *$\mathcal{B}$ is an algebra with the usual sum and product of functions, in particular there exists a constant C_a such that $\|fg\|_{\mathcal{B}} \le C_a \|f\|_{\mathcal{B}} \|g\|_{\mathcal{B}}$ for any $f, g \in \mathcal{B}$.*

Recall that for the system $(\widehat{T}, \hat{\nu})$, if the Lasota-Yorke's inequality (1.2) below is satisfied for any function $f \in \mathcal{B}$, and if the Banach space $\mathcal{B}$ satisfies Assumption B(a), then $\widehat{\mathscr{P}}$ has a fixed point $\hat{h} \in \mathcal{B}$ with $\hat{h} \ge 0$ and $\widehat{\mathscr{P}}\hat{h} = \hat{h}$, and the measure $\hat{\mu}$ defined by $\hat{\mu}(f) = \hat{\nu}(f\hat{h})$ is $\widehat{T}$ invariant.

Denote $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$.

Theorem A. *Let $X \subset \mathbb{R}^m$ be compact subset with $\nu X = 1$ and $\widehat{X} \subset X$ be a compact subset of X.*

Let $T : X \to X$ be a map whose first return map with respect to $\widehat{X}$ is $\widehat{T} = T^\tau$, and $\mathcal{B}$ be a Banach space satisfying Assumption B(a) to (c). We assume the following.

(i) (Lasota-Yorke inequality) *There exist constants $\eta \in (0, 1)$ and $D > 0$ such that for any $f \in \mathcal{B}$,*

$$|\widehat{\mathscr{P}} f|_{\mathcal{B}} \le \eta |f|_{\mathcal{B}} + D\|f\|_1; \tag{1.2}$$

(ii) (Spectral radius) *There exist constants $B, \hat{D} > 0$ and $\hat{\eta} \in (0, 1)$ such that for any $f \in \mathcal{B}$, $z \in \overline{\mathbb{D}}$,*

$$\|R(z)^n f\|_{\mathcal{B}} \le |z^n| \big(B\hat{\eta}^n \|f\|_{\mathcal{B}} + \hat{D}\|f\|_1\big); \tag{1.3}$$

(iii) (Ergodicity) *The measure $\hat{\mu}$ given by $\hat{\mu}(f) = \hat{\nu}(\hat{h}f)$ is ergodic, where $\hat{h}$ is a fixed point of $\widehat{\mathscr{P}}$.*

(iv) (Aperiodicity) *The function $e^{it\tau}$ given by the return time is aperiodic, that is, the only solution for $e^{it\tau} = f/f \circ \hat{T}$ almost everywhere with a measurable function $f : \widehat{X} \to \mathbb{S}$ are f constant almost everywhere and $t = 0$.*

If for any $n \ge 1$, R_n satisfies $\sum_{k=n+1}^{\infty} \|R_k\|_{\mathcal{B}} = O(n^{-\beta})$ for some $\beta > 1$, then there exists a constant $C > 0$ such that for any function $f \in \mathcal{B}$, $g \in L^\infty(X, \nu)$ with $\operatorname{supp} f$, $\operatorname{supp} g \subset \widehat{X}$,

$$\Big|\mathrm{Cov}(f, g \circ T^n) - \Big(\sum_{k=n+1}^{\infty} \mu(\tau > k)\Big) \int f d\mu \int g d\mu\Big| \le C F_\beta(n) \|g\|_\infty \|f\|_{\mathcal{B}}, \tag{1.4}$$

where $F_\beta(n) = 1/n^\beta$ if $\beta > 2$, $(\log n)/n^2$ if $\beta = 2$, and $1/n^{2\beta-2}$ if $2 > \beta > 1$.

Remark 1.1. *By a result of Hennion ([He], see also [HH]), (1.3) implies that the spectral radius and essential spectra radius of $R(z)$ are bounded by $|z|$ and $\hat{\eta}|z|$ respectively.*

Remark 1.2. *Practically, (1.3) usually can be obtained in a similar way as (1.2), (for example, see the proof of Theorem D). On the other hand, since* $\widehat{\mathscr{P}} = R(1)$*, (1.3) implies the Lasota-Yorke inequality for* $\widehat{\mathscr{P}}^n$ *for some* $n > 0$ *with* $B\hat{\eta}^n < 1$*.*

Remark 1.3. *The measure* $\hat{\mu}$ *can be extended to an absolutely continuous invariant measure* μ *on* X *in the usual way (see e.g. [Kk]). It is well known that if* $\hat{\mu}$ *is ergodic, so is* μ*.*

Remark 1.4. *As we said in the Introduction, Assumption (iv) is actually equivalent to the fact that* μ *is weak mixing for* T *(see e.g [PP]). Since decay of correlations implies mixing, we obtain that with Lasota-Yorke inequality, weak mixing implies mixing. This fact is also implied in the theorem of Ionescu-Tulcea and Marinescu ([IM]).*

Assumption (iv) is usually difficult to check. However, for piecewise expanding systems, the condition could be verified and we will give some sufficient conditions in Theorem B below.

The more general version of aperiodicity is as the following. Let $\mathbb{G}$ be a locally compact Abelian polish group. A measurable function $\phi : \widehat{X} \to \mathbb{G}$ is *aperiodic* if the only solutions for $\gamma \circ \phi = \lambda f / f \circ T$ almost everywhere with $\gamma \in \widehat{\mathbb{G}}$, $|\lambda| = 1$ and a measurable function $f : \widehat{X} \to \mathbb{G}$ are $\gamma = 1$, $\lambda = 1$ and f constant almost everywhere. (See [ADSZ] and references therein.) Here we only consider the case $\gamma = \mathrm{id}$, and $\phi = e^{it\tau}$, and $\mathbb{G}$ being the smallest compact subgroup of $\mathbb{S}$ containing e^{it}.

We denote by $B_\varepsilon(\Gamma)$ the ε neighborhood of a set $\Gamma \subset X$.

Assumption T. (a) (Piecewise smoothness) *There are countably many disjoint open sets* $U_1, U_2, \cdots$, *with* $\widehat{X} = \bigcup_{i=1}^{\infty} \overline{U_i}$ *such that for each* i, $\widehat{T}_i := \widehat{T}|_{U_i}$ *extends to a* $C^{1+\alpha}$ *diffeomorphism from* $\overline{U}_i$ *to its image, and* $\tau|_{U_i}$ *is constant; we will use the symbol* $\widehat{T}_i$ *to denote the extension as well.*

(b) (Finite images) $\{\widehat{T}U_i : i = 1, 2, \cdots\}$ *is finite, and* $\nu B_\varepsilon(\partial \widehat{T} U_i) = O(\varepsilon)$ $\forall i = 1, 2, \cdots$.

(c) (Expansion) *There exists* $s \in (0,1)$ *such that* $d(\widehat{T}x, \widehat{T}y) \geq s^{-1} d(x,y)$ $\forall x, y \in \overline{U}_i$ $\forall i \geq 1$.

(d) (Topological mixing) $T : X \to X$ *is topological mixing.*

Remark 1.5. *Conditions (b) and (c) in Assumption T correspond to conditions (F) and (U) in [ADSZ]. There is there a third assumption, (A), which is distortion and which is not necessarily guaranteed in our systems. With this precision, we could regard the systems satisfying Assumption T(a)-(c) as higher dimensional "AFU systems".*

Remark 1.6. *We mention that if T has relatively prime return time on almost all points $x \in \widehat{X}$, then Condition (d) is satisfied.*

Also we put some more assumptions on the Banach space $\mathcal{B}$.

A set $U \subset \widehat{X}$ is said to be *almost open* with respect to $\hat{\nu}$ if for $\hat{\nu}$ almost every point $x \in U$, there is a neighborhood $V(x)$ such that $\hat{\nu}(V(x) \setminus U) = 0$.

Assumption B. (d) (Denseness) *The image of the inclusion $\mathcal{B} \hookrightarrow L^1(\hat{\nu})$ is dense in $L^1(\hat{\nu})$.*

(e) (Lower semicontinuity) *For any sequence $\{f_n\} \subset \mathcal{B}$ with $\lim_{n\to\infty} f_n = f$ $\hat{\nu}$-almost everywhere, $|f|_{\mathcal{B}} \leq \liminf_{n\to\infty} |f_n|_{\mathcal{B}}$.*

(f) (Openness) *For any nonnegative function $f \in \mathcal{B}$, the set $\{f > 0\}$ is almost open with respect to $\hat{\nu}$.*

Remark 1.7. *Assumption B(f) means that functions in $\mathcal{B}$ are not far from continuous functions.*

Take a partition ξ of $\widehat{X}$. Consider a family of skew-products of the form

$$\widetilde{T} = \widetilde{T}_S : \widehat{X} \times Y \to \widehat{X} \times Y \ , \ \widetilde{T}_S(x,y) = \big(\widehat{T}x, \ S(\xi(x))(y)\big), \tag{1.5}$$

where $(Y, \mathcal{F}, \rho)$ is a Lebesgue probability space, $\mathrm{Aut}(Y)$ is the collection of its automorphisms, that is, invertible measure-preserving transformations, and $S : \xi \to \mathrm{Aut}(Y)$ is arbitrary.

Consider functions $\widetilde{f} \in L^1(\hat{\nu} \times \rho)$ and define

$$|\widetilde{f}|_{\widetilde{\mathcal{B}}} = \int_Y |\widetilde{f}(\cdot, y)|_{\mathcal{B}} d\rho(y), \qquad \|\widetilde{f}\|_{\widetilde{\mathcal{B}}} = |\widetilde{f}|_{\widetilde{\mathcal{B}}} + \|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)}.$$

Then we let

$$\widetilde{\mathcal{B}} = \{\widetilde{f} \in L^1(\hat{\nu} \times \rho) : |\widetilde{f}|_{\widetilde{\mathcal{B}}} < \infty\}.$$

It is easy to see that with the norm $\|\cdot\|_{\widetilde{\mathcal{B}}}$, $\widetilde{\mathcal{B}}$ is a Banach space.

The transfer operator $\widetilde{\mathscr{P}} = \widetilde{\mathscr{P}}_{\hat{\nu}\times\rho}$ acting on $L^1(\hat{\nu} \times \rho)$ is defined as the dual of the operator $\widetilde{f} \to \widetilde{f} \circ \widetilde{T}$ from $L^\infty(\hat{\nu} \times \rho)$ to itself. Note that if Y is a space consisting of a single point, then we can identify $\widehat{X} \times Y$, $\widetilde{T}$ and $\widetilde{\mathscr{P}}$ with $\widehat{X}$, $\widehat{T}$ and $\widehat{\mathscr{P}}$ respectively.

Theorem B. *Suppose $\widehat{T}$ satisfies Assumption T(a) to (d) and $\mathcal{B}$ satisfies Assumption B(d) to (f), and $\widetilde{\mathscr{P}}$ satisfies the Lasota-Yorke inequality*

$$|(\widetilde{\mathscr{P}}\widetilde{f})|_{\widetilde{\mathcal{B}}} \leq \widetilde{\eta}|\widetilde{f}|_{\widetilde{\mathcal{B}}} + \widetilde{D}\|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)} \tag{1.6}$$

for some $\widetilde{\eta} \in (0,1)$ and $\widetilde{D} > 0$. Then the absolutely continuous invariant measure $\hat{\mu}$ obtained from the Lasota-Yorke inequality (1.2) is ergodic and $e^{it\tau}$ is aperiodic. Therefore Condition (iii) and (iv) in Theorem A follow.

Remark 1.8. *The theorem is for ergodicity and aperiodicity of μ. As we mentioned in Remark 1.5, aperiodicity of μ is equivalent to weak mixing for μ with respect to T. So practically, if we know that μ is mixing or weak mixing for T, then we do not need to use the theorem.*

Remark 1.9. *Same as for (1.3), the inequality (1.6) may be obtained in a similar way as (1.2). This is because any $S(\xi(x))$ is a measure preserving transformation, and therefore $\widehat{\mathscr{P}}$ and $\widetilde{\mathscr{P}}$ have the same potential function. (See the proof of Theorem D).*

Remark 1.10. *It is well known that for $C^{1+\alpha}$, $\alpha > 1$, uniformly expanding maps or uniformly hyperbolic diffeomorphisms, the absolutely continuous invariant measures μ are ergodic if the maps are topological mixing. However, it is not the case if the conditions on $C^{1+\alpha}$ or uniformty of hyperbolicity fails. In [Qu] the author gives an example of C^1 uniformly expanding maps of the unite circle, and in [HPT] the authors provide an example of C^∞ diffeomorphisms, where Lebesgue measures are preserved and topological mixing does not give ergodicity. In the proof of the theorem we in fact give some additional conditions under which topological mixing implies ergodicity (see Lemma 3.2).*

2 Rates of Decay of Correlations

We prove Theorem A in this section. The proof is based on the results of Sarig [Sr] and Gouëzel [Go]. Here we take Gouëzel's version.

Theorem. *Let T_n be bounded operators on a Banach space $\mathcal{B}$ such that $T(z) = I + \sum_{n\geq 1} z^n T_n$ converges in $\mathrm{Hom}(\mathcal{B}, \mathcal{B})$ for every $z \in \mathbb{D}$. Assume that:*

(1) (Renewal equation) *for every $z \in \mathbb{D}$, $T(z) = (I - R(z))^{-1}$, where $R(z) = \sum_{n\geq 1} z^n R_n$, $R_n \in \mathrm{Hom}(\mathcal{B}, \mathcal{B})$ and $\sum_{n\geq 1} \|R_n\| < +\infty$.*

(2) (Spectral gap) 1 *is a simple isolated eigenvalue of $R(1)$.*

(3) (Aperiodicity) *for every $z \in \overline{\mathbb{D}} \setminus \{1\}$, $I - R(z)$ is invertible.*

Let P be the eigenprojection of $R(1)$ at 1. If $\sum_{k>n} \|R_k\| = O(1/n^\beta)$ for some $\beta > 1$ and $PR'(1)P \neq 0$, then for all n,

$$T_n = \frac{1}{\lambda} P + \frac{1}{\lambda^2} \sum_{k=n+1}^{\infty} P_k + E_n, \tag{2.1}$$

where λ is given by $PR'(1)P = \lambda P$, $P_n = \sum_{k>n} PR_kP$ and $E_n \in \mathrm{Hom}(\mathcal{B}, \mathcal{B})$ satisfies $\|E_n\| = O(1/n^\beta)$ if $\beta > 2$, $O(\log n/n^2)$ if $\beta = 2$, and $O(1/n^{2\beta-2})$ if $2 > \beta > 1$.

Proof of Theorem A. Let T_n and R_n be defined by (1.1).

With the assumption $\sum_{k=n+1}^{\infty} \|R_k\|_{\mathcal{B}} = O(n^{-\beta})$, $\beta > 1$, given in Theorem A, Lemma 2.1 to 2.3 imply conditions (1) to (3) respectively

Let $\hat{h} \in \mathcal{B}$ be the eigenfunction of $\widehat{\mathscr{P}} = R(1)$ at 1 with $\int_{\widehat{X}} \hat{h} d\hat{\nu} = 1$, where $\hat{\nu}$ is the normalized Lebesque measure on $\widehat{X}$. That is, $\hat{\nu}(\widehat{X}) = 1$. Also, let $\hat{\mu}$ be the $\widehat{T}$ invariant measure over $\widehat{X}$ given by $d\hat{\mu} = \hat{h} d\hat{\nu}$. It is well known that $\hat{\mu}$ can be extended to a T invariant absolutely continuous probability ergodic measure μ on X ([Kk]).

Since $Pf = \hat{h} \int f d\hat{\nu}$ and $\int \hat{h} d\hat{\nu} = 1$ for any $f \in \mathcal{B}$, we have $\int Pf d\hat{\nu} = \int f d\hat{\nu}$ and also $\int Pf d\nu = \int f d\nu$. Denote $\bar{f} = \int_{\widehat{X}} f d\hat{\nu}$. By the definition, $R_n P f = 1_{\widehat{X}} \mathscr{P}^n(\bar{f}\hat{h} 1_{\{\tau = n\}})$. Integrating over $\widehat{X}$, we get

$$\int R_n P f d\hat{\nu} = \int 1_{\widehat{X}} \mathscr{P}^n(\bar{f}\hat{h}1_{\{\tau=n\}}) d\hat{\nu} = \int 1_{\widehat{X}} \circ T^n \bar{f}\hat{h} 1_{\{\tau=n\}} d\hat{\nu} = \bar{f}\hat{\mu}(\{\tau = n\}).$$

Hence, by the fact $Pf = \big(\int_{\widehat{X}} f d\hat{\nu}\big)\hat{h}$, we immediately get

$$PR_nPf = \Big(\int R_n P f d\hat{\nu}\Big)\hat{h} = \bar{f}\hat{\mu}(\{\tau = n\})\hat{h} = \hat{\mu}(\{\tau = n\})Pf. \tag{2.2}$$

Note that $\hat{\mu}(\{\tau = n\}) = \mu(\{\tau = n\})/\mu(\widehat{X})$. By the Kac formula and the fact that μ is ergodic,

$$PR(1)'Pf = \sum_{n=1}^{\infty} nPR_nPf = \sum_{n=1}^{\infty} n\hat{\mu}(\tau = n)Pf = \frac{1}{\mu(\widehat{X})} Pf.$$

It follows that $\lambda = (\mu(\widehat{X}))^{-1}$. Also, (2.2) gives

$$P_k f = \sum_{i=k+1}^{\infty} PR_i P f = \sum_{i=k+1}^{\infty} \hat{\mu}(\tau = i) P f = \hat{\mu}(\tau > k)\big(\int f d\hat{\nu}\big)\hat{h}.$$

Note that $T_n f = 1_{\widehat{X}} \mathscr{P}^n(f 1_{\widehat{X}})$. So if $\operatorname{supp} f \subset \widehat{X}$, then (2.1) gives that on $\widehat{X}$,

$$\mathscr{P}^n f = \lambda^{-1}\hat{h}\int f d\hat{\nu} + \lambda^{-2} \sum_{k=n+1}^{\infty} \hat{\mu}(\tau > k)\hat{h} \int f d\hat{\nu} + E_n f.$$

Replacing f by fh and using the fact $\int fh d\nu = \mu(f)$, we get that on $\widehat{X}$,

$$\mathscr{P}^n(fh) = \frac{\mu(\widehat{X})}{\nu(\widehat{X})}\hat{h}\mu(f) + \mu(\widehat{X})^2 \sum_{k=n+1}^{\infty} \frac{\mu(\tau > k)}{\mu(\widehat{X})} \frac{\hat{h}}{\nu(\widehat{X})}\mu(f) + E_n(fh).$$

For any $g \in L^\infty(X, \nu)$ with $\operatorname{supp} g \in \widehat{X}$, $\int (fh) \cdot (g \circ T^n) d\nu = \int (\mathscr{P}^n(fh) \cdot g d\nu$; by observing also that $\frac{\mu(\widehat{X})}{\nu(\widehat{X})} \int g\hat{h} d\nu = \int g d\mu$, we finally get:

$$\int f \cdot (g \circ T^n) d\mu = \mu(f)\mu(g) + \sum_{k=n+1}^{\infty} \mu(\tau > k)\mu(f)\mu(g) + \int \frac{1}{h} E_n(fh) g d\mu.$$

The last term is bounded by

$$\Big|\int \frac{1}{h}E_n(fh)gd\mu\Big| = \Big|\int E_n(fh)gd\nu\Big| \le C_a\|E_n\|_{\mathcal{B}}\|h\|_{\mathcal{B}}\|f\|_{\mathcal{B}}\|g\|_\infty,$$

which ends the proof of the theorem. □

Lemma 2.1. *For every $z \in \mathbb{D}$, $T(z) = (I - R(z))^{-1}$.*

Proof. This is given in Proposition 1 in [Sr]. □

Lemma 2.2. *Suppose that Assumption B(a) is satisfied. If $\widehat{\mathscr{P}}$ satisfies Lasota-Yorke inequality (1.2) and $\hat{\mu}$ is ergodic with respect to $\widehat{T} : \widehat{X} \to \widehat{X}$, then 1 is a simple isolated eigenvalue of $\widehat{\mathscr{P}}$.*

Proof. It is well known that with Assumption B(a) on $\mathcal{B}$, Lasota-Yorke inequality (1.2) implies that $\widehat{\mathscr{P}}$ has at most finitely many eigenvalues in the unit circle, and all other points in the spectrum of $\widehat{\mathscr{P}}$ are contained in a circle of radius strictly smaller than 1. Moreover, since $\hat{\mu}$ is ergodic, 1 is a simple isolated eigenvalue. (See e.g. [BG], [Br], and [HH].) □

Lemma 2.3. *Suppose that Assumption B(a) is satisfied. If $\hat{\mu}$ is ergodic, then $I - R(z)$ is invertible on $\mathcal{B}$ for $z \in \overline{\mathbb{D}} \setminus \{1\}$.*

Proof. We follow the proof of Lemma 6.7 in [Go].

By a theorem of Hennion ([He], see also [HH]), the definition of the norm $\|\cdot\|_{\mathcal{B}}$, the inequality (1.3), and the compactness of the inclusion $\mathcal{B} \hookrightarrow L^1(\widehat{X}, \hat{\nu})$ imply that for any $z \in \overline{\mathbb{D}}$ the spectral and essential spectral radius of $R(z)$ on $\mathcal{B}$ is bounded by $|z| \le 1$ and $|z|\hat{\eta} < 1$ respectively. To obtain the invertibility of $I - R(z)$, it is enough to show that 1 is not an eigenvalue of $R(z)$ for $|z| = 1$ with $z \neq 1$. So we fix $0 < t < 2\pi$ and let $z = e^{it}$.

Suppose that $R(z)f = f$ for some nonzero $f \in \mathcal{B}$. Recall that $\hat{\mu}$ is a $\widehat{T}$ invariant measure given by $\hat{\mu}(g) = \hat{\nu}(\hat{h}g) \quad \forall g \in L^1(\widehat{X}, \hat{\nu})$, where $\hat{h}$ satisfies $\widehat{\mathscr{P}}\hat{h} = \hat{h}$. Define the operator $W : L^\infty(\widehat{X}, \hat{\mu}) \to L^\infty(\widehat{X}, \hat{\mu})$ by $Wu = e^{-it\tau}u \circ \widehat{T}$, where $\tau(x)$ is the returning time of x. By the same arguments as in the proof of the Lemma 6.6 in [Go], we get $\|Wf - f\|_2 = 0$, where $\|\cdot\|_2$ denotes the $L^2(\widehat{X}, \hat{\mu})$ norm. So we have $Wf = f$ $\hat{\mu}$-almost everywhere with respect to the measure $\hat{\mu}$. That is, $e^{-it\tau}f \circ \widehat{T} = f$ almost everywhere. By the aperiodicity condition (iv) we conclude that $t = 0$ and f is a constant $\hat{\mu}$-almost everywhere which is a contradiction. □

3 Aperiodicity

The proof of Theorem B is based on a result in [ADSZ]. We briefly mention the terminology used there.

A *fibered system* is a quintuple $(X,\mathcal{A},\nu,T,\xi)$, where $(X,\mathcal{A},\nu,T)$ is a non-singular transformation on a σ-finite measure space and $\xi\subset\mathcal{A}$ is a finite or countable partition (mod ν) such that:

(1) $\xi_\infty=\bigvee_{i=0}^\infty T^{-i}\xi$ generates $\mathcal{A}$;

(2) every $A\in\xi$ has positive measure;

(3) for every $A\in\xi$, $T|_A:A\to TA$ is bimeasurable invertible with nonsingular inverse.

The transformation given in (1.5) is called the *skew products over* ξ. Consider the corresponding transfer operator of $\widetilde{\mathscr{P}}=\widetilde{\mathscr{P}}_{\nu\times\rho}$. A fibred system $(X,\mathcal{A},\nu,T,\xi)$ with ν finite is called *skew-product rigid* if for every invariant function $\widetilde{h}(x,y)$ of $\widetilde{\mathscr{P}}$ of an arbitrary skew product $\widetilde{T}_S$, the set $\{\widetilde{h}(\cdot,y)>0\}$ is almost open $(\operatorname{mod}\nu)$ for almost every $y\in Y$. In [ADSZ], a set U being almost open $(\operatorname{mod}\nu)$ means that for ν almost every $x\in U$, there is a positive integer n such that $\nu(\xi_n(x)\setminus U)=0$. Since the partition ξ we are interested in satisfies $\nu(\partial A)=0$ for any $A\in\xi_n$ and $\widehat{T}$ is piecewise smooth, the fact that ξ_∞ generates $\mathcal{A}$ implies that the definition given there is the same as we defined for Assumption B(f).

A set that can be expressed in the form $\widehat{T}^n\xi_n(x)$, $n\geq 1$ and $x\in\widehat{X}$, is called an *image set*. A cylinder C of length n_0 is called a *cylinder of full returns*, if for almost all $x\in C$ there exist $n_k\nearrow\infty$ such that $\widehat{T}^{n_k}\xi_{n_k+n_0}(x)=C$. In this case we say that $\widehat{T}^{n_0}(C)$ is a *recurrent image set*.

Our proof of Theorem B is based on a result given in Theorem 2 in [ADSZ]:

Theorem. *Let $(X,\mathcal{A},\mu,T,\xi)$ be a skew-product rigid measure preserving fibered system whose image sets are almost open. Let G be a locally compact Abelian polish group. If $\gamma\circ\phi=\lambda f/f\circ T$ holds almost everywhere, where $\phi:X\to\mathbb{G}$, ξ measurable, $\gamma\in\widehat{\mathbb{G}}$, $\lambda\in\mathbb{S}$, then f is constant on every recurrent image set.*

In the proof of Theorem B and the lemmas below we will work exclusively on the induced space $\widehat{X}$ and with measures $\hat{\nu}$ and $\hat{\mu}$ and density $\hat{h}$. So we will drop the hat on these notations.

Proof of Theorem B. Recall that μ is an $\widehat{T}$ invariant measure with density h, where h is the fixed point of $\widehat{\mathscr{P}}$ in $\mathcal{B}$. By Lemma 3.2 we know that μ is ergodic. So we only need to prove that $e^{it\tau}$ is aperiodic.

Denote by $\mathcal{A}$ the Borel σ-algebra inherited from $\mathbb{R}^m$. Take a countable partition ξ of $\widehat{X}$ into $\{U_i\}$ or finer. We also require that each $A\in\xi$ is almost open, and $\nu B_\varepsilon(\partial\widehat{T}\xi)=O(\varepsilon)$, where $\partial\widehat{T}\xi=\cup_{A\in\xi}\partial(\widehat{T}A)$. The latter is possible because we can take smooth surfaces as the boundary of the elements of ξ, in addition to Assumption T(b). Since $\widehat{T}$ is uniformly expanding by Assumption T(c), we know that each element of $\xi_\infty=\bigvee_{i=0}^\infty\widehat{T}^{-i}\xi$ contains at most one point. So ξ_∞ generates $\mathcal{A}$. We may regard that each $A\in\xi$ has positive measure, otherwise we can use $\widehat{X}\setminus A$ to replace $\widehat{X}$. Also, for every $A \in \xi$, $\widehat{T}|_A : A \to \widehat{T}A$ is a diffeomorphism, and therefore $\widehat{T}|_A$ is bimeasurable invertible with nonsingular inverse. So the quintuple $(\widehat{X}, \mathcal{A}, \mu, \widehat{T}, \xi)$ is a measure preserving fibered system.

The construction of ξ implies that $\mu(\partial\xi) = \nu(\partial\xi) = 0$. Hence, $\mu(\partial\xi_n) = \nu(\partial\xi_n) = 0$ for any $n \geq 1$. Note that the intersection of finite number of almost open sets is still almost open. Differentiability of $\widehat{T}$ on each U_i implies that all elements $\xi_n(x)$ of ξ_n are almost open, and therefore all image sets $\widehat{T}^n\xi_n(x)$ are almost open with respect to μ.

To get skew product rigidity, let us consider the skew product $\widetilde{T}_S$ defined in (1.5) for any $(Y, \mathcal{F}, \rho)$. Let $\widetilde{\mathscr{P}} = \widetilde{\mathscr{P}}_{\nu\times\rho}$ be the transfer operator and $\widetilde{h}$ an invariant function, that is, $\widetilde{\mathscr{P}}\widetilde{h} = \widetilde{h}$. By Proposition 3.3 below we know that $\widetilde{h} \in \widetilde{B}$. Hence, for ρ-almost every $y \in Y$, $\widetilde{h}(\cdot, y) \in \mathcal{B}$. By Assumption B(f), $\{\widetilde{h}(\cdot, y) > 0\}$ is almost open mod ν. This gives the skew product rigidity.

So far we have verified all conditions in the theorem of [ADSZ] stated above. Applying the theorem to the equation $e^{it\tau} = f/f\circ\widehat{T}$ almost everywhere, where $f : \widehat{X} \to \mathbb{S}$ is a measurable function, we get that f is constant on every recurrent image sets J.

Now we prove aperiodicity, by following similar arguments in [Go]. Assume the equation $e^{it\tau} = f/f\circ\widehat{T}$ holds almost everywhere for some real number t and a measurable function $f : \widehat{X} \to \mathbb{S}$. By Lemma 3.1 below we get that $\widehat{X}$ contains a recurrent image set J with $\mu(J) > 0$. By the theorem above, we know that f is constant, say c, almost everywhere on J. By the absolute continuity of μ and the fact that $\{h > 0\}$ is ν-almost open, we can find an open set $J' \subset J$ of positive μ-measure. By Assumption T(d), T is topological mixing. Therefore for all sufficiently large n, we have $T^{-n}J' \cap J' \neq \emptyset$. Since the intersection is open[†] , we get that $\mu(T^{-n}J' \cap J') > 0$. So for any typical point x in $T^{-n}J' \cap J'$, there is $k > 0$ such that $T^n x = \widehat{T}^k x$, and $n = \sum_{i=0}^{k-1} \tau(\widehat{T}^i x)$. Since $e^{it\tau} = f/f\circ\widehat{T}$ along the orbit of x, we have

$$e^{int} = e^{it\sum_0^{k-1}\tau(\widehat{T}^i x)} = \frac{f(x)}{f(\widehat{T}x)}\frac{f(\widehat{T}x)}{f(\widehat{T}^2x)}\cdots\frac{f(\widehat{T}^{k-1}x)}{f(\widehat{T}^kx)} = \frac{f(x)}{f(\widehat{T}^kx)} = \frac{c}{c} = 1.$$

Since this is true for all large n, by replacing n by $n+1$ we get that $e^{it} = 1$. It follows that $t = 0$ and $f = f\circ\widehat{T}$ almost everywhere which implies that f must be a constant almost everywhere since μ is ergodic. □

To prove Lemma 3.1, we need a result from Lemma 2 in Section 4 in [ADSZ]. We state it as the next lemma. The setting for the lemma is a conservative fibered system. So it can be applied directly to our case.

Lemma. *A cylinder $C \in \xi_{n_0}$ is a cylinder of full returns if and only if there exists a set K of positive measure such that for almost every $x \in K$, there are $n_i \to \infty$ with $\widehat{T}^{n_i}\xi_{n_i+n_0}(x) = C$.*

[†]Strictly speaking that intersection contains open sets since T and all its powers, although not continuous, are local diffeomorphisms, on each domain where they are injective.

Lemma 3.1. *There is a recurrent image set J contained in $\widehat{X}$ with $\mu J > 0$.*

Proof. Recall that s is given in Assumption T(c). Take $C_\xi > 0$ such that $\operatorname{diam} D \le C_\xi$ for all $D \in \xi$. Set

$$A'_{k,n_0} = \{x \in \widehat{X} : x \notin B_{C_\xi s^{k+n_0}}(\partial \widehat{T}\xi)\},$$
$$A_{n,n_0} = \bigcap_{k=0}^{n-1} \widehat{T}^{n-k} A'_{k,n_0}.$$

By the construction of ξ, there is $C' > 0$ such that $\nu A'_{k,n_0} \ge 1 - C' C_\xi s^{k+n_0}$. By Assumption B(b), $\|h\|_\infty < \infty$. So if we take $C = C' C_\xi \|h\|_\infty/(1-s)$, then $\mu A_{k,n_0} \ge 1 - C' C_\xi \|h\|_\infty s^{k+n_0} = 1 - C(1-s)s^{k+n_0}$. Since μ is an invariant measure, $\mu A_{n,n_0} \ge 1 - C(1-s)\sum_{i=0}^{n-1} s^{i+n_0} \ge 1 - Cs^{n_0}$. If we choose n_0 large enough, then $\mu A_{n,n_0}$ is bounded below by a positive number for all $n > 0$, and the bound can be chosen arbitrarily close to 1 by taking n_0 sufficiently large.

Note that ξ_n is a partition with at most countably many elements. For each $n_0 > 0$, let B'_{n_0} be the union of finite elements of ξ_{n_0} such that $\mu B'_{n_0} > 1 - Cs^{n_0}/2$. Then set $B_{n,n_0} = B'_{n_0} \cap \widehat{T}^{-n} B'_{n_0}$. Clearly, $\mu B_{n,n_0} \ge 1 - Cs^{n_0}$. Denote $C_{n,n_0} = A_{n,n_0} \cap B_{n,n_0}$. We have $\mu C_{n,n_0} \ge 1 - 2Cs^{n_0}$. Hence, $\sum_{n=0}^{\infty} \mu C_{n,n_0} = \infty$ for all large n_0.

A generalized Borel-Cantelli Lemma by Kochen and Stone ([KS], see also [Ya]) gives that for any given $n_0 > 0$, the set of points that belong to infinitely many C_{n,n_0} has the measure bounded below by

$$\limsup_{n\to\infty} \frac{\sum_{1\le i<k\le n} \mu C_{i,n_0} \mu C_{k,n_0}}{\sum_{1\le i<k\le n} \mu(C_{i,n_0} \cap C_{k,n_0})}.$$

Note that if $n_0 \to \infty$, then both $\mu C_{i,n_0}$ and $\mu C_{k,n_0}$ approach to 1. Hence the upper limit goes to 1 as $n_0 \to 1$. Denote

$$\Gamma_{n_0} = \{x \in \widehat{X} : x \in C_{n,n_0} \text{ infinitely often}\}.$$

The above arguments gives $\mu \Gamma_{n_0} \to 1$ as $n_0 \to \infty$.

Note that for a one to one map T, $T(A \cap T^{-1}B) = B$ if and only if $B \subset TA$. Since $\xi_n(x) = \xi(x) \cap \widehat{T}^{-1}(\xi_{n-1}(\widehat{T}x))$, and $\widehat{T}$ is a local diffeomorphism, we know that $\widehat{T}\xi_n(x) = \xi_{n-1}(\widehat{T}x)$ if and only if $\xi_{n-1}(\widehat{T}x) \subset \widehat{T}\xi(x)$. Inductively, $\widehat{T}^n \xi_{n+n_0}(x) = \xi_{n_0}(\widehat{T}^n x)$ if and only if $\xi_{n-i+n_0}(\widehat{T}^i x) \subset \widehat{T}\xi(\widehat{T}^{i-1}x)$ for $i = 1, \cdots, n$. If $x \in A_{n,n_0}$ for some $n, n_0 > 0$, then $\widehat{T}^{n-i} x \notin B_{C_\xi s^{i+n_0}}(\partial \widehat{T}\xi)$ for all $i = 1, \cdots, n$. Since the diameter of each member of ξ is less then C_ξ, by Assumption T(c), $\operatorname{diam} \xi_n(x) \le C_\xi s^n$ for any $x \in \widehat{X}$ and $n \ge 0$. We get $\xi_{n-i+n_0}(\widehat{T}^i x) \subset \widehat{T}\xi(\widehat{T}^{i-1}x)$ and therefore $\widehat{T}^n \xi_{n+n_0}(x) = \xi_{n_0}(\widehat{T}^n x)$. Consequently, if $x \in \Gamma_{n_0}$, then $x \in C_{n_i,n_0} = A_{n_i,n_0} \cap B_{n_i,n_0}$ for infinitely many n_i. Hence, $\widehat{T}^{n_i} \xi_{n_i+n_0}(x) = \xi_{n_0}(\widehat{T}^{n_i} x)$ and $\widehat{T}^{n_i} x \in B_{n_0}$ for infinitely many n_i, Take $n_0 > 0$ such that $\mu\Gamma_{n_0} > 0$. Since B_{n_0} consists of only finitely many elements in ξ_{n_0}, we know that there is an element $C \in \xi_{n_0}$ with $C \subset B_{n_0}$ such that

$$\mu\{x : \widehat{T}^n \xi_{n+n_0}(x) = \xi_{n_0}(\widehat{T}^n x) = C \text{ infinitely often}\} > 0. \tag{3.1}$$

By the above lemma from [ADSZ], C is a cylinder of full returns. Hence, $J = \widehat{T}^{n_0} C$ is a recurrent image set. Since μ is an invariant measure, (3.1) implies $\mu C > 0$ and therefore $\mu J > 0$. □

Lemma 3.2. *Suppose T and $\mathcal{B}$ satisfies Assumption T(d) and B(f) respectively. Then there is only one absolutely continuous invariant measure μ which is ergodic.*

Proof. Suppose μ has two ergodic components μ_1 and μ_2 whose density functions are h_1 and h_2 respectively. Hence, $\nu(\{h_1 > 0\} \cap \{h_2 > 0\}) = 0$. Since $h_1, h_2 \in \mathcal{B}$, the sets $\{h_1 > 0\}$ and $\{h_2 > 0\}$ are almost open. We can take open sets U_1 and U_2 such that $\nu(U_1 \setminus \{h_1 > 0\}) = 0$ and $\nu(U_2 \setminus \{h_1 > 0\}) = 0$. Since T is topological mixing, there is $n > 0$ such that $T^{-n}U_1 \cap U_2 \neq \emptyset$. Hence, $\nu(T^{-n}U_1 \cap U_2) > 0$ and therefore $\nu(U_1 \cap T^n U_2) > 0$. It follows that there is $k > 0$ such that $\nu(U_1 \cap \widehat{T}^k U_2) > 0$. Since $\widehat{\mathscr{P}} h_2 = h_2$, $h_2(x) > 0$ implies $h_2(\widehat{T}^k x) > 0$. Hence $\nu(\widehat{T}^k U_2 \setminus \{h_2 > 0\}) = 0$. Therefore, $\nu(\{h_1 > 0\} \cap \{h_2 > 0\}) \geq \nu(U_1 \cap \widehat{T}^k U_2) > 0$, which is a contradiction. □

The next proposition is the key step for Lemma **??**. The result was proved for Gibbs-Markov maps in [AD]. We show that it holds in more general cases.

Proposition 3.3. *Suppose that $\mathcal{B}$ satisfies Assumption B(d) and (e), and $\widetilde{\mathscr{P}}$ satisfies Lasota-Yorke inequality (1.6). Then any $L^1(\nu \times \rho)$ function $\widetilde{h}$ on $\widehat{X} \times Y$ that satisfies $\widetilde{\mathscr{P}}_{\nu\times\rho} \widetilde{h} = \widetilde{h}$ belongs to $\widetilde{\mathcal{B}}$.*

Proof. By Assumption B(d), $\mathcal{B}$ is dense in $L^1(\widehat{X}, \nu)$. It is easy to see that $\widetilde{\mathcal{B}}$ is dense in $L^1(\widehat{X} \times Y, \nu \times \rho)$. Hence, for any $\varepsilon > 0$ we can find a nonnegative function $\widetilde{f}_\varepsilon \in \widetilde{\mathcal{B}}$ such that $\|\widetilde{f}_\varepsilon - \widetilde{h}\|_{L^1(\nu\times\rho)} < \varepsilon$. By the stochastic ergodic theorem of Krengel ([Kr]), there exists a nonnegative function $\widetilde{h}_\varepsilon \in L^1(\widehat{X} \times Y, \nu \times \rho)$ and a subsequence $\{n_k\}$ such that

$$\lim_{k\to\infty} \frac{1}{n_k} \sum_{\ell=0}^{n_k-1} \widetilde{\mathscr{P}}^\ell \widetilde{f}_\varepsilon = \widetilde{h}_\varepsilon \qquad \nu \times \rho\text{-a.e.} \tag{3.2}$$

and $\widetilde{\mathscr{P}} \widetilde{h}_\varepsilon = \widetilde{h}_\varepsilon$.

Note that Lasota-Yorke inequality (1.6) implies that for any $\widetilde{f} \in \widetilde{\mathcal{B}}$, $\ell \geq 1$,

$$|\widetilde{\mathscr{P}}^\ell \widetilde{f}|_{\widetilde{\mathcal{B}}} \leq \widetilde{\eta}^\ell |\widetilde{f}|_{\widetilde{\mathcal{B}}} + \widetilde{D}^* \|\widetilde{f}\|_{L^1(\nu\times\rho)} \leq \widetilde{D}_2 \|\widetilde{f}\|_{\widetilde{\mathcal{B}}}, \tag{3.3}$$

where $\widetilde{D}^* = \widetilde{D}\widetilde{\eta}/(1 - \widetilde{\eta}) \geq \widetilde{D}(\widetilde{\eta} + \cdots + \widetilde{\eta}^{\ell-1})$ and $\widetilde{D}_2 = 1 + \widetilde{D}^*$. Denote $\psi_k = \dfrac{1}{n_k} \displaystyle\sum_{\ell=0}^{n_k-1} \widetilde{\mathscr{P}}^\ell f_\varepsilon$. By (3.3) $\psi_k \leq \widetilde{D}_2 \|\widetilde{f}\|_{\widetilde{\mathcal{B}}}$. (3.2) implies that $\liminf\limits_{k\to\infty} \psi_k(x, y) =$

$\widetilde{h}_\varepsilon(x,y)$ for ν-a.e. $x \in \widehat{X}$, ρ-a.e. $y \in Y$. Hence, by Assumption B(e) and Fatou's lemma we obtain

$$\begin{aligned} |\widetilde{h}_\varepsilon|_{\widetilde{\mathcal{B}}} &= \int_Y |\lim_{k\to\infty} \psi_k(\cdot,y)|_{\mathcal{B}} d\rho(y) \le \int_Y \liminf_{k\to\infty} |\psi_k(\cdot,y)|_{\mathcal{B}} d\rho(y) \\ &\le \liminf_{k\to\infty} \int_Y |\psi_k(\cdot,y)|_{\mathcal{B}} d\rho(y) = \liminf_{k\to\infty} |\psi_k|_{\widetilde{\mathcal{B}}} \le \widetilde{D}_2 ||\widetilde{f}_\varepsilon||_{\widetilde{\mathcal{B}}} \end{aligned} \tag{3.4}$$

It means $\widetilde{h}_\varepsilon \in \widetilde{\mathcal{B}}$.

By Fatou's Lemma and the fact that $\widetilde{\mathscr{P}}$ is a contraction on $L^1(\widehat{X} \times Y, \nu \times \rho)$, it follows immediately that (3.2) and the fact $\widetilde{\mathscr{P}}\widetilde{h} = \widetilde{h}$ imply

$$||\widetilde{h} - \widetilde{h}_\varepsilon||_{L^1(\nu\times\rho)} \le \liminf_{k\to\infty} \frac{1}{n_k} \sum_{l=0}^{n_k-1} ||\widetilde{\mathscr{P}}^\ell(\widetilde{h} - \widetilde{f}_\varepsilon)||_{L^1(\nu\times\rho)} \le ||\widetilde{h} - \widetilde{f}_\varepsilon||_{L^1(\nu\times\rho)} \le \varepsilon.$$

By the first inequality of (3.3) we know that for any $n \ge 1$,

$$||\widetilde{h}_\varepsilon||_{\widetilde{\mathcal{B}}} = ||\widetilde{\mathscr{P}}^n \widetilde{h}_\varepsilon||_{\widetilde{\mathcal{B}}} \le \widetilde{\eta}^n ||\widetilde{h}_\varepsilon||_{\widetilde{\mathcal{B}}} + \widetilde{D}^* ||\widetilde{h}_\varepsilon||_{L^1(\nu\times\rho)}.$$

Sending n to infinity we get $||\widetilde{h}_\varepsilon||_{\widetilde{\mathcal{B}}} \le \widetilde{D}^*||\widetilde{h}_\varepsilon||_{L^1(\nu\times\rho)} \le \widetilde{D}^*(||\widetilde{h}||_{L^1(\nu\times\rho} + \varepsilon)$. Replace now ε with a decreasing sequence $c_n \to 0$ as $n \to \infty$. Since $\widetilde{h}_{c_n}$ converges in $L^1(\nu \times \rho)$ to $\widetilde{h}$, there is a subsequence n_i such that $\lim_{i\to\infty} \widetilde{h}_{c_{n_i}} = \widetilde{h}$, $\nu\times\rho$-a.e.. Then by the same arguments as for (3.4), we see

$$|\widetilde{h} - \widetilde{h}_{c_n}|_{\widetilde{\mathcal{B}}} \le \liminf_{i\to\infty} |\widetilde{h}_{c_{n_i}} - \widetilde{h}_{c_n}|_{\widetilde{\mathcal{B}}} \le 2 \sup_{0\le\varepsilon\le1} ||\widetilde{h}_\varepsilon||_{\widetilde{\mathcal{B}}} \le 2\widetilde{D}_1(||\widetilde{h}||_{L^1(\nu\times\rho)} + 1).$$

So we get $\widetilde{h} - \widetilde{h}_{c_n} \in \widetilde{\mathcal{B}}$.

Therefore $h = (h - h_{c_n}) + h_{c_n} \in \widetilde{\mathcal{B}}$ and this completes the proof. □

Part II: Applications to non-Markov Maps

We now apply our results to piecewise expanding non-Markov maps with an indifferent fixed point. We use different Banach spaces for maps in one and higher dimensional spaces.

4 Systems on the interval

The object of this section is twofold: to give an example of a Banach space which fits our assumptions, and to provide the lower bound for the decay of correlations. Moreover, we will use a large space of observables, bounded variation function instead of Hölder continuous functions.

Let $X = I = [0, 1]$ and ν be the Lebesgue measure on X.

Recall that for a map $T : X \to X$ and a subset $\widehat{X} \subset X$, the corresponding first return map is denoted by $\widehat{T} : \widehat{X} \to \widehat{X}$; $\hat{\nu}$ will denote again the normalized Lebesgue measure over $\widehat{X}$.

Assume that $T : X \to X$ is a map satisfying the following assumptions.

Assumption T′. (a) (Piecewise smoothness) *There are points* $0 = a_0 < a_1 < \cdots < a_K = 1$ *such that for each* j, $T_j = T|_{I_j}$ *is a* C^2 *diffeomorphism on its image, where* $I_j = (a_{j-1}, a_j)$.

(b) (Fixed point) $T(0) = 0$.

(c) (Expansion) *There exists* $z \in I_1$ *such that* $T(z) \in I_1$ *and* $\Delta := \inf_{x \in \widehat{X}} |T'(x)| > 2$ *for any* $x \in \widehat{X}$, *where* $\widehat{X} = [z, 1]$.

(d) (Distortion) $\Gamma := \sup_{x \in [z,1]} |\widehat{T}''(x)|/|\widehat{T}'(x)|^2 \leq \infty$.

(e) (Topological mixing) $T : I \to I$ *is topological mixing.*

Denote $J = [0, z]$ and $\widehat{X} = \widehat{X}_J = X \setminus J$. $I_0 = TJ \setminus J \subset I_1$. We also denote the first return map $\widehat{T} = \widehat{T}_J$ by $\widehat{T}_{ij}$ if $\widehat{T} = T_1^i T_j$. Further, we denote $I_{01} = I_1 \setminus J$, $I_{0j} = I_j \setminus T_j^{-1} J$ if $j > 1$, and $I_{ij} = \widehat{T}_{i,j}^{-1} I_0$ for $i > 0$. Hence, $\{I_{ij} : i = 0, 1, 2, \cdots\}$ form a partition of $I_j = (a_j, b_j)$ for $j = 2, \cdots, K$. Also, we denote $\bar{I}_{ij} = [a_{ij}, b_{ij}]$ for any $i = 0, 1, 2, \cdots$ and $j = 1, \cdots, K$.

Recall that the variation of a real or complex valued function f on $[a, b]$ is defined by

$$V_a^b(f) = V_{[a,b]}(f) = \sup_{\xi \in \Xi} \sum_{i=1}^{n} |f(x^{(\ell)}) - f(x^{(\ell-1)})|,$$

where ξ is a finite partition of $[a, b]$ into $a = x^{(0)} < x^{(1)} < \cdots < x^{(n)} = b$ and Ξ is the set of all such partitions. A function $f \in L^1([a, b], \nu)$, where ν denotes the Lebesgue measure, is of bounded variation if $V_{[a,b]}(f) = \inf_g V_{[a,b]}(g) < \infty$, where the infimum is taken over all the function $g = f$ ν-a.e.. Let $\mathcal{B}$ be the set of functions $f \in L^1(\widehat{X}, \hat{\nu})$, $f : \widehat{X} \to \mathbb{R}$ with $V(f) := V_{\widehat{X}}(f) < \infty$. For $f \in \mathcal{B}$, denote by $|f|_{\mathcal{B}} = V(f)$, the total variation of f. Then we define $\|f\|_{\mathcal{B}} = \|f\|_1 + |f|_{\mathcal{B}}$, where the L^1 norm is intended with respect to $\hat{\nu}$. It is well known that $\|\cdot\|_{\mathcal{B}}$ is a norm, and with the norm, $\mathcal{B}$ becomes a Banach space.

To obtain the decay rates, we also assume that there are constants $0 < \gamma < 1$, $\gamma' > \gamma$ and $C > 0$ such that in a neighborhood of the indifferent fixed point $p = 0$,

$$\begin{aligned}
T(x) &= x + Cx^{1+\gamma} + O(x^{1+\gamma'}), \\
T'(x) &= 1 + C(1+\gamma)x^{\gamma} + O(x^{\gamma'}), \qquad (4.1)\\
T''(x) &= C\gamma(1+\gamma)x^{\gamma-1} + O(x^{\gamma'-1}).
\end{aligned}$$

For any sequences of numbers $\{a_n\}$ and $\{b_n\}$, we will denote $a_n \approx b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$, and $a_n \sim b_n$ if $c_1 b_n \le a_n \le c_2 b_n$ for some constants $c_2 \ge c_1 > 0$.

Denote

$$d_{ij} = \sup\{|\widehat{T}_{ij}'(x)|^{-1} : x \in I_{ij}\}, \quad d_n = \max\{d_{n,j} : 2 \le j \le K\}. \tag{4.2}$$

Theorem C. *Let $\widehat{X}$, $\widehat{T}$ and $\mathcal{B}$ are defined as above. Suppose T satisfies Assumption T' (a) to (e). Then Assumption B(a) to (f) and conditions (i) to (iv) in Theorem A are satisfied and $\|R_n\| \le O(d_n)$. Hence, if $d_n = O(n^{-\beta})$ for some $\beta > 1$, then there exists $C > 0$ such that for any functions $f \in \mathcal{B}$, $g \in L^\infty(X,\nu)$ with $\operatorname{supp} f$, $\operatorname{supp} g \subset \widehat{X}$, (1.4) holds.*

In particular, if T satisfies (4.1) near 0, then $\sum_{k=n+1}^{\infty} \mu(\tau > k)$ has the order $n^{-(1/\gamma-1)}$ and d_n has the order $O(n^{-(1/\gamma+1)})$. So we have

$$\operatorname{Cov}(f, g\circ T^n) \approx \sum_{k=n+1}^{\infty} \mu(\tau > k) \int f d\mu \int g d\mu \ \sim 1/n^{1/\gamma-1}.$$

It is well known that if the map T allows a Markov partition, then the rate of decay of correlations is of order $O(n^{-(1/\gamma-1)})$ (see e.g. [Hu], [Sr],[LSV], [PY]). For non-Markov case, the upper bound estimate is given in [Yo2] and [Sr] for observables with some Hölder property. With the methods in [Sr], the lower bound could be obtained by estimating the lower bound of the decay rate of the tower. Since our methods do not require Markov properties, the decay rates can be obtained directly from the size of the sets $\{\tau \ge k\}$. Moreover our observables are functions with bounded variations, which are more general than Hölder functions.

Proof of Theorem C. By Lemma 4.1 below, $\mathcal{B}$ satisfies Assumption B(a) to (f). By Lemma 4.2, we know that condition (i) and (ii) of Theorem A are satisfied. Notice that all requirements of Assumption T are satisfied, since part (a), (c) and (d) follow from Assumption T$'$(a), (c) and (e) directly, and part (b) follows from the definition of $\widehat{T}$. Moreover Lemma 4.2 (iii) gives (1.6). Hence Theorem B can be applied and therefore conditions (iii) and (iv) of Theorem A are satisfied.

The estimate $\|R_n\| = O(d_n)$ follows from Lemma 4.3. Therefore (1.4) is given by Theorem A.

Suppose that T also satisfies (4.1). Denote by $z_n \in I_1$ the point such that $T^n(z_n) = z$. It is well known that $z_n \approx (\gamma n)^{-1/\gamma}$ (see e.g. Lemma 3.1 in [HV]), and then we can obtain $(T_1^{-n})'(x) = O(n^{-1/\gamma-1})$. It follows that $d_n = O(n^{-1/\gamma+1})$. Since the density function h is bounded on $\widehat{X}$, $\mu(\tau > k) \le C_1\nu(\tau > k) \le C_2 z_k$ for some $C_1, C_2 > 0$. Hence $\sum_{k=n+1}^{\infty} \mu(\tau > k) = O(n^{-1/\gamma-1})$. □

Lemma 4.1. *$\mathcal{B}$ is a Banach space satisfying Assumption B(a) to (f) with $C_a = C_b = 1$.*

Proof. These are standard facts, see for instance [Zm]. $\square$

Lemma 4.2. *There exist constants $\eta \in (0,1)$ and $D, \bar{D} > 0$ satisfying*

(i) *for any $f \in \mathcal{B}$, $|\widehat{\mathscr{P}} f|_{\mathcal{B}} \le \eta |f|_{\mathcal{B}} + D\|f\|_{L^1(\hat\nu)}$;*

(ii) *for any $f \in \mathcal{B}$, $\|R(z)f\|_{\mathcal{B}} \le |z|\big(\eta\|f\|_{\mathcal{B}} + \bar{D}\|f\|_{L^1(\hat\nu)}\big)$; and*

(iii) *for any $f \in \widetilde{\mathcal{B}}$, $\|\widetilde{\mathscr{P}}\widetilde{f}\|_{\widetilde{\mathcal{B}}} \le \eta\|\widetilde{f}\|_{\widetilde{\mathcal{B}}} + D\|\widetilde{f}\|_{L^1(\hat\nu\times\rho)}$.*

Proof. (i) Denote $x_{ij} = \widehat{T}_{ij}^{-1}(x)$, and $\widehat{g}(x_{ij}) = |\widehat{T}'_{ij}(x_{ij})|^{-1}$. By the definition, we have

$$\widehat{\mathscr{P}} f(x) = \sum_{j=1}^{K}\sum_{i=0}^{\infty} f(\widehat{T}_{ij}^{-1}x)\widehat{g}(\widehat{T}_{ij}^{-1}x)1_{\widehat{T}I_{ij}}(x).$$

Take a partition ξ of $\widehat{T}I_{ij}$ into $\widehat{T}_{ij}a_{ij} = x^{(0)} < x^{(1)} < \cdots < x^{(k_{ij})} = \widehat{T}_{ij}b_{ij}$, where we assume $\widehat{T}_{ij}a_{ij} < \widehat{T}_{ij}b_{ij}$ without loss of generality. Whenever $\widehat{T}I_{ij}$ may intersect more than one intervals $I_k = (a_k, b_k)$ in the case $i = 0$, then we put the endpoints a_k and b_k into the partition. Denote $x_{ij}^{(\ell)} = \widehat{T}_{ij}^{-1}x^{(\ell)}$. We have

$$\begin{aligned} &\sum_{\ell=1}^{k_{ij}} \big|f(x_{ij}^{(\ell)})\widehat{g}(x_{ij}^{(\ell)}) - f(x_{ij}^{(\ell-1)})\widehat{g}(x_{ij}^{(\ell-1)})\big| \\ \le &\sum_{\ell=1}^{k_{ij}} \widehat{g}(x_{ij}^{(\ell)})\big|f(x_{ij}^{(\ell)}) - f(x_{ij}^{(\ell-1)})\big| + \sum_{\ell=1}^{k_{ij}}\big|f(x_{ij}^{(\ell-1)})\big|\big|\widehat{g}(x_{ij}^{(\ell)}) - \widehat{g}(x_{ij}^{(\ell-1)})\big|. \end{aligned} \tag{4.3}$$

By (4.2), $\widehat{g}(x_{ij}^{(\ell)}) \le d_{ij}$. By definition, $\sum_{\ell=1}^{k_{ij}}\big|f(x_{ij}^{(\ell-1)}) - f(x_{ij}^{(\ell)})\big| \le V_{I_{ij}}(f)$. Also, by the mean value theorem and Assumption T$'$(d),

$$\frac{|g(\widehat{x}_{ij}^{(\ell)}) - \widehat{g}(x_{ij}^{(\ell-1)})|}{x_{ij}^{(\ell)} - x_{ij}^{(\ell-1)}} \le |\widehat{g}'(c_{ij}^{(\ell)})| = |\widehat{T}''(c_{ij}^{(\ell)})|/|\widehat{T}'(c_{ij}^{(\ell)})|^2 \le \Gamma,$$

where $c_{ij}^{(\ell)} \in [x_{ij}^{(\ell-1)}, x_{ij}^{(\ell)}]$. Using the fact that

$$\lim_{\max\{|x_{ij}^{(\ell)} - x_{ij}^{(\ell-1)}|\}\to 0} \sum_{\ell=1}^{k_{ij}} \big|f(x_{ij}^{(\ell-1)})\big|(x_{ij}^{(\ell)} - x_{ij}^{(\ell-1)}) = \int_{a_{ij}}^{b_{ij}} |f| d\hat\nu,$$

we get from (4.3) that

$$V_{\widehat{T}I_{ij}}((f\cdot\widehat{g})\circ\widehat{T}_{ij}^{-1}) \le d_{ij}V_{I_{ij}}(f) + \Gamma\int_{I_{ij}} |f| d\hat\nu. \tag{4.4}$$

Denote $c = \min\{\nu(\widehat{T}I_{ij}) : i = 1, 2, \cdots, 1 \le j \le K\}$, where $c > 0$ because there are only finite number of images $\widehat{T}I_{ij}$. It can be shown that (see e.g. [Br])

$$V(\widehat{\mathscr{P}} f) \le 2\sum_{j=1}^{K}\sum_{i=0}^{\infty} V_{\widehat{T}I_{ij}}((f\cdot\widehat{g})\circ\widehat{T}_{ij}^{-1}) + 2c^{-1}\|f\|_1.$$

By Assumption T′(c), $d_{ij} \le \Delta^{-1}$ for all $i = 1, 2, \cdots$ and $j = 1, \cdots, K$. Hence

$$|\widehat{\mathscr{P}} f|_{\mathcal{B}} = V(\widehat{\mathscr{P}} f) \le 2\Delta^{-1} V(f) + 2\Gamma \int |f| d\hat{\nu} + 2c^{-1} \|f\|_1 = \eta |f|_{\mathcal{B}} + D\|f\|_1,$$

where $\eta = 2\Delta^{-1} < 1$ and $D = 2\Gamma + 2c^{-1} > 0$.

Part (ii) and (iii) can be proved in a similar way for the proof of corresponding part of Lemma 5.2. □

Lemma 4.3. *There exists a constant $C_R > 0$ such that $\|R_n\|_{\mathcal{B}} \le C_R d_n$ for all $n > 0$.*

Proof. For $f \in \mathcal{B}$, denote

$$R_{ij} f = 1_{\widehat{X}} \cdot \mathscr{P}^i (f 1_{I_{ij}})(x). \tag{4.5}$$

Hence $R_i = \sum_{j=1}^{K} R_{ij}$ and $\widehat{\mathscr{P}} = \sum_{i=0}^{\infty} \sum_{j=1}^{K} R_{ij}$ by definition and linearity of $\widehat{\mathscr{P}}$.

Assume $i > 0$; since $\widehat{T}_{ij}[a_{ij}, b_{ij}] = I_0 \subset I$, by (4.2), $\hat{\nu}(I_{ij}) \le d_{ij} \hat{\nu}(I_0) < d_{ij}$. Hence, by Assumption B(b),

$$\int_{I_{ij}} |f| d\hat{\nu} \le \|f\|_\infty \hat{\nu}(I_{ij}) \le C_b \|f\|_{\mathcal{B}} \cdot d_{ij} \hat{\nu}(I_0) \le C_b d_{ij} \|f\|_{\mathcal{B}}. \tag{4.6}$$

Note that $V_{I_{ij}}(f) \le V(f) = |f|_{\mathcal{B}} < \|f\|_{\mathcal{B}}$. By (4.4),

$$V_{\widehat{T} I_{ij}}((f \cdot \widehat{g}) \circ \widehat{T}_{ij}^{-1}) \le d_{ij} \|f\|_{\mathcal{B}} + \Gamma C_b d_{ij} \|f\|_{\mathcal{B}} = (1 + \Gamma C_b) d_{ij} \|f\|_{\mathcal{B}}. \tag{4.7}$$

Since $R_{ij} f(x) = 1_{\widehat{X}}(x) \cdot (f \cdot \widehat{g}) \circ \widehat{T}_{ij}^{-1}(x)$, we have

$$|R_{ij} f|_{\mathcal{B}} \le 2 V_{\widehat{T} I_{ij}}((f \cdot \widehat{g}) \circ \widehat{T}_{ij}^{-1}) + 2 \frac{1}{\hat{\nu}(I_0)} \int_{I_{ij}} |f| d\hat{\nu}.$$

By (4.6) and (4.7),

$$|R_{ij} f|_{\mathcal{B}} \le 2(1 + \Gamma C_b) d_{ij} \|f\|_{\mathcal{B}} + 2 C_b d_{ij} \|f\|_{\mathcal{B}}.$$

On the other hand, by (4.5) and (4.6), we have

$$\|R_{ij} f\|_{L^1} = \int_{\widehat{X}} \widehat{\mathscr{P}}^{i+1} (f 1_{I_{ij}}) d\hat{\nu} = \int_{I_{ij}} f d\hat{\nu} \le \int_{I_{ij}} |f| d\hat{\nu} \le C_b d_{ij} \|f\|_{\mathcal{B}}.$$

Hence, we get

$$\|R_{ij} f\|_{\mathcal{B}} = |R_{ij} f|_{\mathcal{B}} + \|R_{ij} f\|_{L^1} \le [2(1 + \Gamma C_b) + 3C_b] d_{ij} \|f\|_{\mathcal{B}}.$$

By the definition of R_{ij} and d_n, we get

$$\|R_n f\|_{\mathcal{B}} \le \sum_{j=2}^{K} \|R_{n-1,j} f\|_{\mathcal{B}} \le K'(2 + 2\Gamma C_b + 3C_b) d_n,$$

where $K' < K$ is the number of preimages of I_0 that are not in I_1. So the result follows with $C_R = K'(2 + 2\Gamma C_b + 3C_b)$. □

5 Systems on multidimensional spaces: generalities and the role of the derivative

The main difficulty to investigate the statistical properties for systems with an indifferent fixed point p in higher dimensional space is that near p the system could have *unbounded distortion* in the following sense: there are uncountably many points z near p such that for any neighborhood V of z, we can find $\hat{z} \in V$ with the ratio

$$|\det DT_1^{-n}(z)|/|\det DT_1^{-n}(\hat{z})|$$

unbounded as $n \to \infty$ (see Example in Section 2 in [HV]). For this reason we need a more deeper analysis of the expanding features around the neutral fixed point. This has been accomplished in the previous quoted paper and in order to construct an absolutely continuous invariant measure by adding the Assumption T $''$ below, which, together with (5.4), will also be used to get the rate of mixing.

5.1 Setting and Statement of results.

Let $X \subset \mathbb{R}^m$, $m \geq 1$, be again a compact subset with $\overline{\text{int}\, X} = X$, d the Euclidean distance, and ν the Lebesgue measure on X with $\nu X = 1$.

Assume that $T : X \to X$ is a map satisfying the following assumptions.

Assumption T$''$. (a) (Piecewise smoothness) *There are finitely many disjoint open sets $U_1, \cdots, U_K$ with piecewise smooth boundary such that $X = \bigcup_{i=1}^K \overline{U_i}$ and for each i, $T_i := T|_{U_i}$ can be extended to a $C^{1+\hat{\alpha}}$ diffeomorphism $T_i : \widetilde{U_i} \to B_{\varepsilon_1}(T_iU_i)$, where $\widetilde{U_i} \supset U_i$, $\hat{\alpha} \in (0, 1]$ and $\varepsilon_1 > 0$.*

(b) (Fixed point) *There is a fixed point $p \in U_1$ such that $T^{-1}p \notin \partial U_j$ for any $j = 1, \ldots, K$.*

(c) (Topological mixing) *$T : X \to X$ is topologically mixing.*

For any $\varepsilon_0 > 0$, denote

$$G_U(x, \varepsilon, \varepsilon_0) = 2 \sum_{j=1}^{K} \frac{\nu(T_j^{-1} B_\varepsilon(\partial TU_j) \cap B_{(1-s)\varepsilon_0}(x))}{\nu(B_{(1-s)\varepsilon_0}(x))}.$$

Remark 5.1. *We stress that the measure $\nu(T_j^{-1} B_\varepsilon(\partial TU_j)$ usually plays an important role in the study of statistical properties of systems with discontinuities. Here $G_U(x, \varepsilon, \varepsilon_0)$ gives a quantitative measurement of the competition between the expansion and the accumulation of discontinuities near x. We refer to [Ss], Section 2, for more details about its geometric meaning. Furthermore it is proved, still in [Ss] Lemma 2.1, that if the boundary of U_i consists of piecewise C^1 codimension one embedded compact submanifolds, then*

$G_U(\varepsilon,\varepsilon_0) \le 2N_U \frac{\gamma_{m-1}}{\gamma_m} \frac{s\varepsilon}{(1-s)\varepsilon_0}\big(1+o(1)\big)$, *where N_U is the maximal number of smooth components of the boundary of all U_i that meet in one point and γ_m is the volume of the unit ball in $\mathbb{R}^m$.*

From now on we assume that $p=0$.

For any $x \in U_i$, we define $s(x)$ as the inverse of the slowest expansion near x, that is,

$$s(x) = \min\big\{s : d(x,y) \le sd(Tx,Ty), y \in U_i, d(x,y) \le \min\{\varepsilon_1, 0.1|x|\}\ \big\}.$$

where the factor 0.1 makes the ball away from the origin, though any other small factor would work as well.

Take an open neighborhood Q of p such that $TQ \subset U_1$, then let

$$s = s(Q) = \max\{s(x) : x \in X \backslash Q\}. \tag{5.1}$$

Let $\widehat{T} = \widehat{T}_Q$ be the first return map with respect to $\widehat{X} = \widehat{X}_Q = X \setminus Q$. Then for any $x \in U_j$, we have $\widehat{T}(x) = T_j(x)$ if $T_j(x) \notin Q$, and $\widehat{T}(x) = T_1^i T_j(x)$ for some $i>0$ if $T_j(x) \in Q$. Denote $\widehat{T}_{ij} = T_1^i T_j$ for $i \ge 0$.

Further, we take $Q_0 = TQ \setminus Q$. Then we denote $U_{01} = U_1 \setminus Q$, $U_{0j} = U_j \setminus T_j^{-1}Q$ if $j>1$, and $U_{ij} = \widehat{T}_{ij}^{-1}Q_0$ for $i>0$. Hence, $\{U_{ij} : i = 0,1,2,\cdots\}$ form a partition of U_j for $j = 2,\cdots,K$.

For $0 < \varepsilon \le \varepsilon_0$, we denote

$$G_Q(x,\varepsilon,\varepsilon_0) = 2\sum_{j=1}^{K}\sum_{i=0}^{\infty} \frac{\nu(\hat{T}_{ij}^{-1}B_\varepsilon(\partial Q_0) \cap B_{(1-s)\varepsilon_0}(x))}{\nu(B_{(1-s)\varepsilon_0}(x))},$$

and

$$G(x,\varepsilon,\varepsilon_0) = G_U(x,\varepsilon,\varepsilon_0) + G_Q(x,\varepsilon,\varepsilon_0), \quad G(\varepsilon,\varepsilon_0) = \sup_{x\in\widehat{X}} G(x,\varepsilon,\varepsilon_0). \tag{5.2}$$

Remark 5.2. *If $T^{-1}TQ \cap \partial U_j = \emptyset$ for any j, then for any small ε_0, either $G_Q(x,\varepsilon,\varepsilon_0) = 0$ or $G_U(x,\varepsilon,\varepsilon_0) = 0$, and therefore we have $G(x,\varepsilon,\varepsilon_0) = \max\{G_U(x,\varepsilon,\varepsilon_0), G_Q(x,\varepsilon,\varepsilon_0)\}$.*

Remark 5.3. *If T has bounded distortion then G_Q is roughly equal to the ratio between the volume of $B_{\varepsilon_0}(\partial Q_0)$ and the volume of Q_0. Therefore if ε_0 is small enough, then $\sup_{x\in\widehat{X}}\{G_Q(x,\varepsilon,\varepsilon_0)\}$ is bounded by $\sup_{x\in\widehat{X}}\{G_U(x,\varepsilon,\varepsilon_0)\}$.*

Assumption T$''$. (d) (Expansion) *T satisfies $0 < s(x) < 1$ $\forall x \in X \setminus \{p\}$.*

Moreover, there exists an open region Q with $p \in Q \subset \overline{Q} \subset TQ \subset \overline{TQ} \subset U_1$ and constants $\alpha \in (0,\hat{\alpha}]$, $\eta \in (0,1)$, such that for all ε_0 small,

$$s^\alpha + \lambda \le \eta < 1,$$

where s is defined in (5.1) and

$$\lambda = 2 \sup_{0<\varepsilon\le\varepsilon_0} \frac{G(\varepsilon,\varepsilon_0)}{\varepsilon^\alpha}\varepsilon_0^\alpha. \tag{5.3}$$

(e) (Distortion) *For any $b>0$, there exist $J>0$ such that for any small ε_0 and $\varepsilon\in(0,\varepsilon_0)$, we can find $0<N=N(\varepsilon)\le\infty$ with*

$$\frac{|\det DT_1^{-n}(y)|}{|\det DT_1^{-n}(x)|} \le 1+J\varepsilon^\alpha \quad \forall y\in B_\varepsilon(x),\ x\in B_{\varepsilon_0}(Q_0),\ n\in(0,N],$$

and

$$\sum_{n=N}^{\infty} \sup_{y\in B_\varepsilon(x)} |\det DT_1^{-n}(y)| \le b\varepsilon^{m+\alpha} \quad \forall x\in B_{\varepsilon_4}(Q_0),$$

where α is given in part (d).

For sake of simplicity of notations, we may assume $\hat\alpha=\alpha$.

Remark 5.4. *We put Assumption T''(e) since near the fixed point distortion for DT_1 is unbounded in general. It requires that either distortion of DT_1^{-n} is small, or $|\det DT_1^{-n}|$ itself is small.*

Remark 5.5. *There are some sufficient conditions under which Assumption T''(d) and (e) could be easily verified. We refer [HV] for more details, see in particular Theorems B and C in that paper.*

If near p distortion is bounded, then Assumption T$''$(e) is automatically satisfied and it will be stated as follows (it could be regarded as the case $N(\varepsilon)=\infty$ for any $\varepsilon\in(0,\varepsilon_0)$):

Assumption T$''$. (e$'$) (Bounded distortion) *There exist $J>0$ such that for any small ε_0 and $\varepsilon\in(0,\varepsilon_0)$,*

$$\frac{|\det DT_1^{-n}(y)|}{|\det DT_1^{-n}(x)|} \le 1+J\varepsilon^\alpha \quad \forall y\in B_\varepsilon(x),\ x\in B_{\varepsilon_0}(Q_0),\ n\ge 0.$$

Remark 5.6. *It is well known that if $\dim X=m=1$, any system that has the form given by (5.4) below near the fixed point satisfies Assumption T''(e$'$). The systems given in Example 5.1 satisfy it too.*

To estimate the decay rates, we often consider the following special cases: There are constants $\gamma'>\gamma>0$, $C_i, C_i'>0$, $i=0,1,2$, such that in a neighborhood of the indifferent fixed point $p=0$,

$$\begin{aligned} |x|\big(1-C_0'|x|^\gamma+O(|x|^{\gamma'})\big) \le& |T_1^{-1}x| \le |x|\big(1-C_0|x|^\gamma+O(|x|^{\gamma'})\big),\\ 1-C_1'|x|^\gamma+O(|x|^{\gamma'}) \le& \|DT_1^{-1}(x)\| \le 1-C_1|x|^\gamma+O(|x|^{\gamma'}),\\ C_2'|x|^{\gamma-1}+O(|x|^{\gamma'-1}) \le& \|D^2T_1^{-1}(x)\| \le C_2|x|^{\gamma-1}+O(|x|^{\gamma'-1}). \end{aligned} \tag{5.4}$$

We will now define the space of functions particularly adapted to study the action of the transfer operator on the class of maps just introduced. If Ω is a Borel subset of $\widehat{X}$, we define the oscillation of f over Ω by the difference of essential supremum and essential infimum of f over Ω:

$$\operatorname{osc}(f,\Omega) = \operatorname*{Esup}_{\Omega} f - \operatorname*{Einf}_{\Omega} f.$$

If $B_\epsilon(x)$ denotes the ball of radius ϵ about the point x, then we get a measurable function $x \to \operatorname{osc}(f,\ B_\epsilon(x))$.

For $0 < \alpha < 1$ and $\varepsilon_0 > 0$, we define the quasi-Hölder seminorm of f with $\operatorname{supp} f \subset \widehat{X}$ as[‡]

$$|f|_{\mathcal{B}} = \sup_{0<\epsilon\leq\epsilon_0} \epsilon^{-\alpha} \int_{\widehat{X}} \operatorname{osc}(f, B_\epsilon(x)) d\hat{\nu}(x), \tag{5.5}$$

where $\hat{\nu}$ is the normalized Lebsegue measure on $\widehat{X}$, and take the space of the functions as

$$\mathcal{B} = \left\{ f \in L^1(\widehat{X}, \hat{\nu}) : |f|_{\mathcal{B}} < \infty \right\}, \tag{5.6}$$

and then equip it with the norm

$$\|\cdot\|_{\mathcal{B}} = \|\cdot\|_{L^1(\widehat{X},\hat{\nu})} + |\cdot|_{\mathcal{B}}. \tag{5.7}$$

Clearly, the space $\mathcal{B}$ does not depend on the choice of ε_0, though $|\cdot|_{\mathcal{B}}$ does.

Let $s_{ij} = \sup\{\|D\widehat{T}_{ij}^{-1}(x)\| : x \in B_{\varepsilon_0}(Q_0)\}$, and $s_n = \max\{s_{n-1,j} : j = 2, \cdots, K\}$.

Theorem D. *Let $\widehat{X}$, $\widehat{T}$ and $\mathcal{B}$ be defined as above. Suppose T satisfies Assumption T''(a) to (e). Then there exist $\varepsilon_0 \geq \varepsilon_1 > 0$ such that Assumption B(a) to (f) and conditions (i) to (iv) in Theorem A are satisfied and $\|R_n\| \leq O(s_n^\alpha)$. Hence, if $\sum_{k=n+1}^\infty s_n^\alpha \leq O(n^{-\beta})$ for some $\beta > 1$, then there exists $C > 0$ such that for any functions $f \in \mathcal{B}$, $g \in L^\infty(X,\nu)$ with $\operatorname{supp} f$, $\operatorname{supp} g \subset \widehat{X}$, (1.4) holds.*

Remark 5.7. *For Lipschitz observables, the rates of decay of correlation are given by the rates of decay of $\mu\{\tau > n\}$ if the systems have Markov partitions and bounded distortion. It is generally believed that for Hölder observables, the decay rates may be slower if the Hölder exponents become smaller. It is unclear to the authors whether the rates we get are optimal. In the next section, we will put stronger conditions on the systems so that we can get optimal rates for Hölder observables with the Hölder exponents larger than or equal to α.*

[‡]Since the boundary of $\widehat{X}$ is piecewise smooth, we could define the space of the function directly on $\widehat{X}$ instead of $\mathbb{R}^m$ as it was done in [Ss].

Remark 5.8. *For one dimensional systems the rates given in the theorem are optimal, since the decreasing rates given by the norm of derivatives are the same as those given by determinants (see the discussion in the Introduction or Section 6 for more details). So the theorem provides the same decay rates as Theorem C does, but for different sets of observables, since functions with bounded variation are not necessary quasi-Hölder functions and vice versa.*

Before giving the proof, we present an example.

Example 5.1. *Assume that T satisfies Assumption T''(a) to (d), and near the fixed point $p = 0$, the map T satisfies*

$$T(z) = z(1 + |z|^\gamma + O(|z|^{\gamma'})),$$

where $z \in X \subset \mathbb{R}^m$ and $\gamma' > \gamma$.

Denote $z_n = T_1^{-n} z$. It is easy to see that $|z_n| = \dfrac{1}{(\gamma n)^\beta} + O\Big(\dfrac{1}{n^{\beta'}}\Big)$, where $\beta = 1/\gamma$ and $\beta' > \beta$ (see e.g. Lemma 3.1 in [HV]). Using this fact we can check that T satisfies also Assumption T$''$(e$'$). Hence, the theorem can be applied.

If the dimension $m \geq 2$, then $\|DT_1^{-n}\|$ is roughly proportional to $|z_n|$, since, if higher order terms are ignored, T_1^{-n} maps a sphere about the fixed point of radius $|z|$ to a sphere of radius $|z_n|$. So $s_n = O\Big(\dfrac{1}{n^\beta}\Big)$ and $\sum\limits_{k=n+1}^{\infty} s_k^\alpha = O\Big(\dfrac{1}{n^{\alpha\beta-1}}\Big)$. If $\gamma \in (0, 1/2)$ is such that $\alpha\beta > 1$, the series is convergent.

Note that $\nu(\tau > n)$ is of the same order as z_n^m, and therefore $\mu(\tau > n) = O\Big(\dfrac{1}{n^{m\beta-1}}\Big)$. It follows that $\sum\limits_{k=n+1}^{\infty} \mu(\tau > k) = O\Big(\dfrac{1}{n^{m\beta-2}}\Big)$. Since the order is higher, by (1.4), we get

$$\Big|\mathrm{Cov}(f, g \circ T^n)\Big| \leq C/n^{\alpha\beta-1}.$$

for $f \in \mathcal{B}$, $g \in L^\infty(X, \nu)$ with $\operatorname{supp} f$, $\operatorname{supp} g \subset \widehat{X}$.

On the other hand, if $m = 1$, then $\|DT_1^{-n}\|$ is roughly proportional to $|z_n - z_{n+1}|$. So $s_n = O\Big(\dfrac{1}{n^{\beta+1}}\Big)$ and $\sum\limits_{k=n+1}^{\infty} s_k^\alpha = O\Big(\dfrac{1}{n^{\alpha(\beta+1)-1}}\Big)$. If $\gamma \in (0, 1)$ is such that $\alpha(\beta + 1) > 1$, the series is convergent. Also, $\sum\limits_{k=n+1}^{\infty} \mu(\tau > k) = O\Big(\dfrac{1}{n^{\beta-1}}\Big)$. So if $\alpha(\beta + 1) > \beta$, the sum involving s_k^α is of higher order. We get that the decay rate is given by

$$\Big|\mathrm{Cov}(f, g \circ T^n)\Big| = O\Big(\sum_{k=n+1}^{\infty} \mu(\tau > k)\Big) = O\Big(\frac{1}{n^{\beta-1}}\Big).$$

5.2 Proof of Theorem D

Proof of Theorem D. Choose $\varepsilon_0 > 0$ as in Lemma 5.2, and define $\mathcal{B}$ correspondingly by using this ε_0. By Proposition 3.3 in [Ss], $\mathcal{B}$ is complete, and hence is a Banach space. Then Assumption B(a) to (f) follow from Lemma 5.1.

By Lemma 5.2 we know that condition (i) and (ii) of Theorem A are satisfied. Assumption T$''$(a), (d) and (c) imply Assumption T (a), (c) and (d) respectively. Assumption T(b) is implied by the construction of the first return map. Lemma 5.2(iii) give (1.6). So all conditions for Theorem B are satisfied. Hence we obtain condition (iii) and (iv) of Theorem A. The fact $\|R_n\| = O(s_n^\alpha)$ follows from Lemma 5.3. $\square$

In order to deduce the spectral properties of $\hat{\mathcal{P}}$ from the Lasota-Yorke inequality, one needs to verify Assumption B on the space of functions $\mathcal{B}$.

Lemma 5.1. *$\mathcal{B}$ is a Banach space satisfying Assuptions B(a) to (f) with $C_a = 2C_b = 2\gamma_m^{-1}\epsilon_0^{-m}$, where γ_m is the volume of the unit ball in $\mathbb{R}^m$.*

Proof. Parts (a), (b) and (c) are stated in Propositions 3.3 and 3.4 in [Ss] with $C_b = \max\{1, \varepsilon^\alpha\}/\gamma_m\varepsilon_0^m$ and $C_a = 2\max\{1, \varepsilon^\alpha\}/\gamma_m\varepsilon_0^m$. Part (d) follows from the fact that Hölder continuous functions with compact support in $\widehat{X}$ are dense in $L^1(\widehat{X}, \hat{\nu})$.

Let us now assume $f(u) = \lim_{n\to\infty} f_n(u)$ for $\hat{\nu}$-a.e. $u \in \mathbb{R}^m$. Take $x \in \mathbb{R}^m$, and $\varepsilon \in (0, \varepsilon_0)$. It is easy to see that for almost every pair of $y, z \in B_\varepsilon(x)$, we have

$$|f(y) - f(z)| \le \lim_{n\to\infty} |f_n(y) - f_n(z)| \le \liminf_{n\to\infty} \operatorname{osc}(f_n, B_\varepsilon(x)).$$

Hence, $\operatorname{osc}(f, B_\varepsilon(x)) \le \liminf_{n\to\infty} \operatorname{osc}(f_n, B_\varepsilon(x))$. By Fatou's lemma, we have

$$\int \operatorname{osc}(f, B_\varepsilon(x)) d\hat{\nu} \le \liminf_{n\to\infty} \int \operatorname{osc}(f_n, B_\varepsilon(x)) d\hat{\nu}.$$

It implies $|f|_{\mathcal{B}} \le \liminf_{n\to\infty} |f_n|_{\mathcal{B}}$. We get part (e).

It leaves to show part (f). For a function $f \in \mathcal{B}$, denote

$$\mathcal{D}_n(f) = \Big\{x \in \mathbb{R}^m : \liminf_{\varepsilon\to 0} \operatorname{osc}(f, B_\varepsilon(x)) > \frac{1}{n}\Big\}, \quad \mathcal{D}(f) = \bigcup_{n=1}^{\infty} \mathcal{D}_n(f).$$

Clearly $\mathcal{D}(f)$ is the set of discontinuous points of f. If $\hat{\nu}(\mathcal{D}(f)) > 0$, then there exists $N > 0$ such that $\operatorname{Leb}(\mathcal{D}_N(f)) > \iota > 0$. Notice that $\mathcal{D}_N(f) = \bigcup_{k\ge 1} S_k$, where $S_k = \bigcap_{n\ge k}\{x : \operatorname{osc}(f, B_{\frac{1}{n}}(x)) > \frac{1}{N}\}$ is an increasing sequence of measurable sets.

For k big enough we still have $\hat{\nu}(S_k) > \iota$ and therefore, for such a k:

$$|f|_{\mathcal{B}} \ge \sup_{\varepsilon>0} \varepsilon^{-a} \int_{\mathcal{D}_N(f)} \operatorname{osc}(f, B_\varepsilon(x)) d\hat{\nu}(x) \ge \sup_{\varepsilon>0} \varepsilon^{-a} \int_{S_k} \operatorname{osc}(f, B_\varepsilon(x)) d\hat{\nu}(x) = \infty.$$

This means $f \notin \mathcal{B}$; in other words, any $f \in \mathcal{B}$ satisfies $\hat{\nu}(\mathcal{D}(f)) = 0$.

Take any $f \in \mathcal{B}$ with $f \geq 0$ almost everywhere. If $f(x) = 2c > 0$ for some $x \notin \mathcal{D}(f)$, then there is $\varepsilon > 0$ such that $\operatorname{osc}(f, B_\varepsilon(x)) \leq c$. Hence, $f(x') \geq c > 0$ for almost every point $x' \in B_\varepsilon(x)$. So $B_\varepsilon(x) \setminus \{f > 0\}$ has Lebesgue measure zero. This implies that $\{f > 0\}$ is almost open and therefore part (f) follows. $\square$

Before stating the next Lemma, we remind that the space $\mathcal{B}$ depends on the exponent α and the value of the seminorms on ϵ_0: as we did above, we will not index $\mathcal{B}$ with these two parameters. Moreover all the integrals in the next proof will be performed over $\widehat{X}$.

Lemma 5.2. *There exists $\varepsilon_* > 0$ such that for any $\varepsilon_0 \in (0, \varepsilon_*)$, we can find constants $\eta \in (0,1)$ and $D, \hat{D} > 0$ satisfying*

(i) *for any $f \in \mathcal{B}$, $|\widehat{\mathscr{P}} f|_{\mathcal{B}} \leq \eta |f|_{\mathcal{B}} + D\|f\|_{L^1(\hat{\nu})}$;*

(ii) *for any $f \in \mathcal{B}$, $\|R(z)f\|_{\mathcal{B}} \leq |z|\big(\eta\|f\|_{\mathcal{B}} + \hat{D}\|f\|_{L^1(\hat{\nu})}\big)$; and*

(iii) *for any $\widetilde{f} \in \widetilde{\mathcal{B}}$, $\|\widetilde{\mathscr{P}}\widetilde{f}\|_{\widetilde{\mathcal{B}}} \leq \eta\|\widetilde{f}\|_{\widetilde{\mathcal{B}}} + D\|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)}$.*

Proof. By Assumption T$''$ (d), $s^\alpha + \lambda < 1$. Therefore if we first choose b small enough, we obtain $\zeta = J$ according to Assumption T$''$(e), and then we could take ε_0 small enough in order to get

$$\eta := (1 + \zeta\varepsilon_0^\alpha)(s^\alpha + \lambda) + 2\gamma_m^{-1} bK' < 1, \tag{5.8}$$

where K' is the number of j such that $U_{ij} \neq \emptyset$. Clearly, η is decreasing with ε_0. Let us define:

$$D := 2\zeta + 2(1 + \zeta\varepsilon_0^\alpha)\lambda/\varepsilon_0^\alpha + 2\gamma_m^{-1} bK' > 0. \tag{5.9}$$

For any $x \in \widehat{X}$, let us denote $x_{ij} = \widehat{T}_{ij}^{-1} x$, $\widehat{g}_{ij}(x) = |\det D\widehat{T}_{ij}(x)|^{-1}$ and for $f \in \mathcal{B}$:

$$R_{ij} f = 1_{\widehat{X}} \cdot \mathscr{P}^i(f 1_{U_{ij}})(x). \tag{5.10}$$

Clearly,

$$R_{ij} f(x) = f(x_{ij})\hat{g}(x_{ij}) 1_{U_{ij}}(x_{ij}). \tag{5.11}$$

Hence $R_i = \sum_{j=1}^K R_{ij}$ and $\widehat{\mathscr{P}} = \sum_{i=0}^\infty \sum_{j=1}^K R_{ij}$ by definition and the linearity of $\widehat{\mathscr{P}}$. We also define

$$G_{ij}(x, \varepsilon, \varepsilon_0) = 2\frac{\nu(\widehat{T}_{ij}^{-1} B_\varepsilon(\partial\widehat{T}U_{ij}) \cap B_{(1-s)\varepsilon_0}(x))}{\nu(B_{(1-s)\varepsilon_0}(x))}.$$

Clearly, $G(x, \varepsilon, \varepsilon_0) = 2\sum_{i=0}^\infty \sum_{j=1}^K G_{ij}(G(x, \varepsilon, \varepsilon_0)$.

For any $\varepsilon \in (0, \varepsilon_0]$, take $N = N(\varepsilon) > 0$ as in Assumption T$''$(e).

For $i \leq N(\varepsilon)$ and by the proof of Proposition 6.2 in [HV], we know that

$$\begin{aligned} &\operatorname{osc}(R_{ij} f, B_\varepsilon(x)) = \operatorname{osc}\big((f\widehat{g}) \circ \widehat{T}_{ij}^{-1} 1_{\widehat{T}U_{ij}},\ B_\varepsilon(x)\big) \\ &= \operatorname{osc}\big((f\widehat{g}) \circ \widehat{T}_{ij}^{-1}, B_\varepsilon(x)\big) 1_{\widehat{T}U_{ij}}(x) + \big[2 \operatorname*{Esup}_{B_\varepsilon(x)}(f\widehat{g}) \circ \widehat{T}_{ij}^{-1}\big] 1_{B_\varepsilon(\partial\widehat{T}U_{ij})}(x). \end{aligned} \tag{5.12}$$

The computation in the proof also gives

$$\begin{aligned}&\operatorname{osc}\big(f\widehat{g},\ \widehat{T}_{ij}^{-1}B_\varepsilon(x)\cap U_{ij}\big)\\ \le&(1+\zeta\varepsilon^\alpha)\operatorname{osc}\big(f,\ B_{s\varepsilon}(x_{ij})\cap U_{ij}\big)\widehat{g}(x_{ij})+2\zeta\varepsilon^\alpha|f(x_{ij})|\widehat{g}(x_{ij}).\end{aligned}$$

Notice that $\operatorname{osc}\big(f,\ B_{s\varepsilon}(x_{ij})\cap U_{ij}\big)\le\operatorname{osc}\big(f,\ B_{s\varepsilon}(x_{ij})\big)$. By integrating and using (5.11) we get

$$\begin{aligned}&\int\operatorname{osc}\big((f\widehat{g})\circ\widehat{T}_{ij}^{-1},B_\varepsilon(\cdot)\big)1_{\widehat{T}U_{ij}}d\hat{\nu}\\ \le&\int\big[(1+\zeta\varepsilon^\alpha)R_{ij}\operatorname{osc}\big(f,\ B_{s\varepsilon}(\cdot)\big)+2\zeta\varepsilon^\alpha R_{ij}|f|\big]d\hat{\nu}.\end{aligned}\tag{5.13}$$

On the other hand, by the same arguments as in Section 4 of [Ss], we get

$$\begin{aligned}&\int 2\big[\operatorname*{Esup}_{B_{s\varepsilon}(x)}(f\widehat{g})\circ\widehat{T}_{ij}^{-1}\big]1_{B_\varepsilon(\partial\widehat{T}U_{ij})}(x)d\hat{\nu}\\ \le&2(1+\zeta\varepsilon^\alpha)\int_{\widehat{X}}G_{ij}(x,\varepsilon,\varepsilon_0)\big[|f|(x)+\operatorname{osc}(f,B_{\varepsilon_0}(x))\big]d\hat{\nu}.\end{aligned}\tag{5.14}$$

Therefore by (5.12), (5.13) and (5.14),

$$\begin{aligned}|R_{ij}f|_{\mathcal{B}}=&\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}\int\operatorname{osc}\big(R_{ij}f,B_\varepsilon(\cdot)\big)d\hat{\nu}\\ \le&\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}\int\big[(1+\zeta\varepsilon^\alpha)R_{ij}\operatorname{osc}\big(f,\ B_{s\varepsilon}(\cdot)\big)+2\zeta\varepsilon^\alpha R_{ij}|f|\big]d\hat{\nu}\\ +&\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}2(1+\zeta\varepsilon^\alpha)\int_{\widehat{X}}G_{ij}(x,\varepsilon,\varepsilon_0)\big[|f|(x)+\operatorname{osc}(f,B_{\varepsilon_0}(x))\big]d\hat{\nu}.\end{aligned}\tag{5.15}$$

For $i>N(\varepsilon)$, by the definition of oscillation we obtain directly that

$$\operatorname{osc}\big(R_{ij}f,B_\varepsilon(x)\big)\le 2\|f\|_\infty\sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)}\widehat{g}.$$

Hence, by Assumption B(b) with $C_b=\gamma_m^{-1}\varepsilon_0^{-m}$, we have

$$\begin{aligned}|R_{ij}f|_{\mathcal{B}}=&\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}\int\operatorname{osc}\big(R_{ij}f,B_\varepsilon(\cdot)\big)d\hat{\nu}\\ \le&2\|f\|_\infty\varepsilon^{-\alpha}\sup_{0<\varepsilon\le\varepsilon_0}\int\sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)}\widehat{g}\ d\hat{\nu}\\ \le&2(\gamma_m\varepsilon_0^m)^{-1}(|f|_{\mathcal{B}}+\|f\|_1)\ \varepsilon^{-\alpha}\sup_{0<\varepsilon\le\varepsilon_0}\int\sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)}\widehat{g}\ d\hat{\nu}.\end{aligned}\tag{5.16}$$

(i) We first note that for all $0<\varepsilon\le\varepsilon_0$,

$$\begin{aligned}\varepsilon^{-\alpha}\sum_{i=0}^{N(\varepsilon)}\sum_{j=1}^{K}\int R_{ij}\operatorname{osc}\big(f,\ B_{s\varepsilon}(\cdot)\big)d\hat{\nu}\le&\varepsilon^{-\alpha}\int\widehat{\mathscr{P}}\operatorname{osc}\big(f,\ B_{s\varepsilon}(\cdot)\big)d\hat{\nu}\\ \le&s^\alpha(s\varepsilon)^{-\alpha}\int\operatorname{osc}\big(f,\ B_{s\varepsilon}(\cdot)\big)d\hat{\nu}\le s^\alpha|f|_{\mathcal{B}},\end{aligned}\tag{5.17}$$

$$
\begin{aligned}
&\varepsilon^{-\alpha}\sum_{i=0}^{N(\varepsilon)}\sum_{j=1}^{K}\int 2(1+\zeta\varepsilon^{\alpha})G_{ij}(\cdot,\varepsilon,\varepsilon_0)\big[|f|+\operatorname{osc}(f,B_{\varepsilon_0}(\cdot))\big]d\hat{\nu}\\
\le&\varepsilon^{-\alpha}2(1+\zeta\varepsilon^{\alpha})G(\varepsilon,\varepsilon_0)\int\big[|f|+\operatorname{osc}(f,B_{\varepsilon_0}(\cdot))\big]d\hat{\nu}\\
\le&(1+\zeta\varepsilon^{\alpha})\lambda\big[\varepsilon_0^{-\alpha}\|f\|_1+|f|_{\mathcal{B}}\big],
\end{aligned}\tag{5.18}
$$

where we used (5.2) and (5.3). Also, by Assumption T$''$(e) and Assumption B(b) with $C_b=\gamma_m^{-1}\varepsilon_0^{-m+\alpha}$, we have that for all $0<\varepsilon\le\varepsilon_0$:

$$
\varepsilon^{-\alpha}\|f\|_\infty\int\sum_{N(\varepsilon)}^{\infty}\sum_{j=1}^{K'}\sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)}\widehat{g}\;d\hat{\nu}\le\varepsilon^{-\alpha}\|f\|_\infty\cdot bK'\varepsilon^{m+\alpha}\le\gamma_m^{-1}bK'\|f\|_{\mathcal{B}}.\tag{5.19}
$$

Since $\widehat{\mathscr{P}}f(x)=\sum_{i=0}^{\infty}\sum_{j=1}^{K}R_{ij}f(x)$, by (5.15) and (5.16), and using (5.17) to (5.19), we obtain that $|\widehat{\mathscr{P}}f|_{\mathcal{B}}$ is bounded by

$$
\begin{aligned}
&\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}\Big[\int\sum_{i=0}^{\infty}\sum_{j=1}^{K}\operatorname{osc}(R_{ij}f,B_\varepsilon(x))d\hat{\nu}+\int\sum_{i=0}^{\infty}\sum_{j=1}^{K}\operatorname{osc}(R_{ij}f,B_\varepsilon(x))d\hat{\nu}\Big]\\
\le&(1+\zeta\varepsilon_0^{\alpha})s^{\alpha}|f|_{\mathcal{B}}+2\zeta\|f\|_1+(1+\zeta\varepsilon_0^{\alpha})\lambda(\varepsilon_0^{-\alpha}\|f\|_1+|f|_{\mathcal{B}})+2\gamma_m^{-1}bK'\|f\|_{\mathcal{B}}\\
\le&[(1+\zeta\varepsilon_0^{\alpha})(s^{\alpha}+\lambda)+2\gamma_m^{-1}bK']|f|_{\mathcal{B}}+[2\zeta+2(1+\zeta\varepsilon_0^{\alpha})\lambda/\varepsilon_0^{\alpha}+2\gamma_m^{-1}bK']\|f\|_1.
\end{aligned}
$$

By definition of η in (5.8) and D in (5.9) we get the desired inequality.

(ii) Note that for any real valued function f and $z\in\mathbb{C}$, we have $\operatorname{osc}(zf,B_\varepsilon(x))=|z|\operatorname{osc}(f,B_\varepsilon(x))$. Also, note that if $\{a_n\}$ is a sequence of positive numbers and $z\in\overline{\mathbb{D}}$, $|\sum_{n=1}^{\infty}z^na_n|\le|z|\sum_{n=1}^{\infty}a_n$. Hence we have

$$
|R(z)f|_{\mathcal{B}}\le|z|\sup_{0<\varepsilon\le\varepsilon_0}\varepsilon^{-\alpha}\sum_{i=0}^{\infty}\sum_{j=1}^{K}\int\operatorname{osc}(R_{ij}f,B_\varepsilon(x))d\hat{\nu}\le|z||\widehat{\mathscr{P}}f|_{\mathcal{B}}.
$$

By part (i), the inequality becomes

$$
|R(z)f|_{\mathcal{B}}\le|z|(\eta|f|_{\mathcal{B}}+D\|f\|_1).
$$

Also, since $\widehat{\mathscr{P}}$ and R_n are positive operators,

$$
\|R(z)f\|_1\le\sum_{n=1}^{\infty}\big\|z^nR_nf\big\|_1\le|z|\sum_{n=1}^{\infty}\big\|R_n|f|\big\|_1=|z|\big\|\widehat{\mathscr{P}}|f|\big\|_1=|z|\big\|f\big\|_1.
$$

It follows that

$$
\|R(z)f\|_{\mathcal{B}}\le|z|(\eta\|f\|_{\mathcal{B}}+(D+1)\|f\|_1).
$$

Using induction on n, we get the expected result with $\hat{D}=(D+1)/(1-\eta)$.

(iii) The transfer operator $\widetilde{\mathscr{P}}$ has the form (see also [ADSZ])

$$
(\widetilde{\mathscr{P}}\tilde{f})(x,y)=\sum_{n=0}^{\infty}\sum_{j=1}^{K}\tilde{f}(\widehat{T}_{ij}^{-1}x,S(U_{ij})^{-1}(y))g(\widehat{T}_{ij}^{-1}x)1_{\widehat{T}U_{ij}}(x,y),
$$

for any $\widetilde{f} \in \widetilde{\mathcal{B}}$, where $S(U_{ij}) : Y \to Y$ are automorphisms. Let us denote:

$$(\widetilde{R}_{ij}\widetilde{f})(x,y) = \widetilde{f}(\widehat{T}_{ij}^{-1}x, S(U_{ij})^{-1}(y))g(\widehat{T}_{ij}^{-1}x)1_{\widehat{T}U_{ij}}(x,y).$$

Following the same computations as above, we get formulas similar to (5.15) and (5.16) but with R_n and $\widehat{T}_{ij}$ replaced by $\widetilde{R}_n$ and $\widetilde{T}_{ij}$ respectively, and $f(\cdot)$ replaced by $\widetilde{f}(\cdot, y)$. Denote $y_1 = S(U_{ij})^{-1}(y)$. Instead of (5.15) and (5.16), we get that for $i < N(\varepsilon)$,

$$|\widetilde{R}_{ij}\widetilde{f}(\cdot,y)|_{\mathcal{B}} = \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int \operatorname{osc}(\widetilde{R}_{ij}\widetilde{f}(\cdot,y_1), B_\varepsilon(\cdot))\, d\hat{\nu}$$

$$\leq \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int \Big[\Big((1+\zeta\varepsilon^\alpha)\widetilde{R}_{ij} \operatorname{osc}\big(\widetilde{f}(\cdot,y_1),\ B_{s\varepsilon}(\cdot)\big) + 2\zeta\varepsilon^\alpha \widetilde{R}_{ij}|\widetilde{f}(\cdot,y_1)| \Big)$$

$$+2G_{ij}(x,\varepsilon,\varepsilon_0)(1+\zeta\varepsilon^\alpha)\Big(\operatorname{osc}(\widetilde{f}(\cdot,y_1), B_\varepsilon(\cdot)) + |\widetilde{f}(\cdot,y_1)|\Big)\Big] d\hat{\nu},$$

and for $i \geq N(\varepsilon)$,

$$|\widetilde{R}_{ij}\widetilde{f}(\cdot,y)|_{\mathcal{B}} = \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int \operatorname{osc}(\widetilde{R}_{ij}\widetilde{f}(\cdot,y_1), B_\varepsilon(\cdot))\, d\hat{\nu}$$

$$\leq 2(\gamma_m\varepsilon_0^m)^{-1}(|\widetilde{f}(\cdot,y_1)|_{\mathcal{B}} + \|(\widetilde{f}\cdot,y_1)\|_{L^1(\nu)})\varepsilon^{-\alpha} \sup_{0<\varepsilon\leq\varepsilon_0} \int \sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)} \widehat{g}\, d\hat{\nu}.$$

We observe that for any x, $S(U_{ij}) : Y \to Y$ preserves the measure ρ. We set

$$\bar{f}(x) = \int_{\mathbb{S}} \widetilde{f}(x,y_1)\, d\rho(y), \quad \overline{\operatorname{osc}}\big(\widetilde{f}(\cdot), B_\varepsilon(\cdot)\big) = \int_{\mathbb{S}} \operatorname{osc}\big(\widetilde{f}(\cdot,y_1), B_\varepsilon(\cdot)\big)\, d\rho(y).$$

Integrating with respect to y, and using Fubini's theorem, we get

$$|\widetilde{R}_{ij}\widetilde{f}|_{\widetilde{\mathcal{B}}} \leq \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int \Big[\Big((1+\zeta\varepsilon^\alpha)\widetilde{R}_{ij}\overline{\operatorname{osc}}\big(\widetilde{f}(\cdot),\ B_{s\varepsilon}(\cdot)\big) + 2\zeta\varepsilon^\alpha \widetilde{R}_{ij}|\bar{f}(\cdot)| \Big)$$

$$+2G_{ij}(x_{ij},\varepsilon,\varepsilon_0)(1+\zeta\varepsilon^\alpha)\Big(\overline{\operatorname{osc}}(\widetilde{f}(\cdot), B_\varepsilon(\cdot)) + |\bar{f}(\cdot)|\Big)\Big] d\hat{\nu}.$$

and

$$|\widetilde{R}_{ij}\widetilde{f}|_{\widetilde{\mathcal{B}}} \leq 2(\gamma_m\varepsilon_0^m)^{-1}(|\widetilde{f}|_{\widetilde{\mathcal{B}}} + \|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)})\, \varepsilon^{-\alpha} \sup_{0<\varepsilon\leq\varepsilon_0} \int \sup_{\widehat{T}_{ij}^{-1}B_\varepsilon(x)} \widehat{g}\, d\hat{\nu}$$

By Fubini's theorem, we have also $|\widetilde{f}|_{\widetilde{\mathcal{B}}} = \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int \overline{\operatorname{osc}}(\widetilde{f}(\cdot), B_\varepsilon(\cdot))\, d\hat{\nu}$, and $|\widetilde{f}|_{L^1(\hat{\nu}\times\rho)} = \int |\bar{f}(\cdot)|\, d\hat{\nu}$. Using the same arguments as in the proof of part (i) we get

$$|\widetilde{P}\widetilde{f}(\cdot,y)|_{\widetilde{\mathcal{B}}} \leq \sum_{n=0}^{\infty}\sum_{j=1}^{K} |\widetilde{R}_{ij}\widetilde{f}|_{\widetilde{\mathcal{B}}} \leq (1+\zeta\varepsilon_0^\alpha)s^\alpha|\widetilde{f}|_{\widetilde{\mathcal{B}}} + 2\zeta\|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)}$$

$$+(1+\zeta\varepsilon_0^\alpha)\lambda\big(|\widetilde{f}|_{\widetilde{\mathcal{B}}} + \varepsilon_0^{-\alpha}\|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)}\big) + 2\gamma_m^{-1}bK'\big(|\widetilde{f}|_{\widetilde{\mathcal{B}}} + \|\widetilde{f}\|_{L^1(\hat{\nu}\times\rho)}\big),$$

and therefore the result of part (iii) with the same η and D giving in (5.8) and (5.9) respectively. $\square$

Lemma 5.3. *There exists a constant $C_R > 0$ such that $\|R_n\|_{\mathcal{B}} \le C_R s_n^\alpha$ for all $n > 0$.*

Proof. Since $R_i = \sum_j R_{ij}$, we only need to prove the results for R_{ij}.

Take $\varepsilon \in (0, \varepsilon_0]$. Choose any $b > 0$ and let $N(\varepsilon)$ be given by Assumption T$''$(e).

We first consider the case $n = i + 1 \le N(\varepsilon)$.

By the definition of R_{ij} given in (5.10), we have for any $f \in \mathcal{B}$,

$$\int R_{ij} f d\hat{\nu} = \int 1_{\widehat{X}} \cdot \mathscr{P}^{i+1}(f 1_{U_{ij}}) d\hat{\nu} = \int_{\widehat{X}} f 1_{U_{ij}} d\hat{\nu} = \int_{U_{ij}} f d\hat{\nu}. \tag{5.20}$$

We now denote $d_{ij} = \sup\{|\det D\widehat{T}_{ij}^{-1}(x) : x \in B_\varepsilon(Q_0)\}$. Since for any x, $|\det D\widehat{T}_{ij}^{-1}(x)| \le \|D\widehat{T}_{ij}^{-1}(x)\|$, we have $d_{ij} \le s_{ij}$. Since $\widehat{T} U_{ij} = Q_0$,

$$\nu(U_{ij}) \le d_{ij} \nu(Q_0) \le s_{ij} \nu(Q_0). \tag{5.21}$$

Hence by Assumption B(b),

$$\int R_{ij} f d\hat{\nu} \le \|f\|_{L^\infty(\hat{\nu})} \nu(U_{ij}) \le C_b \nu(Q_0) s_{ij} \|f\|_{\mathcal{B}}. \tag{5.22}$$

By similar arguments as for (5.20), we have

$$\int_{\widehat{X}} R_{ij} \operatorname{osc}(f,\ B_{s_{ij}\varepsilon}(\cdot)) d\hat{\nu} \le \int_{\widehat{X}} \operatorname{osc}(f,\ B_{s_{ij}\varepsilon}(\cdot)) 1_{U_{ij}} d\hat{\nu} \le s_{ij}^\alpha \varepsilon^\alpha |f|_{\mathcal{B}}. \tag{5.23}$$

We note that for each j, $\widehat{T} U_{ij} = Q_0$ and the "thickness" of $\widehat{T}_{ij}^{-1} B_\varepsilon(\partial Q_0)$ is of order $s_{ij}\varepsilon$, since ∂Q_0 consists of piecewise smooth surfaces. So $G_{ij}(\varepsilon, \varepsilon_0) \le C_G \varepsilon s_{ij}$ for some C_G independent of i and j. Therefore we have

$$\begin{aligned} &\int_{\widehat{X}} \varepsilon^{-\alpha} 2(1 + \zeta\varepsilon^\alpha) G_{ij}(\cdot, \varepsilon, \varepsilon_0) \big[|f| + \operatorname{osc}(f, B_{\varepsilon_0}(\cdot))\big] d\hat{\nu} \\ \le& 2(1 + \zeta\varepsilon^\alpha) C_G \varepsilon^{1-\alpha} s_{ij} \big[\|f\|_{L^1(\hat{\nu})} + \varepsilon_0^\alpha |f|_{\mathcal{B}}\big], \end{aligned}$$

Hence by (5.15) we get that

$$|R_{ij} f|_{\mathcal{B}} \le C_R' s_{ij}^\alpha \big[\|f\|_{L^1(\hat{\nu})} + |f|_{\mathcal{B}}\big] = C_R' s_{ij}^\alpha \|f\|_{\mathcal{B}}$$

for $C_R' = (1 + \zeta\varepsilon_0^\alpha)(1 + 2C_G \varepsilon_0^{1-\alpha}) + 2\zeta C_b \hat{\nu}(Q_0)$.

We now consider the case $n = i+1 > N(\varepsilon)$. As we mentioned in Remark 5.6, in this case $m \ge 2$. By definition, there is $C_s > 0$ such that $\widehat{g}(x_{ij}) \le C_s^2 s_{ij}^2$ for any $x_{ij} \in \widehat{T}_{ij}^{-1} B_\varepsilon(Q_0)$ with $j = 2, \cdots, K$. By Assumption T$''$(e) we know that for any $x \in B_\varepsilon(Q_0)$,

$$\Big(\sup_{\widehat{T}_{n-1,j}^{-1} B_\varepsilon(x)} \widehat{g}\Big)^{1/2} \le \Big(\sum_{i=N(\varepsilon)}^{\infty} \sup_{\widehat{T}_{ij}^{-1} B_\varepsilon(x)} \widehat{g}\Big)^{1/2} \le \sqrt{b}\varepsilon^{(m+\alpha)/2} \le \sqrt{b}\varepsilon^\alpha.$$

Therefore by (5.16) we have

$$|R_n f|_{\mathcal{B}} \leq C_R'' s_{ij} \|f\|_{\mathcal{B}} \leq C_R'' s_{ij}^{\alpha} \|f\|_{\mathcal{B}}$$

for $C_R'' = 2(\gamma_m \varepsilon_0^m)^{-1}\sqrt{b}C_s$.

Finally, by (5.22), we have

$$\|R_{ij} f\|_1 \leq \int R_{ij}|f| d\hat{\nu} \leq C_b \nu(Q_0) s_{ij} \|f\|_{\mathcal{B}}.$$

Thus we have $\|R_{ij} f\|_{\mathcal{B}} = (C_R' + C_R'' + C_b \nu(Q_0)) s_{ij}^{\alpha} \|f\|_{\mathcal{B}}$. The result of the lemma then follows. $\square$

6 Systems on multidimensional spaces: the role of the determinant

In this section we put additional conditions on the the map T that we studied in the previous chapter in order to get optimal estimates for the decay of correlations.

6.1 Assumptions and statement of the results.

Let us suppose T satisfies Assumption T$''$(a), (d) and (e) in the last section. We replace part (b) and (c) by the following

Assumption T$''$. (b$'$) (Fixed point and a neighborhood) *There is a fixed point $p \in U_1$ and a neighborhood V of p such that $T^{-n}V \notin \partial U_j$ for any $j = 1, \dots, K$ and for any $n \geq 0$.*

(c$'$) (Topological exactness) $T : X \to X$ *is topologically exact, that is, for any $x \in X$, $\varepsilon > 0$, there is an $\widetilde{N} = \widetilde{N}(x, \varepsilon) > 0$ such that $T^{\widetilde{N}} B_\varepsilon(x) = X$.*

Remark 6.1. *It is easy to see that if T has a finite Markov partition, or a finite image structure (see e.g. [Yr]), then T satisfies Assumption T$''$(b$'$) as long as p is not on the boundary of the elements of the partition in the former case and not on the boundary of the images in the latter case.*

Remark 6.2. *Clearly, topological exactness implies topological mixing.*

We rename the seminorm and the Banach space defined in (5.7) and (**??**) by replacing $\mathcal{B}$ with $\mathcal{Q}$ which will therefore depend on α and on ϵ_0, the latter dependence affecting only the value of the seminorms. Then instead of (5.7) we put

$$\|f\|_{\mathcal{Q}} = \|f\|_{L^1(\hat{\nu})} + |f|_{\mathcal{Q}}.$$

Recall that V is a neighborhood of p given in Assumption T$''$(b). We denote the preimages $T_{i_k}^{-1} \dots T_{i_1}^{-1} V$ by $V_{i_1 \dots \imath_k}$ or V_I where $I = i_1 \dots \imath_k$. We also denote with $\mathcal{I}$ the set of all possible words $i_1 \cdots i_k$ such that $T_{i_k}^{-1} \dots T_{i_1}^{-1} V$ is well defined, where $i_k \in \{1, \cdots, K\}$ and $k > 0$.

For an open set O, let $\mathcal{H} := \mathcal{H}_{\varepsilon_1}^{\alpha} = \mathcal{H}_{\varepsilon_1}^{\alpha}(O, H)$ be the set of Hölder functions f over O that satisfies $|f(x) - f(y)| \leq H d(x,y)^{\alpha}$ for any $x, y \in O$ with $d(x,y) \leq \varepsilon_1$.

Let $\hat{h}$ be a fixed point of the transfer operator $\widehat{\mathscr{P}}$, which will be unique under the assumptions of the theorem below. We now define $\mathcal{B}$ by

$$\mathcal{B} := \mathcal{B}_{\varepsilon_0, \varepsilon_1}^{\alpha} = \left\{ f \in \mathcal{Q} : \exists H > 0 \ s.t. \ (f/\hat{h})|_{V_I} \in \mathcal{H}_{\varepsilon_1}^{\alpha}(V_I, H) \ \forall I \in \mathcal{I} \right\}, \tag{6.1}$$

and for any $f \in \mathcal{B}$, let

$$|f|_{\mathcal{H}} := |f|_{\mathcal{H}_{\varepsilon_1}^{\alpha}} = \inf\{H : (f/\hat{h})|_{V_I} \in \mathcal{H}_{\varepsilon_1}^{\alpha}(V_I, H) \ \forall I \in \mathcal{I}\}.$$

Sublemmas 6.3 and 6.4 below imply that $\hat{h} > 0$ on all V_{ij}, and therefore the definition makes sense. Then we take $|\cdot|_{\mathcal{Q}} + |\cdot|_{\mathcal{H}}$ as a seminorm for $f \in \mathcal{B}$ and define the norm in $\mathcal{B}$ by

$$\| \cdot \|_{\mathcal{B}} = \| \cdot \|_1 + | \cdot |_{\mathcal{Q}} + | \cdot |_{\mathcal{H}}. \tag{6.2}$$

Clearly, $\mathcal{B} \subset \mathcal{Q}$ and $\|f\|_{\mathcal{B}} \geq \|f\|_{\mathcal{Q}}$ if $f \in \mathcal{B}$.

Recall that for any sequences of numbers $\{a_n\}$ and $\{b_n\}$, we use $a_n \approx b_n$ if $\lim\limits_{n\to\infty} a_n/b_n = 1$, and $a_n \sim b_n$ if $c_1 b_n \leq a_n \leq c_2 b_n$ for some constants $c_2 \geq c_1 > 0$.

Let $d_{ij} = \sup\{|\det D\widehat{T}_{ij}^{-1}(x) : x \in B_{\varepsilon_0}(Q_0)\}$, and $d_n = \max\{d_{n-1,j} : j = 2, \cdots, K\}$.

Theorem E. *Let $\widehat{X}$, $\widehat{T}$ and $\mathcal{B}$ be defined as above. Suppose T satisfies Assumption T''(a), (b′), (c′), (d) and (e). Then there exist $\varepsilon_0 \geq \varepsilon_1 > 0$ such that Assumption B(a) to (f) and conditions (i) to (iv) in Theorem A are satisfied and $\|R_n\| \leq O(d_n^{m/(m+\alpha)})$. Hence, if $\sum_{k=n+1}^{\infty} d_n^{m/(m+\alpha)} \leq O(n^{-\beta})$ for some $\beta > 1$, then there exists $C > 0$ such that for any functions $f \in \mathcal{B}$, $g \in L^{\infty}(X, \nu)$ with $\operatorname{supp} f$, $\operatorname{supp} g \subset \widehat{X}$, (1.4) holds.*

Moreover, if T satisfies (5.4) near $p = 0$, then $\sum\limits_{k=n+1}^{\infty} \mu(\tau > k)$ has the order $n^{-(m/\gamma-1)}$ or higher. In this case, if $d_n = O(n^{-\beta'})$ for some $\beta' > 1$ and if

$$\beta = \beta' \cdot \frac{m}{m+\alpha} - 1 \geq \max\{2, \frac{m}{\gamma} - 1\}, \tag{6.3}$$

then

$$\operatorname{Cov}(f, g \circ T^n) \approx \sum_{k=n+1}^{\infty} \mu(\tau > k) \int f d\mu \int g d\mu \ \sim 1/n^{m/\gamma - 1}. \tag{6.4}$$

In particular, if Assumption T''(e′) in the last section also holds, then the above statemnets remain true if we replace $m/(m+\alpha)$ by 1.

Remark 6.3. *For the case that T satisfies (5.4) near p, if h is bounded away from 0 on the sets $\{\tau > n\}$, then $\mu(\tau > n)$ and $\nu(\tau > n)$ have the same order, and $\sum_{k=n+1}^{\infty} \mu(\tau > k) = O(n^{-(m/\gamma-1)})$. This is the case in Example 7.1, 7.2 and 7.4 below.*

On the other hand, $\hat{h}$ may be only supported on part of the sets $\{\tau > n\}$, and therefore $\mu(\tau > n)$ may have higher order, like in Example 7.3. In this case, $\sum_{k=n+1}^{\infty} \mu(\tau > k)$ has an order higher than $n^{-(m/\gamma-1)}$.

6.2 Examples

Before giving the proof, we present a few examples. We will always assume that T satisfies Assumption T$''$(a), (b$'$), (c$'$) and (d).

Example 6.1. *Assume $m = 3$, and near the fixed point $p = (0, 0, 0)$, the map T has the form*

$$T(w) = \big(x(1 + |w|^2 + O(|w|^3)),\ y(1 + |w|^2 + O(|w|^3)), z(1 + 2|w|^2 + O(|w|^3))\big)$$

where $w = (x, y, z)$ and $|w| = \sqrt{x^2 + y^2 + z^2}$.

This map is very similar to that in Example 1 in [HV], although it is now in a three dimensional space. We could still use the same arguments to show that Assumption T$''$ (e) is satisfied.

Denote $w_n = T_1^{-n} w$; clearly, $|w| + |w|^3 + O(|w|^4) \le |T(w)| \le |w| + 2|w|^3 + O(|w|^4)$. By standard arguments we know that

$$\frac{1}{\sqrt{4n}} + O\Big(\frac{1}{\sqrt{n^3}}\Big) \le |w_n| \le \frac{1}{\sqrt{2n}} + O\Big(\frac{1}{\sqrt{n^3}}\Big)$$

(see also Lemma 3.1 in [HV]). Since we are in a three dimensional space, we now have $\nu(\tau > k) \sim \dfrac{1}{k^{m/\gamma}} = \dfrac{1}{k^{3/2}}$, and therefore $\displaystyle\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{1/2}}$.

It is easy to see that $\det DT(w) = 1 + 6x^2 + 6y^2 + 8z^2 + O(|w|^3)$. So we have $|\det DT_1^{-1}(w)| \le 1 - 6|w|^2 + O(|w|^3)$. By Lemma 3.2 in [HV] with $r(t) = 1-6t^2+O(t^3)$, $\gamma = 2$, $C' = 6$ and $C = 1$, we get that $|\det DT_1^{-n}(x)| = O(1/n^3)$. Hence we have $\beta' = 3$ and $\beta = 3m/(m + \alpha) - 1 > 3 \cdot 3/(3 + 1) - 1 = 5/4$. Since $m/\gamma - 1 = 1/2$, (6.3) holds, and therefore we have (6.4) with the decay rate of order $1/\sqrt{n}$.

Example 6.2. *Assume $m = 2$, and near the fixed point $p = (0, 0)$, the map T has the form*

$$T(z) = \big(x(1 + |z|^\gamma + O(|z|^{\gamma'})),\ y(1 + 2|z|^\gamma + O(|z|^{\gamma'}))\big)$$

where $z = (x, y)$, $|z| = \sqrt{x^2 + y^2}$, $\gamma \in (0, 1)$ and $\gamma' > \gamma$.

By methods similar to Example 1 in [HV] we can check that Assumption T$''$ (e) is satisfied. Denote $z_n = T_1^{-n} z$. Since $|z| + |z|^{1+\gamma} + O(|z|^{\gamma'}) \le |T(z)| \le |z| + 2|z|^{\gamma+1} + O(|z|^{\gamma'})$, we have

$$\frac{1}{(2\gamma n)^{1/\gamma}} + O\Big(\frac{1}{n^\delta}\Big) \le |z_n| \le \frac{1}{(\gamma n)^{1/\gamma}} + O\Big(\frac{1}{n^\delta}\Big)$$

for some $\delta > 1/\gamma$. So $\nu(\tau > k) \sim \dfrac{1}{k^{2/\gamma}}$, and therefore $\displaystyle\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{2/\gamma - 1}}$.

It is possible to show that $|\det DT(z)| = 1 + \dfrac{(3+\gamma)x^2 + (3+2\gamma)y^2}{|z|^{2-\gamma}} + O(|z|^{\gamma'})$. Therefore $|\det DT_1^{-1}(z)| \le 1 - (3+\gamma)|z|^\gamma + O(|z|^{\gamma'})$, and $|\det DT_1^{-n}(z)| = O(1/n^{1+3/\gamma})$. Hence $\beta' = 1 + \gamma/3$ and $\beta = (1 + 3/\gamma) \cdot 2/(2+\alpha) - 1 > (1 + 3/\gamma) \cdot 2/3 - 1 = 2/\gamma - 1/3 > 2/\gamma - 1$. It means (6.3) holds, and the decay rates is of order $1/n^{2/\gamma - 1}$.

Example 6.3. *Assume $m = 2$, and take the same map as in Example 1 in [HV], namely, near the fixed point $p = (0,0)$, the map T has the form*

$$T(x,y) = \big(x(1 + x^2 + y^2),\ y(1 + x^2 + y^2)^2\big),$$

where $z = (x,y)$ and $|z| = \sqrt{x^2 + y^2}$.

The map allows an infinite absolutely continuous invariant measure. However, the map can be arranged in such a way that there is an invariant component that supports a finite absolutely continuous invariant measure μ. Near the fixed point, the region of this component has the form

$$\{z = (x,y) : |y| < x^2\}.$$

We may regard X as this component, and $T : X \to X$ satisfies the assumptions.

We could check that the map has bounded distortion near the fixed point restricted to this region. Hence, the map satisfies Assumption T$''$(e$'$).

Since $|z_n| = O(1/\sqrt{n})$ and for $z = (x,y)$, $|y| \le x^2$, we get $\nu(\tau > k) \sim \dfrac{1}{k^{3/2}}$, and $\displaystyle\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{1/2}}$.

On the other hand, $|\det DT(z)| = 1 + 5x^2 + 7y^2 + O(|z|^4)$. Since $|y| \le x^2$, $|z| = |x| + O(|z|^2)$; thus $|\det DT(z)| = 1 + 5|z|^2 + O(|z|^4)$, and therefore $|\det DT_1^{-n}(z)| = O(1/n^{5/2})$. So $\beta' = 5/2$ and $\beta = 3/2$. We obtain that the decay rate is of order $1/n^{1/2}$.

Example 6.4. *Assume $m \ge 3$ and near the fixed point $p = (0,0,0)$, the map T has the form*

$$T(z) = z\big(1 + |z|^\gamma + O(|z|^{\gamma+1})\big),$$

where $m > \gamma > 0$.

These examples are comparable with those in Example 6.1, except for the stronger topological assumptions which we now put on the maps. We know that these maps satisfy Assumption, T$''$(e$'$).

Denote $z_n = T_1^{-n}z$. We have $|z_n| = 1/(n\gamma)^{1/\gamma} + O\big(1/(n\gamma)^{1/\gamma+1}\big)$ and $|\det DT(z)| = 1 + (m+\gamma)|z|^\gamma + O\big(|z|^{\gamma+1}\big)$. Hence, we get that $|\det DT_1^{-n}| \sim 1/n^{m/\gamma+1}$. (For the relative computations see Lemma 3.1 and 3.2 in [HV]). Therefore $\beta' = m/\gamma + 1$ and $\beta = m/\gamma$.

On the other hand, we see that $\nu(\tau > k) = O\big(1/k^{m/\gamma}\big)$, and then $\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \dfrac{1}{n^{m/\gamma-1}}$. Since $m > \gamma$, the invariant measure μ is finite and $\beta > 1$. We get that the decay rate is of order $1/n^{m/\gamma-1}$.

6.3 Proof of the Theorem

Proof of Theorem E. Take $\varepsilon_0 > 0$ satisfying Lemma 5.2 in the last section, and then choose $\varepsilon_1 \in (0, \varepsilon_0]$ as in Lemma 6.2 below. We reduce ε_1 further if necessary such that $\eta' := \eta + D_{\mathcal{H}}(\varepsilon_0)\varepsilon_1^\alpha < 1$, where $\eta < 1$ is given in Lemma 5.2 and $D_{\mathcal{H}}(\varepsilon_0) > 0$ is given in Lemma 6.2. Then we take $\mathcal{B} := \mathcal{B}^\alpha_{\varepsilon_0,\varepsilon_1}$ as in (6.1). With the norm given in (6.2), $\mathcal{B}$ satisfies Assumption B(a) to (f) by Lemma 6.1.

By Lemma 5.2 and 6.2, condition (i) of Theorem A is satisfied with constants η and D replaced by η' defined as above and $D + D_{\mathcal{H}}(\varepsilon_0)\varepsilon_1^\alpha$ respectively, where D is the number given in Lemma 5.2. Condition (ii) can be obtained in a similar way. Assumption T$''$(a), (d) and (c$'$) imply Assumption T (a), (c) and (d) respectively. Assumption T(b) follows from the construction of the first return map. Lemma 5.2(iii) and 6.2(iii) give (1.6). Therefore all the conditions for Theorem B are satisfied. Hence we obtain condition (iii) and (iv) of Theorem A.

The facts $\|R_n\| = O(d_n^{m/(m+\alpha)})$, and $\|R_n\| = O(d_n)$ if Assumption T$''$(e$'$) is satisfied, follow from Lemma 6.5. Therefore (1.4) is given by Theorem A.

If T also satisfies (5.4), then we know that for any z close to p, $|T_1^{-n}z|$ is of order $n^{-1/\gamma}$. Hence $\hat\nu\{\tau > k\}$ has the order $k^{-m/\gamma}$, and $\sum_{k=n+1}^\infty k^{-m/\gamma} = O(n^{-m/\gamma+1})$. Then the rest of the theorem is clear. $\square$

Lemma 6.1. *$\mathcal{B}$ is a Banach space satisfying Assumption B(a) to (f) with $C_a = 2C_b = 2\gamma_m^{-1}\varepsilon_0^{-m+\alpha}$, where γ_m is the volume of the unit ball in $\mathbb{R}^m$.*

Proof. We already know that $\mathcal{Q}$ is a Banach space, and the proof of the completeness of $\mathcal{B}$ follows from standard arguments. So $\mathcal{B}$ is a Banach space.

Now we verify Assumption B(a) to (f).

By Lemma 5.1, the unit ball of $\mathcal{Q}$ is compact in $L^1(\widehat{X}, \hat\nu)$. Since $||f||_{\mathcal{B}} \geq ||f||_{\mathcal{Q}}$ for any $f \in \mathcal{B} \subset \mathcal{Q}$, the unit ball of $\mathcal{B}$ is contained in the unit ball of $\mathcal{Q}$. Since $\mathcal{B}$ is closed in $\mathcal{Q}$, the unit ball of $\mathcal{B}$ is also compact. This is Assumption B(a).

Moreover, for any $f \in \mathcal{Q}$, $\|f\|_\infty \le C_b \|f\|_{\mathcal{Q}} \le C_b \|f\|_{\mathcal{B}}$ with $C_b = \gamma_m^{-1} \varepsilon_0^{-m+\alpha}$. We have thus got Assumption B(b).

Invoking again Lemma 5.1, we have for any $f, g \in \mathcal{Q}$, $\|fg\|_{\mathcal{Q}} \le C_a \|f\|_{\mathcal{Q}} \|g\|_{\mathcal{Q}}$, where $C_a = 2\gamma_m^{-1} \varepsilon_0^{-m+\alpha} = 2C_b$. It is easy to check that

$$|fg|_{\mathcal{H}} \le \|f\|_\infty |g|_{\mathcal{H}} + \|g\|_\infty |f|_{\mathcal{H}} \le C_b \|f\|_{\mathcal{Q}} |g|_{\mathcal{H}} + C_b \|g\|_{\mathcal{Q}} |f|_{\mathcal{H}}.$$

Hence,

$$\begin{aligned}\|fg\|_{\mathcal{B}} = \|fg\|_{\mathcal{Q}} + |fg|_{\mathcal{H}} &\le C_a \|f\|_{\mathcal{Q}} \|g\|_{\mathcal{Q}} + C_b \|f\|_{\mathcal{Q}} |g|_{\mathcal{H}} + C_b \|g\|_{\mathcal{Q}} |f|_{\mathcal{H}} \\ &\le C_a \big(\|f\|_{\mathcal{Q}} + |f|_{\mathcal{H}}\big)\big(\|g\|_{\mathcal{Q}} + |g|_{\mathcal{H}}\big) = C_a \|f\|_{\mathcal{B}} \|g\|_{\mathcal{B}}.\end{aligned}$$

Therefore Assumption B(c) follow with $C_a = 2\gamma_m^{-1} \varepsilon_0^{-m+\alpha} = 2C_b$.

Similarly, part (d) of Assumption B follows from the fact that $\mathcal{B}$ contains all Hölder functions, and Hölder functions are dense in $L^1(\widehat{X}, \hat{\nu})$.

Assume $f(x) = \lim_{n\to\infty} f_n(x)$ for $\hat{\nu}$-a.e. $x \in \widehat{X}$. By the proof of Lemma 5.1 we have $|f|_{\mathcal{Q}} \le \liminf_{n\to\infty} |f_n|_{\mathcal{Q}}$. For any $y, z \in V_I$, where $I \in \mathcal{I}$,

$$\frac{|f(y) - f(z)|}{d(y,z)^\alpha} \le \lim_{n\to\infty} \frac{|f_n(y) - f_n(z)|}{d(y,z)^\alpha} \le \liminf_{n\to\infty} |f_n|_{\mathcal{H}}.$$

It gives $|f|_{\mathcal{H}} \le \liminf_{n\to\infty} |f_n|_{\mathcal{H}}$. Since $|f|_{\mathcal{B}} = |f|_{\mathcal{Q}} + |f|_{\mathcal{H}}$, we get part (e).

Since $\mathcal{B} \subset \mathcal{Q}$, part (f) is directly from the fact that $\mathcal{Q}$ satisfies Assumption B(f). □

Lemma 6.2. *Let ε_0 be as in Lemma 5.2. There exists $D_{\mathcal{H}} = D_{\mathcal{H}}(\varepsilon_0), \bar{D}_{\mathcal{H}} = \bar{D}_{\mathcal{H}}(\varepsilon_0) > 0$ and $\varepsilon_- \in (0, \varepsilon_0]$ such that for any $\varepsilon_1 \in (0, \varepsilon_-]$, and by using the notation for the Banach space introduced in (6.1):*

(i) *for any $f \in \mathcal{B}^\alpha_{\varepsilon_0,\varepsilon_1}$, $|\widehat{\mathscr{P}} f|_{\mathcal{H}_{\varepsilon_1}} \le s^\alpha |f|_{\mathcal{H}_{\varepsilon_1}} + D_{\mathcal{H}} \varepsilon_1^\alpha \|f\|_{\mathcal{Q}_{\varepsilon_0}}$;*

(ii) *for any $f \in \mathcal{B}^\alpha_{\varepsilon_0,\varepsilon_1}$, $|R(z) f|_{\mathcal{H}_{\varepsilon_1}} \le |z| \big(s^a |f|_{\mathcal{H}_{\varepsilon_1}} + \bar{D}_{\mathcal{H}} \varepsilon_1^\alpha \|f\|_{\mathcal{Q}_{\varepsilon_0}}\big)$;*

(iii) *and for any $f \in \widetilde{\mathcal{B}}^\alpha_{\varepsilon_0,\varepsilon_1}$ $|\widetilde{\mathscr{P}} \tilde{f}|_{\widetilde{\mathcal{H}}_{\varepsilon_1}} \le s^\alpha |\tilde{f}|_{\widetilde{\mathcal{H}}_{\varepsilon_1}} + D_{\mathcal{H}} \varepsilon_1^\alpha \|\tilde{f}\|_{\widetilde{\mathcal{Q}}_{\varepsilon_0}}$.*

Proof. (i) Let $\varepsilon_* \in (0, \varepsilon_0]$, $J_{\hat{h}} > 0$ as in the proof of Sublemma 6.4 below. Suppose $\varepsilon \in (0, \varepsilon_*]$, and $|f|_{\mathcal{H}_{\varepsilon_1}} = H$ for some f. Take $x, y \in V_I$ for some $I \in \mathcal{I}$ with $d(x,y) = \varepsilon \le \varepsilon_*$. Then by Assumption T$''$(e), we can take $J > 0$, $N = N(\varepsilon) > 0$ for $b = 1$. Note that

$$\begin{aligned}\frac{\widehat{\mathscr{P}} f(x)}{\hat{h}(x)} - \frac{\widehat{\mathscr{P}} f(y)}{\hat{h}(y)} =& \sum_{j=1}^K \sum_{i=1}^\infty \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} \Big(\frac{f(x_{ij})}{\hat{h}(x_{ij})} - \frac{f(y_{ij})}{\hat{h}(y_{ij})} \Big) \\ &+ \sum_{j=1}^K \sum_{i=1}^N \frac{f(y_{ij})}{\hat{h}(y_{ij})} \Big(\frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} - \frac{\hat{g}(y_{ij})\hat{h}(y_{ij})}{\hat{h}(y)} \Big) \\ &+ \sum_{j=1}^K \sum_{i=N+1}^\infty \frac{f(y_{ij})}{\hat{h}(y_{ij})} \Big(\frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} - \frac{\hat{g}(y_{ij})\hat{h}(y_{ij})}{\hat{h}(y)} \Big).\end{aligned} \tag{6.5}$$

Since $|f|_{\mathcal{H}} = H$, we have $f(x_{ij})/\hat{h}(x_{ij}) - f(y_{ij})/\hat{h}(y_{ij}) \le Hd(x_{ij}, y_{ij})^\alpha \le s^\alpha Hd(x,y)^\alpha$. Now, $\widehat{\mathscr{P}}\hat{h} = \hat{h}$ implies

$$\sum_{j=1}^{K}\sum_{i=1}^{\infty} \hat{g}(x_{ij})\hat{h}(x_{ij})/\hat{h}(x) = 1. \tag{6.6}$$

Thus the first sum of the inequality is bounded by $s^\alpha Hd(x,y)^\alpha \le s^\alpha |f|_{\mathcal{H}} d(x,y)^\alpha$.

Note that by our assumption, V_{ij} does not intersect discontinuities. By Sublemma 6.4, $\hat{h}(y)/\hat{h}(x) \le e^{J_{\hat{h}} d(x,y)^\alpha}$, and by Assumption T$''$(e), $\hat{g}(y)/\hat{g}(x) \le e^{Jd(x,y)^\alpha}$ if $i \le N(\varepsilon)$. So $[\hat{g}(y_{ij})\hat{h}(y_{ij})/\hat{h}(y)]/[\hat{g}(x_{ij})\hat{h}(x_{ij})/\hat{h}(x)] \le e^{J' d(x,y)^\alpha}$ for some $J' > 0$. We take $\varepsilon_- \in (0, \varepsilon_*]$ small enough such that $e^{J\varepsilon_1^\alpha} - 1 \le 2J'\varepsilon_1^\alpha$ for any $\varepsilon_1 \le (0, \varepsilon_-]$. Then for $d(x,y) = \varepsilon \le \varepsilon_1$, we have

$$\Big|\frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} - \frac{\hat{g}(y_{ij})\hat{h}(y_{ij})}{\hat{h}(y)}\Big| \le 2J' \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} \cdot d(x,y)^\alpha. \tag{6.7}$$

Therefore by (6.6), the second sum in (6.5) is bounded by

$$\sum_{j=1}^{K}\sum_{i=1}^{N} \frac{f(y_{ij})}{\hat{h}(y_{ij})} \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} \cdot 2J'd(x,y)^\alpha \le 2J'\hat{h}_*^{-1}\|f\|_\infty d(x,y)^\alpha,$$

where $\hat{h}_*$ is the essential lower bound of $\hat{h}$ given by Sublemma 6.3.

By using Assumption T$''$(e), the third sum in (6.5) is bounded by

$$\begin{aligned}\sum_{j=1}^{K}\sum_{i=N+1}^{\infty} \frac{f(y_{ij})}{\hat{h}(y_{ij})} \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} &\le \hat{h}_*^{-2}\|\hat{h}\|_\infty \|f\|_\infty \cdot K'b\varepsilon^{m+\alpha} \\ =&\hat{h}_*^{-2}\|\hat{h}\|_\infty C_b \|f\|_{\mathcal{B}} \cdot K'b\varepsilon^m d(x,y)^\alpha = C_b K' b \varepsilon_1^m \hat{h}_*^{-2}\|\hat{h}\|_\infty \|f\|_{\mathcal{B}} d(x,y)^\alpha,\end{aligned}$$

where C_b is given in Lemma 5.1 which depends on ε_0.

Hence the result of part (1) holds with $D_{\mathcal{H}} = C_b\hat{h}_*^{-1}(2J' + K'b\varepsilon_1^m \hat{h}_*^{-1}\|\hat{h}\|_\infty)$.

Part (ii) and (iii) can be proved by using the same estimates with the same adjustments as in the proof of Lemma 5.2. □

Sublemma 6.3. *There is a $\hat{h}_* > 0$ such that $\hat{h}(x) \ge \hat{h}_*$ for ν-a.e. $x \in \widehat{X}$.*

Proof. By Lemma 3.1 in [Ss], there is a ball $B_\varepsilon(z) \subset \widehat{X}$ such that $\operatorname*{Einf}_{B_\varepsilon(x)} \hat{h} \ge \hat{h}_-$ for some constant $\hat{h}_- > 0$. By Assumption T$''$(c$'$), there is $\widetilde{N} > 0$ such that $T^{\widetilde{N}} B_\varepsilon(z) \supset X$. Then for any $x \in \widetilde{X}$, there is $y_0 \in B_\varepsilon(z)$ such that $T^{\widetilde{N}} y_0 = x$. Since $|\det DT|$ is bounded above, we have $g_* := \inf\{g(y) :\ y \in X\} > 0$. Hence, for $\hat{\nu}$-almost every x,

$$\hat{h}(x) = (\mathscr{P}^{\widetilde{N}}\hat{h})(x) = \sum_{T^{\widetilde{N}}y = x} \hat{h}(y) \prod_{i=0}^{\widetilde{N}-1} g(T^i y) \ge \hat{h}(y_0)\prod_{i=0}^{\widetilde{N}-1} g(T^i y_0) \ge \hat{h}_- g_*^{\widetilde{N}}.$$

The result follows with $\hat{h}_* = \hat{h}_- g_*^{\widetilde{N}}$. □

Sublemma 6.4. *Let ε_0 be as in Lemma 5.2. Then there exists $J_{\hat{h}} > 0$ and $\varepsilon_* \in (0, \varepsilon_0]$ such that for any $x, y \in V_I$ with $d(x,y) \le \varepsilon_*$, $I \in \mathcal{I}$,*

$$\frac{\hat{h}(x)}{\hat{h}(y)} \le e^{J_{\hat{h}} d(x,y)^\alpha}.$$

Proof. Since $\hat{h}$ is the unique fixed point of $\widehat{\mathscr{P}}$, we know that $\hat{h} = \lim_{n\to\infty} \widehat{\mathscr{P}}^n 1_{\widehat{X}}$, where the convergence is in $L^1(\hat{\nu})$. Now we consider the sequence $f_n := \widehat{\mathscr{P}}^n 1_{\widehat{X}}$.

We will prove that there is $J_{\hat{h}} > 0$ and $\varepsilon_* \in (0, \varepsilon_0]$ such that for any $n \ge 0$, for any $x, y \in V_I$, $I \in \mathcal{I}$, with $d(x,y) \le \varepsilon_*$,

$$\frac{f_n(y)}{f_n(x)} \le e^{J_{\hat{h}} d(x,y)^\alpha}. \tag{6.8}$$

Clearly (6.8) is true for $n = 0$ since $f_0(x) = 1$ for any x. We assume that it is true up to f_{n-1}. Consider f_n.

Note that $f_n/\hat{h} = (1/\hat{h})\widehat{\mathscr{P}}^n(h \cdot 1_{\widehat{X}}/\hat{h}) = \widehat{\mathcal{L}}^n(1_{\widehat{X}}/\hat{h})$, where $\widehat{\mathcal{L}}$ is a normalized transfer operator defined by $\widehat{\mathcal{L}}(f) = (1/\hat{h})\widehat{\mathscr{P}}(\hat{h}f)$. Then there are $f_* \ge \hat{h}_*/\hat{h}^*$ and $f^* \le \hat{h}^*/\hat{h}_*$ such that $f_* \le f_n(x) \le f^*$ for every $x \in \widehat{X}$ and $n \ge 0$, where $\hat{h}^*$ and $\hat{h}_*$ are the essential upper and lower bound of $\hat{h}$ respectively. Let also set: $g_* = \inf_x f_1(x) = \inf_x \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(x_{ij})$.

Let us set again $b = 1$. Then put $J > 0$ as in Assumption T$''$(e). Let us take $J_{\hat{h}} > 2Js^\alpha/(1 - s^\alpha)$ so that we have $(J_{\hat{h}} + J)s^\alpha \le J_{\hat{h}}(1 + s^\alpha)/2$. Then we take $\varepsilon_* \in (0, \varepsilon_0]$ small enough such that for any $\varepsilon \in [0, \varepsilon_*]$,

$$e^{J_{\hat{h}}(1+s^\alpha)\varepsilon^\alpha/2} + \frac{f^* K' b \varepsilon^{m+\alpha}}{f_*(g_* - K' b \varepsilon^{m+\alpha})} \le e^{J_{\hat{h}}\varepsilon^\alpha}.$$

For any x, y in the same V_I with $d(x,y) =: \varepsilon \le \varepsilon_*$, we choose $N = N(\varepsilon)$ as in Assumption T$''$(e). Let us denote with $[f_n]_N(x) = \sum_{j=1}^K \sum_{i=0}^N \hat{g}(x_{ij}) f_{n-1}(x_{ij})$ and $\{f_n\}_N(x) = f_n(x) - [f_n]_N(x) = \sum_{j=1}^K \sum_{i=N+1}^\infty \hat{g}(x_{ij}) f_{n-1}(x_{ij})$. We have

$$\begin{aligned}\frac{[f_n]_N(y)}{[f_n]_N(x)} &= \frac{\sum_{j=1}^K \sum_{i=0}^N \hat{g}(y_{ij}) f_{n-1}(y_{ij})}{\sum_{j=1}^K \sum_{i=0}^N \hat{g}(x_{ij}) f_{n-1}(x_{ij})} \\ \le \sup_{1\le j\le K; 0<i\le N} & e^{J d(x_{ij}, y_{ij})^\alpha} e^{J_{\hat{h}} d(x_{ij}, y_{ij})^\alpha} \le e^{(J + J_{\hat{h}}) s^\alpha d(x,y)^\alpha} \le e^{J_{\hat{h}}(1+s^\alpha)\varepsilon^\alpha/2}.\end{aligned}$$

We also get

$$\{f_n\}_N(y) = \sum_{j=1}^K \sum_{i=N+1}^\infty \hat{g}(y_{ij}) f_{n-1}(y_{ij}) \le f^* \sum_{j=1}^K \sum_{i=N+1}^\infty \hat{g}(y_{ij}) \le f^* K' b e^{m+\alpha}.$$

On the other hand,

$$[f_n]_N(x) = \sum_{j=1}^K \sum_{i=N+1}^\infty \hat{g}(y_{ij}) f_{n-1}(y_{ij}) \ge f_* \sum_{j=1}^K \sum_{i=1}^N \hat{g}(y_{ij}) \ge f_*(g_* - K' b e^{m+\alpha}).$$

By the choice of ε_*, we obtain

$$\frac{f_n(y)}{f_n(x)} \le \frac{[f_n]_N(y) + \{f_n\}_N(y)}{[f_n]_N(x)} \le e^{J_{\hat{h}}(1+s^\alpha)\varepsilon^\alpha/2} + \frac{f^* K' b\varepsilon^{m+\alpha}}{f_*(g_* - K' b\varepsilon^{m+\alpha})} \le e^{J_{\hat{h}}\varepsilon^\alpha}.$$

This means (6.8) for n since we have set $\varepsilon = d(x,y)$. □

Lemma 6.5. *There exists a constant $C_R > 0$ such that $\|R_n\|_{\mathcal{B}} \le C_R d_n^{m/(m+\alpha)}$ for all $n > 0$.*

If, moreover, T satisfies Assumption $T''(e')$, then $\|R_n\|_{\mathcal{B}} \le C_R d_n$ for all $n > 0$.

Proof. Since $R_i = \sum_j R_{ij}$, we only need to prove the results for R_{ij}.

Let $s_{ij}(x)$ be the norm of $\|D\widehat{T}_{ij}^{-1}(x)\|$, and $s_{ij} = \max\{s_{i,j}(x) : x \in B_{\varepsilon_0}(Q_0)\}$. Note that $\{\tau > i\} \subset T^{-1}V$ for all large i. We may suppose that i is sufficiently large so that $B_{s_{ij}\varepsilon_1}(U_{ij}) \subset \widehat{T}_{ij}^{-1}V$.

Take $f \in \mathcal{B}$ with $\|f\|_{\mathcal{B}} = 1$.

By using (5.20) and (5.21), we apply arguments similar to (5.22) and get

$$\|R_{ij}f\|_1 = \int_{U_{ij}} |f| d\hat{\nu} \le \|f\|_\infty \hat{\nu}(U_{ij}) \le C_b \hat{\nu}(Q_0) d_{ij} \|f\|_{\mathcal{B}}. \tag{6.9}$$

Next, we consider $|R_{ij}f|_{\mathcal{B}}$. Note that for any $I \in \mathcal{I}$, $f|_{V_I} \in \mathcal{H}^\alpha(V_I, H)$ for some $H \le \|f\|_{\mathcal{B}}$. So $\operatorname{osc}\big(f/\hat{h},\ B_{s\varepsilon}(\cdot)\big) \le 2^\alpha s^\alpha \varepsilon^\alpha H \le 2^\alpha s^\alpha \varepsilon^\alpha \|f\|_{\mathcal{B}}$. Note that Sublemma 6.4 implies $\operatorname{osc}\big(\hat{h},\ B_\varepsilon(x)\big) \le 2^\alpha J'_{\hat{h}} \varepsilon^\alpha$ for all x with $B_\varepsilon(x) \in V_I$ for some $J'_{\hat{h}} \ge J_{\hat{h}} > 0$. By Proposition 3.2(3) in [Ss],

$$\operatorname{osc}\big(f,\ B_{s_{ij}\varepsilon}(\cdot)\big) \le \operatorname{osc}\big(f/\hat{h},\ B_{s_{ij}\varepsilon}(\cdot)\big)\hat{h}_* + \operatorname{osc}\big(\hat{h},\ B_{s_{ij}\varepsilon}(\cdot)\big)\|f\|_\infty/\hat{h}_* \le b_1\varepsilon^\alpha \|f\|_{\mathcal{B}},$$

where $b_1 = 2^\alpha(H\hat{h}_* + J'_{\hat{h}} C_b h_*^{-1}) s_{ij}^\alpha$. By arguments similar to (5.20) and (5.21),

$$\begin{aligned} \int R_{ij} \operatorname{osc}\big(f,\ B_{s_{ij}\varepsilon}(\cdot)\big) d\hat{\nu} &= \int_{U_{ij}} \operatorname{osc}\big(f,\ B_{s_{ij}\varepsilon}(\cdot)\big) d\hat{\nu} \\ &\le b_1\varepsilon^\alpha \|f\|_{\mathcal{B}} \hat{\nu}(U_{ij}) \le b_1 \varepsilon^\alpha d_{ij} \hat{\nu}(Q_0) \|f\|_{\mathcal{B}} \le a_1 \varepsilon^\alpha d_{ij} \|f\|_{\mathcal{B}}, \end{aligned} \tag{6.10}$$

where $a_1 = b_1 \nu(Q_0)$. Also,

$$\hat{\nu}\big(\widehat{T}_{ij}^{-1} B_\varepsilon(\partial \widehat{T} U_{ij})\big) = \int_{B_\varepsilon(\partial \widehat{T} U_{ij})} \hat{g} d\hat{\nu} \le d_{ij} \cdot \hat{\nu}\big(B_\varepsilon(\partial U_0)\big) \le d_{ij} \cdot b_2 \varepsilon,$$

for some $b_2 > 0$ independent of ε. Hence,

$$G_{ij}(x, \varepsilon, \varepsilon_0) = 2d_{ij} \cdot b_2\varepsilon / \hat{\nu}(B_{(1-s)\varepsilon_0}(x)) \le a_2 d_{ij} \varepsilon, \tag{6.11}$$

where $a_2 = 2b_2/\hat{\nu}(B_{(1-s)\varepsilon_0}(x))$. Note that $\int \operatorname{osc}(f, B_{\varepsilon_0}(x_{ij})) d\hat{\nu} \le \varepsilon_0^\alpha |f|_{\mathcal{Q}}$, and $\|f\|_1 + \varepsilon_0^\alpha |f|_{\mathcal{Q}} \le \|f\|_{\mathcal{Q}} \le \|f\|_{\mathcal{B}}$. So for any $\varepsilon \in (0, \varepsilon_0]$ and $i < N(\varepsilon)$, we use (5.15), (6.10), (6.9) and (6.11) to get

$$\begin{aligned} |R_{ij}f|_{\mathcal{Q}} &\le \big[(1+\zeta\varepsilon^\alpha)a_1 + 2\zeta C_b \nu(Q_0) + 2(1+\zeta\varepsilon^\alpha)a_2\varepsilon^{1-\alpha}\big] d_{ij}\|f\|_{\mathcal{B}} \\ &\le C'_2 d_{ij} \|f\|_{\mathcal{B}}, \end{aligned}$$

where $C_2' = (1+\zeta\varepsilon^\alpha)a_1 + 2\zeta C_b \nu(Q_0) + 2(1+\zeta\varepsilon^\alpha)a_2\varepsilon^{1-\alpha}$.

For $\varepsilon \in (0, \varepsilon_0]$ and $i > N(\varepsilon)$, by Assumption T$''$(e) we have $d_{ij} \le b\varepsilon^{m+\alpha}$. Hence, $\varepsilon^{-a} \le (b^{-1}d_{ij})^{-\alpha/(m+\alpha)}$. So by (5.16), we have

$$\begin{aligned} |R_{ij}f|_{\mathcal{Q}} \le & 2(\gamma_m \varepsilon_0^m)^{-1} \cdot \|f\|_{\mathcal{Q}} \cdot \varepsilon^{-\alpha} \cdot d_{ij} \\ \le & 2(\gamma_m \varepsilon_0^m)^{-1} b^{\alpha/(m+\alpha)} d_{ij}^{1-\alpha/(m+\alpha)} \|f\|_{\mathcal{Q}} = C_2'' d_{ij}^{m/m+\alpha} \|f\|_{\mathcal{B}}, \end{aligned} \tag{6.12}$$

where $C_2'' = 2(\gamma_m\varepsilon_0^m)^{-1}b^{\alpha/(m+\alpha)}$. Therefore we get that $|R_{ij}f|_{\mathcal{Q}} \le C_2 d_i^{m/m+\alpha}$, where $C_2 = \max\{C_2', C_2''\}$.

Now we consider $|R_{ij}f|_{\mathcal{H}}$. As in the proof of Lemma 6.2, for any $x, y \in U_{ij}$,

$$\begin{aligned} \Big|\frac{R_{ij}f(x)}{\hat{h}(x)} - \frac{R_{ij}f(y)}{\hat{h}(y)}\Big| \le & \Big|\frac{\hat{g}(x_{ij})f(x_{ij})}{\hat{h}(x)} - \frac{\hat{g}(y_{ij})f(y_{ij})}{\hat{h}(y)}\Big| \\ = & \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)}\Big|\frac{f(x_{ij})}{\hat{h}(x_{ij})} - \frac{f(y_{ij})}{\hat{h}(y_{ij})}\Big| \\ & + \frac{|f(y_{ij})|}{\hat{h}(y_{ij})}\Big|\frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{\hat{h}(x)} - \frac{\hat{g}(y_{ij})\hat{h}(y_{ij})}{\hat{h}(y)}\Big|. \end{aligned} \tag{6.13}$$

Note that $\Big|f(x_{ij})/\hat{h}(x_{ij}) - f(y_{ij})/\hat{h}(y_{ij})\Big| \le |f|_{\mathcal{H}} d(x_{ij}, y_{ij})^\alpha \le \|f\|_{\mathcal{B}} s_{ij}^\alpha d(x,y)^\alpha$. Also, $\hat{g}(x_{ij})\hat{h}(x_{ij})/\hat{h}(x) \le (\hat{h}^*/\hat{h}_*)d_{ij}$. Then the first term in the right hand side of (6.13) is bounded by $a_3 d_{ij}\|f\|_{\mathcal{B}} d(x,y)^\alpha$, where $a_3 = (\hat{h}^*/\hat{h}_*)s_{ij}^\alpha$.

Let us take $\varepsilon = d(x,y)$; if $i \le N(\varepsilon)$, then by (6.7),

$$|\hat{g}(x_{ij})\hat{h}(x_{ij})/\hat{h}(x) - \hat{g}(y_{ij})\hat{h}(y_{ij})/\hat{h}(y)| \le 2J'(\hat{h}^*/\hat{h}_*)d_{ij}d(x,y)^\alpha.$$

Since $f(y_{ij})/\hat{h}(y_{ij}) \le \|f\|_\infty/\hat{h}_* \le C_b\hat{h}_*^{-1}\|f\|_{\mathcal{B}}$, the last term in (6.13) is bounded by $a_4 d_{ij}\|f\|_{\mathcal{B}} d(x,y)^\alpha$, where $a_4 = 2C_b J'(\hat{h}^*/\hat{h}_*^2)$. Therefore we obtain $|R_{ij}f|_{\mathcal{H}} \le C_3' d_{ij}\|f\|_{\mathcal{B}}$, where $C_3' = b_1 + b_2$.

If $i \ge N(\varepsilon)$, then by the first inequality of (6.13), the left side of the inequality is bounded by $\max\{\hat{g}(x_{ij})f(x_{ij})/\hat{h}(x), \hat{g}(y_{ij})\hat{h}(y_{ij})/\hat{h}(y)\} \le d_{ij}\|f\|_\infty/\hat{h}_*$. By the same arguments as for (6.12) we can get that

$$|R_{ij}f|_{\mathcal{H}} \le \varepsilon^{-\alpha} d_{ij}\|f\|_\infty/\hat{h}_* \le C_b\hat{h}_*^{-1}b^{\alpha/(m+\alpha)}d_{ij}^{m/(m+\alpha)}\|f\|_{\mathcal{B}} = C_3'' d_{ij}^{m/(m+\alpha)}\|f\|_{\mathcal{B}},$$

where $C_3'' = C_b\hat{h}_*^{-1}b^{\alpha/(m+\alpha)}\|f\|_{\mathcal{B}}$. Then we conclude that $|R_{ij}f|_{\mathcal{H}} \le C_3 d_{ij}^{m/(m+\alpha)}\|f\|_{\mathcal{B}}$, where $C_3 = \max\{C_3', C_3''\}$.

The conclusion of the first part follows by setting $C_R = C_1 + C_2 + C_3$.

If T satisfies Assumption T$''$(e$'$), then we can regard $N(\varepsilon) = \infty$ for any $\varepsilon > 0$. Hence we get $\|R_{ij}f\|_{\mathcal{B}} \le C_R d_{ij}\|f\|_{\mathcal{B}}$ with $C_R = C_1 + C_2 + C_3'$. □

Acknowledgments

We wish to thank Ian Melbourne for his interest in our maps, for his helpful comments and advices.

We would like to thank also Romain Aimino who was trapped in interminable discussions on aperiodicity and functional spaces.

References

[AD] J. Aaronson and M. Denker, *Local limit theorems for Gibbs-Markov maps*, Mathematisches Institut der Universitä Göttingen, Göttingen, 1997

[ADSZ] J. Aaronson, M. Denker, O. Sarig and R. Zweimüeller, Aperiodicity of cocycles and conditional local limit theorems, *Stoch. & Dynam.*, **4** (2004), 31–62

[ABV] J.F. Alves, C. Bonatti and M. Viana, SRB measures for partially hyperbolic systems whose central direction in mostly expanding, *Invent. Math.*, **140**, (2000), 351–398

[BG] A. Boyarsky and P. Góra, *Laws of Chaos : Invariant Measures and Dynamical Systems in One Dimension*, Probability and its Applications, Birkhauser, 1997

[Br] A. Broise, *Transformations dilatantes de l'intervalle et théorèmes limites*, 5-110, Astérisque, **238** (1996)

[Go] S. Gouëzel, Sharp polynomial estimates for the decay of correlations, *Israel J. Math.*, **139** (2004), 29–65

[He] H. Hennion, Sur un théorème spectral et son application aux noyaux lipchitziens, *Proc. Amer. Math. Soc.*, **118** (1993), 627–634

[HH] H. Hennion and L. Hervé, *Limit theorems for Markov chains and Stochastic Properties of Dynamical Systems by Quasicompactness, Lect. Notes Math.*, 1766, Springer-Verlag, 2001

[Hu] H. Hu, Decay of correlations for piecewise smooth maps with indifferent fixed points, *Ergodic Theory Dynam. Systems*, **24** (2004), 495–524

[HPT] H. Hu, Ya. Pesin and A. Talitskaya, A Volume Preserving Diffeomorphism with Essential Coexistence of Zero and Nonzero Lyapunov Exponents *Comm. Math. Phys.* **319** (2013), 331–378

[HV] H. Hu and S. Vaienti, Absolutely Continuous Invariant Measures for Nonuniformly Expanding Maps, *Ergodic Theory Dynam. Systems*, **29** (2009), 1185–1215

[IM] C.T. Ionescu Tulcea and G. Marinescu, Théorie ergodique pour des classes d'opérations non complètement continues (French), *Ann. of Math.*, **52** (1950), 140–147

[Kk] S. Kakutani, Induced measure preserving transformations, *Proc. Imp. Acad. Tokyo*, **19** (1943), 635–641

[KS] S. Kochen and C. Stone, A note on the Borel-Cantelli lemma, *Ill. J. Math.*, **8** (1964), 248–251

[Kr] U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985.

[LY] A. Lasota and J. Yorke, On the existence of invariant measures for piecewise monotonic transformations, *Trans. Amer. Math. Soc.*, **186**, (1973) 481–488

[LSV] C. Liverani, B.Saussol and S. Vaienti, A probabilistic approach to intermittency, *Ergodic Theory Dynam. Systems*, **19** (1999), 671–685

[PP] W. Parry and M Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque*, **187-188** (1990)

[Ps] Ya. Pesin, Characteristic Ljapunov exponents, and smooth ergodic theory (Russian), *Uspehi Mat. Nauk*, **32** (1977), 55–112

[PY] M. Pollicott and M. Yuri, Statistical Properties of maps with indifferent periodic points, *Comm. Math. Phys.*, **217** (2001), 503–520

[Qu] A. Quas, Non-ergodicity for C^1 expanding maps and g-measures (English summary), *Ergodic Theory Dynam. Systems*, **16** (1996), 531–543

[Sr] O. Sarig, Subexponential decay of correlations, *Invent. Math.*, **150** (2002), 629–653

[Ss] B. Saussol, Absolutely continuous invariant measures for multidimensional expanding maps, *Israel J. Math.*, **116** (2000), 223–248

[Ya] J.-A. Yan, A Simple Proof of Two Generalized Borel-Cantelli Lemmas in *Lecture Notes in Mathematics*, 1874, Springer-Verlag, 2006, 77–79

[Yo1] L.-S. Young, Statistical properties of dynamical systems with some hyperbolicity, *Ann. of Math.*, **147** (1950), 585–650

[Yo2] L.-S. Young, Recurrence times and rates of mixing, *Israel J. Math.*, **110** (1999), 153–188

[Yr] M. Yuri, Invariant measures for certain multi-dimensional maps, *Nonlinearity*, **7** (1994), 1093–1124

[Zm] W. Zeimer, *Weakly Differentiable Functions*, Graduate Text in Mathematics, 120, Springer (1995)

[Z1] R. Zweimüeller, Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points, *Nonlinearity*, **11** (1998), 1263–1276

[Z2] R. Zweimüeller, Ergodic properties of infinite measure-preserving interval maps with indifferent fixed points, *Ergodic Theory Dynam. Systems*, **20** (2000), 1519–1549