RUELLE-PERRON-FROBENIUS SPECTRUM FOR ANOSOV MAPS

MICHAEL BLANK, GERHARD KELLER, AND CARLANGELO LIVERANI

ABSTRACT. We extend a number of results from one dimensional dynamics based on spectral properties of the Ruelle-Perron-Frobenius transfer operator to Anosov diffeomorphisms on compact manifolds. This allows to develop a direct operator approach to study ergodic properties of these maps. In particular, we show that it is possible to define Banach spaces on which the transfer operator is quasicompact. (Information on the existence of an SRB measure, its smoothness properties and statistical properties readily follow from such a result.) In dimension $d = 2$ we show that the transfer operator associated to smooth random perturbations of the map is close, in a proper sense, to the unperturbed transfer operator. This allows to obtain easily very strong spectral stability results, which in turn imply spectral stability results for smooth deterministic perturbations as well. Finally, we are able to implement an Ulam type finite rank approximation scheme thus reducing the study of the spectral properties of the transfer operator to a finite dimensional problem.

1. INTRODUCTION

The aim of the present paper is to extend a number of results from one dimensional dynamics based on spectral properties of the Ruelle-Perron-Frobenius operator (transfer operator, for short) to $\mathcal{C}^3$ Anosov diffeomorphisms on compact manifolds, which allows to develop a direct, and very powerful, operator approach to study quantitative ergodic properties of these maps.

Since the spectrum of an operator depends heavily on the space on which it is defined, the first task is to identify a relevant function space. Note that on function spaces where the norm measures essentially the size of the functions (like $\mathcal{C}^0$ or L^1), the spectrum of a transfer operator carries only little dynamically relevant information: The L^1-spectrum is always contained in $\{|z| = 1\}$, and for other L^p-spaces as well as for $\mathcal{C}^0$ the same is true if the map is volume preserving.[1] Spaces of more regular functions (like $\mathcal{C}^1$ or $W^{1,1}$), though they are well adapted to the dynamics of Anosov diffeomorphisms in unstable directions, are unsuitable to capture the dynamics in stable directions. In fact they would lead to unreasonable spectral radii larger than one. These observations suggest to use strongly anisotropic Banach spaces of generalised functions (where the anisotropy is spatially inhomogeneous as soon as the diffeomorphism is nonlinear). Accordingly, our basic result is:

(1) We define inhomogeniously anisotropic Banach spaces of generalised functions on which the transfer operator is quasicompact and give an explicit bound for the essential spectral radius of the transfer operator in terms of expansion and contraction rates of the diffeomorphism. (Information on the existence of an SRB measure, its smoothness properties and its statistical properties readily follow from such a result.)

In dimension $d = 2$ we also prove the following spectral perturbation results:

(2) Transfer operators associated to certain random perturbations of the map are close, in a proper sense, to the unperturbed transfer operator. This allows to obtain easily very strong spectral stability results which imply, in particular, stability of the invariant measure, of the rate of mixing, of the variance in the central limit theorem, *etc.*
(3) Transfer operators associated to maps that are close to each other in the $\mathcal{C}^2$ topology have "close" spectrum, although they are defined on different Banach spaces.

Date: March 30, 2001.

We acknowledge the partial support of the ESF Programme PRODYN and the DFG priority research program DANSE. The work of M.B. was partially supported by DFG, INTAS, CRDF and RFBR grants. We wish to thank V.Baladi, M.Benedicks, U.Bessi, B.Hasselblatt, A.Katok, F.Ledrappier and J.Schmeling for helpful discussions. G.K. and C.L. acknowledge the hospitality of the I.H.E.S. where this paper was finished and thank D.Ruelle for making this possible.

[1] Note, however, that the Koopman operator (which in the volume preserving case is the adjoint of the transfer operator on L^2) was studied thoroughly, and some qualitative dynamic information like ergodicity, weak mixing and mixing can be read off its spectral measure, see *e.g.* [10].

(4) It is possible to use an Ulam-like discretisation scheme that reduces the computation of the spectral data of the transfer operator to the study of a finite dimensional matrix. This allows to obtain precise quantitative information on the ergodic properties of a given map (e.g. ergodicity, invariant measures, mixing, rate of mixing etc.) and it is a partial, but rather satisfactory, solution to the so called *Ulam Conjecture*.

The randomly perturbed operators from (2) and also the finite rank approximations from (4) are compact operators that can be studied as well on well known function spaces like L^p, $\mathcal{C}^0$, $\mathcal{C}^1$. On all these spaces they have the same pure point spectrum, and a way to interpret (2) and (4) is to say that the *actually observed* (either in real systems where noise is always present or in numerical approximation schemes) dynamically relevant spectrum consists of eigenvalues of finite multiplicity and is rather independent of the employed function space.

The precise choice of the Banach spaces of generalised functions is suggested by two observations: Restricted to an unstable fibre the diffeomorphism acts like a (piecewise) expanding map. Therefore the generalised densities should behave like functions of bounded variation along such fibres. On the other hand, the action of the map on stable fibres is similar to the action of a contractive map. Thus the corresponding Banach space should be equipped with a norm related to the weak convergence of measures. This is the starting point of our construction and according to this idea the norm (2.1.7) consists of two parts – one related to the properties 'along' the unstable foliation and another one related to the stable foliation. As a result the corresponding Banach space heavily depends on the structure of these foliations. This is one of the weaknesses of the approach, because generically different maps lead to different Banach spaces. Nevertheless, it is still possible to "compare" the corresponding transfer operators by means of a random smoothing, as we will see in the discussion of deterministic perturbations.

The Banach spaces introduced are in fact indirectly related to the *cone method* used to investigate rates of mixing in [31]. It is worth mentioning that the idea to study directly spectral properties of smooth hyperbolic systems was proposed earlier by Bakhtin [1] in the analysis of space averagings and also by Rugh [37] and Kitaev [30] in their investigation of dynamical (Ruelle) zeta-functions. On the other hand, the analysis of stochastic stability of spectral properties was earlier done only in the case of one-dimensional piecewise expanding maps with convolution-like random perturbations by Baladi and Young [3] and by Blank, Keller and Liverani for more general classes of perturbations and maps in [4, 27, 32]. A more specific problem about the stochastic stability of the leading eigenfunction – the density of the SRB measure – was firstly studied by Kifer [28] in the case of Gaussian-like random perturbations of smooth hyperbolic maps.

Outline of the work. We start by discussing $\mathcal{C}^3$ Anosov maps on the d-dimensional torus. In this context we *define a proper Banach space* in section 2.1 and prove *quasicompactness of the transfer operator*, see section 2.2. The restriction to $\mathbb{T}^d$ is motivated only by our choice to present the method in the clearest form and not obscured by technical problems. Along the discussion we point out the trivial (for people acquainted with the basic facts of differential geometry) modifications needed to extend the setting to arbitrary compact Riemannian manifolds. The restriction to $\mathcal{C}^3$ Anosov maps could be slightly relaxed by requiring that T and also its Jacobian are of class $\mathcal{C}^2$. At present we do not know how to avoid the additional assumption on the Jacobian. We illustrate the general quasicompactness result by a *simple example*: We give an independent elementary proof of the fact that transfer operators of hyperbolic linear automorphisms of $\mathbb{T}^2$ have no eigenvalue (other than 1) of modulus bigger than the contracting eigenvalue of the defining matrix.

The quasicompactness of the transfer operator allows to derive rather directly the existence and (well known) essential *properties of the Sinai-Ruelle-Bowen measure* of T, see section 2.3.

In section 2.4 we consider *small random perturbations of Anosov maps* of the two-dimensional torus $\mathbb{T}^2$. Here, as above, the choice of the torus rather than any two dimensional compact Riemannian manifold is due to our effort not to cloud the ideas presented by unnecessary technical details (yet, again, we will comment on how to treat the more general setting). On the contrary, the restriction to two dimensions is of a substantial nature. The reason is that in two dimensions the stable and unstable distributions are known to be smooth (at least $\mathcal{C}^{1+\alpha}$) while in higher dimensions distributions can be only Hölder, depending on the extreme expansion rates in the stable and unstable bundle of the tangent space (see [25] or the appendix for precise references). It is not hard to see that, if the distributions are smooth, then our results can easily be extended. For non smooth distributions the situation is more complex. It is possible that such strong stability results are unreasonable in this case and that only weaker results can be obtained. In particular, the spectrum could be stable only far away from the essential spectrum.

In section 2.5 we investigate *smooth deterministic perturbations of Anosov maps* on $\mathbb{T}^2$. Here the problem is that the transfer operators are defined on different Banach spaces, hence no direct comparison is possible. The key to the comparison is given by the fact that the random perturbations are compact, hence their spectrum is rather robust when changing the underlying space. Therefore the strategy is to compare the spectrum of random perturbations (that can be viewed as operators on the same space) and then to use the previous results to conclude the argument.

Finally, we turn to *Ulam finite rank approximation schemes* for the transfer operator $\mathcal{L}$ associated with an Anosov map on $\mathbb{T}^2$. Here the first problem is that the original Ulam scheme consists in taking a conditional expectation with respect to some partition of the manifold. So the natural space of functions to consider must contain functions with discontinuities along unstable fibres, unlike the functions belonging to our Banach space. It turns out, however, that this is not the main problem. Indeed, in section 2.7 we discuss an extension of our space which also contains piecewise constant functions and on which the transfer operator still has basically the same spectral properties as on the original space. Nevertheless it can happen that the discretised operator of the original Ulam scheme has "large" eigenvalues quite far from the spectrum of the given transfer operator. An example of this pathology is presented in section 2.6. The problem can be overcome, similarly as in the previous section, by introducing additional smooth random perturbations (compare to the regularisation by random perturbations in [5, 11, 29]).

In order to facilitate "navigation" in this paper, we provide a detailed list of contents:

Contents

1. Introduction 1
2. Statements and results 4
2.1. Banach spaces of generalised functions 4
2.2. Quasi-compactness of the transfer operator $\mathcal{L}$ on $\mathcal{B}(\mathcal{M})$ 6
2.3. Peripheral spectrum and SRB measures 10
2.4. Spectral stability – Smooth random perturbations 11
2.5. Spectral stability – Deterministic perturbations. 14
2.6. Spectral stability – Ulam finite rank approximation 15
2.7. Remarks on an extension of the space $\mathcal{B}(\mathcal{M})$ 18
3. Proofs 19
3.1. Proofs: Lasota-Yorke type estimate 19
3.2. Proofs: Peripheral spectrum and SRB measures 25
3.3. Proofs: Spectral stability – Smooth random perturbations 28
3.4. Proofs: Spectral stability – Deterministic Perturbations 40
3.5. Proofs: Spectral stability – Ulam finite rank approximation 44
3.6. Proofs: Compact embedding 45
3.7. Proofs: Unstable seminorm 50
Appendix A. On splitting and foliation regularity 54
References 57

Open questions. This article does not attempt to exhaust the possibilities of this new method but it rather exposes the method in some simple but non trivial cases. It seems clear to us that a considerable amount of work is still necessary to understand all the consequences of this approach. To extend the spectral stability results to higher dimensional systems and hence to cases where the stable and unstable foliation are only Hölder continuous is a first task. Then it would be interesting to investigate systems with discontinuities and different types of perturbations (e.g., pure Ulam approximation schemes). In addition, our Banach spaces do not have the amount of symmetry between the stable and the unstable direction that would seem natural. Also, we cannot take advantage of the higher smoothness of the map (in the spirit of Kitaev's work [30]) and this certainly requires a deeper understanding. Finally, we would like to mention that, having at disposal a well defined operator, the question of the relation between its isolated eigenvalues and the poles of the dynamical (Ruelle) zeta function becomes quite a natural one.

2. Statements and results

2.1. Banach spaces of generalised functions. We will consider Anosov diffeomorphisms

$T : \mathcal{M} \to \mathcal{M}$ where T is of class $\mathcal{C}^3$ and $\mathcal{M}$ is a smooth compact Riemannian manifold.

Let us remind that by Anosov we mean (as usual) that there exist a direct sum decomposition of the tangent bundle $\mathcal{TM}$ into continuous sub-bundles E^s and E^u, that is $\mathcal{T}_x\mathcal{M} = E^s_x \oplus E^u_x$, and constants $A \geq 1$ and $0 < \lambda_s < 1 < \lambda_u$ such that[2]

$$
\begin{aligned}
(d_xT)(E^s_x) &= E^s_{Tx}, \quad \|(d_xT^n)_{|E^s_x}\| \leq A\lambda_s^n, \\
(d_xT)(E^u_x) &= E^u_{Tx}, \quad \|(d_xT^{-n})_{|E^u_x}\| \leq A\lambda_u^{-n},
\end{aligned}
\tag{2.1.1}
$$

for all $x \in \mathcal{M}$ and $n \geq 0$. Let $d_{u/s} = \dim(E^{u/s})$. It is well known that for such maps there exist stable and unstable foliations $(W^s(x))_{x\in\mathcal{M}}$ and $(W^u(x))_{x\in\mathcal{M}}$. Each single $W^{s/u}(x)$ is an immersed $\mathcal{C}^3$ sub-manifold of $\mathcal{M}$, and $\mathcal{T}_yW^{s/u}(x) = E^{s/u}_x$ for any $y \in W^{s/u}(x)$. The dependence of $E^{s/u}(x)$ and $W^{s/u}(x)$ on x, however, is only Hölder in general–see the appendix for further precise information on the relevant properties used throughout the paper. We will denote by τ the optimal common Hölder-exponent for both distributions. This exponent depends in a well understood way on various contraction and expansion coefficients of the transformation, and there are a number of cases where the foliations are indeed $\mathcal{C}^{1+\alpha}$ for some $\alpha > 0$.

Some of the results of this paper are proved only under this stronger $\mathcal{C}^{1+\alpha}$-assumption. However, in order to make transparent where in the proofs we really use the $\mathcal{C}^{1+\alpha}$-property, we will argue with the $\mathcal{C}^\tau$-property with $0 < \tau \leq 1$ whenever this is possible without any trouble, and use the $\mathcal{C}^{1+\alpha}$-property with $0 < \alpha \leq 1$ only in more delicate estimates. (This includes to set $\tau = 1$ if necessary.) We believe that, with considerably more effort, one could weaken the $\mathcal{C}^{1+\alpha}$-assumption.

The map T induces naturally a map on the Borel measures on $\mathcal{M}$ defined by

$$T^*\mu(f) := \mu(f \circ T) \quad \forall \mu \in M(\mathcal{M}, \mathbb{R}) \text{ and } f \in \mathcal{C}^0(\mathcal{M}, \mathbb{R}).$$

It is immediate to verify that, if μ is absolutely continuous with respect to the Riemannian volume m, then $T^*\mu$ is absolutely continuous with respect to m. Given this fact, it is possible to define the evolution of the corresponding densities:

$$\mathcal{L}\frac{d\mu}{dm} := \frac{d(T^*\mu)}{dm}.$$

The above defined operator $\mathcal{L}$ is usually called the Perron-Frobenius or the Ruelle-Perron-Frobenius or the transfer operator. There are numerous papers where such operators are used to investigate statistical properties of expanding systems. The aim of this paper is to show that such a strategy can be successfully extended to Anosov systems.

A direct computation shows that $\mathcal{L}$ has the following representation:

$$\mathcal{L}f = f \circ T^{-1} \cdot \frac{d(T^*m)}{dm} =: f \circ T^{-1} \cdot g \quad \forall f \in L^1(\mathcal{M}, m) \tag{2.1.2}$$

where $g \in \mathcal{C}^2(\mathcal{M})$.[3]

To study the spectrum of the operator $\mathcal{L}$ we need to specify the space on which such an operator acts. Indeed it is quite clear that the spectrum of $\mathcal{L}$ on $L^1(\mathcal{M}, m)$ carries only little interesting information, *cf.* the discussion in the introduction and footnote [1]. Instead we construct a space of generalised functions which behave more or less like functions of bounded variation along unstable manifolds whereas they look like signed measures (equipped with a weak topology) along stable manifolds.

We start by defining a suitable set of test functions to control the stable direction. For points $x, y \in \mathcal{M}$ with $y \in W^s(x)$ we define $d^s(x, y)$ as the distance between x and y within the Riemannian manifold $W^s(x)$ (which inherits its Riemannian structure from $\mathcal{M}$). Fix some $\delta > 0$. For $0 < \beta \leq 1$

---

[2] By dT we denote the differential of T, clearly $d_xT : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_{Tx}\mathcal{M}$. Similarly, if $f : \mathcal{M} \to \mathbb{R}$ is differentiable, then $d_xf : \mathcal{T}_x\mathcal{M} \to \mathbb{R}$.

[3] Note that, if $\mathcal{M} = \mathbb{T}^d$, then in (2.1.2) $g(x) = |\det(D_xT^{-1})|$.

and bounded measurable $\varphi : \mathcal{M} \to \mathbb{R}$, we define[4]

$$H^s_\beta(\varphi) := \sup_{d^s(x,y)\le\delta} \frac{|\varphi(x)-\varphi(y)|}{d^s(x,y)^\beta} . \tag{2.1.3}$$

Here the supremum is taken over all pairs of points x and y such that $y \in W^s(x)$. Clearly, H^s_β is a seminorm, and we use it to define

$$\mathcal{D}_\beta := \{\varphi : \mathcal{M} \to \mathbb{R} \,:\, \varphi \text{ measurable}, |\varphi|_\infty \le 1, H^s_\beta(\varphi) \le 1\}. \tag{2.1.4}$$

In order to control the unstable direction we provide a set $\mathcal{V}$ of measurable test vector fields $v : \mathcal{M} \to \mathcal{TM}$ adapted to the unstable foliation in the sense that $v(x) \in E^u_x$ for all $x \in \mathcal{M}$.

Given the fact that $x \mapsto E^u_x$ is, in general, only τ-Hölder for some $\tau < 1$ we cannot ask the vector fields to be globally more regular than that. By a slight abuse of notation we define[5]

$$H^s_\beta(v) := \sup_{d(x,y)^s\le\delta} \frac{\|v(x)-v(y)\|}{d^s(x,y)^\beta} . \tag{2.1.5}$$

Then we will consider the vector fields

$$\mathcal{V}_\beta := \{v \in \mathcal{V} \,:\, |v|_\infty \le 1;\; H^s_\beta(v) \le 1\} . \tag{2.1.6}$$

From now on we will always assume that

$$0 < \beta < \gamma \le 1 .$$

Using the above defined classes of test functions and test vector fields we now define the norms that will describe our Banach spaces of generalised functions. For $f \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$ let

$$\begin{aligned} \|f\|_s &:= \sup_{\varphi\in\mathcal{D}_\beta} \textstyle\int_{\mathcal{M}} f\varphi\, dm \\ \|f\|_u &:= \sup_{v\in\mathcal{V}_\beta} \textstyle\int_{\mathcal{M}} df(v)dm \\ \|f\| &:= \|f\|_u + b\|f\|_s \\ \|f\|_w &:= \sup_{\varphi\in\mathcal{D}_\gamma} \textstyle\int_{\mathcal{M}} f\varphi\, dm. \end{aligned} \tag{2.1.7}$$

where $\int_{\mathcal{M}} df(v)dm$ is short hand for $\int_{\mathcal{M}} d_x f(v(x))m(dx)$. The constant $b \ge 1$ will be specified later. Except for $\|\cdot\|_u$, which is only a seminorm, all these expressions define norms on $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ and $\|f\|_w \le \|f\|_s \le b^{-1}\|f\|$. Note that the above norms are inhomogeniously anisotropic because the stable and unstable directions are treated differently and may change from point to point.

Definition 2.1.1. *$\mathcal{B}(\mathcal{M})$ and $\mathcal{B}_w(\mathcal{M})$ denote the completions of $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ w.r.t. the norms $\|\cdot\|$ and $\|\cdot\|_w$, respectively.*

Each $f \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$ naturally gives rise to a bounded linear functional on $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ by virtue of

$$\langle f, \varphi\rangle := \int_{\mathcal{M}} f\varphi\, dm .$$

Obviously, $\|f\|^*_{\mathcal{C}^1} \le \|f\|_w \le \|f\| \le \|f\|_{\mathcal{C}^1}$. Therefore there exist canonical continuous embedding (not necessarily one-to-one)

$$\mathcal{C}^1(\mathcal{M},\mathbb{R}) \to \mathcal{B}(\mathcal{M}) \to \mathcal{B}_w(\mathcal{M}) \to \mathcal{C}^1(\mathcal{M},\mathbb{R})^* .$$

In fact, each $f \in \mathcal{B}_w$ defines a bounded linear functional on $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ by $\langle f,\varphi\rangle := \lim_{n\to\infty}\langle f_n,\varphi\rangle$ where $f_n \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$ and $\lim_{n\to\infty}\|f - f_n\|_w = 0$. In the same way one can embed $\mathcal{B}(\mathcal{M})$ into $\mathcal{B}_w(\mathcal{M})$.

[4]The δ in the definition is fixed once and for all, yet it must satisfy various smallness requirements that will be specified as we proceed in the discussion.

[5]To compute the difference between two tangent vectors at different (close) points we parallel transport one of them to the tangent space of the other along the geodesic. We will not mention this explicitly, since it is completely trivial on $\mathcal{M} = \mathbb{T}^d$ and it is a routine operation on general Riemannian manifolds, see *e.g.* [13, Section 2.3].

Remark 2.1.2. *Note that although the embedding of $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ into each of the other three spaces is one-to-one this need not be true for the other embedding. We have a closer look at the embedding of $\mathcal{B}(\mathcal{M})$ into $\mathcal{B}_w(\mathcal{M})$, but the same remarks apply to the embedding into $\mathcal{C}^1(\mathcal{M},\mathbb{R})^*$.*

As a functional on the dense subspace $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ of $\mathcal{B}(\mathcal{M})$ the map $f \mapsto \|f\|_w$ is Lipschitz continuous, so it extends to all of $\mathcal{B}(\mathcal{M})$. This extension may be only a seminorm, however. Therefore, if we consider $\mathcal{B}(\mathcal{M})$ as a linear subspace of $\mathcal{B}_w(\mathcal{M})$, we really mean $\mathcal{B}(\mathcal{M})/N_w \hookrightarrow \mathcal{B}_w(\mathcal{M})$ where N_w is the closed linear subspace $\{f \in \mathcal{B}(\mathcal{M}) : \|f\|_w = 0\}$.

Nevertheless, with an abuse of notation, we write $\mathcal{B}(\mathcal{M}) \subset \mathcal{B}_w(\mathcal{M})$, where the inclusion must be interpreted in the above sense. In particular, for $f \in \mathcal{B}(\mathcal{M})$, $\|f\|_w$ can denote the seminorm $\|\cdot\|_w$ of f (which generally will be the case), or it can be the norm of the element $f + N_w \in \mathcal{B}_w(\mathcal{M})$. Both interpretations are equivalent for our purposes.

Remark 2.1.3. *Since the space $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ is separable, since the norms $\|\cdot\|$ and $\|\cdot\|_w$ are weaker than the $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ norm, and since $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ is dense both in $\mathcal{B}(\mathcal{M})$ and $\mathcal{B}_w(\mathcal{M})$ by construction, it follows that the Banach spaces $\mathcal{B}(\mathcal{M})$ and $\mathcal{B}_w(\mathcal{M})$ are separable.*

Note that the unstable seminorm is given by a formula that does not extend to all of $\mathcal{B}(\mathcal{M})$, since for general elements of that space there is no way to make sense of the differential. The obvious way to compute $\|f\|_u$ for general $f \in \mathcal{B}(\mathcal{M})$ would be to approximate f by functions from $\mathcal{C}^1(\mathcal{M},\mathbb{R})$. Since this is not always desirable, it is sometimes convenient to write $\|\cdot\|_u$ in a way that applies directly to any element of $\mathcal{B}(\mathcal{M})$. The following Lemma is step in this direction.

Lemma 2.1.4. *Let $U \subset \mathcal{M}$ be a sufficiently small open set. There are τ-Hölder vector fields $\{V_i\}_{i=1}^{d_u}$ in U which, at each point $x \in U$, form a basis of $E^u(x)$ and which, restricted to each unstable manifold, are $\mathcal{C}^1$ vector fields.*

These vector fields can be constructed such that there exist constants $C > 0$ and $\beta' \in (0,\tau)$ with the following property: Given any $\varphi \in \mathcal{C}^\beta(\mathcal{M},\mathbb{R}^{d_u})$ with $\operatorname{supp}\varphi \subset U$ which is $\mathcal{C}^1$ when restricted to any unstable manifold, let $V_\varphi := \sum_i \varphi_i V_i$. Then $C^{-1}V_\varphi \in \mathcal{V}_\beta$, and there exists a measurable function $A : \mathcal{M} \to \mathbb{R}^{d_u}$ such that $C^{-1}A_i \in \mathcal{D}_{\beta'}(\mathcal{M})$ and such that, for each $f \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$,

$$\int_{\mathcal{M}} df(V_\varphi)\, dm = -\int_{\mathcal{M}} f \sum_i d\varphi_i(V_i)\, dm + \int_{\mathcal{M}} f \sum_i A_i\varphi_i\, dm.$$

Observe that $d\varphi_i(V_i)$ is well defined because $V_i(x) \in E^u(x)$ for all x, although the φ_i need only be Hölder continuous in stable direction. This point is further clarified in the proof.

Remark 2.1.5. *On the torus $\mathbb{T}^2$ there is the normalised (and oriented) unstable vector field $V_1 = v^u$ which is even of class $\mathcal{C}^{1+\alpha}$ for some $\alpha > 0$ (see the appendix for details). Hence $A_1 = -\operatorname{div}(V_1)$ in this case.*

Since it is easy to check that the choice of a different basis changes only the functions A_i in the above formula one could define the unstable norm - locally - by

$$\sup_{\varphi\in\mathcal{C}^1(U,\mathbb{R}^{d_u})} \frac{\int_{\mathcal{M}} f \sum_i d\varphi_i(V_i)\, dm}{|\varphi|_\infty + H^s_\beta(\varphi)}$$

provided one is willing to choose $\beta \leq \beta'$. Yet, for this reason and since it would be a bit cumbersome to extend globally the above definition, we prefer to work with the norm (2.1.7). For the special case $\mathcal{M} = \mathbb{T}^2$ a related problem is encountered in the proof of Lemma 2.7.1.

The proof of Lemma 2.1.4 is best done in local coordinates introduced in section 3.6 and so it is postponed to section 3.7.

2.2. **Quasi-compactness of the transfer operator $\mathcal{L}$ on $\mathcal{B}(\mathcal{M})$.** To simplify the exposition we will give precise statements and complete proofs only for the case $\mathcal{M} = \mathbb{T}^d$. Nevertheless we will indicate for the results of this subsection and in the course of their proofs the modifications necessary in order to deal with general manifolds.

Our assumption that T is of class $\mathcal{C}^3$ implies that the *Jacobian* g of T (see (2.1.2)) is of class $\mathcal{C}^2$; in fact, we will use only that is of Class $\mathcal{C}^{1+\beta}$.

Our first result is a Lasota-Yorke type inequality.

Lemma 2.2.1. *Suppose $\beta < \min\{\tau, 1\}$ and $\gamma \in (\beta, 1]$. Then $\mathcal{L}$ extends naturally to a bounded linear operator on both $\mathcal{B}_w(\mathcal{M})$ and $\mathcal{B}(\mathcal{M})$. In addition, for each $\sigma > \max\{\lambda_u^{-1}, \lambda_s^\beta\}$, we can* *choose constants b and δ in (2.1.3) - (2.1.7) for which there exists $B>0$ such that, for each $f \in \mathcal{B}(\mathcal{M})$, we have*

$$\|\mathcal{L}^n f\|_w \le A\|f\|_w \qquad \text{and} \qquad \|\mathcal{L}^n f\| \le 3A^2\sigma^n\|f\| + B\|f\|_w \qquad \text{for } n=1,2,\dots$$

where A is the constant from (2.1.1).

This clearly implies that the spectral radius of $\mathcal{L}$ is bounded by one both on $\mathcal{B}_w(\mathcal{M})$ and $\mathcal{B}(\mathcal{M})$. Thus $\mathrm{sp}(\mathcal{L}) \subseteq \{z \in \mathbb{C} : |z| \le 1\}$. To exploit Lemma 2.2.1 further, we need the following fact.

Proposition 2.2.2. *If $\gamma \cdot \min\{\tau, 1\} > \beta$, then the ball $\mathcal{B}_1 := \{f \in \mathcal{B}(\mathcal{M}) : \|f\| \le 1\}$ is relatively compact in $\mathcal{B}_w(\mathcal{M})$.*[6]

The proof of Lemma 2.2.1 is given in section 3.1, that of Proposition 2.2.2 in section 3.6.
From these two results the quasicompactness of $\mathcal{L}$ follows along classical lines [22]. A more recent argument, which is based on the Nussbaum formula, also provides an estimate on the essential spectral radius:

Theorem 1. *Suppose that $\gamma \cdot \min\{\tau, 1\} > \beta$. Then, for each $\sigma > \max\{\lambda_u^{-1}, \lambda_s^{\beta}\}$, the operator $\mathcal{L} : \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{M})$ has essential spectral radius bounded by σ and is thus quasicompact.*

An immediate consequence of Theorem 1 is that for each constant $r \in (\sigma, 1)$ the portion[7]

$$\mathrm{sp}_r(\mathcal{L}) := \mathrm{sp}(\mathcal{L}) \setminus \{z \in \mathbb{C} : |z| \le r\}$$

of the spectrum consists of finitely many eigenvalues $\lambda_1, \dots, \lambda_p$ of finite multiplicity. Denote by $\Pi_1, \dots, \Pi_p$ the corresponding spectral projectors and let $\Pi := \mathrm{Id}_{\mathcal{B}} - \sum_{j=1}^{p} \Pi_j$. Then $\mathrm{rank}(\Pi_j) < \infty$ $(j = 1, \dots, p)$, and the spectral radius $\rho_{\mathrm{sp}}(\mathcal{L}\Pi) \le r$. The next lemma shows that $\mathrm{sp}_r(\mathcal{L})$ and the spectral projectors Π_j are the same on $\mathcal{B}(\mathcal{M})$ and on $\mathcal{B}(\mathcal{M})/N_w$, *cf.* Remark 2.1.2.

Lemma 2.2.3. *Restricted to $(\mathrm{Id}_{\mathcal{B}} - \Pi)(\mathcal{B}(\mathcal{M}))$, the seminorm $\|\cdot\|_w$ dominates the norm $\|\cdot\|$ (so that its restriction is actually a norm).*

Proof. Since $\|\mathcal{L}^{-n}(\mathrm{Id}_{\mathcal{B}} - \Pi)\| \le \mathrm{const} \cdot r^{-n}$, we have for $f = (\mathrm{Id}_{\mathcal{B}} - \Pi)f$ and each $n \in \mathbb{N}$

$$\|f\| = \|\mathcal{L}^{-n}\mathcal{L}^n f\| \le \mathrm{const} \cdot r^{-n}\,\|\mathcal{L}^n f\| \le \mathrm{const} \cdot A^2 \left(\frac{\sigma}{r}\right)^n \|f\| + \mathrm{const} \cdot r^{-n} B\|f\|_w$$

where we used Lemma 2.2.1 for the last inequality. Choosing n such that $\mathrm{const} \cdot A^2 \left(\frac{\sigma}{r}\right)^n \le \frac{1}{2}$ it follows that $\|f\| \le 2\,\mathrm{const} \cdot r^{-n} B\|f\|_w$. □

Remark 2.2.4. *Although the norms $\|\cdot\|$ and $\|\cdot\|_w$ depend on constants b and δ which will be determined in the course of the proofs, they are equivalent for any two choices of these constants. So the spectral assertions in Theorem 1 do not depend on them.*

Proof of Theorem 1. For the reader's convenience we recall the brief argument from [20] which deduces the estimate on the essential spectral radius from Nussbaum's formula [34]. Let $\epsilon > 0$. Since $\mathcal{B}_1$ is relatively compact in $\mathcal{B}_w$, there are $f_1, \dots, f_N \in \mathcal{B}_1$ such that $\mathcal{B}_1 \subseteq \bigcup_{i=1}^{N} U_\epsilon(f_i)$, where $U_\epsilon(f_i) = \{f \in \mathcal{B} : \|f - f_i\|_w < \epsilon\}$. For $f \in \mathcal{B}_1 \cap U_\epsilon(f_i)$, Lemma 2.2.1 implies that

$$\|\mathcal{L}^n(f - f_i)\| \le 3A^2\sigma^n\,\|f - f_i\| + \frac{3A^2B}{1-\sigma}\|f - f_i\|_w \le 6A^2\sigma^n + \frac{3A^2B\epsilon}{1-\sigma}\ .$$

Choosing $\epsilon = \sigma^n$ we can conclude that for each $n \in \mathbb{N}$ the set $\mathcal{L}^n(\mathcal{B}_1)$ can be covered by a finite number of $\|\cdot\|$–balls of radius $C\sigma^n$, $C = 6A^2 + \frac{3A^2B}{1-\sigma}$. This is precisely what is needed to apply Nussbaum's formula, which states that the essential spectral radius of $\mathcal{L}$ is at most $\liminf_{n\to\infty} \sqrt[n]{C\sigma^n} = \sigma$. □

The above result allows to obtain a quite precise picture of the spectrum of $\mathcal{L}$, yet some more can be easily seen. First of all notice that $\mathcal{L}$ is invertible and $\mathcal{L}^{-1}$ is nothing else than the transfer operator associated to the map T^{-1}. Then the following bound (the proof can be found at the end of section 3.1) holds.

---

[6] This assertion can be interpreted in either of the two equivalent ways discussed in Remark 2.1.2.

[7] By $A \setminus B$ we mean the symmetric difference of the sets A and B.

Lemma 2.2.5. *Let μ_s, μ_u, $\mu_s \le \lambda_s < 1 < \lambda_u \le \mu_u$, the maximal contraction and expansion rate of T, respectively.*[8] *Then we have the following bound for the spectral radius of $\mathcal{L}^{-1}$:*

$$\rho_{sp}(\mathcal{L}^{-1}) \le \mu_u \mu_s^{-1} \ .$$

Accordingly, the essential spectrum of $\mathcal{L}$ must be contained in the annulus $\{z \in \mathbb{C} : \mu_u\mu_s^{-1} \le |z| \le \sigma\}$ as shown in figure 1.

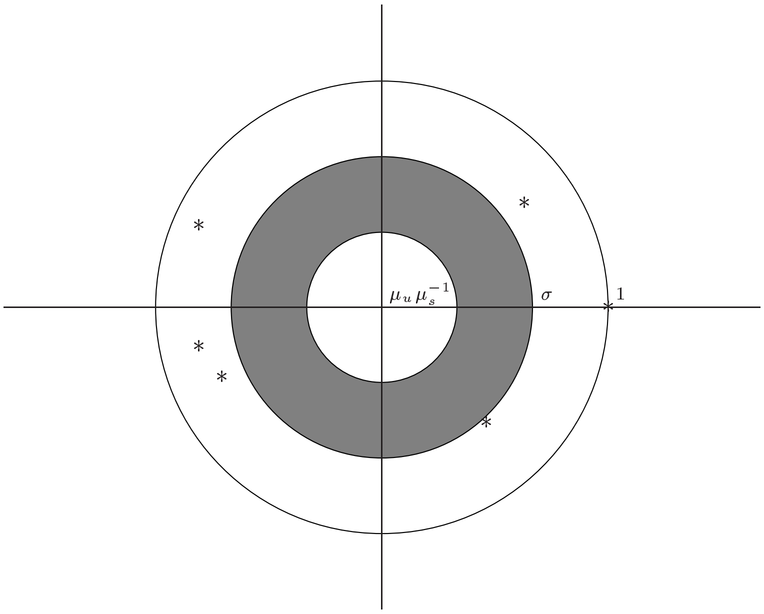

FIGURE 1. Region containing the spectrum of the transfer operator

Example 2.2.6 (*Eigenvalues of linear automorphisms of $\mathbb{T}^2$*)**.** Let $\mathbb{A}$ be a 2×2 integer matrix with $\det(\mathbb{A}) = 1$ and trace different from $\{-2, \dots, 2\}$. Then $\mathbb{A}$ has eigenvalues $\lambda_\mathbb{A}$ and $\lambda_\mathbb{A}^{-1}$ with $|\lambda_\mathbb{A}| > 1$, and $T(x) := \mathbb{A}x \mod \mathbb{Z}^2$ defines a hyperbolic automorphism of $\mathbb{T}^2$. The stable and unstable fibres are just the straight lines in direction of the eigenvectors for $\lambda_\mathbb{A}^{-1}$ and $\lambda_\mathbb{A}$, respectively. Therefore $\tau = 1+\alpha \ge 2$, and we can choose for the definition of our norms $\gamma = \alpha = 1$ and $\beta \in (0,1)$ as close to one as we wish. Denote by $\mathcal{L}$ the transfer operator for T. We show here that for any choice of $\beta \in (0,1)$

$$\text{eigenvalues}(\mathcal{L}) \subseteq \{z \in \mathbb{C} : |z| \le \lambda_\mathbb{A}^{-1}\} \cup \{1\} \tag{2.2.1}$$

and that 1 is a simple eigenvalue. In view of the bound for the essential spectral radius of $\mathcal{L}$ established in Theorem 1 it follows that

$$\text{sp}(\mathcal{L}) \subseteq \left\{z \in \mathbb{C} : |z| \le \lambda_\mathbb{A}^{-\beta}\right\} \cup \{1\} \ .$$

This fact is not only a concrete illustration of our Theorem 1, but also its proof introduces in a highly simplified setting some of the ideas that will be worked out later in more generality. Therefore, in spite of the relative simplicity, we provide many details.

So suppose that $\mathcal{L}f = \lambda f$ with $|\lambda| > \lambda_\mathbb{A}^{-1}$ and $f \in \mathcal{B}_{\beta_0}$ (where the index of $\mathcal{B}$ indicates the actual choice of β we make). Since $\mathcal{D}_{\beta_0} \supseteq \mathcal{D}_\beta$ for each $\beta \in [\beta_0, 1)$, we have $\mathcal{B}_{\beta_0} \subseteq \mathcal{B}_\beta$ and hence also $f \in \mathcal{B}_\beta$ for each such β. In particular we can choose $\beta \in [\beta_0, 1)$ such that $|\lambda| > \lambda_\mathbb{A}^{-\beta}$.

Since $\int \mathcal{L}h\, dm = \int h\, dm$ for all $h \in \mathcal{C}^1(\mathbb{T}^2)$, it follows that $\lambda\langle f, 1\rangle = \langle \mathcal{L}f, 1\rangle = \langle f, 1\rangle$, so that either $\langle f, 1\rangle = 0$ or $\lambda = 1$. In the latter case, if 1 is not a simple eigenvalue, there nevertheless exists some $f \in \mathcal{B}(\mathbb{T}^2)$ with $\mathcal{L}f = f$ and $\langle f, 1\rangle = 0$. (Observe that because of $\|\mathcal{L}\|_w \le 1$ the eigenvalue 1 is semisimple.) So we will assume $\langle f, 1\rangle = 0$ from now on.

Our goal is to show that $\|f\| = 0$. As a first step we consider $\|f\|_u$. Let $v \in \mathcal{V}_\beta$ (see (2.1.6)) and observe that v is in direction of the unstable eigenvector of $\mathbb{A}$. Observing the fact that T leaves

[8]By this we mean $\|DT^{-n}\| \le A\mu_s^n$, $\|DT^n\| \le A\mu_u^n$.

the Lebesgue measure on $\mathbb{T}^2$ invariant, we have for each $h \in \mathcal{C}^1(\mathbb{T}^2)$

$$\begin{aligned}\int_{\mathbb{T}^2} d(\mathcal{L}h)(v)\,dm &= \int_{\mathbb{T}^2} d(h\circ T^{-1})(v)\,dm = \int_{\mathbb{T}^2} dh_{T^{-1}x}(\mathbb{A}^{-1}(v(x)))\,m(dx)\\ &= \int_{\mathbb{T}^2} dh_x(\mathbb{A}^{-1}(v(Tx)))\,m(dx) = \lambda_{\mathbb{A}}^{-1}\int_{\mathbb{T}^2} dh(v\circ T)\,dm\end{aligned}$$

Since $|v\circ T|_\infty = |v|_\infty$ and since $H^s_\beta(v\circ T) \le H^s_\beta(v)$ (T contracts in stable direction!), also $v\circ T \in \mathcal{V}_\beta$, and it follows that $\|\mathcal{L}h\|_u \le \lambda_{\mathbb{A}}^{-1}\|h\|_u$. Of course, this estimate extends to all elements of $\mathcal{B}(\mathbb{T}^2)$ and in particular to f. So $\|f\|_u = |\lambda^{-1}|\cdot\|\mathcal{L}f\|_u \le |\lambda\lambda_{\mathbb{A}}|^{-1}\cdot\|f\|_u$, which implies $\|f\|_u = 0$ because $|\lambda\lambda_{\mathbb{A}}| > 1$ by assumption.

In order to show that also $\|f\|_s = 0$, we use the flow S_t which moves points at unit speed along unstable fibres. Then, for each $h \in \mathcal{C}^1(\mathbb{T}^2)$ and $x \in \mathbb{T}^2$, integration by parts yields[9]

$$\frac{1}{2r}\int_{-r}^{r} h(S_t x)\,dt = h(x) + \frac{1}{2r}\int_{-r}^{r} h_1(S_t x)\,\theta(t,r)\,dt$$

where $h_1(x) = d_x h(v^u)$, v^u denoting the normalised unstable eigenvector, and $\theta(t,r) = -t + r\cdot\mathrm{sign}(t)$. So, if $\varphi \in \mathcal{D}_\beta$, then

(2.2.2)

$$\begin{aligned}&\langle h,\varphi\rangle\\
&= \int_{\mathbb{T}^2}\left(\frac{1}{2r}\int_{-r}^{r} h(S_t x)\,dt\right)\varphi(x)\,m(dx) - \int_{\mathbb{T}^2}\left(\frac{1}{2r}\int_{-r}^{r} h_1(S_t x)\,\theta(t,r)\,dt\right)\varphi(x)\,m(dx)\\
&= \frac{1}{2r}\int_{-r}^{r}\left(\int_{\mathbb{T}^2} h(S_t x)\,\varphi(x)\,m(dx)\right)dt - \frac{1}{2r}\int_{-r}^{r}\left(\int_{\mathbb{T}^2} h_1(S_t x)\,\theta(t,r)\,\varphi(x)\,m(dx)\right)dt\\
&= \frac{1}{2r}\int_{-r}^{r}\left(\int_{\mathbb{T}^2} h(x)\,\varphi(S_{-t}x)\,m(dx)\right)dt - \frac{1}{2r}\int_{-r}^{r}\left(\int_{\mathbb{T}^2} h_1(x)\,\theta(t,r)\,\varphi(S_{-t}x)\,m(dx)\right)dt\\
&= \int_{\mathbb{T}^2}\left(\frac{1}{2r}\int_{-r}^{r}\varphi(S_{-t}x)\,dt\right)h(x)\,m(dx) - 2r\int_{\mathbb{T}^2} d_x h\left(\int_{-r}^{r}\varphi(S_{-t}x)\,\frac{1}{(2r)^2}\theta(t,r)\,dt\cdot v^u\right)m(dx)\end{aligned}$$

Suppose now that $w:\mathbb{R}\to\mathbb{R}$ is bounded with $\int_{-\infty}^{\infty}|w(t)|\,dt \le 1$. Define $\varphi_w:\mathbb{T}^2\to\mathbb{R}$ by $\varphi_w(x) = \int_{-\infty}^{\infty}\varphi(S_{-t}x)w(t)\,dt$. Then, if $y \in W^s(x)$ with $d(x,y) < \delta$,

$$|\varphi_w(y) - \varphi_w(x)| \le \int_{-\infty}^{\infty}|\varphi(S_{-t}x) - \varphi(S_{-t}y)|\,w(t)\,dt \le H^s_\beta(\varphi)\,d(x,y)^\beta\ ,$$

because $S_{-t}x \in W^s(S_{-t}y)$ and $d(S_{-t}x, S_{-t}y) = d(x,y)$. Since also $|\varphi_w|_\infty \le |\varphi|_\infty$, it follows that $\varphi_w \in \mathcal{D}_\beta$. Therefore, the integrals in brackets in the last line of (2.2.2) are, as functions of x, elements of $\mathcal{D}_\beta$. So, the second one multiplied with v^u is a vector field from $\mathcal{V}_\beta$. Hence

$$\left|\int_{\mathbb{T}^2} d_x h\left(\int_{-r}^{r}\varphi(S_{-t}x)\,\frac{1}{(2r)^2}\theta(t,r)\,dt\cdot v^u\right)m(dx)\right| \le \|h\|_u.$$

Accordingly, setting

$$\varphi_r := \frac{1}{2r}\int_{-r}^{r}\varphi(S_{-t}x)\,dt,$$

for each sequence $\{f_n\}\subset\mathcal{C}^1$ such that $f_n\to f$ in $\mathcal{B}(\mathbb{T}^2)$, holds

$$|\langle f,\varphi\rangle| = |\lim_{n\to\infty}\langle f_n,\varphi\rangle| \le \lim_{n\to\infty}|\langle f_n,\varphi_r\rangle| + 2r\lim_{n\to\infty}\|f_n\|_u = |\langle f,\varphi_r\rangle|. \tag{2.2.3}$$

On the other hand

$$\langle f,\varphi_r\rangle = \lambda^{-n}\langle\mathcal{L}^n f,\varphi_r\rangle = \langle f,\lambda^{-n}\varphi_r\circ T^n\rangle.$$

To conclude just notice that

$$\begin{aligned}H^s_\beta(\lambda^{-n}\varphi_r\circ T^n) &\le |\lambda|^{-n}\lambda_{\mathbb{A}}^{-n\beta}\\ |D^u(\lambda^{-n}\varphi_r\circ T^n)| &\le \frac{|\lambda|^{-n}\lambda_{\mathbb{A}}^{n}}{r},\end{aligned}$$

[9] Just integrate separately from $-r$ to 0 and from 0 to $+r$.

where D^u stands for the derivative along the unstable direction. Since estimate (2.2.3) is uniform in r we can choose $r = r_n := \lambda_{\mathbb{A}}^{2n}$, which implies that the β-Hölder constant (in all directions!) of $\hat{\varphi}_{r_n} := \lambda^{-n}\varphi_{r_n} \circ T^n$ is at most $2|\lambda|^{-n}\lambda_{\mathbb{A}}^{-n\beta}$. Accordingly

$$\left| \hat{\varphi}_{r_n} - \int_{\mathbb{T}^2} \hat{\varphi}_{r_n}\, dm \right|_\infty \le 2\delta^{-1+\beta}(|\lambda|\lambda_{\mathbb{A}}^{\beta})^{-n}.$$

The above estimate means that $\frac{1}{2}(\hat{\varphi}_{r_n} - \int \hat{\varphi}_{r_n}\, dm)\lambda^n\lambda_{\mathbb{A}}^{-n\beta}\delta^{1-\beta} \in \mathcal{D}_\beta$ for each $n \in \mathbb{N}$. Observing also that $\left|\int \hat{\varphi}_{r_n}\, dm\right| < \infty$ for all n and that $|\lambda|\lambda_{\mathbb{A}}^{\beta} > 1$ we conclude

$$|\langle f, \varphi\rangle| \le |\langle f, \hat{\varphi}_{r_n}\rangle| \le 2\delta^{-1+\beta}(|\lambda|\lambda_{\mathbb{A}}^{\beta})^{-n}\, \|f\|_s + \underbrace{|\langle f, 1\rangle|}_{=0} \left| \int_{\mathbb{T}^2} \hat{\varphi}_{r_n}\, dm \right| \to 0$$

as $n \to \infty$. Since this holds for any $\varphi \in \mathcal{D}_\beta$, this implies $\|f\|_s = 0$ and finishes the proof.

2.3. **Peripheral spectrum and SRB measures.** Recall that $\mathrm{sp}_r(\mathcal{L}) := \mathrm{sp}(\mathcal{L})\backslash\{z \in \mathbb{C} : |z| \le r\}$. If r is sufficiently close to 1, then $\mathrm{sp}_r(\mathcal{L})$ contains only eigenvalues of modulus 1, and since $\|\mathcal{L}\|_w \le 1$ (see Lemma 2.2.1), these eigenvalues are semisimple so that $\mathcal{L}\Pi_j = \lambda_j\Pi_j$ for $j = 1, \dots, p$. We will call *peripheral spectrum* the set $\{\lambda \in \mathrm{sp}(\mathcal{L}) : |\lambda| = 1\}$. This spectral decomposition allows to reproduce easily some well known facts from the ergodic theory of Anosov diffeomorphisms.

First of all, if $|\lambda| = 1$, then

$$\Pi_\lambda := \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \lambda^{-k}\mathcal{L}^k = \sum_{j=1}^{p} \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left(\frac{\lambda_j}{\lambda}\right)^k \Pi_j = \begin{cases} \Pi_j & \text{if } \lambda = \lambda_j \\ 0 & \text{otherwise} \end{cases} \tag{2.3.1}$$

exists in the uniform $\|\cdot\|$ - operator norm. In particular $\Pi_j = \Pi_{\lambda_j}$. Since, by Lemma 2.2.1,

$$\|\Pi_{\lambda_j} f\| \le \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \|\mathcal{L}^k f\| \le \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left(3A^2\sigma^k\|f\| + B\|f\|_w\right) = B\|f\|_w \tag{2.3.2}$$

for all $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$, the finite rank operators Π_j extend continuously to an operator from $\mathcal{B}_w(\mathcal{M})$ to $\mathcal{B}(\mathcal{M})$ and hence *a fortiori* from $\mathcal{C}^0(\mathcal{M}, \mathbb{R})$ to $\mathcal{B}(\mathcal{M})$ and also from $L^1(\mathcal{M}, \mathbb{R})$ to $\mathcal{B}(\mathcal{M})$. Since $\mathcal{C}^1(\mathcal{M}, \mathbb{R})$ is a dense linear subspace of $\mathcal{B}_w(\mathcal{M})$ and since each Π_j has finite rank, we have

$$\Pi_j(\mathcal{B}_w(\mathcal{M})) = \Pi_j(\mathcal{C}^1(\mathcal{M}, \mathbb{R})) = \Pi_j(\mathcal{B}(\mathcal{M}))\ . \tag{2.3.3}$$

In the following proposition we characterise these projections more precisely.

Proposition 2.3.1. *For each $h \in \Pi_j(\mathcal{B}(\mathcal{M}))$ there is a finite signed Borel measure μ_h on $\mathcal{M}$ such that $\langle h, \varphi\rangle = \int \varphi\, d\mu_h$ for all $\varphi \in \mathcal{D}_\gamma$. Moreover, $\lambda_1 := 1 \in sp(\mathcal{L})$, $\mu := \mu_{\Pi_1 1}$ is a positive measure, $\mu(\mathcal{M}) = m(\mathcal{M})$, and all μ_h are absolutely continuous with respect to μ and $|\frac{d\mu_h}{d\mu}| \le \|f\|_\infty$ μ-almost surely for each $f \in \mathcal{C}^1(\mathcal{M})$ such that $\Pi_1 f = h$. Finally,*

a) $T^*\mu_h = \mu_h$ for all $h \in \Pi_1(\mathcal{B}(\mathcal{M}))$ and
b) if $\lambda_1 = 1$ is a simple eigenvalue, then $\Pi_1 f = \langle f, 1\rangle \cdot \Pi_1 1$ for all $f \in \mathcal{B}_w(\mathcal{M})$.

In addition, it is possible to further characterise the invariant measures. The proofs of the above and the following proposition can be found in section 3.2.

Proposition 2.3.2. *The measure μ is a* Sinai-Ruelle-Bowen (SRB) *measure in the following sense:*

a) The measure μ is the weak Cesaro-limit of the Riemannian measure under the dynamics;
b) The ergodic decomposition of the measure μ is determined by the eigenspace of $\mathcal{L}$ associated to the eigenvalue 1. More precisely, there exists a finite basis $\{h_i\}$ of this eigenspace and measurable T- invariant sets $A_i \subseteq \mathcal{M}$ with $\mu(\bigcup_i A_i) = 1$, such that the measures $\mu_i(A) := \mu(A \cap A_i)$ are ergodic and, for each $\varphi \in \mathcal{D}_\gamma$, $\int \varphi\, d\mu_i = \langle h_i, \varphi\rangle$.
c) For all $\varphi \in \mathcal{C}^0(\mathcal{M}, \mathbb{R})$, the limit $\varphi^+(x) := \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n-1} \varphi \circ T^i(x)$ exists m-a.e. and takes only finitely many different values. If μ is ergodic, then $\varphi^+(x) = \int \varphi\, d\mu$ for m-almost every x.
d) For each $h \in \Pi_j(\mathcal{B}(\mathcal{M}))$ the measure μ_h, locally conditioned to the unstable foliation, is absolutely continuous on μ-almost every fibre with respect to the Riemannian measure on the fibre. In addition, the density is a function of bounded variation.

Observe that although the spectral assertions on $\mathcal{L}$ stated in Theorem 1 do not depend on the constants δ and b entering the norm $\|\cdot\|$ (see Remark 2.2.4), quantities like the norms of $\|\mathcal{L}\|$, $\|\mathcal{L}\Pi\|$ and $\|\Pi_j\|$ are affected.

2.4. **Spectral stability – Smooth random perturbations.** It is well known that the Sinai-Ruelle-Bowen measure of a mixing Anosov diffeomorphism is stable (in the sense of the weak topology on measures) under sufficiently smooth small random perturbations, see *e.g.* [28], [2] for the case of Gaussian-like random perturbations. Further results and a list of references on stochastic stability of SRB measures can be found in [5]. In terms of our spectral decomposition this means that the eigenvector $h = \Pi_1 1$ of $\mathcal{L}$ is stable under perturbations. In this section we assert much more: The full spectral structure of $\mathcal{L}$ outside the disk with radius σ (see Lemma 2.2.1) does not change much under small perturbations. As we noted already in the introduction, the smoothness of the stable and unstable foliations, which plays an essential rôle in the proofs, is automatically guaranteed only in the two-dimensional case where the foliations are $\mathcal{C}^{1+\alpha}$ for some $\alpha > 0$. Therefore, and in order to avoid some other technical problems, we give the full proof only for the case $\mathcal{M} = \mathbb{T}^2$. Nevertheless we describe the setting and prove some partial results in greater generality in order to point out more clearly where problems due to higher dimensions do arise.

We start by introducing a class of small random perturbations.

Definition 2.4.1. *a) A* kernel family *is a family of measurable functions $q_\varepsilon : \mathbb{T}^2 \times \mathbb{T}^2 \to \mathbb{R}^+$, $\varepsilon \in (0,1]$, such that*

$$q_\varepsilon(x,y) = 0 \text{ if } |x-y| > \varepsilon \ .$$

For $(x,\xi) \in \mathbb{T}^2 \times \mathbb{R}^2$ we denote[10]

$$\tilde{q}_\varepsilon(x,\xi) := q_\varepsilon(x, x+\xi) \ .$$

b) A stochastic kernel family *is a kernel family satisfying additionally*

$$\int_{\mathcal{M}} q_\varepsilon(x,y)\, m(dy) = 1 \quad \text{for all } x \in \mathbb{T}^2 \ .$$

c) An admissible kernel family *is a kernel family of continuously differentiable functions q_ε which satisfies some additional regularity properties. To state them we denote $\partial_1\tilde{q}_\varepsilon(x,\xi) := \frac{\partial}{\partial x}\tilde{q}_\varepsilon(x,\xi)$ and $\partial_2\tilde{q}_\varepsilon(x,\xi) := \frac{\partial}{\partial \xi}\tilde{q}_\varepsilon(x,\xi)$. We require that there is a stochastic kernel family $(p_\varepsilon)_{0<\varepsilon\le 1}$ such that for each $\varepsilon > 0$ the functions $\tilde{q}_\varepsilon$, $\partial_1\tilde{q}_\varepsilon$ and $\varepsilon \cdot \partial_2\tilde{q}_\varepsilon$ are ε^{-1}-*dominated *by $\tilde{p}_\varepsilon$ in the sense of the following definition.*

Definition 2.4.2. *Let $\tilde{q}_\varepsilon : \mathbb{T}^2 \times \mathbb{T}^2 \to \mathbb{R}^k$, $k \ge 1$, and let $(p_\varepsilon)_{0<\varepsilon\le 1}$ be a stochastic kernel family with the property that for each $0 < \varepsilon' \le \varepsilon$ there exists a constant $C > 0$ such that $\tilde{p}_{\varepsilon'}(x,\xi) \le C\,\tilde{p}_\varepsilon(x,\xi)$. We say that $\tilde{q}_\varepsilon$ is ε^{-1}-*dominated *by $\tilde{p}_\varepsilon$ if there are some $a, M > 0$ such that, for each $0 < \varepsilon \le 1$,*

$$|\tilde{q}_\varepsilon(x,\xi)| \le M\, \tilde{p}_{a\varepsilon}(x,\xi) \tag{2.4.1}$$

$$|\tilde{q}_\varepsilon(x,\xi) - \tilde{q}_\varepsilon(z,\xi)| \le M\, d(x,z)^\alpha\, \tilde{p}_{a\varepsilon}(x,\xi) \tag{2.4.2}$$

$$|\tilde{q}_\varepsilon(x,\xi) - \tilde{q}_\varepsilon(x,\zeta)| \le M\, \varepsilon^{-1}\, d(x+\xi, x+\zeta) \cdot \left(\tilde{p}_{a\varepsilon}(x,\xi) + \tilde{p}_{a\varepsilon}(x,\zeta)\right) \tag{2.4.3}$$

Remark 2.4.3. *Requirements (2.4.1) and (2.4.2) imply that $\tilde{q}_\varepsilon(z,\xi) \le 2M\tilde{p}_{a\varepsilon}(x,\xi)$ if $d(x,z) \le 1$.*

Remark 2.4.4. *A typical example, of which the present setting constitutes an obvious generalisation, is the perturbation on $\mathbb{T}^2$ given by*

$$q_\varepsilon(x,y) = \varepsilon^{-2}\bar{q}(\varepsilon^{-1}(y-x))$$

where $\bar{q}$ is nonnegative and continuously differentiable, $\operatorname{supp}(\bar{q}) \subset \{\xi \in \mathbb{R}^2 : |\xi| \le 1\}$, $\int \bar{q}(\xi)\, d\xi = 1$ *and $\inf\{\bar{q}(\xi) : |\xi| \le \frac{1}{2}\} =: \omega > 0$. In this case $\tilde{q}_\varepsilon(x,\xi) = \varepsilon^{-2}\bar{q}(\varepsilon^{-1}\xi)$, $\partial_1\tilde{q}_\varepsilon = 0$ and $\partial_2\tilde{q}(x,\xi) = \varepsilon^{-3}\bar{q}'(\varepsilon^{-1}\xi)$ so that $\tilde{q}_\varepsilon$, $\partial_1\tilde{q}_\varepsilon$ and $\varepsilon^{-1}\partial_2\tilde{q}_\varepsilon$ are ε^{-1}-dominated by $\tilde{q}_\varepsilon$ with $a = 2$ and $M = \omega^{-1}\|\bar{q}\|_{\mathcal{C}^2}$.*

[10] This definition can easily be extended to general manifolds $\mathcal{M}$. Just introduce, in a neighbourhood U of x, normal coordinates via the exponential map $\exp_x$ and identify isometrically the tangent space $\mathcal{T}_x\mathcal{M}$ with $\mathbb{R}^d$. Then define:

$$\tilde{q}_\varepsilon(z,\xi) := q_\varepsilon(z, \exp_x(\exp_x^{-1}(z) + \xi)).$$

Combined with a deterministic transformation T, any stochastic kernel q_ε gives rise to a Markov process $(X_n^\varepsilon)_{n\geq 0}$ on $\mathbb{T}^2$ where X_0 can have any probability distribution and

$$P(X_{n+1}^\varepsilon \in A \mid X_n^\varepsilon) = \int_A q_\varepsilon(T(X_n^\varepsilon), y)\, m(dy)\ .$$

Definition 2.4.5. *A family $(X_n^\varepsilon)_{n\geq 0}$ associated to an admissible stochastic kernel family is called a* small random perturbation *of T.*

If the distribution of X_n^ε has density f with respect to m, then the distribution of X_{n+1}^ε has density $Q_\varepsilon \mathcal{L} f$ where

$$Q_\varepsilon f(y) := \int_{\mathbb{T}^2} f(x) q_\varepsilon(x,y)\, m(dx)\ .$$

Obviously Q_ε extends to $\mathcal{B}_w(\mathbb{T}^2)$ by continuity. If $\varphi \in \mathcal{C}^0(\mathbb{T}^2,\mathbb{R})$ is an "observable", then the conditional expectation of $\varphi(X_{n+1}^\varepsilon)$ given X_n^ε is $Q_\varepsilon^*\varphi(TX_n^\varepsilon)$ where

$$Q_\varepsilon^*\varphi(x) = \int_{\mathbb{T}^2} q_\varepsilon(x,y)\varphi(y)\, m(dy)\ .$$

In particular, $Q_\varepsilon^* 1 = 1$. Both operators are dual in the sense that

$$\langle Q_\varepsilon f, \varphi\rangle = \langle f, Q_\varepsilon^*\varphi\rangle\ .$$

We will call the operators Q_ε *smooth averaging operators.*[11]

As already mentioned, the results of this section make much stronger use of regularity properties of the foliations than did those of the previous section. Indeed, we assume that the unstable distribution is $\mathcal{C}^\tau$, for some $\tau - 1 \geq \alpha \geq \gamma > \beta > 0$.[12]

The first basic estimate is stated in the following lemma.

Lemma 2.4.6. *Suppose that the unstable foliation is $\mathcal{C}^\tau$, $\tau - 1 \geq \alpha \geq \gamma > \beta > 0$. Then any smooth averaging operator can be extended in a unique way to a continuous operator on $\mathcal{B}(\mathcal{M})$ and one can choose δ (small) and b (large) in such a way that there exists $K > 0$ with*

$$\|Q_\varepsilon\|_w \leq K \qquad \text{and} \qquad \|Q_\varepsilon f\| \leq 3\|f\| + K\|f\|_w$$

for all sufficiently small $\varepsilon > 0$ and all $f \in \mathcal{B}(\mathcal{M})$.

Remember from the previous discussion that the transfer operator associated to the randomly perturbed system is defined by

$$\mathcal{L}_\varepsilon := Q_\varepsilon \mathcal{L}\ .$$

The operators $\mathcal{L}_\varepsilon$ satisfy a uniform Lasota-Yorke type inequality.

Lemma 2.4.7. *Assume that $\mathcal{M} = \mathbb{T}^2$. There is a constant $B' > 0$ such that for sufficiently small $\varepsilon > 0$ and each $n \in \mathbb{N}$*

$$\|\mathcal{L}_\varepsilon^n\|_w \leq B' \quad \text{and} \quad \|\mathcal{L}_\varepsilon^n f\| \leq 9A^2\sigma^n\|f\| + B'\|f\|_w$$

for all $f \in \mathcal{B}(\mathcal{M})$.

To state the next relevant fact it is helpful to introduce a special norm for operators that involves the strong and the weak norm and was used previously in [26, 27].

$$|||Q_\varepsilon||| := \sup_{\{f\in\mathcal{B}(\mathcal{M})\,:\,\|f\|\leq 1\}} \|Q_\varepsilon f\|_w.$$

Lemma 2.4.8. *Assume that $\mathcal{M} = \mathbb{T}^2$. There exists a constant $K > 0$ such that*

$$|||Q_\varepsilon - \mathrm{Id}\,||| \leq K\varepsilon^{\gamma-\beta}.$$

[11] It is interesting to notice - and it will be used in the following - that the product of smooth averaging operators is again a smooth averaging operator, *cf.* Lemma 3.3.5.

[12] This is always the case if the unstable foliation has codimension one or if the "gap" between the weakest and the strongest contraction rates in the stable sub-bundle is sufficiently small, see the appendix for more details.

Lemma 2.4.8 implies

$$|||\mathcal{L}_\varepsilon - \mathcal{L}||| \le \|\mathcal{L}\|\, K\, \varepsilon^{\gamma-\beta}\ . \tag{2.4.4}$$

The proofs of the last three lemmas are furnished in section 3.3. Estimate (2.4.4) together with Lemma 2.4.7 allow to supplement Theorem 1 with a rather strong result on spectral stability under smooth random perturbations. Indeed, the following is a direct consequence of [27].

To formulate the result precisely we need some preparations. Fix any $r \in (\sigma, 1)$ and denote by $\mathrm{sp}(\mathcal{L})$ the spectrum of $\mathcal{L} : \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{M})$. Since the essential spectral radius of $\mathcal{L}$ does not exceed σ, the set $\mathrm{sp}(\mathcal{L}) \cap \{z \in \mathbb{C} : |z| \ge r\}$ consists of a finite number of eigenvalues $\lambda_1, \dots, \lambda_p$ of finite multiplicity. Changing r slightly we may assume that $\mathrm{sp}(\mathcal{L}) \cap \{z \in \mathbb{C} :\ |z| = r\} = \emptyset$. Hence there is $\delta_* \in (0, r-\sigma)$ such that

$$\begin{gathered}|\lambda_i - \lambda_j| > \delta_* \quad (i \ne j)\\ \mathrm{dist}(\mathrm{sp}(\mathcal{L}), \{|z| = r\}) > \delta_*\end{gathered}$$

It is shown in [27, Theorem 1] that, given $\delta \le \delta_*$, for sufficiently small ε the spectral projectors

$$\begin{aligned}\Pi_\varepsilon^{(j,\delta)} &:= \frac{1}{2\pi\imath}\int_{\{|z-\lambda_j|=\delta\}} (z-\mathcal{L}_\varepsilon)^{-1}\, dz\\ \Pi_\varepsilon^{(r)} &:= \frac{1}{2\pi\imath}\int_{\{|z|=r\}} (z-\mathcal{L}_\varepsilon)^{-1}\, dz\end{aligned} \tag{2.4.5}$$

are well defined (and do not depend on δ). Finally we set $\eta := 1 - \frac{\log r}{\log \sigma}$ and observe that $0 < \eta < 1$.

Further results that follow directly from [27] can now be summarised as follows.

Theorem 2. *Let $T : \mathbb{T}^2 \to \mathbb{T}^2$ be a $\mathcal{C}^3$ Anosov diffeomorphism. For each $\delta \in (0, \delta_*]$ there exists $\varepsilon_0 = \varepsilon_0(\delta, r)$ such that for $\varepsilon \in [0, \varepsilon_0]$ holds:*

a) There is $K_1 = K_1(\delta, r) > 0$ such that $|||\Pi_\varepsilon^{(j,\delta)} - \Pi_0^{(j,\delta)}||| \le K_1\, \varepsilon^{\eta(\gamma-\beta)}$.
b) $\mathrm{rank}(\Pi_\varepsilon^{(j,\delta)}) = \mathrm{rank}(\Pi_0^{(j,\delta)})$
c) $\lim_{\varepsilon\to 0} |||\Pi_\varepsilon^{(r)} - \Pi_0^{(r)}||| = 0$.
d) There is $K_2 = K_2(\delta, r) > 0$ such that $\|\mathcal{L}_\varepsilon^n \Pi_\varepsilon^{(r)}\| \le K_2\, r^n$ for all $n \in \mathbb{N}$.

For later use we provide the key estimates from [27] that are behind the previous theorem.

Proposition 2.4.9. *In the situation of Theorem 2 there are constants $\varepsilon_0 = \varepsilon_0(\delta, r)$, $a = a(r)$, $b = b(\delta, r)$ and $c = c(\delta, r)$ such that for $\varepsilon \in [0, \varepsilon_0]$ and $z \in \mathbb{C}$ with $r < |z| \le 2$ and $\mathrm{dist}(z, sp(\mathcal{L})) > \delta$*

$$\|(z-\mathcal{L}_\varepsilon)^{-1} f\| \le a\, \|f\| + b\, \|f\|_w \quad \textit{for all } f \in \mathcal{B}(\mathcal{M}) \tag{2.4.6}$$

and

$$|||(z-\mathcal{L}_\varepsilon)^{-1} - (z-\mathcal{L})^{-1}||| \le c\, \varepsilon^{\eta(\gamma-\beta)}\, \|(z-\mathcal{L})^{-1}\|\ . \tag{2.4.7}$$

Remark 2.4.10. *The constants ε_0, a, b and c are explicitly determined in the proofs of* [27]*. For the present setting one can deduce that a and c depend on the map T only through the constants σ and B' from Lemma 2.4.7. (Working through the proof of that lemma one notices that σ and B' change continuously under small $\mathcal{C}^2$–perturbations of T.) The constants ε_0 and b depend on T through σ and B', and also through the quantity*

$$H(\delta, r) := \sup\left\{\|(z-\mathcal{L})^{-1}\| :\ z \in \mathbb{C}, |z| > r, \mathrm{dist}(z, sp(\mathcal{L})) > \delta\right\}\ . \tag{2.4.8}$$

Remark 2.4.11. *In the next section we will have to deal with perturbations of the type $Q_\varepsilon \mathcal{L} Q_\varepsilon$. Proposition 2.4.9 applies to these operators as well. To see this note that*

$$|||Q_\varepsilon \mathcal{L} Q_\varepsilon - \mathcal{L}||| \le \|\mathcal{L}_\varepsilon\|_w \cdot |||Q_\varepsilon - \mathrm{Id}\,||| + |||Q_\varepsilon - \mathrm{Id}\,||| \cdot \|\mathcal{L}\| \le \mathrm{const}\cdot \varepsilon^{\gamma-\beta}$$

and that the estimates from Lemma 2.4.7 carry over from $\mathcal{L}_\varepsilon$ to $Q_\varepsilon \mathcal{L} Q_\varepsilon$ because $(Q_\varepsilon \mathcal{L} Q_\varepsilon)^n = Q_\varepsilon \mathcal{L} (Q_\varepsilon^2 \mathcal{L})^{n-1} Q_\varepsilon$, and Q_ε^2 is a again smooth averaging operator, cf. *Lemma 3.3.5.*

Quasi-compactness in dimension $d = 2$. We finish this subsection with a short proof of how to deduce the quasicompactness and the estimate for the essential spectral radius of $\mathcal{L}$ in the two-dimensional case without relying on Proposition 2.2.2 (compact embedding). Instead we use Lemmas 2.4.6 and 2.4.8 (hence assuming that $\tau - 1 \ge \gamma > \beta$). Since, by Lemma 2.2.1,

$\|\mathcal{L}^n - \mathcal{L}^n Q_\varepsilon\| \le \sigma^n \|\operatorname{Id} - Q_\varepsilon\| + \frac{B}{a-\sigma}|||\operatorname{Id} - Q_\varepsilon|||$, the two above mentioned lemmas guarantee that one can choose ε so small that $\|\mathcal{L}^n - \mathcal{L}^n Q_\varepsilon\| \le \text{const}\ \sigma^n$. But the $\mathcal{L}^n Q_\varepsilon$ are compact:

$$\mathcal{B}(\mathcal{M}) \xrightarrow{Q_\varepsilon} \mathcal{C}^1(\mathcal{M},\mathbb{R}) \overset{\text{compact}}{\hookrightarrow} \mathcal{C}^0(\mathcal{M},\mathbb{R}) \overset{\text{conts.}}{\hookrightarrow} \mathcal{B}_w(\mathcal{M}) \xrightarrow{\mathcal{L}^n} \mathcal{B}(\mathcal{M}) , \tag{2.4.9}$$

so [14, Lemma VIII.8.2] implies that $\mathcal{L}$ is quasicompact with essential spectral radius bounded by $\liminf_{n\to\infty} \sqrt[n]{\text{const}\ \sigma^n} = \sigma$.

2.5. **Spectral stability – Deterministic perturbations.** The next natural question concerns the relation between the spectrum associated to two Anosov maps T and $\tilde{T}$ when the two maps are $\mathcal{C}^1$ close. Clearly the associated operators $\mathcal{L}_T$ and $\mathcal{L}_{\tilde{T}}$ live on different Banach spaces $\mathcal{B}(\mathcal{M})$ and $\tilde{\mathcal{B}}(\mathcal{M})$, respectively, so their spectra cannot be compared directly. However, denoting $R(z) = (z - \mathcal{L}_T)^{-1} : \mathcal{B}(\mathcal{M}) \to \mathcal{B}(\mathcal{M})$ and $\tilde{R}(z) = (z - \mathcal{L}_{\tilde{T}})^{-1} : \tilde{\mathcal{B}}(\mathcal{M}) \to \tilde{\mathcal{B}}(\mathcal{M})$ (at points z where they are well defined), both these resolvents can be considered as linear operators from $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ to $\mathcal{C}^1(\mathcal{M},\mathbb{R})^*$, because[13]

$$\|f\|^*_{\mathcal{C}^1} \le \|f\|_w \le \|f\| \le \|f\|_{\mathcal{C}^1} \quad \text{and also} \quad \|f\|^*_{\mathcal{C}^1} \le \|f\|^\sim_w \le \|f\|^\sim \le \|f\|_{\mathcal{C}^1} \tag{2.5.1}$$

for each $f \in \mathcal{C}^1$, where $\|\cdot\|_w$, $\|\cdot\|$ and $\|\cdot\|^\sim_w$, $\|\cdot\|^\sim$ denote the norms associated to T and $\tilde{T}$, respectively. By the same reason also the corresponding spectral projectors (see (2.4.5)) can be interpreted as operators from $\mathcal{C}^1(\mathcal{M},\mathbb{R})$ to $\mathcal{C}^1(\mathcal{M},\mathbb{R})^*$, and since their ranges are finite-dimensional, there is even no big loss in doing so. Therefore the clue to comparing the spectra of $\mathcal{L}_T$ and $\mathcal{L}_{\tilde{T}}$ is a good bound on $\|R(z) - \tilde{R}(z)\|_{\mathcal{C}^1\to(\mathcal{C}^1)^*}$, which in turn is obtained by approximating both transfer operators by suitable randomly perturbed operators.

Proposition 2.5.1. *Given an Anosov map $T : \mathbb{T}^2 \to \mathbb{T}^2$ and a constant $K > \|T\|_{\mathcal{C}^3}$, a real number r bigger than the constant σ associated with T in Lemma 2.2.1 and some $\delta > 0$, there are constants $\kappa = \kappa(T)$, $C_i = C_i(T,K,r,\delta)$ and $\ell_0 = \ell_0(T,K,r,\delta)$ such that for each map $\tilde{T} : \mathbb{T}^2 \to \mathbb{T}^2$ with $\operatorname{dist}_{\mathcal{C}^1}(\tilde{T},T) \le \ell_0$ and $\|\tilde{T}\|_{\mathcal{C}^3} \le K$* [14]

$$\|\tilde{R}(z)\|^\sim \le C_0 \quad \textit{and} \tag{2.5.2}$$

$$\|\tilde{R}(z) - R(z)\|_{\mathcal{C}^1\to(\mathcal{C}^1)^*} \le C_1 \cdot \operatorname{dist}_{\mathcal{C}^1}(\tilde{T},T)^{\kappa\eta(\gamma-\beta)} \tag{2.5.3}$$

provided that $|z| > r$ and $\operatorname{dist}(z, sp(\mathcal{L}_T)) > \delta$ where again $\eta = 1 - \frac{\log r}{\log \sigma}$.

The proof of this proposition can be found in section 3.4 and rests on two observations:

1) If $\operatorname{dist}_{\mathcal{C}^1}(T,\tilde{T})$ is bounded by a suitable power of ε, then $\|Q_\varepsilon f\| \le \text{const} \cdot \|f\|^\sim$. This is proved in Lemma 3.4.4 and leads directly to (2.5.2).

2) In view of (2.5.1) we have

$$\begin{aligned} &\|\tilde{R}(z) - R(z)\|_{\mathcal{C}^1\to(\mathcal{C}^1)^*} \\ &\qquad \le |||\tilde{R}(z) - \tilde{R}_\varepsilon(z)|||^\sim + \|\tilde{R}_\varepsilon(z) - R_\varepsilon(z)\|_{\mathcal{C}^1\to(\mathcal{C}^1)^*} + |||R_\varepsilon(z) - R(z)||| \end{aligned} \tag{2.5.4}$$

where $R_\varepsilon(z) := (z - Q_\varepsilon \mathcal{L}_T Q_\varepsilon)^{-1}$ and $\tilde{R}_\varepsilon(z) := (z - Q_\varepsilon \mathcal{L}_{\tilde{T}} Q_\varepsilon)^{-1}$ for a suitable smooth averaging operator Q_ε with a $\mathcal{C}^2$ kernel. In section 3.4 we estimate these three terms. For the third one we use Proposition 2.4.9, the second one will turn out to be small if $\operatorname{dist}_{\mathcal{C}^1}(T,\tilde{T})$ is of the order of a certain power of ε, and the first one can again be estimated by help of Proposition 2.4.9, but with the rôles of $\mathcal{L}$ and $\mathcal{L}_\varepsilon$ interchanged, which requires an estimate on $\|\tilde{R}_\varepsilon(z)\|$ in terms of quantities that depend only on the unperturbed map T, see Remark 2.4.10. In a different context this idea was already used in [32].

Observing the integral representation (2.4.5) of the spectral projectors in terms of the resolvent, the following theorem is an easy corollary:

Theorem 3. *Given an Anosov map $T : \mathbb{T}^2 \to \mathbb{T}^2$, for each map $\tilde{T}$ sufficiently close to T in $\mathcal{C}^1$–distance and satisfying uniform $\mathcal{C}^3$-bounds the non-essential spectra $sp_r(\mathcal{L}_T, \mathcal{B}_T(\mathcal{M},\mathbb{R}))$ and $sp_r(\mathcal{L}_{\tilde{T}}, \mathcal{B}_{\tilde{T}}(\mathcal{M},\mathbb{R}))$ are close with multiplicities and the $\|\cdot\|_{\mathcal{C}^1\to(\mathcal{C}^1)^*}$–distance of corresponding spectral projectors is small.*

[13] This statement must be interpreted in the sense of Remark 2.1.2.

[14] Note that also $\tilde{T}$ is an Anosov diffeomorphism if $\operatorname{dist}_{\mathcal{C}^1}(\tilde{T},T)$ is small enough.

2.6. **Spectral stability – Ulam finite rank approximation.** An Ulam approximation is constructed in the following way: Let $\mathcal{A} = \{A_i\}$ be a partition of $\mathbb{T}^2$ into convex polytopes (*e.g.* into squares). Define the operator, in $L^1(\mathbb{T}^2, \mathbb{R})$,

$$\Pi_{\mathcal{A}} f(x) := \sum_{A \in \mathcal{A}} \frac{1}{m(A)} \ \mathbf{1}_A(x) \int_A f \, dm \tag{2.6.1}$$

Ulam's idea is then to look at the finite rank operator

$$P_{\mathcal{A}} := \Pi_{\mathcal{A}} \mathcal{L} \Pi_{\mathcal{A}} \tag{2.6.2}$$

and to prove that, in some sense, its spectral properties are close to the ones of the original transfer operator; this is usually referred to as the *Ulam conjecture*. Since $P_{\mathcal{A}}$ is essentially a finite matrix, its spectral properties are numerically accessible. Obviously these spectral properties are the same as those of $\mathcal{L}_{\mathcal{A}} := \Pi_{\mathcal{A}} \mathcal{L}$.

Clearly we cannot compare directly the spectra of $\mathcal{L}$ and $\mathcal{L}_{\mathcal{A}}$, since the space $\Pi_{\mathcal{A}}(\mathcal{B}(\mathbb{T}^2))$ consists of piecewise constant functions which are, in particular, discontinuous along unstable fibres, in contrast to the functions belonging to our space $\mathcal{B}(\mathbb{T}^2)$. This problem could be overcome by extending the space $\mathcal{B}(\mathbb{T}^2)$ to a space including the indicator functions $\mathbf{1}_{A_i}$ as we will discuss briefly in the next section. It turns out, however, that this would not really help. At the end of this section we provide examples of linear hyperbolic torus automorphisms for which straightforward Ulam discretisations $\mathcal{L}_{\mathcal{A}}$ produce eigenvalues far away from all possible eigenvalues of $\mathcal{L}$.

Another possibility to avoid this problem is to replace the indicator functions $\mathbf{1}_A(x)$ in (2.6.1) by a smooth partition of unity. For example one could use a convolution kernel

$$q_\varepsilon(x, y) := \varepsilon^{-2} \bar{q}(\varepsilon^{-1}(x - y)); \quad \bar{q} \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R}); \quad \int \bar{q}(x, y) m(dy) = 1,$$

and its associated smooth averaging operator operator Q_ε and consider the partition of unity $\phi_A := Q_\varepsilon \mathbf{1}_A$ with ε of the order of the diameter of the sets A. With this choice of ε, however, it does not seem possible to prove spectral stability of the resulting discretisation $Q_\varepsilon \Pi_{\mathcal{A}} \mathcal{L}$, the obstacles being basically the same as in the case of the original Ulam procedure. A much bigger ε, roughly of the order of $\operatorname{diam}(\mathcal{A})^{1/3}$, is needed to achieve a spectral stability result where the approximation error for the resolvent is again of the order $\varepsilon^{\eta(\gamma - \beta)}$ as it was the case for smooth perturbations.

Since we will use $L^1_m(\mathbb{T}^2)$ as an auxiliary space which contains the ranges of both Q_ε and $\Pi_{\mathcal{A}}$ (the price to pay for this is the big ε!), the discontinuity of the indicator functions $\mathbf{1}_A(x)$ plays absolutely no rôle in our proof, and for technical reasons we prefer to work with a slightly different discretisation. Namely, we define the finite rank operators

$$P_{\mathcal{A},\varepsilon} := \Pi_{\mathcal{A}} Q_\varepsilon \mathcal{L} \Pi_{\mathcal{A}} \quad \text{and} \quad \mathcal{L}_{\mathcal{A},\varepsilon} := \Pi_{\mathcal{A}} Q_\varepsilon \mathcal{L} \ . \tag{2.6.3}$$

Both operators have the same spectral properties. So we will compare the resolvent of $\mathcal{L}_{\mathcal{A},\varepsilon}$ with that of $\mathcal{L}$, while $P_{\mathcal{A},\varepsilon}$ can be used for numerical investigations, since it can be considered as a matrix acting on the same space as $P_{\mathcal{A}}$. So it can serve as a tool to study the spectral properties of $\mathcal{L}$ on the same footing as the standard Ulam discretisation. For results on the numerical approximation of SRB-measures such an approach was previously used in [11, 12][15]; see also [29] for stabilising effects of noise on numerical simulations of dynamical systems.

As an intermediate step of our program it turns out to be necessary to consider $\Pi_{\mathcal{A}}$, Q_ε and $\mathcal{L}$ as operators acting on the Banach space of functions of bounded variation $\mathrm{BV}(\mathbb{T}^2, \mathbb{R})$. Observe that $\mathrm{BV}(\mathbb{T}^2, \mathbb{R})$ consists exactly of those functions $f \in L^1(\mathbb{T}^2, \mathbb{R})$ that can be approximated in L^1–norm by functions f_n from the Sobolev space $W^{1,1}$ (integrable functions with integrable derivatives) in such a way that $\|f\|_{\mathrm{BV}} = \lim_{n\to\infty} \|f_n\|_{1,1}$. Recall also that for any sequence (f_n) from $W^{1,1}$ with $\lim_{n\to\infty} \|f - f_n\|_1 = 0$ holds: $\|f\|_{\mathrm{BV}} \leq \liminf_{n\to\infty} \|f_n\|_{1,1}$. Therefore it suffices to prove the following lemma for $f \in W^{1,1}$ (and hence only for $f \in \mathcal{C}^1$) with the $\|\cdot\|_{1,1}$–norm instead of the $\|\cdot\|_{\mathrm{BV}}$–norm. But that is a simple exercise.

[15] In these papers the discretisation of smoothly perturbed Anosov diffeomorphisms serves as a theoretical justification for numerical discretisations. The effectively implemented procedures, however, do not make use of random noise. In this context observe the example at the end of the current section.

Lemma 2.6.1. *There exists a constant $K > 0$ such that, for each $\varepsilon \in (0,1)$ and $f \in \mathrm{BV}(\mathbb{T}^2, \mathbb{R})$,*

$$\begin{aligned}\|Q_\varepsilon f\|_{\mathrm{BV}} &\le \|f\|_{\mathrm{BV}}\\ \|Q_\varepsilon f\|_{\mathrm{BV}} &\le K\varepsilon^{-1}\|f\|_1\\ \|\mathcal{L}f\|_{\mathrm{BV}} &\le K\|f\|_{\mathrm{BV}}\ .\end{aligned}$$

This lemma shows in particular that Q_ε and $\mathcal{L}$ are well defined bounded operator on $\mathrm{BV}(\mathbb{T}^2, \mathbb{R})$. The next lemma shows that the same is true for $\Pi_{\mathcal{A}}$.

Lemma 2.6.2 ([9])**.** *If the partition $\mathcal{A}$ is made of convex polytopes of size*[16] *at most N^{-1}, then, for each $f \in \mathrm{BV}(\mathbb{T}^2, \mathbb{R})$,*

$$\begin{aligned}\|\Pi_{\mathcal{A}} f\|_{\mathrm{BV}} &\le 2\|f\|_{\mathrm{BV}}\\ \|f - \Pi_{\mathcal{A}} f\|_1 &\le 2N^{-1}\|f\|_{\mathrm{BV}}\end{aligned}$$

The first inequality of this lemma is proven in Lemma 10 of [9], the second one in Corollary 3 of the same paper.

The following proposition contains the key estimates of this section.

Proposition 2.6.3. *Let $\mathcal{A}$ be a partition of $\mathbb{T}^2$ into convex polytopes of size at most N^{-1}. If $N^{-1}\varepsilon^{-3}$ is smaller than some constant*[17]*, then $(z - \mathcal{L}_{\mathcal{A},\varepsilon})$ is invertible as an operator on $L^1(\mathbb{T}^2, \mathbb{R})$ and*

$$\|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1} f\|_1 \le \mathrm{const}\,(\varepsilon^{-1}\,\|Q_\varepsilon \mathcal{L} f\| + \|f\|_1) \tag{2.6.4}$$

Furthermore,

$$\|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1} - (z - Q_\varepsilon\mathcal{L})^{-1}\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*} \le \mathrm{const}\ N^{-1}\,\varepsilon^{-3}\ . \tag{2.6.5}$$

Combined with Proposition 2.4.9 this yields at once that under the conditions of the last proposition

$$\|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1} - (z - \mathcal{L})^{-1}\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*} \le \mathrm{const}\ \left(N^{-1}\,\varepsilon^{-3} + \varepsilon^{\eta(\gamma-\beta)}\right)\ , \tag{2.6.6}$$

where the constant depends on z through the quantity $H(\delta, r)$ from (2.4.8). Recall from section 2.4 that $\eta = 1 - \frac{\log r}{\log \sigma}$ where r must be chosen such that $|z| > r > \sigma$.

The main result of this section now follows again from spectral calculus.

Theorem 4. *Let $\mathcal{A}$ be a partition of $\mathbb{T}^2$ into convex polytopes of size at most N^{-1} where $N = N(\varepsilon) \approx \varepsilon^{-3-\eta(\gamma-\beta)}$. Then the assertions of Theorem 2 hold for $\mathcal{L}_{\mathcal{A},\varepsilon}$ instead of $\mathcal{L}_\varepsilon$. In particular, the non-essential spectra $sp_r(\mathcal{L}, \mathcal{B}(\mathbb{T}^2))$ and $sp_r(\mathcal{L}_{\mathcal{A},\varepsilon}, L^1(\mathbb{T}^2, \mathbb{R}))$ are close with multiplicities and the $\|\cdot\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*}$-distance of corresponding spectral projectors is of order $\varepsilon^{\eta(\gamma-\beta)}$.*

Since $\mathcal{L}_{\mathcal{A},\varepsilon}$ is a finite rank operator, hence its spectrum and eigenvalues can be explicitly computed, the above theorem provides a constructive tool to study the invariant measure, the rate of decay of correlations etc. of the map. Some related numerical implementations (using the original Ulam method modified by an adaptive choice of cell sizes) are reviewed in [18, 12]. It seems worth to mention that our rate of convergence compares favourably to the one proved in [24] for the SRB measure, *i.e.* for the "eigenfunction" of the eigenvalue 1.

Just to have a more precise idea on how the business really works let us briefly, and not optimally, discuss the determination of the invariant measure in cases where the system is ergodic, i.e. where 1 is a simple eigenvalue of $\mathcal{L}$. For a more detailed discussion of this type of questions see [32].

Let μ be the invariant measure and $h \in \mathcal{B}(\mathbb{T}^2)$ the corresponding eigenvector of $\mathcal{L}$ (that is $\mathcal{L}h = h$ and $\mu(\varphi) = \langle h, \varphi\rangle$ for all $\varphi \in \mathcal{D}_\gamma$). Let $h_\varepsilon, h_{\mathcal{A},\varepsilon}, \in L^1(\mathbb{T}^2, \mathbb{R})$ be such that $\mathcal{L}_{\mathcal{A},\varepsilon} h_{\mathcal{A},\varepsilon} = h_{\mathcal{A},\varepsilon}$ and $Q_\varepsilon\mathcal{L}h_\varepsilon = h_\varepsilon$, and suppose that $N \approx \varepsilon^{-3-\eta(\gamma-\beta)}$. Then there exists $\rho > 0$ such that

$$\begin{aligned}\|h_\varepsilon - h_{\mathcal{A},\varepsilon}\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*} &\le \int_{\{z \in \mathbb{C}:\ |z-1| = \rho\}} \|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}(1) - (z - Q_\varepsilon\mathcal{L})^{-1}(1)\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*}\\ &\le \mathrm{const}\ \varepsilon^{\eta(\gamma-\beta)}\ .\end{aligned}$$

In addition, the results of subsection 2.4 together with [27] imply

$$\|h_\varepsilon - h\|_w \le \mathrm{const}\ \varepsilon^{\eta(\gamma-\beta)}\ .$$

---

[16] The size here is the side length of the smallest coordinate cube which contains the polytope.

[17] This constant depends on a and b from Proposition 2.4.9 and hence from the distance of z to $\mathrm{sp}(\mathcal{L})$.

Hence, calling $\mu_{\mathcal{A},\varepsilon}(\varphi) := \int \varphi\, h_{\mathcal{A},\varepsilon}\, dm$, for each $\varphi \in \mathcal{C}^1(\mathbb{T}^2, \mathbb{R})$ holds

$$|\mu(\varphi) - \mu_{\mathcal{A},\varepsilon}(\varphi)| \le C_2 \varepsilon^{\eta(\gamma-\beta)} |\varphi|_{\mathcal{C}^1} \ . \tag{2.6.7}$$

Since we are focusing on the eigenvalue 1, the radius r can be chosen arbitrarily close to 1 in which way also η can be made as close to 1 as we wish. A closer look at the proofs of [27] reveals in fact that convergence takes place at a rate $\varepsilon^{\gamma-\beta}\log\varepsilon^{-1}$.

Inequality (2.6.7) implies, by standard approximation arguments, that the measures $\mu_{\mathcal{A},\varepsilon}$ converge weakly to μ, yet it provides much more informations since it tells us exactly how much the two measures differ on a wide class of observables.

Counterexamples to spectral stability of the pure Ulam method. As we already mentioned in the introduction the coexistence of expansion and contraction in hyperbolic maps is a serious obstacle to the applicability of the original Ulam scheme. Some results related to this problem were discussed earlier in [4, 5], where in particular problems with the discretisation of Arnold's cat map were pointed out. Here we discuss a family of torus automorphisms whose transfer operators are all poorly approximated by Ulam discretisations.

Let $T_{a,k} : \mathbb{T}^2 \to \mathbb{T}^2$ be the linear torus automorphisms defined by the following two-parameter family of 2×2 matrices

$$\mathbb{A}_{a,k} := \begin{pmatrix} a & ka-1 \\ a+1 & k(a+1)-1 \end{pmatrix}, \qquad a, k \in \mathbb{Z}^1_+,\ a \ge 2.$$

Trivially, the $T_{a,k}$ have stable and unstable foliations of class $\mathcal{C}^2$, so $\alpha = \tau - 1 = 1$. Therefore we may choose $\gamma = 1$ and any $0 < \beta < 1$ as parameters for our Banach spaces. Denote by $\lambda_{a,k} > 1$ and $\lambda_{a,k}^{-1}$ the two eigenvalues of $\mathbb{A}_{a,k}$. Then $\lambda_u = \lambda_{a,k}$ and $\lambda_s = \lambda_{a,k}^{-1}$. In particular $\max\{\lambda_u^{-1}, \lambda_s^\beta\} = \lambda_{a,k}^{-\beta}$.

Denote by $\mathcal{L}_{a,k}$ the transfer operator of $T_{a,k}$. We know from Example 2.2.6 that

$$\mathrm{sp}(\mathcal{L}_{a,k}) \subseteq \{z \in \mathbb{C} : |z| \le \max\{\lambda_u^{-1}, \lambda_s^\beta\}\} \cup \{1\} = \left\{z \in \mathbb{C} : |z| \le \lambda_{a,k}^{-\beta}\right\} \cup \{1\} \ .$$

Consider Ulam discretisations $P_{\mathcal{A}_n}$ of $\mathcal{L}_{a,k}$ using partitions $\mathcal{A}_n$ into squares with sides of length n^{-1} parallel to the coordinate axes, see (2.6.2). Mixed numerical and theoretical investigations that we line out below show that for many choices of a and k the spectra $\mathrm{sp}(P_{\mathcal{A}_n})$ (for arbitrarily large even n) contain numbers with modulus strictly between $\lambda_{a,k}^{-1}$ and 1 and hence between $\lambda_{a,k}^{-\beta}$ and 1 when β is chosen close enough to 1. Hence the most straightforward Ulam procedure, based on squares with equal size and sides parallel to the axes, produces fake eigenvalues at arbitrarily fine grid sizes n^{-1}. It is not clear to us whether this effect is an "arithmetic artifact" of torus *linear* automorphisms or if it is a more common phenomenon for hyperbolic diffeomorphisms.

The task to determine (at least some) eigenvalues that occur in $P_{\mathcal{A}_n}$ for arbitrarily large n is greatly simplified by the following observation:

Proposition 2.6.4. [4] *Let $\mathbb{A}$ be a $d \times d$ matrix with nonnegative integer entries, and define $T : \mathbb{T}^d \to \mathbb{T}^d$ by $Tx := \mathbb{A}x \mod 1$. Denote by $\mathcal{A}_n$ the partition of $\mathbb{T}^d$ into cubes of side length n^{-1} with vertices in $n^{-1}\mathbb{Z}^d$. As before let $P_{\mathcal{A}_n} = \Pi_{\mathcal{A}_n}\mathcal{L}_T\Pi_{\mathcal{A}_n}$. Then*

$$sp(P_{\mathcal{A}_m}) \subseteq sp(P_{\mathcal{A}_{mN}})$$

for all positive integers N and m.

According to this proposition it suffices to show that $P_{\mathcal{A}_2}$ has an eigenvalue with modulus strictly between $\lambda_{a,k}^{-1}$ and 1, since then this number belongs to $\mathrm{sp}(P_{\mathcal{A}_{2N}})$ for all $N \in \mathbb{N}$.

In case $a = 2$ and $k = 1$ the matrix representing the operator $P_{\mathcal{A}_2}$ can be figured out by hand. It is

$$\begin{pmatrix} 1/6 & 1/3 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/6 & 1/3 & 1/3 & 1/6 \end{pmatrix} .$$

This matrix has the four eigenvalues $1, -1/3, 0, 0$. In particular, $-1/3$ is a common eigenvalue of all $P_{\mathcal{A}_{2N}}$. On the other hand, the matrix $\mathbb{A}_{2,1}$ defining the automorphism has the leading eigenvalue $\lambda_{2,1} = 2 + \sqrt{3}$ so that $\lambda_{2,1}^{-1} = 2 - \sqrt{3} \approx 0.268$. With $\beta = 0.9$ we have indeed $\lambda_{2,1}^{-\beta} < \frac{1}{3} < 1$.

This, in fact, is not the worst case and other choices of the parameters a, k demonstrate even more striking differences between the eigenvalues of the Ulam approximation by square partitions and the spectrum of $\mathcal{L}_{a,k}$. We presented the case $a = 2$, $k = 1$ in some detail, because in this case everything can be calculated analytically, while for general values of the parameters only numerical approximations are available. By means of a program written in "Mathematica" we have calculated eigenvalues for several other choices of the parameter values.

n	a	k	r_2	$\lambda_{a,k}^{-1}$	a	k	r_2	$\lambda_{a,k}^{-1}$
2	20	1	0.3400	0.0250	20	3	0.1115	0.0122
4	20	1	0.3400	0.0250	20	3	0.2240	0.0122
6	20	1	0.3400	0.0250	20	3	0.1115	0.0122
8	20	1	0.3400	0.0250	20	3	0.2240	0.0122
2	500	1	0.5000	0.0010	500	3	0.2500	0.0005
4	500	1	0.5000	0.0010	500	3	0.2500	0.0005
6	500	1	0.5000	0.0010	500	3	0.2500	0.0005
8	500	1	0.5000	0.0010	500	3	0.2500	0.0005

Here r_2 is the modulus of the second largest eigenvalue of $P_{\mathcal{A}_n}$. It is to be compared to $\lambda_{a,k}^{-1}$, which we consider to be the "true" second eigenvalue.

What is still possible is that the leading eigenfunction – the SRB measure – may be stable for the class of maps we consider (the above example does not contradict to this – the SRB measure is preserved). For piecewise expanding interval maps this is known to be true, see for example [5] and further references therein. Moreover numerous numerical studies confirm this stability for a much broader classes of dynamical systems, see *e.g.* [12, 18], but presently we do not have an adequate explanation for this numerical effectiveness of the Ulam method.

For Ulam discretisations based on Markov partitions the situation is much better: In the case of linear automorphisms of $\mathbb{T}^2$ Brini et al.[8] showed that there are arbitrarily fine Markov partitions $\mathcal{A}$ for which $\mathrm{sp}(P_{\mathcal{A}}) \subseteq \{1, \lambda_s, \lambda_s^2, 0\}$. In particular they have no eigenvalue with modulus in $(\lambda_{a,k}^{-1}, 1)$. Froyland [17] proved for general Anosov maps of $\mathbb{T}^2$ that the eigenvector to the eigenvalue 1 of Ulam discretisations based on suitable Markov partitions approximates the SRB measure if the partitions get finer and finer.

2.7. **Remarks on an extension of the space $\mathcal{B}(\mathcal{M})$.** Following one of the paths that lead from the Sobolev space $W^{1,1}$ of differentiable functions with integrable derivative to the space BV of functions of bounded variation, we show in this section how to extend the norm on our space $\mathcal{B}(\mathcal{M})$ to a larger subspace $\bar{\mathcal{B}}(\mathcal{M})$ of $\mathcal{B}_w(\mathcal{M})$ that contains, in particular, the space $\mathrm{BV}(\mathcal{M})$. Most of this program is abstract and works equally well in any dimension. Only to show that $\mathcal{B}(\mathcal{M})$ is a *closed* subspace of $\bar{\mathcal{B}}(\mathcal{M})$ needs much more effort, and since we do not really make use of it except in the discussion of Ulam discretisations on the two-dimensional torus, we will prove that only in the case $\mathcal{M} = \mathbb{T}^2$.

For $f \in \mathcal{B}_w(\mathcal{M})$ let

$$\|f\|^{\bar{}} := \lim_{\epsilon \to 0} \inf \{\|g\| \,:\, g \in \mathcal{B}(\mathcal{M}),\, \|f - g\|_w \le \epsilon\} \ . \tag{2.7.1}$$

Observe that one can replace the limit by a supremum over $\varepsilon > 0$. It is easy to see that $\|\cdot\|^{\bar{}}$ is a norm on

$$\bar{\mathcal{B}}(\mathcal{M}) := \{f \in \mathcal{B}_w(\mathcal{M}) \,:\, \|f\|^{\bar{}} < \infty\} \ .$$

Then $(\bar{\mathcal{B}}(\mathcal{M}), \|\cdot\|^{\bar{}})$ is clearly a normed linear space, and since $\|f\|^{\bar{}} \le \|f\|$ for $f \in \mathcal{B}(\mathcal{M})$ by definition, we have $\mathcal{B}(\mathcal{M}) \subseteq \bar{\mathcal{B}}(\mathcal{M})$.[18] We list some simple consequences:

i) The unit ball $\bar{\mathcal{B}}_1 := \{f \in \bar{\mathcal{B}}(\mathcal{M}) \,:\, \|f\|^{\bar{}} \le 1\}$ is the $\|\cdot\|_w$–closure of the unit ball $\mathcal{B}_1 := \{f \in \mathcal{B}(\mathcal{M}) \,:\, \|f\| \le 1\}$.
ii) $\mathcal{B}_1$ is closed in $\mathcal{B}_w(\mathcal{M})$ if and only if the norms $\|\cdot\|$ and $\|\cdot\|^{\bar{}}$ are equivalent on $\mathcal{B}$.
iii) Since $\mathcal{B}_1$ is a relatively compact subset of $\mathcal{B}_w(\mathcal{M})$, $\bar{\mathcal{B}}_1$ is a compact subset of $\mathcal{B}_w(\mathcal{M})$.
iv) For $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$ we have $\|f\| \le \|f\|_{1,1}$ and $\|f\|_w \le \|f\|_1$. Therefore

$$\mathrm{BV}_1 = \|\cdot\|_1\text{-closure}((W^{1,1})_1) \subseteq \|\cdot\|_w\text{-closure}((W^{1,1})_1) \subseteq \|\cdot\|_w\text{-closure}(\mathcal{B}_1) = \bar{\mathcal{B}}_1 \ .$$

(Here the lower index 1 denotes the unit ball of the respective space.) In particular, $\bar{\mathcal{B}}(\mathcal{M})$ contains indicator functions of sufficiently regular sets.
v) $\mathcal{L}$ extends to a bounded operator on $\bar{\mathcal{B}}(\mathcal{M})$.

[18] Recall Remark 2.1.2 for a discussion of the embedding of $\mathcal{B}(\mathcal{M})$ into $\mathcal{B}_w(\mathcal{M})$.

vi) The Lasota Yorke type inequality carries over to $\bar{\mathcal{B}}(\mathcal{M})$: $\|\mathcal{L}^n f\|^{\bar{}} \leq 3A^2\sigma^n\|f\|^{\bar{}} + K\|f\|_w$.
vii) In particular, $\mathcal{L}$ acts on $\bar{\mathcal{B}}(\mathcal{M})$ as a quasicompact operator with essential spectral radius bounded by σ.

What we cannot conclude at this point is that $\mathcal{B}(\mathcal{M})$ embeds as a *closed* linear subspace into $\bar{\mathcal{B}}(\mathcal{M})$. This is the case if and only if there is some $C > 0$ such that

$$\|f\| \leq C \cdot \|f\|^{\bar{}} \quad \text{for } f \in \mathcal{B}(\mathcal{M}), \tag{2.7.2}$$

i.e. if the two norms are equivalent.[19] In section 3.7 we will prove the following lemma.

Lemma 2.7.1. *The estimate (2.7.2) holds for $\mathcal{M} = \mathbb{T}^2$.*

It turns out that for torus automorphisms (in any dimension, in fact) the constant $C = 1$ so that both norms coincide.

An important consequence of this observation is:

Lemma 2.7.2.
If $\mathcal{L}f = \lambda f$ for some $f \in \bar{\mathcal{B}}(\mathbb{T}^2)$ and $|\lambda| > \sigma$, then $f \in \mathcal{B}(\mathbb{T}^2)$.

Proof. Since $f \in \bar{\mathcal{B}}(\mathbb{T}^2)$, there are $g_n \in \mathcal{B}(\mathbb{T}^2)$ with $\|f - g_n\|_w \leq \frac{1}{n}$ and $\|g_n\| \leq \|f\|^{\bar{}}$. Hence

$$\begin{aligned}\|f - \lambda^{-n}\mathcal{L}^n g_n\|^{\bar{}} = \|\lambda^{-n}\mathcal{L}^n(f - g_n)\|^{\bar{}} &\leq 3A^2\left(\frac{\sigma}{|\lambda|}\right)^n \|f - g_n\|^{\bar{}} + C\,\|f - g_n\|_w \\ &\leq 3A^2\left(\frac{\sigma}{|\lambda|}\right)^n 2\|f\|^{\bar{}} + \frac{C}{n} \to 0\end{aligned}$$

so that f belongs to the $\|\cdot\|^{\bar{}}$–closure of $\mathcal{B}(\mathbb{T}^2)$. But since $\mathcal{B}(\mathbb{T}^2)$ is closed in $\bar{\mathcal{B}}(\mathbb{T}^2)$, this implies that $f \in \mathcal{B}(\mathbb{T}^2)$. □

With an only slightly more complicated argument one shows that also if $(\lambda - \mathcal{L})^k f = 0$ for some $k > 1$ and $f \in \bar{\mathcal{B}}(\mathbb{T}^2)$ and λ as above, then $f \in \mathcal{B}(\mathbb{T}^2)$.

3. Proofs

In the sequel we will mostly write $\mathcal{B}$ and $\mathcal{B}_w$ instead of $\mathcal{B}(\mathcal{M})$ and $\mathcal{B}_w(\mathcal{M})$.

3.1. Proofs: Lasota-Yorke type estimate.

Proof of Lemma 2.2.1. Let us start by considering only $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$.

Recall from (2.1.2) that $\mathcal{L}^n f(x) = f \circ T^{-n} \cdot g_n$ where the Jacobian g_n of T^{-n} is of class $\mathcal{C}^{1+\beta}$. Then, for $v \in \mathcal{V}_\beta$, holds

$$d_x(\mathcal{L}^n f)(v(x)) = d_x g_n(v(x)) \cdot f \circ T^{-n}(x) + g_n(x) \cdot d_{T^{-n}(x)} f(d_x T^{-n}(v(x))). \tag{3.1.1}$$

To continue we need to note the following.

Sub-lemma 3.1.1. *For each $n \in \mathbb{N}$ there is $\delta(n) > 0$ such that if $0 < \delta < \delta(n)$ (δ from the definition of $H^s_\beta(v)$ in (2.1.5)), then*

$$R_n v(x) := A^{-2}\lambda_u^n \, d_{T^n x} T^{-n}(v(T^n x)) \in \mathcal{V}_\beta$$

for each $v \in \mathcal{V}_\beta$.

Proof. Obviously, $|R_n v|_\infty \leq A^{-2}\lambda_u^n A \lambda_u^{-n}|v|_\infty \leq 1$.

To investigate the Hölder norm of $R_n v$ it is more convenient to introduce normal coordinates. Let us be more precise. Given an arbitrary point $x \in \mathcal{M}$ let us consider a neighbourhood $x \in U \subset \mathcal{M}$ so small to be contained in the domain of injectivity of the exponential map.[20] We assume that δ is sufficiently small that the ball of centre x and radius δ is always contained in U. We consider an isometric identification of $\mathcal{T}_x\mathcal{M}$ with $\mathbb{R}^d$ and then consider the chart $\exp_x : V \subset \mathbb{R}^d \to U \subset \mathcal{M}$ given by the exponential map.

On the torus $\mathcal{M} = \mathbb{T}^d$ the situation is rather trivial. Here all tangent spaces $\mathcal{T}_x\mathcal{M}$ are canonically identified with each other and with $\mathbb{R}^d$ and $\exp_x(\eta) = x + \eta \mod \mathbb{Z}^d$. Since there is no danger of confusion, we will even skip the "mod $\mathbb{Z}^d$" henceforth. (It suffices to choose $\delta < \frac{1}{2}$.) Then, for

[19] The "if" implication is obvious. For the "only if" implication observe that if $\mathcal{B}(\mathcal{M})$ is closed in $\bar{\mathcal{B}}(\mathcal{M})$, then $(\mathcal{B}(\mathcal{M}), \|\cdot\|^{\bar{}})$ is a Banach space. Since $\|\cdot\|^{\bar{}} \leq \|\cdot\|$, the open mapping theorem implies (2.7.2).

[20] Recall that the exponential map is obtained by flowing along the geodesics, [7].

each $y \in W^s_\delta(x)$, we can identify $d_{T^n y}T^{-n}$ with a $d \times d$ matrix $\mathbb{A}_y$, so $R_n v(y) = A^{-2}\lambda_u^n \mathbb{A}_y v(T^n y)$. Since T^n is of class $\mathcal{C}^3$ there are $c_n > 0$ such that

$$\|\mathbb{A}_x - \mathbb{A}_y\| \le c_n\, d(x,y) \quad \text{for } y \in U.$$

It follows that

$$\begin{aligned} \|R_n v(x) - R_n v(y)\| &= A^{-2}\lambda_u^n \|\mathbb{A}_x v(T^n x) - \mathbb{A}_y v(T^n y)\| \\ &\le A^{-2}\lambda_u^n \|\mathbb{A}_x (v(T^n x) - v(T^n y))\| + c_n d(x,y)\ . \end{aligned} \tag{3.1.2}$$

For manifolds other than $\mathbb{T}^d$ more care has to be taken to derive an analogous estimate.[21] To continue we need to recall that the vectors v are adapted to the unstable foliation. Given a vector w in $\mathbb{R}^d$ and a point $x \in \mathcal{M}$ we can write w as $w = \xi + \eta$, with $\xi \in E^u(x)$ and $\eta \in E^s(x)$. Therefore, for each $n \ge 0$, an unstable subspace $E^u(T^n y)$ can be represented by an operator $\mathbb{V}_y^n : E^u(T^n x) \to E^s(T^n x)$; the subspace is just $\{\xi + \mathbb{V}_y^n \xi\}_{\xi \in E^u(T^n x)}$. Clearly $\mathbb{V}_x^n = 0$.[22] In terms of the operators $\mathbb{V}_y^n$ $(y \in U)$ the τ-Hölder continuity of the unstable distribution implies [25, 19.1.6]: There is a constant $C > 0$, independent of x and y, such that

$$\|\mathbb{V}_y^n\| \le C d(T^n x, T^n y)^{\tau'};$$

where $\tau' = \min\{\tau, 1\}$. For $y \in W^s_\delta(x)$, let $v(T^n y) = \zeta + \mathbb{V}_y^n \zeta$, $\zeta \in E^u(T^n x)$. Denote $c_n' := \sup_{x \in \mathcal{M}} \|\mathbb{A}_x\|$. Then, from (3.1.2) we have

$$\begin{aligned} \|R_n v(x) - R_n v(y)\| &\le A^{-2}\lambda_u^n \|\mathbb{A}_x (v(T^n x) - \zeta)\| + A^{-2}\lambda_u^n \|\mathbb{A}_x (\mathbb{V}_y^n \zeta)\| + c_n d(x,y) \\ &\le A^{-2}\lambda_u^n A \lambda_u^{-n} \|v(T^n x) - \zeta\| + A^{-2}\lambda_u^n c_n' \|\mathbb{V}_y^n \zeta\| + c_n d(x,y) \\ &\le A^{-1}\|v(T^n x) - v(T^n y)\| + (A^{-2}\lambda_u^n c_n' + A^{-1})\|\mathbb{V}_y^n \zeta\| + c_n d(x,y). \end{aligned}$$

As

$$\|\mathbb{V}_y^n \zeta\| \le C d(T^n x, T^n y)^{\tau'} \|\zeta\| \le C \lambda_s^{n\tau'} d(x,y)^{\tau'} \|\zeta\| \le C \lambda_s^{n\tau'} \delta^{\tau'} \|\zeta\|\ ,$$

we find for sufficiently small δ (such that $C\lambda_s^{n\tau'}\delta^{\tau'} \le \frac{1}{2}$) that $\|\zeta\| \le 2\|v(T^n y)\| \le 2$ and hence $\|\mathbb{V}_y^n \zeta\| \le 2C d^s(x,y)^{\tau'} |v|_\infty$. So we conclude

$$\|R_n v(x) - R_n v(y)\| \le \left(\lambda_s^{\beta n} + 2C\lambda_s^{\tau' n}(A^{-2}\lambda_u^n c_n' + A^{-1})\delta^{\tau' - \beta} + c_n \delta^{1-\beta}\right) d^s(x,y)^\beta \le d^s(x,y)^\beta$$

provided δ is chosen small enough. This finishes the proof of the sub-lemma. □

It is convenient to introduce the *distortion*

$$\Delta_n^v(x) = \frac{d_x g_n(v(x))}{g_n(x)}. \tag{3.1.3}$$

---

[21] The first fact to notice for general $\mathcal{M}$ is that, in normal coordinates, the parallel transport of a tangent vector from $y \in U$ to x, along the geodesic, coincides with the vector obtained by simply identifying the tangent space by the Cartesian structure of $\mathbb{R}^d$. More precisely: Let us call $P_{yx} : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_y\mathcal{M}$ the above mentioned parallel transport. Let V be a neighbourhood of the origin in $\mathcal{T}_x\mathcal{M}$ which is mapped by the exponential map $\exp_x$ diffeomorphically to the neighbourhood U of x. For $\xi \in V$ let $y = \exp_x(\xi)$. Then

$$\tilde{P}_{\xi 0} \circ d_x \exp_x^{-1} \circ P_{xy} = d_y \exp_x^{-1} : \mathcal{T}_y\mathcal{M} \to \mathcal{T}_\xi(\mathcal{T}_x\mathcal{M})$$

where $\tilde{P}_{\xi 0} : \mathcal{T}_0(\mathcal{T}_x\mathcal{M}) \to \mathcal{T}_\xi(\mathcal{T}_x\mathcal{M})$ denotes the trivial parallel transport by the vector ξ.

Next, we introduce normal coordinates in a neighbourhood U' of Tx as well. (For simplicity of notation we consider only the case $A = 1$, $n = 1$.) Let $\exp_{Tx} : V' \to U'$ be diffeomorphic, V' be a neighbourhood of the origin in $\mathcal{T}_{Tx}\mathcal{M}$. Let $\xi' = \exp_{Tx}^{-1}(Ty)$. Then

$$\mathbb{A}_y := d_y \exp_x^{-1} \circ d_{Ty} T^{-1} \circ d_{\xi'} \exp_{Tx} : \mathcal{T}_{\xi'}(\mathcal{T}_x\mathcal{M}) \to \mathcal{T}_\xi(\mathcal{T}_x\mathcal{M})$$

can be viewed as a matrix – due to the canonical identification of the spaces $\mathcal{T}_{\xi'}(\mathcal{T}_x\mathcal{M})$ and $\mathcal{T}_\xi(\mathcal{T}_x\mathcal{M})$ with $\mathbb{R}^d$. Clearly, due to the smoothness properties of T

$$\|\mathbb{A}_x - \mathbb{A}_y\| \le c_1 d(x,y).$$

With these conventions we have, just as in (3.1.2),

$$\begin{aligned} \|d_x \exp_x^{-1}(Rv(x)) - d_y \exp_x^{-1}(Rv(y))\| =& \lambda_u \|\mathbb{A}_x \cdot d_{Tx} \exp_{Tx}^{-1}(v(Tx)) - \mathbb{A}_y \cdot d_{Ty} \exp_{Tx}^{-1}(v(Ty))\| \\ \le& \lambda_u \|\mathbb{A}_x \cdot (d_{Tx} \exp_{Tx}^{-1}(v(Tx)) - d_{Ty} \exp_{Tx}^{-1}(v(Ty)))\| + \lambda_u c_1 d(x,y). \end{aligned}$$

[22] On general manifolds $\mathcal{M}$ we use $d_x \exp_x^{-1}$ to lift $E^u(y)$ and $E^s(y)$ from $\mathcal{T}_y\mathcal{M}$ to $\mathcal{T}_0(\mathcal{T}_x\mathcal{M})$, the latter being identified with $\mathbb{R}^d$. Then we proceed along the same lines.

We can then write

$$
\begin{aligned}
\int_{\mathcal{M}} d(\mathcal{L}^n f)(v)\, dm &= \int_{\mathcal{M}} d(f\circ T^{-n})(v)\cdot g_n\, dm + \int_{\mathcal{M}} f\circ T^{-n}\cdot \Delta_n^v \cdot g_n\, dm \\
&= A^2\lambda_u^{-n}\int_{\mathcal{M}} \mathcal{L}^n(df(R_n v))\, dm + \int_{\mathcal{M}} \mathcal{L}^n f\cdot \Delta_n^v\, dm \\
&= A^2\lambda_u^{-n}\int_{\mathcal{M}} df(R_n v)\, dm + \int_{\mathcal{M}} f\cdot(\Delta_n^v\circ T^n)\, dm
\end{aligned}
\tag{3.1.4}
$$

Since the Jacobian g_n is of class $\mathcal{C}^{1+\beta}$ by assumption and since $\inf g_n > 0$, the following estimate is immediate:

Sub-lemma 3.1.2. *For each $n\in\mathbb{N}$ there exists a constant $K_n\in\mathbb{R}^+$ such that, for each $v\in\mathcal{V}_\beta$, $K_n^{-1}\Delta_n^v\circ T^n\in\mathcal{D}_\beta$.*

Thus

$$
\|\mathcal{L}^n f\|_u \le A^2\lambda_u^{-n}\|f\|_u + K_n\|f\|_s \tag{3.1.5}
$$

To estimate the stable norm we write

$$
\int_{\mathcal{M}} \mathcal{L}^n f\cdot\varphi\, dm = \int_{\mathcal{M}} f\cdot(\varphi\circ T^n)\, dm.
$$

But $|\varphi\circ T^n|_\infty = |\varphi|_\infty$ and $H^s_\beta(\varphi\circ T^n)\le A\lambda_s^{\beta n}H^s_\beta(\varphi)$. Applied with γ instead of β this yields at once

$$
\|\mathcal{L}^n f\|_w \le \max\{1, A\lambda_s^{\beta n}\}\cdot\|f\|_w. \tag{3.1.6}
$$

To obtain strict contraction for the $\|\cdot\|_s$-norm, we must approximate $\varphi\in\mathcal{D}_\beta$ by a suitable function $\tilde\varphi\in\mathcal{D}_\gamma$.

Sub-lemma 3.1.3. *There are $\delta_0>0$ and $B_1>0$ which do not depend on the parameter δ from the definition of $H^s_\beta(.)$ and $H^s_\gamma(.)$ such that for any choice of $\delta\in(0,\delta_0]$ and for each $\varphi\in\mathcal{D}_\beta$ there exists $\tilde\varphi:\mathcal{M}\to\mathbb{R}$ such that*

$$
\begin{aligned}
&|\tilde\varphi|_\infty \le 1 && |\varphi-\tilde\varphi|_\infty \le 2^{-\beta}\,\delta^\beta \\
&H^s_\gamma(\tilde\varphi)\le B_1\,\delta^{-\gamma} && H^s_\beta(\tilde\varphi)\le 1 + B_1\,\delta^{1-\beta}\ .
\end{aligned}
$$

We postpone the proof of the above result to the end of the section and finish the proof of Lemma 2.2.1 before. So we choose $\sigma>\max\{\lambda_u^{-1},\lambda_s^\beta\}$ and fix an integer N such that

$$
(\sigma\,\min\{\lambda_u,\lambda_s^{-\beta}\})^N > 9A^2\ . \tag{3.1.7}
$$

Then we choose $\delta\in(0,\delta_0]$ so small that $\delta\le 2\lambda_s^N$ and $B_1\delta^{1-\beta}\le 1$. The sub-lemma implies now

$$
B_1^{-1}\left(\delta^{-1}A\lambda_s^n\right)^{-\gamma}\cdot\tilde\varphi\circ T^n\in\mathcal{D}_\gamma \quad\text{and}\quad \frac{1}{3A}\lambda_s^{-\beta n}(\varphi-\tilde\varphi)\circ T^n\in\mathcal{D}_\beta \quad\text{for all } n=1,\dots,N\ .
$$

Using such a decomposition and denoting $K'_n := B_1\left(\delta^{-1}A\lambda_s^n\right)^\gamma$,

$$
\int_{\mathcal{M}}\mathcal{L}^n f\cdot\varphi\, dm = 3A\lambda_s^{\beta n}\int_{\mathcal{M}} f\cdot\frac{1}{3A}\lambda_s^{-\beta n}(\varphi-\tilde\varphi)\circ T^n\, dm + K'_n\int_{\mathcal{M}} f\cdot({K'_n}^{-1}\cdot\tilde\varphi\circ T^n)\, dm
$$

so that

$$
\|\mathcal{L}^n f\|_s \le 3A\lambda_s^{\beta n}\,\|f\|_s + K'_n\,\|f\|_w \quad\text{for } n=1,\dots,N.
$$

Combined with (3.1.5) this yields for $\|\cdot\| = \|\cdot\|_u + b\,\|\cdot\|_s$ with a sufficiently large $b=b(N)$

$$
\|\mathcal{L}^n f\| \le A^2\cdot\max\{3\lambda_s^{\beta n},\lambda_u^{-n}\}\,\|f\| + K'_n\,\|f\|_w \quad\text{for } n=1,\dots,N \tag{3.1.8}
$$

and, by choice of N,

$$
\|\mathcal{L}^N f\| \le \frac{1}{3}\sigma^N\,\|f\| + K'_N\|f\|_w\ . \tag{3.1.9}
$$

Combining and iterating the last two estimates (and observing (3.1.6)) finally yields

$$
\|\mathcal{L}^n f\| \le 3A^2\sigma^n\,\|f\| + B\,\|f\|_w \quad\text{for all } n=1,2,\dots \tag{3.1.10}
$$

for a suitable constant B that depends on N. To conclude the proof of the Lemma note that (3.1.6) and (3.1.10) extend by continuity to $\mathcal{B}_w(\mathcal{M})$ and $\mathcal{B}(\mathcal{M})$ respectively. The boundedness of $\mathcal{L}$ is then obvious. □

Proof of Sub-lemma 3.1.3: The basic idea is to define $\tilde{\varphi}(x)$ as an average of φ on a small "ball", centred at x, in $W^s(x)$.

Since only uniform Hölder continuity along stable leaves is demanded from the approximating function $\tilde{\varphi}$ (except for global measurability, of course), its construction can be carried out "leaf-wise". We fix a stable leaf $W = W^s(x_0)$. Since W is a stable leaf, it is a d^s-dimensional $\mathcal{C}^3$ manifold with the Riemannian structure inherited from $\mathcal{M}$ (as smooth as the map T). In order to simplify the exposition we restrict again to the case $\mathcal{M} = \mathbb{T}^d$ where all tangent spaces $\mathcal{T}_x W$ can be naturally immersed into $\mathbb{T}^d$, yet the general case can be treated in essentially the same way.

For each $x, y \in W$ we will define an isometric identification $I_{x,y} : \mathcal{T}_x W \to \mathcal{T}_y W$ of the respective tangent spaces.[23] In addition, for each $x \in W$ we consider a chart $\Upsilon_x : B_{\rho_*} \subseteq \mathcal{T}_x W \to W$ where $B_{\rho_*} = \{\xi \in \mathbb{R}^{d_s} : \|\xi\| < \rho_*\}$ with a radius $\rho_* > 0$ that can be chosen the same for all W and all $x \in W$. We will use Υ_x to define "balls centred at x" in W.[24] The charts Υ_x enjoy the following useful properties: There exists a constant $C_{\rho_*} > 0$ (also independent of W and of x) such that for $0 < \rho < \frac{\rho_*}{4}$ and sufficiently small $\delta > 0$ holds:

A) For all $x \in W$ and all $\xi \in B_\rho$,
$$(1 + C_{\rho_*}\rho)^{-1} \le J\Upsilon_x(\xi) \le (1 + C_{\rho_*}\rho)$$
where $J\Upsilon_x$ denotes the Jacobian determinant of Υ_x.

B) For all $x, y \in W$ with $d^s(x, y) < \delta < \frac{\rho_*}{2}$ and all $\xi \in B_\rho$,
$$d^s(\Upsilon_x(\xi), \Upsilon_y(I_{x,y}\xi)) \le (1 + C_{\rho_*}\rho)\, d^s(x, y)$$
and
$$|J\Upsilon_x(\xi) - J\Upsilon_y(I_{x,y}\xi)| \le C_{\rho_*}\, d^s(x, y)\ .$$

C) For all $x, y \in W$ with $d^s(x, y) < \delta < \frac{\rho_*}{2}$,
$$\frac{\mathrm{vol}(\Upsilon_x^{-1}\Upsilon_y I_{x,y} B_\rho \setminus B_\rho)}{\mathrm{vol}(B_\rho)} \le C_{\rho_*}\, \rho^{-1} \cdot d^s(x, y)\ .$$

Before defining the maps Υ_x let us see how such maps allow to construct the wanted smooth approximation of φ.

Let $\rho \le \frac{\delta}{2}$, denote by m^s the Riemannian volume on W and define for $\varphi \in \mathcal{D}_\beta$ and $x \in W$
$$\tilde{\varphi}(x) := \frac{\int_{\Upsilon_x B_\rho} \varphi(y)\, m^s(dy)}{m^s(\Upsilon_x B_\rho)} = \frac{\int_{B_\rho} \varphi(\Upsilon_x\xi)\, J\Upsilon_x(\xi)\, d\xi}{\int_{B_\rho} J\Upsilon_x(\xi)\, d\xi} = \frac{1}{m(B_\rho)} \int_{B_\rho} \varphi(\Upsilon_x\xi)\, K(x, \xi)\, d\xi$$
where $K(x, \xi) := \frac{J\Upsilon_x(\xi)}{m(B_\rho)^{-1} \int_{B_\rho} J\Upsilon_x(\zeta)\, d\zeta}$. Then
$$K(x, \xi) \le (1 + C_{\rho_*}\rho)^2 \qquad \text{and} \qquad |K(x, \xi) - K(y, I_{x,y}\xi)| \le \mathrm{const} \cdot d^s(x, y)$$
because $I_{x,y}$ maps the ball $B_\rho \subset \mathcal{T}_x W$ onto the ball $B_\rho \subset \mathcal{T}_y W$ isometrically. It follows
$$|\tilde{\varphi}|_\infty \le |\varphi|_\infty \le 1 \quad \text{and} \quad |\tilde{\varphi} - \varphi|_\infty \le 2^{-\beta}\, \delta^\beta\, H^s_\beta(\varphi)\ .$$
We estimate the Hölder constant of $\tilde{\varphi}$. To this end consider $x, y \in W$ with $d^s(x, y) < \delta$. Then
$$\begin{aligned}
|\tilde{\varphi}(y) - \tilde{\varphi}(x)| &\le \frac{1}{m(B_\rho)} \left| \int_{B_\rho} [\varphi(\Upsilon_x\xi) - \varphi(\Upsilon_y I_{x,y}\xi)]\ K(x, \xi)\, d\xi \right| \\
&\quad + \frac{1}{m(B_\rho)} \left| \int_{B_\rho} \varphi(\Upsilon_y I_{x,y}\xi)\ [K(x, \xi) - K(y, I_{x,y}\xi)]\ d\xi \right| \\
&\le H^s_\beta(\varphi)\ (1 + C_{\rho_*}\rho)^{2+\beta}\ d^s(x, y)^\beta + \mathrm{const} \cdot d^s(x, y) \\
&\le (1 + C_{\rho_*}\rho)^3\ d^s(x, y)^\beta + \mathrm{const} \cdot \delta^{1-\beta}\, d^s(x, y)^\beta\ .
\end{aligned} \tag{3.1.11}$$

[23] This can be done in many ways. On general manifolds the natural choice is to do it by parallel transport. Here we will make a simpler and more naïve choice that uses explicitly the possibility to identify locally $\mathbb{T}^d$ with its tangent space.

[24] For general manifolds the best choice seems to be to consider the exponential map $\Upsilon_x := \exp^s_x : B_{\rho_*} \subseteq \mathcal{T}_x W \to W$. We do not do this in the following since this seems to yield weaker estimates (although sufficient for our present needs) and since the proof of the following properties (A-C) would need a more heavy machinery from differential geometry.

Choose now $\rho = \frac{\delta}{2}$. Then this is the desired estimate on $H^s_\beta(\tilde\varphi)$.

In order to estimate $H^s_\gamma(\tilde\varphi)$ we proceed differently. To this end denote $\tilde B_\rho = \Upsilon_x^{-1}\Upsilon_y I_{x,y} B_\rho$. Then

$$\left|\int_{B_\rho} \varphi(\Upsilon_x\xi)\, J\Upsilon_x(\xi)\, d\xi - \int_{B_\rho} \varphi(\Upsilon_y I_{x,y}\xi)\, J\Upsilon_y(I_{x,y}\xi)\, d\xi\right| \\ \leq \int_{B_\rho\setminus\tilde B_\rho} |\varphi(\Upsilon_x\xi)|\, J\Upsilon_x(\xi)\, d\xi \leq (1 + C_{\rho_*}\rho)\ \mathrm{vol}(B_\rho\setminus\tilde B_\rho)$$

Hence

$$\begin{aligned}
|\tilde\varphi(y) - \tilde\varphi(x)| &\leq \left|\int_{B_\rho} J\Upsilon_x(\xi)\, d\xi - \int_{B_\rho} J\Upsilon_y(I_{x,y}\xi)\, d\xi\right| \cdot \left|\frac{\int_{B_\rho}\varphi(\Upsilon_y I_{x,y}\xi)\, J\Upsilon_y(I_{x,y}\xi)\, d\xi}{\int_{B_\rho} J\Upsilon_y(I_{x,y}\xi)\, d\xi\cdot\int_{B_\rho} J\Upsilon_x(\xi)\, d\xi}\right| \\
&\quad + \frac{1}{\int_{B_\rho} J\Upsilon_x(\xi)\, d\xi}\cdot\left|\int_{B_\rho}\varphi(\Upsilon_y I_{x,y}\xi)\, J\Upsilon_y(I_{x,y}\xi)\, d\xi - \int_{B_\rho}\varphi(\Upsilon_x\xi)\, J\Upsilon_x(\xi)\, d\xi\right| \\
&\leq \mathrm{const}\cdot|\varphi|_\infty \frac{\mathrm{vol}(B_\rho\setminus\tilde B_\rho)}{\mathrm{vol}(B_\rho)} \leq \mathrm{const}\cdot\rho^{-1}\, d^s(x,y) \qquad (3.1.12)\\
&\leq \mathrm{const}\cdot\delta^{-\gamma}\, d^s(x,y)^\gamma
\end{aligned}$$

with a constant that depends only on C_{ρ^*}, because $\rho = \frac{\delta}{2}$ and $d^s(x,y) \leq \delta$. It follows that $H^s_\gamma(\tilde\varphi) \leq \mathrm{const}\ \delta^{-\gamma}$.

To conclude we need to finally define the charts Υ_x and the identifications $I_{x,y}$ and prove properties (A-C). Throughout the rest of this proof various constants C will appear that we do not specify precisely, but that all depend at most on the first and second derivatives of the stable manifolds W which in turn are uniformly bounded, see the appendix for a reference. In addition, they depend on a suitably small choice of the radius ρ_*.

To begin with, we define for each pair $x, y \in W$ a linear map $U_{x,y} : E^s(x) \to E^s(x)^\perp$ by $\xi + U_{x,y}\xi \in E^s(y)$ for each $\xi \in E^s(x)$. Clearly

$$\|U_{x,y}\| \leq C\, d(x,y), \qquad (3.1.13)$$

where C depends only on the second derivative (curvature) of W. In addition, locally we can view W as a graph over $E^s(x)$. Let $F_x : E^s(x) \to E^s(x)^\perp$ such that $x + \zeta + F_x(\zeta) \in W$ for each $\zeta \in B_{\rho_*}(0) \subset E^s(x)$, and define $\Upsilon_x : E^s(x) \to W$, [25]

$$\Upsilon_x(\zeta) := x + \zeta + F_x(\zeta)\ . \qquad (3.1.14)$$

It is easy to see that $d_\zeta F_x = U_{x,\Upsilon_x(\zeta)}$ and hence $d_\zeta\Upsilon_x = \mathrm{Id} + U_{x,\Upsilon_x(\zeta)}$ so that $\|F_x(\zeta)\| \leq C\|\zeta\|$, see also Figure 2. Similarly, $d(x, \Upsilon_x(\zeta)) \leq C\|\zeta\|$. This implies in particular that Υ_x is a diffeo-

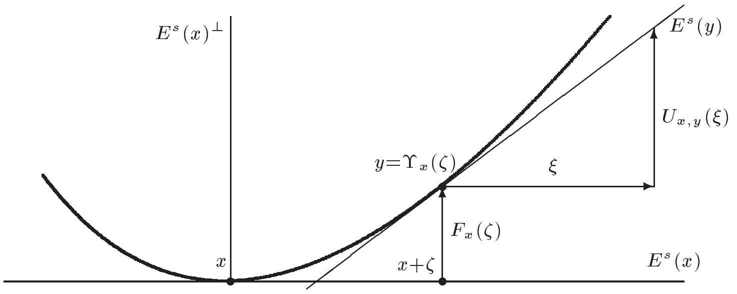

FIGURE 2. The chart Υ_x and related quantities

morphism close to the identity. Note also that

$$J_\zeta\Upsilon_x := \sqrt{\det(\mathrm{Id} + U^*_{x,\Upsilon_x(\zeta)} U_{x,\Upsilon_x(\zeta)})}. \qquad (3.1.15)$$

[25]Remark that here, and in the following, we are using explicitly the (in $\mathbb{T}^d$) trivial identification of tangent vectors and vectors in the space.

The last ingredient needed is a way to identify the tangent space of W at different points. Choose a basis $\{e_i\}$ of $E^s(x)$, then $\{e_i + U_{x,y}e_i\} =: \{\tilde{b}_i\}$ is a basis of $E^s(y)$, unfortunately it may not be orthonormal. Yet,

$$\langle \tilde{b}_i, \tilde{b}_j\rangle = \delta_{ij} + \langle e_i, U^*_{x,y}U_{x,y}e_j\rangle =: (\mathbf{1} + A)_{ij}.$$

This means that setting $\Lambda := (\mathrm{Id} + A)^{-\frac{1}{2}}$ the new basis $b_i := \sum_j \Lambda_{ij}\tilde{b}_j$ is orthonormal. We then choose the isometric identification $I_{x,y}e_i := b_i$. We are now in the position to prove properties (A-C).

A) This is nothing else than (3.1.13) and (3.1.15).

B) Consider $x, y \in W$ and let $\zeta \in E^s(x)$ such that $\Upsilon_x(\zeta) = y$. Consider the map $\Xi_{x,y} : E^s(x) \to E^s(x)$ defined by

$$\Upsilon_y \circ I_{x,y}(\xi) = \Upsilon_x(\zeta + \Xi_{x,y}(\xi)). \tag{3.1.16}$$

We can apply the orthogonal projection Π_x on $E^s(x)$ to both sides of the above equation to obtain

$$\Xi_{x,y}(\xi) = \Pi_x I_{x,y}\xi + \Pi_x F_y(I_{x,y}\xi). \tag{3.1.17}$$

Note that if we consider $I_{x,y}$ as a map in the ambient space, as above, then for $\xi = \sum_i \xi_i e_i$ we have

$$I_{x,y}\xi = \sum_i \xi_i b_i = \sum_{ij} \Lambda_{ij}\xi_i\tilde{b}_j$$

and thus it immediately follows $\|I_{x,y}\xi - \xi\| \leq C\, d(x,y)\|\xi\|$, which implies

$$\|\Xi_{x,y}(\xi) - \xi\| \leq \|\Pi_x(I_{x,y}\xi - \xi)\| + \|(\Pi_x - \Pi_y)(F_y I_{x,y}\xi)\| \leq C\,\|\xi\| d(x,y)$$

because $x \mapsto \Pi_x$ is a $\mathcal{C}^1$ function. By using the above facts and recalling the definition of $\Xi_{x,y}$ in equation (3.1.16) we can compute

$$\begin{aligned}\|\Upsilon_y \circ I_{x,y}(\xi) - \Upsilon_x(\xi)\| &= \|\Upsilon_x(\zeta + \Xi_{x,y}(\xi)) - \Upsilon_x(\xi)\| \\ &\leq \sup_{\xi\in B_\rho} \|d_\xi \Upsilon_x\| \cdot \|\zeta + \Xi_{x,y}(\xi) - \xi\| \\ &\leq (1 + C\,\rho)\,(\|\zeta\| + \|\Xi_{x,y}(\xi) - \xi\|) \leq (1 + C\,\rho) d(x,y),\end{aligned}$$

where all the constants C depend only on the curvature of W.

To obtain the estimate on the Jacobian we need to differentiate (3.1.16) which yields as before (observing that $JI_{x,y} = 1$)

$$\begin{aligned}|J\Upsilon_y(I_{x,y}\xi) - J\Upsilon_x(\xi)| &= |J\Upsilon_x(\zeta + \Xi_{x,y}(\xi))J\Xi_{x,y}(\xi) - J\Upsilon_x(\xi)| \\ &\leq (1 + C\,\rho)\, d(x,y) + C\,|1 - J\Xi_{x,y}(\xi)|\ .\end{aligned}$$

The wanted estimate follows now from $d_\xi \Xi_{x,y} = \Pi_x I_{x,y} + \Pi_x U_{y,\Upsilon_y(I_{x,y}\xi)} I_{x,y}$.

C) Again we use (3.1.16) to have

$$\Upsilon_x^{-1} \circ \Upsilon_y(\xi) = \zeta + \Xi_{x,y}(I^{-1}_{x,y}\xi)$$

this means that

$$\big|\|\Upsilon_x^{-1} \circ \Upsilon_y(\xi)\| - \|\xi\|\big| \leq C\, d(x,y),$$

from which C) immediately follows. [26] □

Proof of Lemma 2.2.5. Since $\mathcal{L}^{-n}$ is the transfer operator of T^{-n}, we have $\int \mathcal{L}^{-n} f\,\varphi\, dm = \int f\,(\varphi \circ T^{-n})\, dm$ for each $\varphi \in \mathcal{D}_\beta$ and $|\varphi \circ T^{-n}|_\infty = |\varphi|_\infty \leq 1$. In order to estimate $H^s_\beta(\varphi \circ T^{-n})$ let $x, y \in \mathcal{M}$ with $y \in W^s(x)$ and $d^s(x,y) < \delta$. Since T^{-n} expands $W^s(x)$, $d^s(T^{-n}x, T^{-n}y)$ may be as large as $A\mu_s^{-1}\delta > \delta$. Therefore let $\ell = \ell(n)$ be the smallest integer such that $\ell d^s(x,y) >$

[26] In the case of general manifolds, using for $I_{x,y}$ and Υ_x the choices specified in Footnotes 23 and 24, the same results can be obtained, *i.e.* $d(\Upsilon_x \circ I_{x,y})$ can be computed, as usual, via Jacobi fields, [13, Chapter 5]. Doing so the constants would depend also on the derivative of the curvature, which would make no difference, however, since in view of our $\mathcal{C}^3$ hypothesis on T and of the results in the appendix, the curvature of the stable and unstable manifolds is uniformly $\mathcal{C}^1$.

$d^s(T^{-n}x, T^{-n}y)$ and fix a chain of points $z_0 = T^{-n}x, z_1, \dots, z_\ell = T^{-n}y$ such that $d^s(z_{i-1}, z_i) \le d^s(x, y)$ for all i. Since $d^s(T^{-n}x, T^{-n}y) \le A\mu_s^{-n} d^s(x, y)$, it follows that $\ell \le A\mu_s^{-n} + 1$ and hence

$$\frac{|\varphi(T^{-n}x) - \varphi(T^{-n}y)|}{d^s(x,y)^\beta} \le \sum_{i=1}^{\ell} \frac{|\varphi(z_{i-1}) - \varphi(z_i)|}{d^s(z_{i-1}, z_i)^\beta} \cdot \frac{d^s(z_{i-1}, z_i)^\beta}{d^s(x,y)^\beta} \le (A\mu_s^{-n} + 1) \cdot H_\beta^s(\varphi) \; .$$

It follows that $\|\mathcal{L}^{-n} f\|_s \le (A\mu_s^{-n} + 1)\|f\|_s$. Exactly the same estimate holds also for the $\|\cdot\|_w$–norm.

For the unstable norm one proceeds as in (3.1.4) and in Sub-lemma 3.1.1, taking the same care when estimating the Hölder constant of $v \circ T^{-n}$ along stable fibres that we took above in the case of $\varphi \circ T^{-n}$. For sufficiently small $\delta > 0$ this leads to the estimate

$$\|\mathcal{L}^{-n} f\|_u \le A\mu_u^n\,(A\mu_s^{-n} + 1)\|f\|_u + K_n\|f\|_s$$

and for large enough b we obtain with arguments analogous to those leading to (3.1.10) that

$$\|\mathcal{L}^{-n} f\| \le A\mu_u^n\,(A\mu_s^{-n} + 2)\|f\| + K_n'\|f\|_w \; .$$

Iterating this estimate for fixed n and recalling that $\|\mathcal{L}^{-n} f\|_w \le (A\mu_s^{-n} + 1)\|f\|_w$ leads to the estimate $\rho_{\mathrm{sp}}(\mathcal{L}^{-n}) = \limsup_{k\to\infty} \sqrt[kn]{\|\mathcal{L}^{-kn}\|} \le \sqrt[n]{A\mu_u^n\,(A\mu_s^{-n} + 2)}$. Since we can choose $n \in \mathbb{N}$ as large as we like, $\rho_{\mathrm{sp}}(\mathcal{L}^{-1}) \le \mu_u\mu_s^{-1}$. □

3.2. Proofs: Peripheral spectrum and SRB measures.

Proof of Proposition 2.3.1. Let us start with $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$. For $\varphi \in \mathcal{C}^0(\mathcal{M}, \mathbb{R}) \cap \mathcal{D}_\gamma$, and $|\lambda| = 1$,

$$\langle \Pi_\lambda f, \varphi\rangle = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \lambda^{-k} \langle \mathcal{L}^k f, \varphi\rangle = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \lambda^{-k} \int \varphi \circ T^k \cdot f\, dm \; . \tag{3.2.1}$$

Hence $\langle \Pi_\lambda f, \varphi\rangle \le |\varphi|_\infty \cdot \int |f|\, dm$, so that $\Pi_\lambda f$ extends to a continuous linear functional on $\mathcal{C}^0(\mathcal{M}, \mathbb{R})$, and by Riesz' theorem there exists $\mu_{\Pi_\lambda f}$ such that $\langle \Pi_\lambda f, \varphi\rangle = \int \varphi\, d\mu_{\Pi_\lambda f}$ for each $\varphi \in \mathcal{C}^0(\mathcal{M}, \mathbb{R}) \cap \mathcal{D}_\gamma$.

The equality for all the functions in $\mathcal{D}_\gamma$ follows by an approximation argument. The idea is to approximate a function in $\mathcal{D}_\gamma$ by a smooth function but keeping under control the regularity along the stable direction. To this end it is convenient to introduce the following singular convolution:

$$\mathbb{A}_\epsilon(\varphi)(x) := \frac{1}{Z_\epsilon(x)} \int_{W_\epsilon^s(x)} \varphi(y) m_s(dy), \tag{3.2.2}$$

where $W_\epsilon^s(x)$ is the intersection of the local stable manifold with the ball of radius ϵ centred at x, m_s is the Riemannian measure restricted to the stable manifold, and $Z_\epsilon(x)$ is chosen so that $\mathbb{A}_\epsilon 1 = 1$. Obviously,

$$\|\mathbb{A}_\epsilon \varphi - \varphi\|_\infty \le \epsilon^\gamma \quad \text{for } \varphi \in \mathcal{D}_\gamma \; . \tag{3.2.3}$$

Sub-lemma 3.2.1. *There exist constants $c, \epsilon_* > 0$ such that for each $\epsilon \le \epsilon_*$ we have $c\epsilon\mathbb{A}_\epsilon : \mathcal{C}^0(\mathcal{M}, \mathbb{R})_1 \to \mathcal{C}^0(\mathcal{M}, \mathbb{R}) \cap \mathcal{D}_\gamma$.*

Proof. Clearly $\mathbb{A}_\epsilon$ is bounded with respect to the sup norm. If $y \in W_\delta^s(x)$ then, by the uniform smoothness of the stable manifold it follows that there exists $c_0 > 0$ such that, for ϵ small enough

$$\left| \int_{W_\epsilon^s(x)\setminus W_\epsilon^s(y)} \varphi(z) m_s(dz) \right| \le |\varphi|_\infty m_s\left(\{z \in W_{2\epsilon}^s(x) \;:\; |\epsilon - d(z,x)| \le c_0 d(x,y)\}\right).$$

From the above equation, remembering that the stable manifolds have uniformly bounded curvature, it follows that there exists $c > 0$ such that $c\epsilon\mathbb{A}_\epsilon : L^\infty(\mathcal{M}, \mathbb{R})_1 \to \mathcal{D}_\gamma$. The continuity of $c\epsilon A_\epsilon(\varphi)$ is an immediate consequence of the continuity of the stable foliation. □

Notice that $\mu_{\lambda,f,\epsilon}(\varphi) := \mu_{\Pi_\lambda f}(\mathbb{A}_\epsilon\varphi)$ and $m_\epsilon(\varphi) := m(\mathbb{A}_\epsilon\varphi)$ are measures. By Lusin's theorem, each $\varphi \in \mathcal{D}_\gamma$ can be approximated $(m + \mu_{\lambda,f,\epsilon})$-almost everywhere by a sequence $\{\varphi_n\} \subset \mathcal{C}^0(\mathcal{M}, \mathbb{R})$, $|\varphi_n|_\infty \le 1$, and in view of Lebesgue's dominated convergence theorem also the m– and $\mu_{\lambda,f,\epsilon}$–integrals of $|\varphi - \varphi_n|$ tend to zero.

Now fix $\varphi \in \mathcal{D}_\gamma$ and $\epsilon > 0$. Because of (2.3.1) there is $N = N(\epsilon) \in \mathbb{N}$ such that

$$\left\| \frac{1}{N} \sum_{k=0}^{N-1} \lambda^{-k} \mathcal{L}^k f - \Pi_\lambda f \right\|_w \le \epsilon^2 .$$

Hence, recalling (3.2.3) and Sub-lemma 3.2.1, we have

$$|\mu_{\Pi_\lambda f}(\varphi) - \mu_{\Pi_\lambda f}(\mathbb{A}_\epsilon \varphi)| \le \epsilon^\gamma$$

$$|\mu_{\Pi_\lambda f}(\mathbb{A}_\epsilon \varphi) - \langle \Pi_\lambda f, \mathbb{A}_\epsilon \varphi_n \rangle| = |\mu_{\Pi_\lambda f}(\mathbb{A}_\epsilon \varphi) - \mu_{\Pi_\lambda f}(\mathbb{A}_\epsilon \varphi_n)| \le \int |\varphi - \varphi_n| \, d\mu_{\lambda,f,\epsilon}$$

$$\left| \langle \Pi_\lambda f, \mathbb{A}_\epsilon \varphi_n \rangle - \frac{1}{N} \sum_{k=0}^{N-1} \lambda^{-k} \langle \mathcal{L}^k f, \mathbb{A}_\epsilon \varphi_n \rangle \right| \le (c\epsilon)^{-1} \epsilon^2 = c^{-1}\epsilon$$

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} \lambda^{-k} \langle \mathcal{L}^k f, \mathbb{A}_\epsilon (\varphi_n - \varphi) \rangle \right| \le \max_{k=0,\dots,N-1} \|\mathcal{L}^k f\|_\infty \int |\varphi_n - \varphi| \, dm_\epsilon$$

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} \lambda^{-k} \langle \mathcal{L}^k f, \mathbb{A}_\epsilon \varphi - \varphi \rangle \right| \le \|f\|_1 \, \epsilon^\gamma$$

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} \lambda^{-k} \langle \mathcal{L}^k f, \varphi \rangle - \langle \Pi_\lambda f, \varphi \rangle \right| \le \epsilon^2$$

The second and the fourth expression tend to 0 as $n \to \infty$ for fixed ϵ and N. Hence

$$|\mu_{\Pi_\lambda f}(\varphi) - \langle \Pi_\lambda f, \varphi \rangle| \le \epsilon^\gamma (1 + \|f\|_1) + c^{-1}\epsilon + \epsilon^2 .$$

Hence $\langle \Pi_\lambda f, \varphi \rangle = \mu_{\Pi_\lambda f}(\varphi)$ for all $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$ and all $\varphi \in \mathcal{D}_\gamma$.

If $\lambda = 1$ and $f, \varphi \ge 0$, then $\int \varphi \, d\mu_{\Pi_1 f} = \langle \Pi_1 f, \varphi \rangle \ge 0$ and $\langle \Pi_1 f, 1 \rangle = \int f \, dm$ by (3.2.1). It follows that $\mu_{\Pi_1 1}$ is a positive measure and $\lambda = 1$ is an eigenvalue of $\mathcal{L}$. Finally equation (3.2.1) implies that for $\varphi \ge 0$

$$|\mu_{\Pi_j f}(\varphi)| = |\langle \Pi_j f, \varphi \rangle| \le \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \int \varphi \circ T^k \cdot |f| \, dm \le \|f\|_\infty \langle \Pi_1 1, \varphi \rangle = \|f\|_\infty \int \varphi \, d\mu .$$

It follows that $\mu_{\Pi_j f}$ is absolutely continuous with respect to μ and $\left| \frac{d\mu_{\Pi_j f}}{d\mu} \right| \le \|f\|_\infty$ μ-almost surely. Observe also that, because of (2.3.3), each $h \in \Pi_j \mathcal{B}$ can be represented as $h = \Pi_j f$ for some $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$.

Next, if $h = \Pi_1 f$, then $T^* \mu_h = \mu_h$ follows since

$$\int \varphi \, d(T^* \mu_h) = \int \varphi \circ T \, d\mu_h = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \int \varphi \circ T^{k+1} f \, dm = \int \varphi \, d\mu_h$$

for all $\varphi \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$. The last statement follows immediately from the observation $\langle \Pi_1 f, 1 \rangle = \langle f, 1 \rangle$, see (3.2.1). □

Proof of Proposition 2.3.2. a) For each $\varphi \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$,

$$\lim_{n \to \infty} \int \varphi \, d\left(\frac{1}{n} \sum_{k=0}^{n-1} T^{*k} m \right) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \int \varphi \circ T^k \, dm = \langle \Pi_1 1, \varphi \rangle = \int \varphi \, d\mu .$$

b) Let $|\lambda| = 1$. Recall from (2.3.2) and (2.3.3) that $\Pi_1 : \mathcal{B}_w \to \mathcal{B}$ is a bounded linear operator with $\Pi_1(\mathcal{B}_w) = \Pi_1(\mathcal{B})$.

Introduce the map $\Phi : \Pi_1 \mathcal{B} \to L^\infty_{\mu,\mathrm{inv}} := \{h \in L^\infty_\mu(\mathcal{M}, \mathbb{R}) : h \circ T = h\}$ by the relation

$$\Phi(f) = \frac{d\mu_f}{d\mu}$$

with μ_f and μ as in Proposition 2.3.1.[27] Then

- Φ is linear by inspection of the definition
- Φ is injective, because

$$\Phi(f) = 0 \Rightarrow \mu_f = 0 \Rightarrow \langle f, \phi \rangle = 0 \, \forall \, \phi \in \mathcal{D}_\gamma \Rightarrow \|f\|_w = 0,$$

[27] $\Phi(f) \circ T = \Phi(f)$ because both measures, μ_f and μ, are T-invariant.

and since $f \in \Pi_1\mathcal{B}$, *i.e.* $f = \Pi_1 f$, it follows from (2.3.2) that also $\|f\| = 0$ and hence $f = 0$.

- Φ is surjective, as we show below.

Hence $\dim(L^\infty_{\mu,\mathrm{inv}}) = \dim(\Pi_1\mathcal{B})$ is the number of ergodic components of μ.

Proof of the surjectivity of Φ: Let $h \in L^\infty_{\mu,\mathrm{inv}}$ and choose $h_n \in \mathcal{C}^1(\mathcal{M},\mathbb{C}) \subseteq \mathcal{D}_\gamma$ [28] such that $\lim_{n\to\infty}\int |h - h_n|\,d\mu = 0$. Then, for each $\varphi \in \mathcal{D}_\gamma$,

$$\begin{aligned}\langle h_n\cdot\Pi_1 1 - h_k\cdot\Pi_1 1,\varphi\rangle = \langle \Pi_1 1,(h_n-h_k)\cdot\varphi\rangle &= \int (h_n-h_k)\cdot\varphi\,d\mu\\ &\le \int |h_n-h_k|\,d\mu\end{aligned}$$

so that $\|h_n\cdot\Pi_1 1 - h_k\cdot\Pi_1 1\|_w \le \int |h_n-h_k|\,d\mu \to 0$ when $n,k\to\infty$. It follows that

$$f := \|\cdot\|_w\text{-}\lim_{n\to\infty} h_n\cdot\Pi_1 1 \in \mathcal{B}_w$$

exists.

To show that $\mathcal{L}f = f$, consider

$$\begin{aligned}\|\mathcal{L}f - f\|_w &= \lim_{n\to\infty}\|\mathcal{L}(h_n\cdot\Pi_1 1) - h_n\cdot\Pi_1 1\|_w\\ &= \lim_{n\to\infty}\sup_{\varphi\in\mathcal{D}_\gamma}\left(\langle\mathcal{L}(h_n\cdot\Pi_1 1),\varphi\rangle - \langle h_n\cdot\Pi_1 1,\varphi\rangle\right)\\ &= \lim_{n\to\infty}\sup_{\varphi\in\mathcal{D}_\gamma}\left(\langle\Pi_1 1, h_n\cdot(\varphi\circ T)\rangle - \langle\Pi_1 1, h_n\cdot\varphi\rangle\right)\\ &= \lim_{n\to\infty}\sup_{\varphi\in\mathcal{D}_\gamma}\int h_n\cdot(\varphi\circ T-\varphi)\,d\mu\\ &\le \sup_{\varphi\in\mathcal{D}_\gamma}\int h\cdot(\varphi\circ T-\varphi)\,d\mu + \lim_{n\to\infty}\underbrace{\sup_{\varphi\in\mathcal{D}_\gamma}\int (h_n-h)\cdot(\varphi\circ T-\varphi)\,d\mu}_{\le 2\int|h_n-h|\,d\mu}\\ &= \sup_{\varphi\in\mathcal{D}_\gamma}\left(\int (h\cdot\varphi)\circ T\,d\mu - \int h\cdot\varphi\,d\mu\right)\quad(\text{since } h\circ T = h)\\ &= 0.\end{aligned}$$

It follows that $f = \Pi_1 f \in \Pi_1\mathcal{B}_w$. Hence

$$\int \varphi\cdot h\,d\mu = \lim_{n\to\infty}\int\varphi\cdot h_n\,d\mu = \lim_{n\to\infty}\langle h_n\cdot\Pi_1 1,\varphi\rangle = \langle f,\varphi\rangle = \int\varphi\,d\mu_f$$

for all $\varphi \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$, so that $h = \frac{d\mu_f}{d\mu} = \Phi(f)$.

Since $\dim(L^\infty_{\mu,\mathrm{inv}}) < \infty$, this space has a basis consisting of indicator functions of pairwise disjoint sets A_i. Let $f_i := \Phi^{-1}\,\mathbf{1}_{A_i}$ and $\mu_i := \mu_{f_i}$. Then $\frac{d\mu_i}{d\mu} = \mathbf{1}_{A_i}$.

c) The third statement follows from Birkhoff's Theorem. Indeed, if $\varphi \in \mathcal{C}^0(\mathcal{M},\mathbb{R})$ there exists a measurable T-invariant set $\Omega_\varphi \subset \mathcal{M}$, $\mu(\Omega_\varphi) = 1$, such that

$$\varphi^+(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\varphi\circ T^i(x) \text{ exists } \forall x\in\Omega_\varphi.$$

The characteristic function $\mathbf{1}_{\Omega_\varphi}$ of Ω_φ is a measurable function, moreover, if $x \in \Omega_\varphi$ and $y \in W^s(x)$, $d^s(x,y) \le \delta$, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\varphi\circ T^i(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\varphi\circ T^i(y)$$

hence $y \in \Omega_\varphi$. Accordingly, $\mathbf{1}_{\Omega_\varphi} \in \mathcal{D}_\gamma$. Consequently

$$m(\Omega_\varphi) = m(\mathbf{1}_{\Omega_\varphi}\circ T^k) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}T^{*k}m(\mathbf{1}_{\Omega_\varphi}) = \langle\Pi_1 1,\mathbf{1}_{\Omega_\varphi}\rangle = \mu(\Omega_\varphi) = 1,$$

where we have used Proposition 2.3.1. We can then define, for each $t \in \mathbb{R}$, the sets $\Omega_{\varphi,t} = \{x \in \Omega_\varphi : \varphi^+(x) \le t\}$. Clearly these are invariant sets as well and their characteristic functions belong to $\mathcal{D}_\gamma$, hence we have $\mu(\Omega_{\varphi,t}) = m(\Omega_{\varphi,t})$. But, in view of what we just proved in b), there can be only finitely many invariant sets of positive μ measure. If μ is ergodic it follows that $m(\Omega_{\varphi,t}) \in \{0,1\}$

[28] We complexify $\mathcal{D}_\gamma$ tacitly.

for each $t \in \mathbb{R}$ so that $\varphi^+(x) = \int_{\mathcal{M}} \varphi \, d\mu$ for m-almost all x.

d) The proof of this part depends heavily on Lemma 2.1.4 and its proof. So we postpone it to section 3.7. □

3.3. **Proofs: Spectral stability – Smooth random perturbations.** Recall that we are assuming now that $d = \dim \mathcal{M} = 2$, hence the stable and unstable foliations are $\mathcal{C}^\tau$ with, at least, $1 < \tau = 1 + \alpha \leq 2$ and $0 < \beta < \gamma \leq \alpha$. Furthermore we give a detailed proof only for the case $\mathcal{M} = \mathbb{T}^2$.

What makes the $\mathcal{C}^{1+\alpha}$ setting so special, is the following lemma.

Lemma 3.3.1. *Each point $x \in \mathcal{M}$ has a neighbourhood U together with a $\mathcal{C}^{1+\alpha}$-diffeomorphism Γ from U into $\mathbb{R}^2$ which straightens both foliations simultaneously in the following sense:*

1) $\Gamma(x) = 0$,

2) $\Gamma([y, y']) = (\Gamma_1(y), \Gamma_2(y'))$ for $y, y' \in U$, where $[y, y']$ denotes the unique point in $W^u(y) \cap W^s(y')$.

3) Let $\Gamma(y) = (\xi, \eta)$. Then $\{\Gamma^{-1}(t, \eta)\}$ and $\{\Gamma^{-1}(\xi, t)\}$ are the stable and unstable manifold of y, respectively.

4) $\{\Gamma^{-1}(t, 0)\}$ and $\{\Gamma^{-1}(0, t)\}$ are the stable and unstable manifold of x, respectively, parametrised by arc-length.

Proof. Since the stable and unstable distributions are $\mathcal{C}^{1+\alpha}$ (see Remark A.1 in the appendix) the regularity statement follows like Theorem 6.1 in [35] from the main result of [23] which says in our context that since Γ^{-1} is $\mathcal{C}^{1+\alpha}$ in each of its two real variables separately, it is jointly $\mathcal{C}^{1+\alpha}$ in both variables.[29] The third assertion is a restatement of (2) which implies that $\{\Gamma^{-1}(t, \eta)\}$ and $\{\Gamma^{-1}(\xi, t)\}$ are the stable and unstable manifold of $\Gamma^{-1}(\xi, \eta)$. In addition, it is always possible to reparametrise the two coordinate axis so that (4) is satisfied. □

As in the previous section we start by studying the action of smooth averaging operators on test functions. So let $\tilde{q}_\varepsilon : \mathbb{T}^2 \times \mathbb{R}^2 \to \mathbb{R}^k$ be ε^{-1}-dominated by $\tilde{p}_\varepsilon$ as in Definition 2.4.2. In analogy to the definition of Q^*_ε in section 2.4 we define the operator Q^*_ε here by

$$Q^*_\varepsilon \varphi(x) = \int_{\mathbb{R}^2} \tilde{q}_\varepsilon(x, \xi) \, \varphi(x + \xi) \, d\xi \ .$$

Here $\tilde{q}_\varepsilon$ is a $\mathbb{R}^k$-valued function (where $k > 1$ is possible), φ is $\mathbb{R}$-valued. Therefore $Q^*_\varepsilon \varphi$ is also $\mathbb{R}^k$-valued.

Lemma 3.3.2. *Let $q_\varepsilon : \mathbb{T}^2 \times \mathbb{R}^2 \to \mathbb{R}^k$ be such that $\tilde{q}_\varepsilon$ is ε^{-1}-dominated by the stochastic kernel $\tilde{p}_\varepsilon$ in the sense of Definition 2.4.2. Then there exist $K > 0$ and $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0]$ and for each smooth function φ holds*

$$H^s_\beta(Q^*_\varepsilon \varphi) \leq (1 + K\varepsilon^\alpha)^\beta H^s_\beta(\varphi) + K\delta^{\alpha-\beta} |\varphi|_\infty .$$

The same holds with γ instead of β.

Proof. Given $x \in \mathcal{M}$ let us consider a neighbourhood U that is foliated on the one hand by the leaves $W^s_{\mathrm{loc}}(y)$, $y \in U$, and, on the other hand, by the leaves $W^u_{\mathrm{loc}}(y)$, $y \in U$. For each $z \in W^s_{\mathrm{loc}}(x)$, define the two functions $\Sigma_z, \Psi_z : U \subset \mathcal{M} \to \mathcal{M}$ by[30]

$$\Sigma_z(y) := z + y - x$$

$$\Psi_z(y) = [\Sigma_z(y), y].$$

where $[y, y']$ denotes the unique point in $W^u_{\mathrm{loc}}(y) \cap W^s_{\mathrm{loc}}(y')$, see Figure 3.

Sub-lemma 3.3.3. *The map Ψ_z is injective and there is a constant $C > 0$, which can be chosen the same for all x and z, such that for all $y \in U$*

$$d(\Sigma_z(y), \Psi_z(y)) \leq C \min \{ d(x, z) \, d(x, y)^\alpha, \, d(x, y) \, d(x, z)^\alpha \} \ . \tag{3.3.1}$$

[29] G.K. thanks Boris Hasselblatt for pointing out to him the relevance of the argumentation in [35] for the present setting. Note that the above argument holds in any dimension provided the distributions are $\mathcal{C}^{1+\alpha}$.

[30] In general Riemannian manifolds the same can be done by simply defining

$$\Sigma_z(y) := \exp_x(\exp_x^{-1}(y) + \exp_x^{-1}(z))$$

.

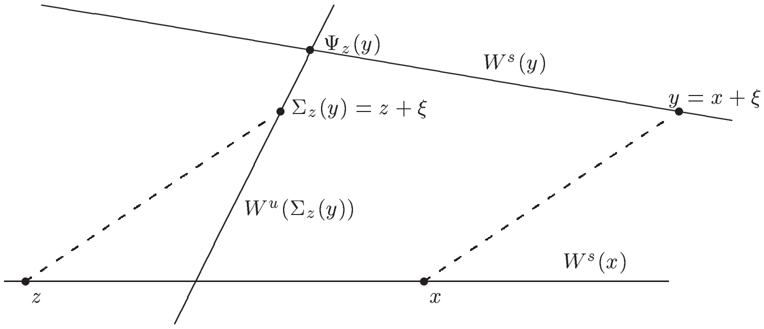

FIGURE 3

In addition Ψ_z is absolutely continuous and its Jacobian $J\Psi_z$ satisfies[31]

$$\sup_{y\in U} |J\Psi_z(y) - 1| \leq C\, d(x,z)^\alpha \ . \tag{3.3.2}$$

Proof. The injectivity of Ψ_z is a consequence of the transversality of the defining foliations $\mathcal{W}^u$ and $\mathcal{W}^s$. The uniformity of all constants is due to the compactness of $\mathcal{M} = \mathbb{T}^2$.

Because of Lemma 3.3.1, in a suitable neighbourhood U of x, both foliations can be straightened simultaneously by some $\mathcal{C}^{1+\alpha}$-diffeomorphism Γ from U onto $\Gamma(U) \subseteq \mathbb{R}^2$. This means in particular that for any $y, z \in U$

$$\begin{aligned} y \in W^u(z) \ &\Rightarrow\ \Gamma_1(y) = \Gamma_1(z) \\ y \in W^s(z) \ &\Rightarrow\ \Gamma_2(y) = \Gamma_2(z) \end{aligned}$$

Without loss of generality we can assume that there is a constant $C_1 > 0$ such that $e^{-C_1} \leq \|D\Gamma_{|y}\| \leq e^{C_1}$ for all $y \in U$.

Now let $z \in W^s(x)$ and $y \in U$. Then

$$\Gamma_1(\Psi_z y) = \Gamma_1(\Sigma_z y), \quad \Gamma_2(\Psi_z y) = \Gamma_2(y) \ . \tag{3.3.3}$$

It follows that

$$\begin{aligned} d(\Psi_z y, \Sigma_z y) &= d\left(\Gamma^{-1}(\Gamma(\Psi_z y)), \Gamma^{-1}(\Gamma(\Sigma_z y))\right) \leq e^{C_1}\ \|\Gamma(\Psi_z y) - \Gamma(\Sigma_z y)\| \\ &= e^{C_1}\ |\Gamma_2(y) - \Gamma_2(\Sigma_z y)| \ . \end{aligned}$$

Let $\xi := y - x$. There are two ways to estimate $|\Gamma_2(y) - \Gamma_2(\Sigma_z y)|$, and we will need them both. Since $\Gamma_2(x) = \Gamma_2(z)$ we have

$$\begin{aligned} |\Gamma_2(y) - \Gamma_2(\Sigma_z y)| &= |(\Gamma_2(x+\xi) - \Gamma_2(z+\xi)) - (\Gamma_2(x) - \Gamma_2(z))| \\ &\leq \sup_{t\in[0,1]} \|D\Gamma_2|_{z+t\xi} - D\Gamma_2|_{x+t\xi}\| \cdot \|\xi\| \\ &\leq C_2\, d(x,z)^\alpha\, d(x,y), \end{aligned} \tag{3.3.4}$$

because Γ is of class $\mathcal{C}^{1+\alpha}$. Similarly,

$$\begin{aligned} |\Gamma_2(y) - \Gamma_2(\Sigma_z y)| &= |(\Gamma_2(x+\xi) - \Gamma_2(x)) - (\Gamma_2(z+\xi) - \Gamma_2(z))| \\ &\leq \sup_{t\in[0,1]} \left\|D\Gamma_2|_{x+\xi+t(z-x)} - D\Gamma_2|_{x+t(z-x)}\right\| \cdot \|z - x\| \\ &\leq C_2\, \|\xi\|^\alpha\, d(x,z) = C_2\, d(x,y)^\alpha\, d(x,z) \ . \end{aligned} \tag{3.3.5}$$

This yields (3.3.1).

[31] We note that, with some more effort and using Lemma A.3 from the appendix, it is possible to replace the exponent α by 1 in this and in the previous estimate.

We turn to the estimate for the Jacobian $J\Psi_z$. First note that Ψ_z is continuously differentiable. Indeed, the implicit function theorem gives in view of (3.3.3) for all $y \in U$

$$D_y\Psi_z = (D_{\Psi_z}\Gamma)^{-1}\begin{pmatrix} D_{\Sigma_z y}\Gamma_1 \\ D_y\Gamma_2 \end{pmatrix} = 1 + (D_{\Psi_z}\Gamma)^{-1}\begin{pmatrix} D_{\Sigma_z y}\Gamma_1 - D_{\Psi_z y}\Gamma_1 \\ D_y\Gamma_2 - D_{\Psi_z y}\Gamma_2 \end{pmatrix}$$

from which it follows that

$$\begin{aligned}|J\Psi_z(y) - 1| &\le \text{const}\,(d(\Psi_z y, \Sigma_z y)^\alpha + d(\Psi_z y, y)^\alpha) \le \text{const}\,(d(\Psi_z y, \Sigma_z y)^\alpha + d(\Sigma_z y, y)^\alpha) \\ &\le \text{const}\; d(x,z)^\alpha\end{aligned}$$

by the previous estimate and since $\Sigma_z y = (z-x)+y$. □

We return to the proof of Lemma 3.3.2. In our estimates, in order to compare $Q_\varepsilon^*\varphi(x)$ and $Q_\varepsilon^*\varphi(z)$, we will have to compare $\varphi(y)$ to $\varphi(\Sigma_z(y))$. But since we have control over φ only along stable manifolds, we will compare $\varphi(y)$ to $\varphi(\Psi_z(y))$ and care for the remainder by estimating the difference of $q_\varepsilon(z,\Sigma_z(y))$ and $q_\varepsilon(z,\Psi_z(y))$.

We start by noticing that

$$\begin{aligned}|\tilde q_\varepsilon(x, y-x) - \tilde q_\varepsilon(z, \Sigma_z(y)-z)| &= |\tilde q_\varepsilon(x,y-x) - \tilde q_\varepsilon(z,y-x)| \\ &\le M\, d(x,z)^\alpha\, \tilde p_{a\varepsilon}(x, y-x) \\ &= M\, d(x,z)^\alpha\, p_{a\varepsilon}(x,y),\end{aligned}$$

in view of (2.4.2) so that

$$\int_{\mathcal M} |\tilde q_\varepsilon(x,y-x) - \tilde q_\varepsilon(z,\Sigma_z(y)-z)|\, m(dy) \le M\, d(x,z)^\alpha\ .$$

Similarly, denoting $B_{2\varepsilon}(x) = \{y \in \mathcal M : d(x,y) \le 2\varepsilon\}$ we have

$$\begin{aligned}&\int_{B_{2\varepsilon}(x)} |\tilde q_\varepsilon(z,\Sigma_z(y)-z) - \tilde q_\varepsilon(z,\Psi_z(y)-z)|\, m(dy) \\ &\quad\le M\int_{B_{2\varepsilon}(x)} \varepsilon^{-1}\, d(\Psi_z(y),\Sigma_z(y))\;(p_{a\varepsilon}(z,\Sigma_z(y)) + p_{a\varepsilon}(z,\Psi_z(y)))\; m(dy) \\ &\quad\le (1 + \|J\Psi_z^{-1}\|_\infty)\, 2MC\, d(x,z)^\alpha \le 6MC\, d(x,z)^\alpha\end{aligned}$$

where we used (2.4.3), the second estimate in (3.3.1) and the fact that $J\Sigma_z = 1$. Accordingly,

$$\int_{B_{2\varepsilon}(x)} |\tilde q_\varepsilon(x,y-x) - \tilde q_\varepsilon(z,\Psi_z(y)-z)|\, m(dy) \le (M+6MC)\, d(x,z)^\alpha\ . \tag{3.3.6}$$

Now we can estimate the Hölder constant of $Q_\varepsilon^*\varphi$:

$$\begin{aligned}|Q_\varepsilon^*\varphi(x) - Q_\varepsilon^*\varphi(z)| &= \left|\int_{\mathbb R^2} (\tilde q_\varepsilon(x,\xi)\varphi(x+\xi) - \tilde q_\varepsilon(z,\xi)\varphi(z+\xi))\; d\xi\right| \\ &= \left|\int_{\mathcal M} (\tilde q_\varepsilon(x,y-x)\varphi(y) - \tilde q_\varepsilon(z,y-z)\varphi(y))\; m(dy)\right| \\ &\le \int_{\mathcal M} |\tilde q_\varepsilon(x,y-x)|\,|\varphi(y) - \varphi(\Psi_z y)|\, m(dy) \\ &\qquad + \int_{\mathcal M} |\tilde q_\varepsilon(x,y-x) - \tilde q_\varepsilon(z,\Psi_z y - z)J\Psi_z|\,|\varphi(\Psi_z y)|\, m(dy) \\ &\le M\int_{\mathcal M} p_{a\varepsilon}(x,y)\, m(dy)\, H_\beta^s(\varphi) \sup_{|y-x|\le\varepsilon} d(y,\Psi_z(y))^\beta \\ &\qquad + \int_{B_{2\varepsilon}(x)} |\tilde q_\varepsilon(x,y-x) - \tilde q_\varepsilon(z,\Psi_z y - z)|\,|\varphi(\Psi_z y)|\, m(dy) \\ &\qquad + M\int_{\mathcal M} p_{a\varepsilon}(x,y)\, m(dy)\, |\varphi|_\infty \sup_{y\in U} |J\Psi_z^{-1}(y) - 1|\end{aligned}$$

where we used in the last step that $|\Psi_z y - z| \le |\Psi_z y - \Sigma_z y| + |\Sigma_z y - z| \le d(x,y)(C\delta^\alpha + 1) \le 2\varepsilon$ if δ is small enough. Therefore

$$\begin{aligned}|Q_\varepsilon^*\varphi(x) - Q_\varepsilon^*\varphi(z)| &\le H_\beta^s(\varphi)(1 + C\varepsilon^\alpha)^\beta\, d(x,z)^\beta + (M+6MC)\, d(x,z)^\alpha\, |\varphi|_\infty \\ &\qquad + M\,|\varphi|_\infty\, C\, d(x,z)^\alpha \\ &\le d(x,z)^\beta\left[H_\beta^s(\varphi)(1+C\varepsilon^\alpha)^\beta + |\varphi|_\infty\big(M + 6MC + MC\big)\delta^{\alpha-\beta}\right]\end{aligned}$$

where we used the first estimate in (3.3.1), estimate (3.3.6) and also (3.3.2). This proves the lemma. □

Proof of Lemma 2.4.6. Let $(q_\varepsilon)_{0<\varepsilon\leq 1}$ be an admissible stochastic kernel family. We draw first conclusions for this family from Lemma 3.3.2. It is immediately seen that $\sup_{0<\varepsilon\leq\varepsilon_0}\|Q_\varepsilon\|_w<\infty$; this is the first assertion of Lemma 2.4.6.

Combining Lemma 3.3.2 with Sub-lemma 3.1.3 we see that there is a constant $B_2\geq 1$ such that for each $\varphi\in\mathcal{D}_\beta$ there is $\tilde\varphi$ with

$$H^s_\gamma(Q^*_\varepsilon\tilde\varphi)\leq B_2\ ,\quad |Q^*_\varepsilon\tilde\varphi|_\infty\leq 1$$

and

$$H^s_\beta(Q^*_\varepsilon(\varphi-\tilde\varphi))\leq(1+1+\frac14)(1+K\varepsilon^\alpha)^\beta+K2^{-\beta}\delta^\alpha\ ,\quad |Q^*_\varepsilon(\varphi-\tilde\varphi)|_\infty\leq 2^{-\beta}\delta^\beta\ .$$

Our first conclusion is that for sufficiently small ε and δ,

$$\frac25 Q^*_\varepsilon(\varphi-\tilde\varphi)\in\mathcal{D}_\beta\ .$$

Since

$$\int_{\mathcal{M}}Q_\varepsilon f\cdot\varphi\,dm=\frac52\int_{\mathcal{M}}f\cdot\frac25 Q^*_\varepsilon(\varphi-\tilde\varphi)\,dm+B_2\int_{\mathcal{M}}f\cdot B_2^{-1}Q^*_\varepsilon\tilde\varphi\,dm$$

it follows that

$$\|Q_\varepsilon f\|_s\leq\frac52\|f\|_s+B_2\|f\|_w\ . \tag{3.3.7}$$

We turn to the estimate for $\|Q_\varepsilon f\|_u$. Remember that, in dimension two, the stable and unstable distributions are of class $\mathcal{C}^{1+\alpha}$, (see Remark A.1 in the appendix). Therefore they can be described by $\mathcal{C}^{1+\alpha}$ fields of unit tangent vectors $v^s(x)$ and $v^u(x)$. Each vector field $v\in\mathcal{V}_\beta$ can be written as $v=\varphi v^u$ where φ is measurable, $|\varphi|_\infty\leq 1$ and $H^s_\beta(\varphi)\leq 2H^s_\beta(v)$. One just has to make sure by choosing δ sufficiently small that the angle between $v(x)$ and $v(y)$ is small enough if $d(x,y)<\delta$.

So let $v=\varphi v^u\in\mathcal{V}_\beta$. Then

$$\begin{aligned}\int_{\mathcal{M}}d(Q_\varepsilon f)(v)\,dm&=\int_{\mathcal{M}}\int_{\mathcal{M}}f(x)\,\frac{\partial}{\partial z}\tilde q_\varepsilon(x,z-x)\cdot v^u(z)\,\varphi(z)\,m(dx)\,m(dz)\\&=\int_{\mathcal{M}}f(x)\,(\Phi_1(x)+\Phi_2(x))\ m(dx)\end{aligned}$$

where

$$\begin{aligned}\Phi_1(x)&:=\int_{\mathcal{M}}\frac{\partial}{\partial z}\tilde q_\varepsilon(x,z-x)\cdot(v^u(z)-v^u(x))\,\varphi(z)\,m(dz)\ ,\\\Phi_2(x)&:=\int_{\mathcal{M}}\frac{\partial}{\partial z}\tilde q_\varepsilon(x,z-x)\cdot v^u(x)\,\varphi(z)\,m(dz)\ .\end{aligned}$$

Observe that we interpret $\frac{\partial}{\partial z}\tilde q_\varepsilon$ as a vector so that the dot-products under the integral signs are inner products. Let $w(x,\xi):=v^u(x+\xi)-v^u(x)$. Then $|w(x,\xi)|\leq\text{const}\ |\xi|$ and

$$\begin{aligned}\Phi_1(x)=&\int_{\mathcal{M}}\partial_2\tilde q_\varepsilon(x,z-x)\cdot w(x,z-x)\ \varphi(z)\,m(dz)\\\leq&M|\varphi|_\infty\varepsilon^{-1}\int_{\mathcal{M}}\tilde p_{a\varepsilon}(x,z-x)|w(x,z-x)|m(dz).\end{aligned}$$

Since this integral extends only over points $z\in B_\varepsilon(x)$, we have

$$|\Phi_1|_\infty\leq\text{const}\,|\varphi|_\infty$$

for sufficiently small $\varepsilon>0$.

In order to estimate $H^s_\beta(\Phi_1)$ we apply Lemma 3.3.2 to the $\mathbb{R}^2$-valued function $\partial_2\tilde q_\varepsilon\cdot w$. Since w is $\mathcal{C}^{1+\alpha}$ in both coordinates and since q_ε is admissible, $\partial_2\tilde q_\varepsilon\cdot w$ is ε^{-1}-dominated by the same $\tilde p_\varepsilon$ that dominates $\tilde q_\varepsilon$.[32] Hence

$$H^s_\beta(\Phi_1)\leq(1+K\varepsilon^\alpha)^\beta H^s_\beta(\varphi)+K\delta^{\alpha-\beta}|\varphi|_\infty\leq(1+K\varepsilon^\alpha)^\beta+K\delta^{\alpha-\beta}\leq 2 \tag{3.3.8}$$

[32]The proof is a direct computation that uses the same ideas employed in (3.3.4), and (3.3.5). The details are left to the reader.

for sufficiently small ε and δ. So there exists a constant $M_1 > 0$ such that $\frac{1}{M_1}\Phi_1 \in \mathcal{D}_\beta$.

We turn to the estimate of Φ_2.

$$\begin{aligned}
\Phi_2(x) &= -\int_{\mathcal{M}} \frac{\partial}{\partial x}\tilde{q}_\varepsilon(x,z-x)\cdot v^u(x)\ \varphi(z)\,dm(z) + \int_{\mathcal{M}} \partial_1\tilde{q}_\varepsilon(x,z-x)\cdot v^u(x)\ \varphi(z)\,dm(z)\\
&= -\operatorname{div}\left(v^u\, Q_\varepsilon^*\varphi\right)(x) + (\operatorname{div} v^u)(x)\, Q_\varepsilon^*\varphi(x) + v^u(x)\cdot\int_{\mathcal{M}} \partial_1\tilde{q}_\varepsilon(x,z-x)\,\varphi(z)\,m(dz)\\
&=: -\operatorname{div}\left(v^u\, Q_\varepsilon^*\varphi\right)(x) + \Phi_{2,1}(x) + \Phi_{2,2}(x).
\end{aligned}$$

Since $\frac{1}{2}\varphi \in \mathcal{D}_\beta$, it follows from Lemma 3.3.2 that $B_3^{-1}\Phi_{2,1} \in \mathcal{D}_\beta$ for a suitable constant $B_3 > 0$ that depends on the $\mathcal{C}^{1+\alpha}$-norm of $v^u(x)$. Since $\partial_1\tilde{q}_\varepsilon$ is ε^{-1}-dominated by by $\tilde{p}_\varepsilon$ (see the definition of admissible kernels in Definition 2.4.1), we see as above that $\frac{1}{M_1}\Phi_{2,2} \in \mathcal{D}_\beta$.

In order to collect the last estimates we write

$$\begin{aligned}
\int_{\mathcal{M}} d(Q_\varepsilon f)(v)\,dm &= \int_{\mathcal{M}} f\cdot\left(\Phi_1 + \Phi_{2,1} + \Phi_{2,2} - \operatorname{div}(v^u Q_\varepsilon^*\varphi)\right)\,dm\\
&= \int_{\mathcal{M}} f\cdot\left(\Phi_1 + \Phi_{2,1} + \Phi_{2,2}\right)dm + \int_{\mathcal{M}} df(v^u\, Q_\varepsilon^*\varphi)\,dm\ .
\end{aligned} \tag{3.3.9}$$

Observe that $H_\beta^s(v^u\, Q_\varepsilon^*\varphi) \le H_\beta^s(Q_\varepsilon^*\varphi)+1$ when δ is small enough because $|v^u|_\infty = 1$ and $|Dv^u|_\infty < \infty$. We invoke Lemma 3.3.2 once more and conclude that $H_\beta^s(v^u\, Q_\varepsilon^*\varphi) \le 3$ when δ and ε are sufficiently small. Note also that $|v^u Q_\varepsilon^*\varphi|_\infty \le |\varphi|_\infty \le 1$. Hence $\frac{1}{3}v^u Q_\varepsilon^*\varphi \in \mathcal{V}_\beta$ so that

$$\|Q_\varepsilon f\|_u \le 3\|f\|_u + (2M_1 + B_3)\|f\|_s\ .$$

Combined with the estimate (3.3.7) for the stable norm this yields for sufficiently large b

$$\|Q_\varepsilon f\| \le 3\|f\| + bB_2\|f\|_w$$

which finishes the proof of Lemma 2.4.6. □

Proof of Lemma 2.4.7. We start by proving the lemma not for $\mathcal{L}_\varepsilon = Q_\varepsilon\mathcal{L}$ but for $Q_\varepsilon\mathcal{L}^N$ where N is the iterate from (3.1.7) for which $(\sigma\,\min\{\lambda_u,\lambda_s^{-\beta}\})^N > 9A^2$. The reduction to the case $N=1$ is then accomplished in Lemmas 3.3.4 and 3.3.5.

Let $\varphi \in \mathcal{D}_\gamma$. Since $(Q_\varepsilon\mathcal{L}^N)^*\varphi = (\mathcal{L}^N)^*(Q_\varepsilon^*\varphi) = (Q_\varepsilon^*\varphi)\circ T^N$, it follows from Lemma 3.3.2 that

$$H_\gamma^s\left((Q_\varepsilon\mathcal{L}^N)^*\varphi\right) \le A\lambda_s^{\gamma N} H_\gamma^s(Q_\varepsilon^*\varphi) \le A\lambda_s^{\gamma N}(1+K\varepsilon^\alpha)^\gamma H_\gamma^s(\varphi) + A\lambda_s^{\gamma N}K\delta^{\alpha-\gamma}|\varphi|_\infty$$

whence, for sufficiently small ε,

$$H_\gamma^s\left((Q_\varepsilon\mathcal{L}^N)^*\varphi\right) \le \sigma^N H_\gamma^s(\varphi) + K\delta^{\alpha-\gamma}|\varphi|_\infty\ .$$

By induction, it follows

$$H_\gamma\left((Q_\varepsilon\mathcal{L}^N)^{*k}\varphi\right) \le \sigma^{kN}H_\gamma^s(\varphi) + \frac{K}{1-\sigma^N}\delta^{\alpha-\gamma}|\varphi|_\infty\ .$$

(Observe that $|(Q_\varepsilon\mathcal{L}^N)^*\varphi|_\infty \le |\varphi|_\infty$.) Hence there is $K_2 > 0$ such that $K_2^{-1}(Q_\varepsilon\mathcal{L}^N)^{*k}\varphi \in \mathcal{D}_\gamma$ for all $k \in \mathbb{N}$ so that

$$\|(Q_\varepsilon\mathcal{L}^N)^k f\|_w \le K_2\|f\|_w \qquad \text{for all } k. \tag{3.3.10}$$

We turn to the second estimate from Lemma 2.4.7. Again we consider first $Q_\varepsilon\mathcal{L}^N$ instead of $\mathcal{L}_\varepsilon = Q_\varepsilon\mathcal{L}$. For each $n = 1,\dots,N$ we have

$$\begin{aligned}
&\|(Q_\varepsilon\mathcal{L}^n)^k f\|\\
&\le 3\|\mathcal{L}^n(Q_\varepsilon\mathcal{L}^n)^{k-1}f\| + K\|\mathcal{L}^n(Q_\varepsilon\mathcal{L}^n)^{k-1}f\|_w\\
&\le 3A^2\,\max\{3\lambda_s^{\beta n},\lambda_u^{-n}\}\|(Q_\varepsilon\mathcal{L}^n)^{k-1}f\| + 3B\|(Q_\varepsilon\mathcal{L}^n)^{k-1}f\|_w + KA\|(Q_\varepsilon\mathcal{L}^n)^{k-1}f\|_w\\
&\le 3A^2\,\max\{3\lambda_s^{\beta n},\lambda_u^{-n}\}\|(Q_\varepsilon\mathcal{L}^n)^{k-1}f\| + K_2\big(3B+K\big)\|f\|_w
\end{aligned} \tag{3.3.11}$$

where we used Lemma 2.4.6 for the first, equations (3.1.6) and (3.1.8) for the second, and equation (3.3.10) for the third inequality. For $n = N$ this leads, by induction on k and observing the choice of N, to

$$\|(Q_\varepsilon\mathcal{L}^N)^k f\| \le \sigma^{kN}\|f\| + \frac{K_2(3B+K)}{1-\sigma^N}\|f\|_w\ .$$

It remains to carry over these results for $Q_\varepsilon\mathcal{L}^N$ to $\mathcal{L}_\varepsilon^N = (Q_\varepsilon\mathcal{L})^N$. For this we use repeatedly the next two lemmas to show that there exists a smooth averaging operator $Q_{\varepsilon_N}^{(N)}$ such that $\mathcal{L}_\varepsilon^N =$

$(Q_\varepsilon\mathcal{L})^N = Q^{(N)}_{\varepsilon_N}\mathcal{L}^N$ with $\varepsilon_N \le C_N\varepsilon$ for some $C_N > 0$. Applying (3.3.10) and the last estimate to this operator instead of Q_ε and using estimate (3.3.11) for $n = 1$ at most $N-1$ times then finishes the proof of Lemma 2.4.7. □

Lemma 3.3.4. *Denote by Λ^{-1} the Lipschitz constant of T^{-1}. For each smooth averaging operator Q_ε there exists another smooth averaging operator Q'_ε such that $\mathcal{L}Q_\varepsilon = Q'_{\Lambda^{-1}\varepsilon}\mathcal{L}$.*

Lemma 3.3.5. *If Q_ε and Q'_ε are smooth averaging operators, then also $Q''_\varepsilon := Q'_{(1-p)\varepsilon}Q_{p\varepsilon}$ are smooth averaging operators for each $p \in (0,1)$.*

Proof of Lemma 3.3.4. We define $Q'_\varepsilon := \mathcal{L}Q_{\Lambda\varepsilon}\mathcal{L}^{-1}$. Then $Q'_\varepsilon f(y) = \int_{\mathcal{M}} f(x)q'_{\Lambda\varepsilon}(x,y)\,m(dx)$ with

$$\begin{aligned} &q'_\varepsilon(x,y) = q_{\Lambda\varepsilon}(T^{-1}x,T^{-1}y)\cdot g(y) \qquad \text{and hence} \\ &\tilde q'_\varepsilon(x,\xi) = \tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot g(x+\xi) \quad \text{where} \quad G(x,\xi) := (T^{-1}x, T^{-1}(x+\xi)-T^{-1}x)\ . \end{aligned}$$

It follows immediately that $q'_\varepsilon(x,y) = 0$ if $d(x,y) > \varepsilon$, thus q'_ε is a kernel family. Clearly, $\int_{\mathbb{T}^2} q'_\varepsilon(x,y)\,m(dy) = 1$, so q'_ε is a stochastic kernel family. Moreover,

$$\begin{aligned} \partial_1\tilde q'_\varepsilon(x,\xi) &= \underbrace{\partial_1\tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot DT^{-1}_{|x}\cdot g(x+\xi)}_{=:A_1} \\ &\quad + \underbrace{\Lambda\varepsilon\,\partial_2\tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot(\Lambda\varepsilon)^{-1}\left(DT^{-1}_{|x+\xi}-DT^{-1}_{|x}\right)\cdot g(x+\xi)}_{=:A_2} + \underbrace{\tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot Dg_{|x+\xi}}_{=:A_3} \\ \varepsilon\,\partial_2\tilde q'_\varepsilon(x,\xi) &= \underbrace{\varepsilon\,\partial_2\tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot DT^{-1}_{|x+\xi}\cdot g(x+\xi)}_{=:A_4} + \underbrace{\varepsilon\,\tilde q_{\Lambda\varepsilon}(G(x,\xi))\cdot Dg_{|x+\xi}}_{=:A_5} \end{aligned}$$

Observe that

- G is Lipschitz continuous,
- DT^{-1}, g and Dg are bounded and Lipschitz continuous, because T is of class $\mathcal{C}^3$.

Therefore the verification of (2.4.1) – (2.4.3) for the kernel q'_ε and for the terms A_i, $i \ne 2$, is straightforward.[33] The term A_2 needs some more care. First note that

$$\left\|DT^{-1}_{|x+\xi}-DT^{-1}_{|x}\right\| \le \text{const}\cdot\|\xi\| \le \text{const}\cdot\varepsilon\ .$$

This controls the A_2-contribution to (2.4.1) and reduces the Hölder estimates for A_2 in (2.4.2) and (2.4.3) to the corresponding estimates for $\varepsilon^{-1}(DT^{-1}_{|x+\xi}-DT^{-1}_{|x})$. But

$$\begin{aligned} \varepsilon^{-1}&\left(DT^{-1}_{|x+\xi}-DT^{-1}_{|x}\right)-\varepsilon^{-1}\left(DT^{-1}_{|z+\xi}-DT^{-1}_{|z}\right) \\ &= \varepsilon^{-1}\left(DT^{-1}_{|x+\xi}-DT^{-1}_{|z+\xi}\right)-\varepsilon^{-1}\left(DT^{-1}_{|x}-DT^{-1}_{|z}\right) \\ &= \frac{d}{dt}\left(DT^{-1}_{|x+t\xi}-DT^{-1}_{|z+t\xi}\right)_{t=t_0}\cdot\varepsilon^{-1}\xi \end{aligned}$$

for a suitable $t_0 \in (0,1)$ so that this term is bounded by $\text{const}\cdot d(x+t_0\xi, z+t_0\xi) = \text{const}\cdot d(x,z)$, and finally

$$\varepsilon^{-1}\left(DT^{-1}_{|x+\xi}-DT^{-1}_{|x}\right)-\varepsilon^{-1}\left(DT^{-1}_{|x+\zeta}-DT^{-1}_{|x}\right) = \varepsilon^{-1}\left(DT^{-1}_{|x+\xi}-DT^{-1}_{|x+\zeta}\right)$$

so that this term is bounded by $\text{const}\cdot\varepsilon^{-1}\,d(x+\xi,x+\zeta)$. Hence q'_ε is an admissible kernel family. □

Proof of Lemma 3.3.5. Let $q''_\varepsilon(x,y) := \int_{\mathcal{M}} q_{p\varepsilon}(x,z)q'_{(1-p)\varepsilon}(z,y)\,m(dz)$. Obviously $q''_\varepsilon(x,y) = 0$ if $d(x,y) > \varepsilon$. The associated kernel $\tilde q''_\varepsilon(x,\xi)$ is

$$\tilde q''_\varepsilon(x,\xi) = \int_{\mathbb{R}^2}\tilde q_{p\varepsilon}(x,\zeta)\tilde q_{(1-p)\varepsilon}(x+\zeta,\xi-\zeta)\,d\zeta$$

It follows that

$$\begin{aligned} \partial_1\tilde q''_\varepsilon(x,\xi) &= \int_{\mathbb{R}^2}\partial_1\tilde q_{p\varepsilon}(x,\zeta)\tilde q_{(1-p)\varepsilon}(x+\zeta,\xi-\zeta)\,d\zeta \\ &\quad + \int_{\mathbb{R}^2}\tilde q_{p\varepsilon}(x,\zeta)\partial_1\tilde q_{(1-p)\varepsilon}(x+\zeta,\xi-\zeta)\,d\zeta \\ \varepsilon\,\partial_2\tilde q''_\varepsilon(x,\xi) &= \int\tilde q_{p\varepsilon}(x,\zeta)\cdot\varepsilon\,\partial_2\tilde q_{(1-p)\varepsilon}(x+\zeta,\xi-\zeta)\,d\zeta\ , \end{aligned}$$

[33] They are all ε^{-1}-dominated by the same stochastic kernel family that dominates q_ε.

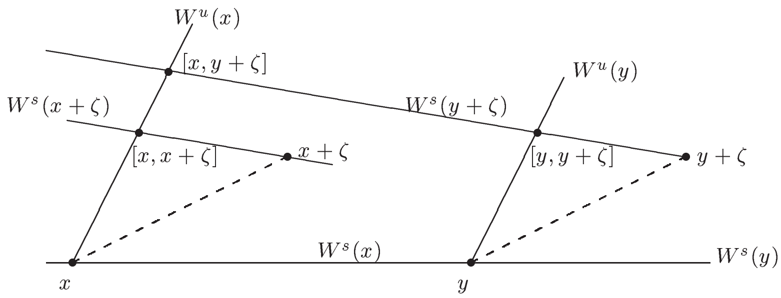

FIGURE 4

and (2.4.1) – (2.4.3) for $\tilde q''_\varepsilon$, $\partial_1\tilde q''_\varepsilon$ and $\partial_2\tilde q''_\varepsilon$ can be deduced easily from the corresponding properties for $\tilde q_\varepsilon$, $\partial_1\tilde q_\varepsilon$ and $\partial_2\tilde q_\varepsilon$. $\square$

Proof of Lemma 2.4.8. The last step in the discussion of smooth random perturbations consists in showing that there exists a constant $K>0$ such that $|||Q_\varepsilon - \mathrm{Id}\,||| \le K\varepsilon^{\gamma-\beta}$. To this end consider $f\in\mathcal C^1(\mathcal M,\mathbb R)$.

Let $\varphi\in\mathcal D_\gamma$. Denoting as before $[x,y] := W^u(x)\cap W^s(y)$ we have

$$\begin{aligned}
&\int_{\mathcal M}(Q_\varepsilon f - f)\,\varphi\,dm\\
&=\int_{\mathcal M}\int_{\mathbb R^2} f(x)\tilde q_\varepsilon(x,\zeta)\varphi(x+\zeta)\,d\zeta\,m(dx) - \int_{\mathcal M} f(x)\,\varphi(x)\,m(dx)\\
&=\int_{\mathcal M}\int_{\mathbb R^2}(f(x)-f([x,x+\zeta]))\tilde q_\varepsilon(x,\zeta)\varphi(x+\zeta)\,d\zeta\,m(dx)\\
&\quad+\int_{\mathcal M}\left(\int_{\mathbb R^2}\tilde q_\varepsilon(x,\zeta)f([x,x+\zeta])\varphi(x+\zeta)\,d\zeta - f(x)\varphi(x)\right)\,m(dx)\\
&=: I_1 + I_2\ .
\end{aligned}\tag{3.3.12}$$

Our goal is to show that $|I_1|,|I_2|\le \mathrm{const}\cdot\varepsilon^{\gamma-\beta}\cdot\|f\|$.

We denote by U_t the flow at unit speed along the unstable manifolds. Let $r(x,\zeta)$ be the signed distance between x and $[x,x+\zeta]$ within $W^u(x)$. The direction is chosen such that $U_{r(x,\zeta)}(x) = [x,x+\zeta]$ (see Fig. 4). We first look at I_1, a term which involves a difference of f-values along an unstable fibre. It is natural to try to estimate this difference in terms of $d_xf(v^u(x))$ where $v^u(x)$, as in the preceding proof, denotes the unit tangent field in unstable direction.

$$I_1 = -\int_{\mathcal M}\int_{\mathbb R^2}\tilde q_\varepsilon(x,\zeta)\int_0^{r(x,\zeta)} d_{U_tx}f(v^u(U_t(x)))\,\varphi(x+\zeta)\,dt\,d\zeta\,m(dx)$$

Changing variables to $\hat x = U_t(x)$ and setting

$$\Omega(x,\zeta) := \{t\in\mathbb R\,:\,\mathrm{sign}(t) = \mathrm{sign}(r(U_{-t}x,\zeta))\,;\,|t|\le|r(U_{-t}x,\zeta)\}$$

we get

$$\begin{aligned}
I_1 &= \int_{\mathcal M}\int_{\mathbb R^2}\int_{\mathbb R}\tilde q_\varepsilon(U_{-t}\hat x,\zeta)\,1_{\Omega(\hat x,\zeta)}(t)\,d_{\hat x}f(v^u(\hat x))\,\varphi(U_{-t}\hat x+\zeta)\,JU_{-t}(\hat x)\,dt\,d\zeta\,m(d\hat x)\\
&=\int_{\mathcal M} d_{\hat x}f(\Phi(\hat x)\,v^u(\hat x))\,m(d\hat x)
\end{aligned}\tag{3.3.13}$$

where $JU_{-t} = JU_t^{-1}$ denotes the Jacobian of U_{-t}, $\log JU_t(x) = \int_0^t(\mathrm{div}\,v^u)(U_sx)\,ds$, and

$$\Phi(x) := \int_{\mathbb R^2}\int_{\mathbb R}\tilde q_\varepsilon(U_{-t}x,\zeta)\,1_{\Omega(x,\zeta)}(t)\,\varphi(U_{-t}x+\zeta)\,JU_{-t}(x)\,dt\,d\zeta.$$

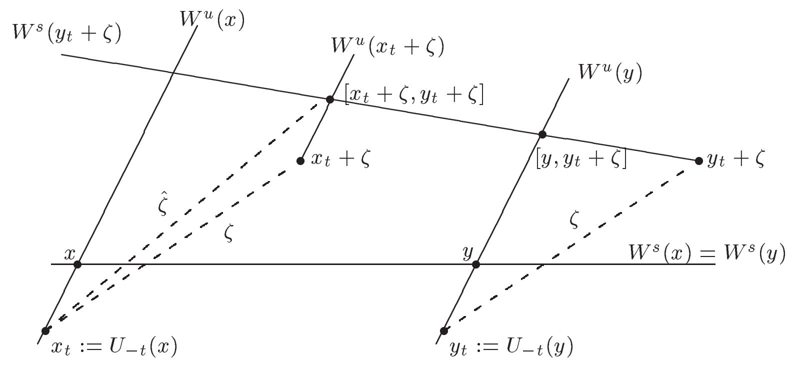

FIGURE 5

To proceed we will show that $C^{-1}\varepsilon^{-1}\Phi\, v^u$ is a test vector field in $\mathcal{V}_\beta$ for a suitable $C>0$. To this end it suffices to check that $C^{-1}\varepsilon^{-1}\Phi \in \mathcal{D}_\beta$ for some $C>0$.

It is easy to estimate $|\Phi|_\infty$. Because of the uniform transversality of the stable and unstable foliation, $r(U_{-t}x,\zeta) \le \text{const}\ \|\zeta\|$ so that, for each ζ with $\tilde q_\varepsilon(U_{-t}x,\zeta)\neq 0$, the set $\{t\in\mathbb{R} : 0\le t\le r(U_{-t}x,\zeta)\}$ is contained in the interval $[0,\text{const}\ \varepsilon]$. Hence $|\Phi|_\infty \le C\varepsilon|\varphi|_\infty \le C\varepsilon$ with a constant $C>0$ which depends only on the stable and unstable foliation.

It remains to estimate $H^s_\beta(\Phi)$. Various points and corresponding stable and unstable fibres which we use in the following computation are shown schematically in Figure 5. Here are some estimates on distances between related points.

We abbreviate $U_{-t}x$ as x_t *etc.* Then, if ε is sufficiently small,

$$
\begin{aligned}
d(x_t+\zeta, y_t+\zeta) &= d(x_t,y_t) \le 2d(x,y)\\
d(x_t+\zeta, [x_t+\zeta, y_t+\zeta]) &\le \text{const}\ d(x_t+\zeta, y_t+\zeta) \le \text{const}\ d(x,y)\\
d(y_t+\zeta, [x_t+\zeta, y_t+\zeta]) &\le \text{const}\ d(x_t+\zeta, y_t+\zeta) \le \text{const}\ d(x,y)
\end{aligned}
\tag{3.3.14}
$$

for suitable constants that depend on the lower bound of the angles of intersection of the stable and unstable foliation and other quantities associated with them.[34]

We turn to the Hölder estimate for Φ.

$$
\begin{aligned}
&|\Phi(x)-\Phi(y)|\\
&\le \left|\int_{\mathbb{R}^2}\int_{\mathbb{R}} \tilde q_\varepsilon(x_t,\zeta)\, 1_{\Omega(x,\zeta)}(t)\,\big(\varphi(x_t+\zeta)-\varphi([x_t+\zeta,y_t+\zeta])\big)\, JU_{-t}(x)\,dt\,d\zeta\right|\\
&\quad+\left|\int_{\mathbb{R}^2}\int_{\mathbb{R}} \tilde q_\varepsilon(x_t,\zeta)\big\{1_{\Omega(x,\zeta)}(t)\,\varphi([x_t+\zeta,y_t+\zeta])\,JU_{-t}(x)\right.\\
&\qquad\qquad\left.-\,1_{\Omega(y,\zeta)}(t)\,\varphi(y_t+\zeta)\,JU_{-t}(y)\big\}\,dt\,d\zeta\right|\\
&\quad+\left|\int_{\mathbb{R}^2}\int_{\mathbb{R}} (\tilde q_\varepsilon(x_t,\zeta)-\tilde q_\varepsilon(y_t,\zeta))\,1_{\Omega(y,\zeta)}(t)\,\varphi(y_t+\zeta)\,JU_{-t}(y)\,dt\,d\zeta\right|\\
&=: A_1+A_2+A_3
\end{aligned}
\tag{3.3.15}
$$

Notice that according to our assumptions JU_{-t} is $\mathcal{C}^\alpha$-close to 1.

The A_3-part in (3.3.15) is most easily estimated, since

$$|\tilde q_\varepsilon(x_t,\zeta)-\tilde q_\varepsilon(y_t,\zeta)| \le M\,2^\alpha\, d(x,y)^\alpha\,\tilde p_{a\varepsilon}(x_t,\zeta)$$

by Definition 2.4.1 and (3.3.14), and since $0\le|t|\le|r(y_t,\zeta)|\le \text{const}\,\|\zeta\|$. This implies

$$A_3 \le \text{const}\ \varepsilon\,|\varphi|_\infty\, d(x,y)^\alpha \le \text{const}\ \varepsilon\, d(x,y)^\beta\delta^{\gamma-\beta}\ . \tag{3.3.16}$$

[34]On general manifolds a term of higher order in $d(x,y)$ must be added on the r.h.s. of these estimates, but this has no consequence if δ is chosen small enough.

Denote the expression in curly brackets in A_2 by $G(t,\zeta)$. For each fixed ζ, the Lebesgue measure of the set of the points $t \in \mathbb{R}$ that belong to exactly one of the two indicator sets is of order

$$\begin{aligned}
&|r(x_t,\zeta) - r(y_t,\zeta)| \\
&\leq |(\Gamma_2(x_t+\zeta) - \Gamma_2(x_t)) - (\Gamma_2(y_t+\zeta) - \Gamma_2(y_t))| + C\, d(x_t,y_t)^\alpha\ |\Gamma_2(y_t+\zeta) - \Gamma_2(y_t)| \\
&= |(\Gamma_2(x_t+\zeta) - \Gamma_2(y_t+\zeta)) - (\Gamma_2(x_t) - \Gamma_2(y_t))| + C\, d(x_t,y_t)^\alpha\ |\Gamma_2(y_t+\zeta) - \Gamma_2(y_t)| \\
&\leq \|\zeta\|\, \|D\Gamma_2(x_t+q\zeta) - D\Gamma_2(y_t+q\zeta)\| + C\, d(x_t,y_t)^\alpha\, \|\zeta\|\, \|D\Gamma_2(y_t+\tilde{q}\zeta)\| \\
&\leq \|\zeta\|\, C\, d(x_t,y_t)^\alpha \leq \|\zeta\|\, C\, d(x,y)^\beta \delta^{\alpha-\beta}
\end{aligned} \tag{3.3.17}$$

where Γ is a local $\mathcal{C}^{1+\alpha}$ diffeomorphisms that straightens the stable and unstable foliation simultaneously, see Lemma 3.3.1. Therefore the corresponding part of the integral can be estimated as $C\,\varepsilon\, d(x,y)^\beta \delta^{\alpha-\beta} |\varphi|_\infty \leq C\,\varepsilon\, d(x,y)^\beta \delta^{\gamma-\beta}$.

Now consider for fixed ζ the set Y_ζ of those $t \in \mathbb{R}$ which belong to both indicator sets. Then $Y_\zeta \subseteq [0, \text{const}\ \varepsilon]$, and for $t \in Y_\zeta$ we have

$$\begin{aligned}
G(t,\zeta) = &\ \big(\varphi([x_t+\zeta, y_t+\zeta]) - \varphi(y_t+\zeta)\big)\, JU_{-t}(y) + \varphi([x_t+\zeta, y_t+\zeta])\, \big(JU_{-t}(x) - JU_{-t}(y)\big) \\
=: &\ A_{2,1}(t,\zeta) + A_{2,2}(t,\zeta)\ .
\end{aligned}$$

$A_{2,1}$ and $A_{2,2}$ are easy to estimate:

$$\begin{aligned}
|A_{2,1}(t,\zeta)| &\leq H^s_\gamma(\varphi)\, d([x_t+\zeta, y_t+\zeta], y_t+\zeta)^\gamma\, |JU_{-t}|_\infty \\
&\leq \text{const}\ H^s_\gamma(\varphi)\, d(x,y)^\gamma \leq \text{const}\ d(x,y)^\beta \delta^{\gamma-\beta} \\
|A_{2,2}(t,\zeta)| &\leq \text{const}\ |\varphi|_\infty\, d(x,y)^\alpha \leq \text{const}\ d(x,y)^\beta \delta^{\gamma-\beta},
\end{aligned}$$

where we have used $\varphi \in \mathcal{D}_\gamma$. It follows that

$$A_2 \leq \text{const}\ \delta^{\gamma-\beta}\, \varepsilon\, d(x,y)^\beta\ . \tag{3.3.18}$$

To estimate A_1 we shift the difference in A_1 to the kernel $\tilde{q}_\varepsilon$ using the following change of variables: $\Psi_t : \{\zeta \in \mathbb{R}^2 : \|\zeta\| < \varepsilon\} \to \mathcal{M}$, $\zeta \mapsto \hat{\zeta} := [x_t+\zeta, y_t+\zeta] - x_t$.

Sub-lemma 3.3.6. *The map Ψ_t is injective and there is a constant $C > 0$, which can be chosen the same for all x, y and $t \leq r(U_{-t}x,\zeta)$, such that*

$$\sup_{\|\zeta\|<\varepsilon} \|\Psi_t(\zeta) - \zeta\| \leq C\ \min\{\varepsilon\, d(x,y)^\alpha, d(x,y)\}\ . \tag{3.3.19}$$

In addition Ψ_t is absolutely continuous and its Jacobian $J\Psi_t$ satisfies

$$\sup_{\|\zeta\|<\varepsilon} |J\Psi_t(\zeta) - 1| \leq C\, d(x,y)^\alpha\ . \tag{3.3.20}$$

Proof. The proof is very similar to that of Sub-lemma 3.3.3.

The rôle of the points z and x in that Sub-lemma is played by the points x_t and y_t here, that of $\Sigma_z(y)$ and $\Psi_z(y)$ by $x_t+\zeta$ and $[x_t+\zeta, y_t+\zeta]$. The only difference is that $\Gamma_2(z) = \Gamma_2(x)$ in that proof, while in the present situation $|\Gamma_2(x_t) - \Gamma_2(y_t)| \leq C\ \min\{d(x_t,y_t), t\, d(x_t,y_t)^\alpha\}$. Therefore the estimate (3.3.4) does not change at all (except for the constant) whereas the estimate in (3.3.5) is replaced by the weaker one $\ldots \leq \text{const}\ d(x_t,y_t)$. This leads to

$$\|\Psi_t(\zeta) - \zeta\| = \|[x_t+\zeta, y_t+\zeta] - (x_t+\zeta)\| \leq C\ \min\{\varepsilon\, d(x,y)^\alpha, d(x,y)\}\ .$$

□

Armed with these estimates we continue to estimate A_1 from (3.3.15). Observe first that

$$\begin{aligned}
&\int_{\mathbb{R}^2} \tilde{q}_\varepsilon(x,\zeta)\, 1_{\Omega(x,\zeta)}(t)\, \varphi([x_t+\zeta, y_t+\zeta])\, d\zeta \\
&= \int_{\mathbb{R}^2} \tilde{q}_\varepsilon(x, \Psi_t^{-1}\hat{\zeta})\, 1_{\Omega(x,\Psi_t^{-1}\hat{\zeta})}(t)\, \varphi(x_t+\hat{\zeta})\, J\Psi_t^{-1}(\hat{\zeta})\, d\hat{\zeta}
\end{aligned}$$

and that, as in (3.3.17),

$$\left|r(x_t,\zeta) - r(x_t,\Psi_t^{-1}\zeta)\right| \leq \|\zeta\|\, C\, d(x,y)^\alpha\ .$$

This implies

$$
\begin{aligned}
A_1 \le & \int_{\mathbb{R}^2}\int_{\mathbb{R}} |\varphi(x_t+\zeta)|\left|\tilde q_\varepsilon(x,\zeta)-\tilde q_\varepsilon(x,\Psi_t^{-1}\zeta)\right| JU_{-t}(x)\,1_{\Omega(x,\zeta)}(t)\,dt\,d\zeta\\
&+\int_{\mathbb{R}^2}\int_{\mathbb{R}} |\varphi(x_t+\zeta)|\tilde q_\varepsilon(x,\Psi_t^{-1}\zeta)\left|1_{\Omega(x,\zeta)}(t)-1_{\Omega(x,\Psi_t^{-1}\hat\zeta)}(t)\right| JU_{-t}(x)\,dt\,d\zeta\\
&+\int_{\mathbb{R}^2}\int_{\mathbb{R}} |\varphi(x_t+\zeta)|\tilde q_\varepsilon(x,\Psi_t^{-1}\zeta)\,1_{\Omega(x,\Psi_t^{-1}\hat\zeta)}(t)\left|J\Psi_t^{-1}(\zeta)-1\right| JU_{-t}(x)dt\,d\zeta\\
&\le \text{const }|\varphi|_\infty\left(\varepsilon\, M\sup_{0\le t\le r(x_t,\zeta)}\varepsilon^{-1}\|\zeta-\Psi_t^{-1}(\zeta)\|+2\cdot C\,\varepsilon\, d(x,y)^\alpha\right)\\
&\le \text{const }\varepsilon\, d(x,y)^\beta\delta^{\gamma-\beta}
\end{aligned}
$$

where we used (3.3.19) for the last step.

Combining this with (3.3.15), (3.3.16) and (3.3.18) we find that, for $\varphi\in\mathcal{D}_\gamma$,

$$|\Phi(x)-\Phi(y)|\le \text{const }\varepsilon\,\delta^{\gamma-\beta}\,d(x,y)^\beta$$

so that $\varepsilon^{-1}\Phi\in\mathcal{D}_\beta$ for sufficiently small δ. It follows that $\varepsilon^{-1}\Phi\, v^u\in\mathcal{V}_\beta$. Therefore,

$$|I_1|\le\varepsilon\,\|f\|_u\ , \tag{3.3.21}$$

see (3.3.12) and (3.3.13) to recall the meaning of I_1.

It remains to estimate the integral

$$I_2=\int_{\mathcal{M}}\left(\int_{\mathbb{R}^2}\tilde q_\varepsilon(x,\zeta)f([x,x+\zeta])\varphi(x+\zeta)\,d\zeta-f(x)\varphi(x)\right)\,m(dx)\ ,$$

see (3.3.12). The strategy is similar to that for the estimate of I_1. For each ζ, we use the coordinate change $\Psi_\zeta(x):=[x,x+\zeta]$.

Sub-lemma 3.3.7. *The map Ψ_ζ is injective and there is a constant $C>0$, which can be chosen the same for all x and ζ and $y\in W^s(x)$, such that*

$$\|\Psi_\zeta^{-1}(x)-x\|\le C\,\|\zeta\| \tag{3.3.22}$$

$$\|\Psi_\zeta^{-1}(x)-\Psi_\zeta^{-1}(y)\|\le C\,d(x,y)\ . \tag{3.3.23}$$

In addition Ψ_ζ is continuously differentiable and its Jacobian $J\Psi_\zeta$ satisfies

$$|J\Psi_\zeta(x)-1|\le C\,\|\zeta\|^\alpha \tag{3.3.24}$$

$$|J\Psi_\zeta(x)-J\Psi_\zeta(y)|\le C\,d(x,y)^\alpha\ . \tag{3.3.25}$$

We postpone the proof of this Sub-lemma to the end of this section.

Using the coordinate change Ψ_ζ,

$$
\begin{aligned}
I_2=&\int_{\mathcal{M}}\left(\int_{\mathbb{R}^2}\left(\tilde q_\varepsilon(\Psi_\zeta^{-1}x,\zeta)\varphi(\Psi_\zeta^{-1}x+\zeta)J\Psi_\zeta^{-1}(x)-\tilde q_\varepsilon(x,\zeta)\varphi(x)\right)\,d\zeta\right)\,f(x)\,m(dx)\\
=&\int_{\mathcal{M}}\left(\int_{\mathbb{R}^2}\tilde q_\varepsilon(\Psi_\zeta^{-1}x,\zeta)\big(\varphi(\Psi_\zeta^{-1}x+\zeta)J\Psi_\zeta^{-1}(x)-\varphi(x)\big)\,d\zeta\right)\,f(x)\,m(dx)\\
&\quad+\int_{\mathcal{M}}\left(\int_{\mathbb{R}^2}\left(\tilde q_\varepsilon(\Psi_\zeta^{-1}x,\zeta)-\tilde q_\varepsilon(x,\zeta)\right)\varphi(x)\,d\zeta\right)\,f(x)\,m(dx)\\
=:&\int_{\mathcal{M}}\Phi_1(x)f(x)\,m(dx)+\int_{\mathcal{M}}\Phi_2(x)f(x)\,m(dx)\ ,
\end{aligned}
$$

and we have to estimate $|\Phi_i|_\infty$ and $H^s_\beta(\Phi_i)$ in order to bound I_2 in terms of $\|f\|_s$.

We start with the estimate for Φ_1. In view of Remark 2.4.3 and of (3.3.22) we have $\tilde q_\varepsilon(\Psi_\zeta^{-1}x,\zeta)\le$ const $\tilde p_{a\varepsilon}(x,\zeta)$ So it suffices to show that, for each fixed ζ with $\|\zeta\|<\varepsilon$, the supremum and the β–Hölder constant of

$$\Phi_{1,1}(x,\zeta):=\varphi(\Psi_\zeta^{-1}x+\zeta)-\varphi(x)\quad\text{ and }\quad\Phi_{1,2}(x,\zeta):=\varphi(\Psi_\zeta^{-1}x+\zeta)(J\Psi_\zeta^{-1}(x)-1)$$

are bounded (uniformly in ζ) by const $\varepsilon^{\gamma-\beta}$. The relevant points are sketched in Figure 6.

Recall that $\varphi\in\mathcal{D}_\gamma$. Then

$$|\Phi_{1,1}|_\infty\le((C+1)\varepsilon)^\gamma H^s_\gamma(\varphi)\le\text{const }\varepsilon^{\gamma-\beta}$$

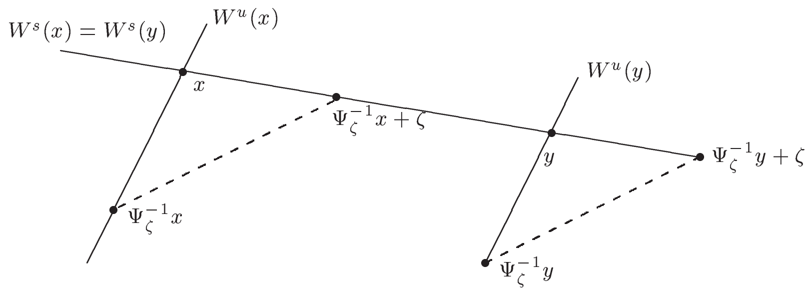

FIGURE 6

because of (3.3.22). Similarly, invoking (3.3.24),

$$|\Phi_{1,2}|_\infty \le |\varphi|_\infty\, C\,\varepsilon^\alpha \le \text{const } \varepsilon^{\gamma-\beta}\ .$$

To analyse the Hölder constant of $\Phi_{1,1}$ we consider two different cases which may happen for points x and $y \in W^s(x)$. First, if $d(x,y) \le \varepsilon$, we use (3.3.23) to estimate

$$\begin{aligned}|\Phi_{1,1}(x,\zeta) - \Phi_{1,1}(y,\zeta)| &\le |\varphi(\Psi_\zeta^{-1}x+\zeta) - \varphi(\Psi_\zeta^{-1}y+\zeta)| + |\varphi(x)-\varphi(y)|\\ &\le \text{const } d(x,y)^\gamma\, H_\gamma^s(\varphi)\\ &\le \text{const } \varepsilon^{\gamma-\beta}\, d(x,y)^\beta\end{aligned}$$

(Observe that $\Psi_\zeta^{-1}(x)+\zeta \in W^s(x) = W^s(y) \ni \Psi_\zeta^{-1}(y)+\zeta$.)

Otherwise, if $d(x,y) > \varepsilon$ we proceed as follows:

$$|\Phi_{1,1}(x,\zeta) - \Phi_{1,1}(y,\zeta)| \le 2|\Phi_{1,1}|_\infty \le \text{const } \varepsilon^\gamma \le \text{const } \varepsilon^{\gamma-\beta}\, d(x,y)^\beta\ .$$

In any case,

$$H_\beta^s(\Phi_{1,1}(\,.\,,\zeta)) \le \text{const } \varepsilon^{\gamma-\beta}\ .$$

The Hölder constant of $\Phi_{1,2}$ is estimated along the same lines using (3.3.25) – the details are left to the reader.

We turn to the estimates for Φ_2. Because of (2.4.2) and (3.3.22) we have

$$|\Phi_2|_\infty \le |\varphi|_\infty\, M(C\|\zeta\|)^\alpha \le \text{const } \varepsilon^\alpha \le \text{const } \varepsilon^{\gamma-\beta}$$

In order to estimate the Hölder constant of Φ_2 we consider the cases $d(x,y) \le \varepsilon$ and $d(x,y) > \varepsilon$ separately as in the estimate for $H_\beta^s(\Phi_{1,1})$. This yields

$$\begin{aligned}H_\beta^s(\Phi_2) &\le \max\left\{\text{const } H_\gamma^s(\varphi)\,\varepsilon^{\gamma-\beta} + \text{const } \varepsilon^{\alpha-\beta}, \text{const } \varepsilon^{\alpha-\beta}\right\}\\ &\le \text{const } \varepsilon^{\gamma-\beta}\ .\end{aligned}$$

□

Proof of Sub-lemma 3.3.7. Let Γ be a $\mathcal{C}^{1+\alpha}$-diffeomorphism as in the proof of Sub-lemma 3.3.3. This means that

$$\Gamma_1(\Psi_\zeta^{-1}x) = \Gamma_1(x)\ ,\quad \Gamma_2(\Psi_\zeta^{-1}x+\zeta) = \Gamma_2(x)\ .$$

Hence

$$\begin{aligned}\left\|\Gamma(\Psi_\zeta^{-1}x) - \Gamma(x)\right\| = \left|\Gamma_2(\Psi_\zeta^{-1}x) - \Gamma_2(x)\right| &= \left|\Gamma_2(\Psi_\zeta^{-1}x) - \Gamma_2(\Psi_\zeta^{-1}x+\zeta)\right|\\ &\le \left\|D\Gamma_{2|\Psi_\zeta^{-1}x+t\zeta}\right\| \cdot \|\zeta\|\end{aligned}$$

for a suitable $t \in (0,1)$. This implies (3.3.22).

To proceed further, we need the estimate

$$\left\|D(\Gamma_2 \circ \Psi_\zeta^{-1} - \Gamma_2)_{|x}\right\| \le \text{const } \|\zeta\|^\alpha\ . \tag{3.3.26}$$

This is proved as follows: Since $\Gamma_2(x) = \Gamma_2(\Psi_\zeta^{-1}x + \zeta)$, we have

$$\begin{aligned}\left\|D(\Gamma_2 \circ \Psi_\zeta^{-1})_{|x} - D\Gamma_{2|x}\right\| &= \left\|D\Gamma_{2|\Psi_\zeta^{-1}x} \cdot D\Psi_\zeta^{-1}{}_{|x} - D(\Gamma_2 \circ (\Psi_\zeta^{-1} + \zeta))_{|x}\right\| \\ &= \left\|\left(D\Gamma_{2|\Psi_\zeta^{-1}x} - D\Gamma_{2|\Psi_\zeta^{-1}x+\zeta}\right) \cdot D\Psi_\zeta^{-1}{}_{|x}\right\| \\ &\le \text{const } \|\zeta\|^\alpha \sup_{x,\zeta} \left\|D\Psi_\zeta^{-1}{}_{|x}\right\|\end{aligned}$$

and

$$\sup_{x,\zeta} \left\|D\Psi_\zeta^{-1}{}_{|x}\right\| < \infty \tag{3.3.27}$$

since the manifold $\mathcal{M}$ is compact and since Ψ_ζ is determined by the implicit function theorem.

Now we apply (3.3.26) to prove estimate (3.3.23). Since $\Gamma_1(\Psi_\zeta^{-1}y) - \Gamma_1(\Psi_\zeta^{-1}x) = \Gamma_1(y) - \Gamma_1(x)$, it suffices to estimate $\Gamma_2(\Psi_\zeta^{-1}y) - \Gamma_2(\Psi_\zeta^{-1}x)$. To this end let $t_0 := \Gamma_1(y) - \Gamma_1(x)$ and assume without loss of generality that $t_0 > 0$. For $0 \le t \le t_0$ denote $x_t := \Gamma^{-1}(\Gamma x + t\,\mathbf{e}_1)$. Then $x_0 = x$, $x_{t_0} = y$, and

$$\begin{aligned}\left\|\Gamma_2(\Psi_\zeta^{-1}y) - \Gamma_2(\Psi_\zeta^{-1}x)\right\| &= \left\|\left(\Gamma_2(\Psi_\zeta^{-1}x_{t_0}) - \Gamma_2 x_{t_0}\right) - \left(\Gamma_2(\Psi_\zeta^{-1}x_0) - \Gamma_2 x_0\right)\right\| \\ &\le |t_0| \left\|D(\Gamma_2 \circ \Psi_\zeta^{-1} - \Gamma_2)_{|x_t}\right\| \\ &\le \text{const } \|\zeta\|^\alpha\, d(x,y)\ .\end{aligned}$$

We turn to the proof of (3.3.24). Obviously it suffices to estimate $\|D\Psi_\zeta^{-1}(x) - \mathbf{1}\|$ where $\mathbf{1}$ denotes the identity matrix.

$$\begin{aligned}\left\|D\Psi_\zeta^{-1}{}_{|x} - \mathbf{1}\right\| &= \left\|D\Gamma^{-1}{}_{|\Gamma(\Psi_\zeta^{-1}x)} \cdot D(\Gamma \circ \Psi_\zeta^{-1})_{|x} - D\Gamma^{-1}{}_{|\Gamma x} \cdot D\Gamma_{|x}\right\| \\ &\le \left\|D\Gamma^{-1}{}_{|\Gamma(\Psi_\zeta^{-1}x)}\right\| \left\|D(\Gamma \circ \Psi_\zeta^{-1})_{|x} - D\Gamma_{|x}\right\| \\ &\qquad + \left\|D\Gamma^{-1}{}_{|\Gamma(\Psi_\zeta^{-1}x)} - D\Gamma^{-1}{}_{|\Gamma x}\right\| \left\|D\Gamma_{|x}\right\| \\ &\le \text{const } \left\|D(\Gamma_2 \circ \Psi_\zeta^{-1})_{|x} - D\Gamma_{2|x}\right\| + \text{const } d(\Psi_\zeta^{-1}x, x)^\alpha \\ &\le \text{const } \|\zeta\|^\alpha\end{aligned}$$

where we used (3.3.26) and (3.3.22) for the last step.

The proof of estimate (3.3.25) is more subtle. To simplify the notation denote by G and H the maps $x \mapsto \Psi_\zeta^{-1}(x)$ and $x \mapsto \Psi_\zeta^{-1}(x) + \zeta$, respectively. Then $DG_{|x} = DH_{|x}$ and hence

$$D\Gamma^{-1}{}_{|\Gamma(Gx)}\, D(\Gamma \circ G)_{|x} = D\Gamma^{-1}{}_{|\Gamma(Hx)}\, D(\Gamma \circ H)_{|x}$$

so that

$$D(\Gamma \circ G)_{|x} = \left(D\Gamma_{|Gx}\, (D\Gamma_{|Hx})^{-1}\right)\, D(\Gamma \circ H)_{|x} =: (\mathbf{1} + \Delta_{\zeta,x})\, D(\Gamma \circ H)_{|x}\ .$$

Since $d(Gx, Hx) = \|\zeta\|$, it follows that $\|\Delta_{\zeta,x}\| \le \text{const } \|\zeta\|^\alpha$.

Now consider again x and y with $\Gamma_2(x) = \Gamma_2(y)$. For such points

$$D(\Gamma \circ G)_{|y} - D(\Gamma \circ G)_{|x} = D(\Gamma \circ H)_{|y} - D(\Gamma \circ H)_{|x} + \Delta_{\zeta,y} \cdot D(\Gamma \circ H)_{|y} - \Delta_{\zeta,x} \cdot D(\Gamma \circ H)_{|x}$$

Multiplying this equation with the row vector $\mathbf{e}_2' = (0\ 1)$ from the left we obtain

$$D(\Gamma_2 \circ G)_{|y} - D(\Gamma_2 \circ G)_{|x} = D\Gamma_{2|y} - D\Gamma_{2|x} + \mathbf{e}_2' \cdot \left(\Delta_{\zeta,y} \cdot D(\Gamma \circ H)_{|y} - \Delta_{\zeta,x} \cdot D(\Gamma \circ H)_{|x}\right)$$

because $\Gamma_2 \circ H = \Gamma_2$. Observing now that also $\Gamma_1 \circ G = \Gamma_1$ we can conclude that

$$\begin{aligned}&\left\|D(\Gamma \circ G)_{|y} - D(\Gamma \circ G)_{|x}\right\| \\ &\quad \le \left\|D\Gamma_{|y} - D\Gamma_{|x}\right\| + \left\|\Delta_{\zeta,y} \cdot D(\Gamma \circ H)_{|y} - \Delta_{\zeta,x} \cdot D(\Gamma \circ H)_{|x}\right\| \\ &\quad \le \text{const } d(x,y)^\alpha + \|\Delta_{\zeta,y} - \Delta_{\zeta,x}\| \left\|D(\Gamma \circ H)_{|y}\right\| + \|\Delta_{\zeta,x}\| \left\|D(\Gamma \circ H)_{|y} - D(\Gamma \circ H)_{|x}\right\| \\ &\quad \le C_1\, d(x,y)^\alpha + C_2\, \varepsilon^\alpha \left\|D(\Gamma \circ H)_{|y} - D(\Gamma \circ H)_{|x}\right\|\end{aligned}$$

for suitable $C_1, C_2 > 0$, where we used the boundedness of $\left\|D(\Gamma \circ H)_{|y}\right\|$, the estimate $\|\Delta_{\zeta,x}\| \le \text{const } \|\zeta\|^\alpha$ and the fact that $\|\Delta_{\zeta,y} - \Delta_{\zeta,x}\| \le \text{const } d(x,y)^\alpha$ in the last step.

So we proved

$$\left\|D(\Gamma \circ G)_{|y} - D(\Gamma \circ G)_{|x}\right\| - C_2\, \varepsilon^\alpha \left\|D(\Gamma \circ H)_{|y} - D(\Gamma \circ H)_{|x}\right\| \le C_1\, d(x,y)^\alpha\ .$$

Interchanging the rôles of G and H we obtain in the same way

$$\left\|D(\Gamma\circ H)_{|y}-D(\Gamma\circ H)_{|x}\right\|-C_2\,\varepsilon^\alpha\,\left\|D(\Gamma\circ G)_{|y}-D(\Gamma\circ G)_{|x}\right\|\le C_1\,d(x,y)^\alpha$$

with the same constants. Therefore we can add both inequalities and arrive at

$$\left\|D(\Gamma\circ G)_{|y}-D(\Gamma\circ G)_{|x}\right\|+\left\|D(\Gamma\circ H)_{|y}-D(\Gamma\circ H)_{|x}\right\|\le\frac{2C_1\,d(x,y)^\alpha}{1-C_2\,\varepsilon^\alpha}$$

provided $\varepsilon>0$ is sufficiently small. Hence

$$\begin{aligned}&\left\|D\Psi_\zeta^{-1}{}_{|y}-D\Psi_\zeta^{-1}{}_{|x}\right\|\\&=\left\|DG_{|y}-DG_{|x}\right\|\\&=\left\|D\Gamma^{-1}_{|\Gamma(Gy)}\cdot D(\Gamma\circ G)_{|y}-D\Gamma^{-1}_{|\Gamma(Gx)}\cdot D(\Gamma\circ G)_{|x}\right\|\\&\le\left\|D\Gamma^{-1}_{|\Gamma(Gy)}-D\Gamma^{-1}_{|\Gamma(Gx)}\right\|\left\|D(\Gamma\circ G)_{|y}\right\|+\left\|D\Gamma^{-1}_{|\Gamma(Gx)}\right\|\left\|D(\Gamma\circ G)_{|y}-D(\Gamma\circ G)_{|x}\right\|\\&\le\text{const}\;d(\Psi_\zeta^{-1}y,\Psi_\zeta^{-1}x)^\alpha\left\|D\Psi_\zeta^{-1}{}_{|y}\right\|+\text{const}\;d(x,y)^\alpha\\&\le\text{const}\;d(x,y)^\alpha\end{aligned}$$

where we used (3.3.23) and (3.3.27) in the last step. From this estimate (3.3.25) follows immediately. □

3.4. **Proofs: Spectral stability – Deterministic Perturbations.**

Let T be a $\mathcal{C}^3$ Anosov diffeomorphism of the 2-dimensional torus $\mathbb{T}^2$. Our goal is to prove that the spectral properties of the associated transfer operator do not change much under $\mathcal{C}^1$–small perturbations as long as the $\mathcal{C}^3$ norm of the perturbed maps remains smaller than some constant $K>\|T\|_{\mathcal{C}^3}$. From now on we will use $\tilde T$ to denote any Anosov map such that $d_{\mathcal{C}^1}(T,\tilde T)=:\ell<<1$ and $\|\tilde T\|_{\mathcal{C}^3}\le K$.

Since we are working now with spaces associated to T and also with spaces associated to $\tilde T$, we will denote the space of test functions related to $\tilde T$ by $\tilde{\mathcal{D}}_\beta$, the resulting norm by $\|\cdot\|^\sim$ *etc.*

The essential geometric fact that makes it possible to compare $\mathcal{L}_T:\mathcal{B}\to\mathcal{B}$ and $\mathcal{L}_{\tilde T}:\tilde{\mathcal{B}}\to\tilde{\mathcal{B}}$ is contained in the following Lemma.

Lemma 3.4.1. *There exist constants $C,\rho>0$ such that, for each $f\in\mathcal{C}^2$,*

$$\|f\|\le C\ell^\rho\|f\|_{2,1}+C\|f\|^\sim$$

where $\|\cdot\|_{k,1}$ denotes the L^1-Sobolev norms.

Proof. By using a smooth partition of unity we can always restrict ourselves to a situation in which the functions are all supported in small balls of radius r.[35] In such a ball one can consider two changes of variables Γ and $\tilde\Gamma$ that straighten the foliations of the maps T and $\tilde T$, respectively. As seen in the previous section these are $\mathcal{C}^{1+\alpha}$.

Accordingly, setting $F:=\Gamma^{-1}\circ\tilde\Gamma$, F is $\mathcal{C}^{1+\alpha}$. More can be said.

Sub-lemma 3.4.2. *There exist constants $C>0$ and $\rho>0$ such that*

$$\|F-\mathrm{Id}\,\|_\infty\le C\ell^\rho\quad\textit{and}\quad\|dF-\mathbf{1}\,\|_\infty\le C\ell^\rho.$$

Since the above result is hardly surprising and its proof is a bit technical we postpone it to the end of the present section.

For each $\varphi\in\mathcal{D}_\beta$ and $f\in\mathcal{C}^1$ holds

$$\begin{aligned}\left|\int_{\mathbb{T}^2}f\varphi\,dm\right|&\le\left|\int_{\mathbb{T}^2}(f-f\circ F^{-1})\cdot\varphi\,dm\right|+\left|\int_{\mathbb{T}^2}f\cdot(\varphi\circ F)JF\,dm\right|\\&\le\|f\|_{1,1}\|\,\mathrm{Id}-F^{-1}\|_\infty+\left|\int_{\mathbb{T}^2}f\,(\varphi\circ F)\,JF\,dm\right|\;.\end{aligned}$$

Clearly there exists $C>0$ such that $C^{-1}\varphi\circ F\cdot JF\in\tilde{\mathcal{D}}_\beta$, hence Sub-Lemma 3.4.2 implies

$$\|f\|_s\le C\|f\|_{1,1}\ell^\rho+C\|f\|_s^\sim.$$

Analogously,

$$\int_{\mathbb{T}^2}df(\varphi v^u)\,dm=\int_{\mathbb{T}^2}d_Ff[(v^u\circ F)(\varphi\circ F)]JF\,dm=\int_{\mathbb{T}^2}d(f\circ F)[((DF^{-1}v^u)\circ F)(\varphi\circ F)]JF\,dm\;.$$

[35] If in doubt look at Lemma 3.6.1.

Since F sends the stable and unstable manifolds of $\tilde{T}$ to the corresponding ones of T and is $\mathcal{C}^{1+\alpha}$ it follows that there exists an α-Hölder function ω such that $(DF^{-1}v^u)\circ F = \omega\tilde{v}^u$. Accordingly, $\tilde{v} := C^{-1}(DF^{-1}v^u)\circ F\cdot\varphi\circ F\cdot JF \in \tilde{\mathcal{V}}_\beta$, for some constant $C>0$. Therefore the following estimate finishes the proof:

$$\begin{aligned}\int_{\mathbb{T}^2} |d(f\circ F)(\tilde{v}) - df(\tilde{v})|\,dm &= \int_{\mathbb{T}^2} |d_F f(dF(\tilde{v})-\tilde{v})|\,dm + \int_{\mathbb{T}^2} |(d_F f - df)(\tilde{v})|\,dm\\ &\le \|f\|_{1,1}\|dF-\mathbf{1}\|_\infty + \|f\|_{2,1}\|F-\mathrm{Id}\,\|_\infty\\ &\le C\,\ell^\rho\,\|f\|_{2,1}\end{aligned}$$

where the last inequality follows from Sub-Lemma 3.4.2. □

Some further estimates are collected in the next lemma.

Lemma 3.4.3. *Under the above assumptions on q_ε, T and $\tilde{T}$, we have for all $f\in\mathcal{C}^1(\mathbb{T}^2,\mathbb{R})$*

$$\|Q_\varepsilon f\|^{(\sim)} \le \text{const}\ \varepsilon^{-1}\,\|f\|_1 \tag{3.4.1}$$

$$\|Q_\varepsilon f\|_1 \le \text{const}\ \varepsilon^{-1}\,\|f\|^{(\sim)} \tag{3.4.2}$$

$$\|Q_\varepsilon f\|_{1,1} \le \text{const}\ \varepsilon^{-2}\,\|f\|^{(\sim)} \tag{3.4.3}$$

$$\|Q_\varepsilon f\|_{2,1} \le \text{const}\ \varepsilon^{-3}\,\|f\|^{(\sim)} \tag{3.4.4}$$

Furthermore,

$$\|(\mathcal{L}_T - \mathcal{L}_{\tilde{T}})f\|_1 \le \text{const}\cdot\text{dist}_{\mathcal{C}^1}(T,\tilde{T})\,\|f\|_{1,1}\ . \tag{3.4.5}$$

Proof. For $\varphi\in\mathcal{D}_\beta$, $\int_{\mathbb{T}^2} Q_\varepsilon f\cdot\varphi\,dm \le \|Q_\varepsilon f\|_1 \le \|f\|_1$, so $\|Q_\varepsilon f\|_s\le\|f\|_1$. In order to estimate $\|Q_\varepsilon f\|_u$, let $v\in\mathcal{V}_\beta$. Observing that q_ε is a convolution kernel, one verifies easily that for $f\in\mathcal{C}^1(\mathbb{T}^2,\mathbb{R})$

$$\begin{aligned}\int_{\mathbb{T}^2} d(Q_\varepsilon f)(v)\,dm &= \int_{\mathbb{T}^2} D(Q_\varepsilon f)\cdot v\,dm = \int_{\mathbb{T}^2} Q_\varepsilon(Df)\cdot v\,dm = \int_{\mathbb{T}^2} Df\cdot Q_\varepsilon v\,dm\\ &= -\int_{\mathbb{T}^2} f\ \mathrm{div}(Q_\varepsilon v)\,dm \le \|f\|_1\cdot\|\,\mathrm{div}(Q_\varepsilon v)\|_\infty\ ,\end{aligned}$$

and $|\,\mathrm{div}(Q_\varepsilon v)|\le\text{const}\ \varepsilon^{-1}\,|v|_\infty\le\text{const}\ \varepsilon^{-1}$. Hence $\|Q_\varepsilon f\|\le\text{const}\ \varepsilon^{-1}\,\|f\|_1$, and since the same arguments apply to the $\|\cdot\|^\sim$ norm, this finishes the proof of (3.4.1).

Next we estimate $\|Q_\varepsilon f\|_1$. To this end consider $\varphi\in\mathcal{C}^1(\mathbb{T}^2,\mathbb{R})$ with $|\varphi|_\infty\le 1$. Then

$$\int_{\mathbb{T}^2}\varphi\,Q_\varepsilon f\,dm = \int_{\mathbb{T}^2} Q_\varepsilon^*\varphi\,f\,dm \le \max\{|Q_\varepsilon^*\varphi|_\infty, H_\beta^s(Q_\varepsilon^*\varphi)\}\cdot\|f\|_s \le \text{const}\ \varepsilon^{-1}\,\|f\|\ .$$

As $\|Q_\varepsilon f\|_1$ is the supremum of all such integrals, this proves (3.4.2). Again the same estimate applies to the $\|\cdot\|^\sim$ norm.

The proof of (3.4.3) is similar to that of (3.4.2). Consider $v\in\mathcal{C}^1(\mathbb{T}^2,\mathbb{R}^2)$ with $|v|_\infty\le 1$. Then

$$\begin{aligned}\int_{\mathbb{T}^2} D(Q_\varepsilon f)\cdot v\,dm &= -\int_{\mathbb{T}^2} Q_\varepsilon f\ \mathrm{div}(v)\,dm = -\int_{\mathbb{T}^2} f\ \mathrm{div}(Q_\varepsilon^* v)\,dm\\ &\le \max\{|\,\mathrm{div}(Q_\varepsilon^* v)|_\infty, H_\beta^s(\mathrm{div}(Q_\varepsilon^* v))\}\cdot\|f\|_s \le \text{const}\ \varepsilon^{-2}\,\|f\|\ ,\end{aligned}$$

and the argument is concluded as before. Estimate (3.4.4) is proved in the same way.

We turn to (3.4.5). To this end we consider T^{-1} and $\tilde{T}^{-1}$ as maps from $[0,1]^2$ to $\mathbb{R}^2$ and denote $w(x) := \tilde{T}^{-1}(x) - T^{-1}(x)$. For $t\in[0,1]$ define

$$S_t(x) := (1-t)\,T^{-1}(x) + t\,\tilde{T}^{-1}(x)\quad\text{ and }\quad g_t(x) := JS_t(x)\ ,$$

so $S_0 = T^{-1}$ and $S_1 = \tilde{T}^{-1}$ and g_0, g_1 are the Jacobians of T^{-1} and $\tilde{T}^{-1}$, respectively. Therefore, not writing the variable x explicitly and extending f periodically to the covering $\mathbb{R}^2$ of $\mathbb{T}^2$, we have

$$\begin{aligned}|\mathcal{L}_{\tilde{T}}f - \mathcal{L}_T f| &= |f\circ S_1\cdot g_1 - f\circ S_0\cdot g_0|\\ &\le \int_0^1\left|\frac{\partial}{\partial t}(f\circ S_t\cdot g_t)\right|dt\\ &\le \int_0^1 \left|Df_{|S_t}\cdot w\cdot g_t\right|dt + \int_0^1 |f\circ S_t|\cdot\left|\frac{\partial g_t}{\partial t}\right|dt\\ &\le \|w\|_\infty\cdot\int_0^1 |Df|\circ S_t\cdot JS_t\,dt + \left\|\frac{\partial}{\partial t}\ln g_t\right\|_\infty\cdot\int_0^1 |f|\circ S_t\cdot JS_t\,dt\end{aligned}$$

Fix $a>0$ such that $K_t := S_t([0,1]^2) \subset [-a,a]^2$ for all $t \in [0,1]$. Then

$$\begin{aligned}\|\mathcal{L}_{\tilde T} f - \mathcal{L}_T f\|_1 &\le \|w\|_\infty \cdot \int_0^1 \left(\int_{K_t} |Df|\,dm\right) dt + \sup_t \left\|\frac{\partial}{\partial t}\ln g_t\right\|_\infty \cdot \int_0^1 \left(\int_{K_t} |f|\,dm\right) dt \\ &\le a^2 \left(\|w\|_\infty + \sup_t \left\|\frac{\partial}{\partial t}\ln g_t\right\|_\infty\right) \cdot \|f\|_{1,1}\ .\end{aligned}$$

Clearly, $\|w\|_\infty \le \mathrm{const}\cdot \mathrm{dist}_{\mathcal{C}^0}(\tilde T, T)$, and a routine calculation shows that

$$\begin{aligned}\left|\frac{\partial}{\partial t}\ln g_t\right| &= \left|\sum_{k=1}^\infty (-t)^{k-1}\,\mathrm{trace}\left((DS_0)^{-1}DS_1 - 1\right)^k\right| \\ &\le \frac{2\left|(DS_0)^{-1}DS_1 - 1\right|}{1-\left|(DS_0)^{-1}DS_1 - 1\right|} \le \mathrm{const}\cdot \mathrm{dist}_{\mathcal{C}^1}(\tilde T, T)\end{aligned}$$

provided this distance is small enough to make the denominator of the fraction not smaller than $\frac12$, say. □

Introducing smooth averaging operators Q_ε we can get rid of the Sobolev norms in Lemma 3.4.1.

Lemma 3.4.4. *Suppose that T and $\tilde T$ are $\mathcal{C}^3$–Anosov diffeomorphisms of $\mathbb{T}^2$, $\|T\|_{\mathcal{C}^3}, \|\tilde T\|_{\mathcal{C}^3} \le K$. Let $q_\varepsilon(x,y) = \varepsilon^{-2}\bar q(\varepsilon^{-1}(y-x))$ be $\mathcal{C}^2$. Then there is a constant $C>0$ that depends only on $\bar q$ and on T such that for sufficiently small $\varepsilon>0$ holds: If $\mathrm{dist}_{\mathcal{C}^1}(T,\tilde T) \le \varepsilon^{\frac{3}{\rho}}$, then*

$$\|Q_\varepsilon f\| \le C\,\|f\|^\sim \quad \textit{and} \quad \|Q_\varepsilon f\|^\sim \le C\,\|f\|$$

for any $f \in \mathcal{C}^1(\mathbb{T}^2, \mathbb{R})$.

Proof. ¿From Lemmas 3.4.1 and 3.4.3 follows

$$\|Q_\varepsilon f\| \le C\ell^\rho \|Q_\varepsilon f\|_{2,2} + C\|Q_\varepsilon f\|^\sim \le C\ell^\rho\varepsilon^{-3}\|f\| + C\|Q_\varepsilon f\|^\sim\ ,$$

and the first inequality follows by Lemma 2.4.6. The second inequality is proven in the same way since the rôles of T and $\tilde T$ are perfectly symmetric. □

Proof of Proposition 2.5.1. We fix $r>\sigma$ and $\delta>0$ and consider only z with $r<|z|\le 2$ and $\mathrm{dist}(z, \mathrm{sp}(\mathcal{L})) > \delta$.[36] Several constants introduced in the sequel will depend on r, δ and $R(z)$ in various ways, but we will not precisely keep track of them in this section. So "const" denotes an arbitrary constant that depends in particular on z, but not on ε.

Recall from Remark 2.4.11 that Proposition 2.4.9 applies also to $Q_\varepsilon\mathcal{L}Q_\varepsilon$ and hence gives estimates on $R_\varepsilon(z) = (z - Q_\varepsilon\mathcal{L}Q_\varepsilon)^{-1}$. In particular, $\|R_\varepsilon(z)\| \le \mathrm{const}$. Therefore

$$\begin{aligned}\|R_\varepsilon(z)Q_\varepsilon(\mathcal{L}_T - \mathcal{L}_{\tilde T})Q_\varepsilon f\| &\le \mathrm{const}\ \|Q_\varepsilon(\mathcal{L}_T - \mathcal{L}_{\tilde T})Q_\varepsilon f\| \\ &\le \mathrm{const}\ \varepsilon^{-1}\,\|(\mathcal{L}_T - \mathcal{L}_{\tilde T})Q_\varepsilon f\|_1 \\ &\le \mathrm{const}\ \varepsilon^{-1}\ \mathrm{dist}_{\mathcal{C}^1}(T,\tilde T)\,\|Q_\varepsilon f\|_{1,1} \\ &\le \mathrm{const}\ \varepsilon^{-3}\ \mathrm{dist}_{\mathcal{C}^1}(T,\tilde T)\,\|f\|\end{aligned}$$

so that

$$\left\|(1 + R_\varepsilon(z)Q_\varepsilon(\mathcal{L}_T - \mathcal{L}_{\tilde T})Q_\varepsilon)^{-1}\right\| \le 2$$

provided that $\mathrm{dist}_{\mathcal{C}^1}(T,\tilde T) \le \frac12\,\mathrm{const}\ \varepsilon^3$. Since $\tilde R_\varepsilon(z) = \left[1 + R_\varepsilon(z)Q_\varepsilon(\mathcal{L}_T - \mathcal{L}_{\tilde T})Q_\varepsilon\right]^{-1} R_\varepsilon(z)$, it follows that

$$\|\tilde R_\varepsilon(z)\| \le 2\|R_\varepsilon(z)\|\ . \tag{3.4.6}$$

Observe next that

$$\tilde R_\varepsilon(z) = z^{-1} + z^{-2}Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon + z^{-2}Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon\tilde R_\varepsilon(z)Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon\ .$$

Then it follows from (3.4.6) and from Lemma 3.4.4 that

$$\|\tilde R_\varepsilon(z)\|^\sim \le \mathrm{const}\ (1 + \|R_\varepsilon(z)\|)\ . \tag{3.4.7}$$

The next step is to apply Proposition 2.4.9 to $\mathcal{L}_{\tilde T}$ and $Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon$ with the rôles of these two operators interchanged, *i.e.* interpreting $\mathcal{L}_{\tilde T}$ as a perturbation of $Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon$. This is possible, because of the following two observations:

1) The operators $Q_\varepsilon\mathcal{L}_{\tilde T}Q_\varepsilon$ satisfy the Lasota-Yorke type inequality from Lemma 2.4.7 with uniform

[36] Since $\rho_{\mathrm{sp}}(\tilde{\mathcal{L}}) \le 1$, estimate (2.5.2) is trivially verified for $|z|>2$.

constants that are close to those for $\mathcal{L}_T$ and $Q_\varepsilon \mathcal{L}_T Q_\varepsilon$, *cf.* Remark 2.4.11. (The dependence of these constants on the map T is continuous under $\mathcal{C}^2$ changes of T.)

2) In view of (3.4.7) and our convention on the use of "const", $\|\tilde{R}_\varepsilon(z)\|^\sim$ is bounded in terms of quantities depending only on T and on the kernel $\bar{q}(x,y)$. Therefore Remark 2.4.10 guarantees that

$$\|\tilde{R}(z)\|^\sim \leq \text{const} \quad \text{and} \tag{3.4.8}$$

$$|||\tilde{R}_\varepsilon(z) - \tilde{R}(z)|||^\sim \leq \text{const}\,\varepsilon^{\eta(\gamma-\beta)} \tag{3.4.9}$$

with constants that do not depend on $\tilde{T}$ and on ε if ε and $\text{dist}_{\mathcal{C}^1}(T,\tilde{T})$ are small enough.

Estimate (3.4.8) is exactly assertion (2.5.2) of Proposition 2.5.1.

For assertion (2.5.3) of this proposition – as discussed in section 2.5 – we must bound the three terms

$$|||R(z) - R_\varepsilon(z)|||, \quad \|\tilde{R}_\varepsilon(z) - R_\varepsilon(z)\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*}, \quad |||\tilde{R}_\varepsilon(z) - \tilde{R}(z)|||^\sim \tag{3.4.10}$$

where $R(z)$, $\tilde{R}(z)$, $R_\varepsilon(z)$ and $\tilde{R}_\varepsilon(z)$ denote the resolvents of $\mathcal{L}_T$, $\mathcal{L}_{\tilde{T}}$, $Q_\varepsilon \mathcal{L}_T Q_\varepsilon$ and $Q_\varepsilon \mathcal{L}_{\tilde{T}} Q_\varepsilon$, respectively. Because of (3.4.9), the first and the third term are bounded by const $\varepsilon^{\eta(\gamma-\beta)}$. It remains to estimate second one. Observing the resolvent estimate (3.4.7), the norm comparisons (2.5.1) and Lemma 3.4.3, we have

$$\begin{aligned}
\left\| \left(\tilde{R}_\varepsilon(z) - R_\varepsilon(z) \right) f \right\|_{(\mathcal{C}^1)^*} &\leq \left\| \tilde{R}_\varepsilon(z) Q_\varepsilon (\mathcal{L}_{\tilde{T}} - \mathcal{L}_T) Q_\varepsilon R_\varepsilon(z) f \right\|^\sim \\
&\leq \text{const} \cdot \|Q_\varepsilon(\mathcal{L}_{\tilde{T}} - \mathcal{L}_T) Q_\varepsilon R_\varepsilon(z) f\|^\sim \\
&\leq \text{const} \cdot \varepsilon^{-1} \, \|(\mathcal{L}_{\tilde{T}} - \mathcal{L}_T) Q_\varepsilon R_\varepsilon(z) f\|_1 \\
&\leq \text{const} \cdot \varepsilon^{-1} \cdot \text{dist}_{\mathcal{C}^1}(T,\tilde{T}) \, \|Q_\varepsilon R_\varepsilon(z) f\|_{1,1} \\
&\leq \text{const} \cdot \varepsilon^{-3} \cdot \text{dist}_{\mathcal{C}^1}(T,\tilde{T}) \, \|R_\varepsilon(z) f\| \\
&\leq \text{const} \cdot \varepsilon^{-3} \cdot \text{dist}_{\mathcal{C}^1}(T,\tilde{T}) \, \|f\|_{\mathcal{C}^1} \ .
\end{aligned}$$

Collecting the estimates for the three terms in (3.4.10) and setting $\varepsilon = \text{dist}_{\mathcal{C}^1}(T,\tilde{T})^{\rho/(3+\eta(\gamma-\beta))}$ we conclude that

$$\|\tilde{R}(z) - R(z)\|_{\mathcal{C}^1 \to (\mathcal{C}^1)^*} \leq \text{const} \cdot \text{dist}_{\mathcal{C}^1}(\tilde{T},T)^{\frac{\rho\eta(\gamma-\beta)}{3+\eta(\gamma-\beta)}} \ .$$

This is (2.5.3). □

Proof of Sub-Lemma 3.4.2. If $v^u(x)$ and $\tilde{v}^u(x)$ are the unitary vectors fields in the unstable directions for the two maps, it is clear that there exists $\rho_0 > 0$ such that[37]

$$\|v^u(x) - \tilde{v}^u(x)\| \leq C\ell^{\rho_0}. \tag{3.4.11}$$

The estimate $\|F - \text{Id}\,\|_\infty \leq C\ell^{\rho_0}$ is an immediate consequence of (3.4.11) and the definition of F. The second is more subtle and requires some work.

We start by studying the derivative of the map $\Gamma : U \to \mathbb{R}^2$. From the definition of Γ is follows immediately that

$$\begin{aligned}
\frac{\partial}{\partial \xi}\Gamma^{-1}(\xi,\eta) &= v^s(\Gamma^{-1}(\xi,\eta)) J^u \circ \Gamma^{-1}(\xi,\eta) \\
\frac{\partial}{\partial \eta}\Gamma^{-1}(\xi,\eta) &= v^u(\Gamma^{-1}(\xi,\eta)) J^s \circ \Gamma^{-1}(\xi,\eta),
\end{aligned}$$

where $J^s \circ \Gamma^{-1}(\xi,\eta)$ is the Jacobian of the holonomy H^u_ξ from the unstable fibre $\Gamma^{-1}\{(0,\cdot)\}$ to the fibre $\Gamma^{-1}\{(\xi,\cdot)\}$ at the point $(0,\eta)$, and $J^u \circ \Gamma^{-1}(\xi,\eta)$ is the Jacobian of the holonomy H^s_η from the unstable fibre $\Gamma^{-1}\{(\cdot,0)\}$ to the fibre $\Gamma^{-1}\{(\cdot,\eta)\}$ at the point $(\xi,0)$. While v^s, v^u are the stable and unstable unit vector fields respectively.

[37] This follows from the standard construction of the unstable distribution by cone field contraction. Given a strictly invariant continuous cone field $\mathcal{C}(x)$ for the map T, it will be strictly invariant for $\tilde{T}$ as well, provided the two maps are sufficiently close in the $\mathcal{C}^1$ topology. Accordingly, for each $x \in \mathcal{M}$ the cones $\mathcal{C}_n(x) := D_{T^{-n}x}T^n\mathcal{C}(T^{-n}x)$ and the analogously defined $\tilde{\mathcal{C}}_n(x)$ shrink at an exponential rate towards the respective unstable directions $v^u(x)$ and $\tilde{v}^u(x)$. Hence there exist $\nu \in (0,1)$ and $C > 0$ such that for each vector $v \in \mathcal{C}(T^{-n}x)$ holds $\|v^u(x) - D_{T^{-n}x}T^n v\| \leq C\nu^n$ and the same for $\tilde{T}$. On the other hand

$$\|D_{T^{-n}x}T^n - D_{\tilde{T}^{-n}x}\tilde{T}^n\| \leq \Lambda^n \ell$$

for some $\Lambda > 1$, from which, by choosing n appropriately, the statement follows.

The above formulae imply that

$$D\Gamma^{-1} = A\begin{pmatrix} J^s\circ\Gamma^{-1} & 0\\ 0 & J^u\circ\Gamma^{-1}\end{pmatrix}\quad ;\quad A:=(v^u,\, v^s). \tag{3.4.12}$$

As the same formulae apply to the change of coordinates $\tilde\Gamma$ associated to the map $\tilde T$, we have

$$DF = A\begin{pmatrix} \frac{J^s\circ\Gamma^{-1}}{\tilde J^s\circ\tilde\Gamma^{-1}} & 0\\ 0 & \frac{J^u\circ\Gamma^{-1}}{\tilde J^u\circ\tilde\Gamma^{-1}}\end{pmatrix}\tilde A^{-1}, \tag{3.4.13}$$

where the tilde designate the quantities relative to $\tilde\Gamma$. It is immediate to see that $A,\tilde A$ are $\mathcal{C}^{1+\alpha}$ matrix valued functions. Moreover, the fact that the distributions are uniformly transversal, equation (3.4.11) and the fact that $\|F-\mathrm{Id}\,\|_\infty\le C\ell^{\rho_0}$ show that there exists a constant $C>0$ such that

$$\begin{aligned}&\|A\|_\infty+\|\tilde A\|_\infty+\|A^{-1}\|_\infty+\|\tilde A^{-1}\|_\infty\le C\\ &\|A\tilde A^{-1}-\mathbf{1}\,\|_\infty\le C\ell^{\rho_0}.\end{aligned}$$

Hence, to see that DF is close to unity we need only to control the ratio between the Jacobians of the corresponding holonomies. Since the stable and the unstable one are treated exactly in the same way we will consider only the unstable ones. We want to prove that there exist $C,\rho>0$ such that

$$\left\|\frac{J^u\circ\Gamma^{-1}}{\tilde J^u\circ\tilde\Gamma^{-1}}-1\right\|_\infty\le C\ell^\rho. \tag{3.4.14}$$

The above fact is a ready consequence of the well known formula for the Jacobian of holonomies [33]:

$$J^u(x)=\prod_{n=0}^\infty\frac{D_{y_n}T|_{E^s(y_n)}}{D_{x_n}T|_{E^s(x_n)}},$$

where $x_0:=x=\Gamma^{-1}(\xi,\eta)$, $y_0:=H^u_\xi(x)$ and $x_n:=T^{-n}x$, $y_n:=T^{-n}y$. Since x_0 and y_0 are on the same unstable fibre, $|x_n-y_n|\le\text{const}\cdot\lambda_u^{-n}|x_0-y_0|$, so that there is some constant $C>0$ such that, for each $N\in\mathbb{N}$,

$$\left|J^u(x)\prod_{n=0}^N\frac{D_{x_n}T|_{E^s(x_n)}}{D_{y_n}T|_{E^s(y_n)}}-1\right|\le C\,\lambda_u^{-N}\ .$$

Accordingly, letting $\tilde x:=F^{-1}(x)$,

$$\begin{aligned}\left|\frac{J^u(x)}{\tilde J^u(\tilde x)}-1\right|&=\left|\prod_{n=0}^\infty\frac{D_{y_n}T|_{E^s(y_n)}\cdot D_{\tilde x_n}\tilde T|_{\tilde E^s(\tilde x_n)}}{D_{x_n}T|_{E^s(x_n)}\cdot D_{\tilde y_n}\tilde T|_{\tilde E^s(\tilde y_n)}}-1\right|\\ &\le C\,\lambda_u^{-N}+\left|\prod_{n=0}^N\frac{D_{y_n}T|_{E^s(y_n)}\cdot D_{\tilde x_n}\tilde T|_{\tilde E^s(\tilde x_n)}}{D_{x_n}T|_{E^s(x_n)}\cdot D_{\tilde y_n}\tilde T|_{\tilde E^s(\tilde y_n)}}-1\right|\end{aligned}$$

Denote by Λ the maximum of the derivatives of the maps in the neighbourhood under consideration. Since we know already that $|x-\tilde x|=|F(\tilde x)-\tilde x|\le C\ell^{\rho_0}$, it follows that $|x_n-\tilde x_n|\le\ell^{\rho_0}C\Lambda^n$, and the same estimate holds for $|y_n-\tilde y_n|$. Remembering (3.4.11) and the fact that the foliations are $\mathcal{C}^{1+\alpha}$ this yields immediately

$$\left|\frac{J^u(x)}{\tilde J^u(\tilde x)}-1\right|\le C\,\lambda_u^{-N}+C\ell^{\rho_0}\Lambda^N\|T\|_{\mathcal{C}^2}\ .$$

The result follows then by choosing $N=\frac{\rho_0|\ln\ell|}{\ln(\lambda_u\Lambda)}$, which yields $\rho=\rho_0\frac{\ln\lambda_u}{\ln(\lambda_u\Lambda)}$. □

3.5. **Proofs: Spectral stability – Ulam finite rank approximation.**

Proof of Proposition 2.6.3. Our starting point is the formal identity

$$(z-\mathcal{L}_{\mathcal{A},\varepsilon})^{-1}=\underbrace{\left(1+(z-Q_\varepsilon\mathcal{L})^{-1}(1-\Pi_{\mathcal{A}})Q_\varepsilon\mathcal{L}\right)^{-1}}_{=:\mathcal{N}}(z-Q_\varepsilon\mathcal{L})^{-1}\ .$$

We show that $\mathcal{N}$ is indeed invertible: For $f \in \mathcal{C}^1(\mathbb{T}^2, \mathbb{R})$ we have

$$\begin{aligned}
&\left\|(z - Q_\varepsilon\mathcal{L})^{-1}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\right\|_1 \\
&\quad\le \left\|(z^{-1} + z^{-2}Q_\varepsilon\mathcal{L})(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\right\|_1 + \left\|z^{-2}Q_\varepsilon\mathcal{L}(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\right\|_1 \\
&\quad\le (r^{-1} + r^{-2})\,\frac{2}{N}\,\varepsilon^{-1}\,\|f\|_1 + \text{const }\varepsilon^{-1}\left\|(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\right\|
\end{aligned}$$

where we used (3.4.2) and Lemma 2.6.1. Using Proposition 2.4.9 the second summand is further estimated by

$$\begin{aligned}
\left\|(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\right\| &\le (a + b)\;\|Q_\varepsilon\mathcal{L}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}f\| \\
&\le \text{const }\varepsilon^{-1}\,N^{-1}\,\|Q_\varepsilon\mathcal{L}f\|_{\mathrm{BV}} \\
&\le \text{const }N^{-1}\,\varepsilon^{-2}\,\|f\|_1
\end{aligned}$$

where we also used (3.4.1) and Lemmas 2.6.2 and 2.6.1. Combining both estimates we see that

$$\left\|(z - Q_\varepsilon\mathcal{L})^{-1}(1 - \Pi_\mathcal{A})Q_\varepsilon\mathcal{L}\right\|_1 \le \text{const }N^{-1}\,\varepsilon^{-3}\;.$$

This proves that $\mathcal{N}$ is invertible with $\|\mathcal{N}^{-1}\|_1 \le 2$ if $N^{-1}\varepsilon^{-3}$ is sufficiently small. Hence

$$\begin{aligned}
\left\|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_1 &\le 2\;\left\|(z^{-1} + z^{-2}Q_\varepsilon\mathcal{L} + z^{-2}Q_\varepsilon\mathcal{L}(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L})f\right\|_1 \\
&\le \text{const }\left(\|f\|_1 + \varepsilon^{-1}\left\|(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L}f\right\|\right) \\
&\le \text{const }\left(\|f\|_1 + \varepsilon^{-1}\,\|Q_\varepsilon\mathcal{L}f\|\right)
\end{aligned}$$

where we used (3.4.2). This proves (2.6.4).

We turn to the proof of (2.6.5).

$$\begin{aligned}
&\left\|\left((z - Q_\varepsilon\mathcal{L})^{-1} - (z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}\right)f\right\|_w \\
&= \left\|(z - Q_\varepsilon\mathcal{L})^{-1}(\Pi_\mathcal{A} - 1)Q_\varepsilon\mathcal{L}(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_w \\
&\le \left\|z^{-1}(\Pi_\mathcal{A} - 1)Q_\varepsilon\mathcal{L}(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_1 + \left\|z^{-1}(z - Q_\varepsilon\mathcal{L})^{-1}Q_\varepsilon\mathcal{L}(\Pi_\mathcal{A} - 1)Q_\varepsilon\mathcal{L}(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\| \\
&\le \text{const }N^{-1}\;\left\|Q_\varepsilon\mathcal{L}(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_{\mathrm{BV}} + \text{const }N^{-1}\,\varepsilon^{-1}\;\left\|Q_\varepsilon\mathcal{L}(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_{\mathrm{BV}} \\
&\le \text{const }N^{-1}\,\varepsilon^{-1}\,\varepsilon^{-1}\;\left\|(z - \mathcal{L}_{\mathcal{A},\varepsilon})^{-1}f\right\|_1 \\
&\le \text{const }N^{-1}\,\varepsilon^{-2}\;\left(\|f\|_1 + \varepsilon^{-1}\,\|Q_\varepsilon\mathcal{L}f\|\right) \\
&\le \text{const }N^{-1}\,\varepsilon^{-2}\;\left(\|f\|_1 + \varepsilon^{-1}\,\|f\|\right) \\
&\le \text{const }N^{-1}\,\varepsilon^{-3}\,\|f\|_{1,1}\;.
\end{aligned}$$

This is even a bit stronger than (2.6.5). □

3.6. **Proofs: Compact embedding.** In this section we prove Proposition 2.2.2. We start with a simple remark on products of functions.

Lemma 3.6.1. *For each $\phi \in \mathcal{C}^2(\mathcal{M}, \mathbb{R})$ there exists C_ϕ such that, for each $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$, holds*

$$\|f\phi\|_w \le C_\phi\|f\|_w\;, \quad \|f\phi\|_s \le C_\phi\|f\|_s\;, \quad \|f\phi\|_u \le C_\phi\|f\|\;.$$

Proof. A direct computation yields, for each $\varphi \in \mathcal{D}_{\tilde\beta}$, $\tilde\beta \in (0, 1]$, $v \in \mathcal{V}_\beta$,

$$\frac{\phi\varphi}{|\phi|_{\mathcal{C}^1}} \in \mathcal{D}_{\tilde\beta}; \quad \frac{\phi v}{|\phi|_{\mathcal{C}^1}} \in \mathcal{V}_\beta; \quad \frac{d\phi(v)}{|\phi|_{\mathcal{C}^2}} \in \mathcal{D}_\beta.$$

From this the lemma follows immediately with $C_\phi = 2|\phi|_{\mathcal{C}^2}$. □

The above Lemma allows us to use a smooth partition of unity $\{\phi_i\}$ and write $f = \sum_i f\phi_i$. The advantage consists in the possibility of working locally and using explicit coordinates which makes the argument more transparent.

Proof of Proposition 2.2.2. Let us set $\tau' := \min\{\tau, 1\}$. The idea of the proof is to represent the embedding $\mathcal{B} \to \mathcal{B}_w$ as a composition of compact and continuous linear maps:

$$\begin{array}{ccccccc}
\mathcal{B} & \longrightarrow & (\mathcal{C}^\beta)^* \times (\mathcal{C}^\beta)^* & \overset{\text{compact}}{\hookrightarrow} & (\mathcal{C}^{\gamma\tau'})^* \times (\mathcal{C}^{\gamma\tau'})^* & \longrightarrow & \mathcal{B}_w \\
f & & (f, df) & & (f, df) & & f
\end{array}$$

The first embedding is continuous, because $\|f\|^*_{\mathcal{C}^\beta} \le \|f\|_s$, $\|df\|^*_{\mathcal{C}^\beta} \le \|f\|_u$. The compactness of the second embedding is equivalent to the compact embedding of $\mathcal{C}^{\gamma\tau'}(\mathcal{M}, \mathbb{R})$ into $\mathcal{C}^\beta(\mathcal{M}, \mathbb{R})$, and this is well known because $\gamma\tau' > \beta$. The continuity of the last embedding is an immediate consequence of the next lemma. □

Lemma 3.6.2. *There exists $K > 0$ such that, for each $\varphi \in \mathcal{D}_\gamma$, there are $\phi^\varphi \in \mathcal{C}^{\gamma\tau'}(\mathcal{M}, \mathbb{R})$ and $V^\varphi \in \mathcal{C}^{\gamma\tau'}(\mathcal{M}, \mathbb{R}) \cap \mathcal{V}_\beta$ with*

$$\|\phi^\varphi\|_{\mathcal{C}^{\gamma\tau'}} + \|V^\varphi\|_{\mathcal{C}^{\gamma\tau'}} \leq K$$

and such that for each $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$

$$\int_{\mathcal{M}} f\varphi \, dm = \int_{\mathcal{M}} f\phi^\varphi \, dm + \int_{\mathcal{M}} df(V^\varphi) \, dm \ .$$

The rest of this section is devoted to the proof of this lemma. We need some more preparations. By Lemma 3.6.1 it follows that we can restrict ourselves to functions supported in some small open set U. In particular we require that U belongs to only one chart $\Phi : U \to \mathbb{R}^d$. Using the Cartesian structure of the chart we can identify all the tangent spaces in U with the tangent space of a chosen point $\bar{z} \in U$. We can do a linear change of coordinates in $\Phi(U)$ such that in the new coordinates $E^u(\bar{z}) = \{(x, 0) \in \mathbb{R}^d \ : \ x \in \mathbb{R}^{d_u}\}$ and $E^s(\bar{z}) = \{(0, y) \in \mathbb{R}^d \ : \ y \in \mathbb{R}^{s_s}\}$. In addition, without loss of generality we can assume that $\Phi(\bar{z}) = (0, 0)$ and $\Phi(U)$ is the product of two balls of radius $4r$ centred at zero.

From now on we will use the Cartesian coordinates (x, y) without further comments and we will identify U with $\Phi(U)$. Note that, since Φ is smooth, the Riemannian distance is uniformly equivalent to the Cartesian distance in the chart.

Let us specify more precisely how small the set U needs to be. The first smallness assumption on U results from the requirement that all unstable manifolds in U can be viewed as a graph of a function of x. We can then introduce the maps $H : \mathbb{R}^{2d_u + d_s} \to \mathbb{R}^{d_s}$ defined by the following requirements

- $H(x, x, y) = y$;
- for each $(x, y) \in B_{3r}(0) \times B_{3r}(0)$, $\{(\xi, H(x, \xi, y)) \ : \ \xi \in B_{3r}(0)\}$ is the local unstable manifold containing the point (x, y).

The point $(\xi, H(x, \xi, y))$ is nothing else than the intersection of the unstable manifold of the point (x, y) with the subspace $\{(w, z) \in \mathbb{R}^{d_u + d_s} \ : \ w = \xi\}$. This implies the obvious formula

$$H(\zeta, \xi, H(x, \zeta, y)) = H(x, \xi, y). \tag{3.6.1}$$

Depending on which of its three arguments we fix, H can serve various purposes and have different smoothness properties, which can all be inferred from the information on stable and unstable foliation collected in the appendix.

(3.6.2) For each $(x, y) \in B_{3r}(0) \times B_{3r}(0)$ the function $\xi \mapsto H(x, \xi, y)$ is of class $\mathcal{C}^3$, since it is just a straightforward parametrisation of the unstable manifold through (x, y).

This leads to a second smallness requirement, namely $\sup_{(x,y)} \|H(x, \cdot, y)\|_{\mathcal{C}^1(B_{3r})} \leq \frac{1}{6}$.[38] Accordingly, $\{H(x, \xi, y) : \ \xi \in B_{3r}(0)\} \subset B_{4r}(0)$. Moreover,

(3.6.3) For fixed (x, ξ) the map $y \mapsto H(x, \xi, y)$ is the holonomy along unstable fibres from the subspace $\{(w, z) \in \mathbb{R}^{d_u + d_s} \ : \ w = x\}$ to the subspace $\{(w, z) \in \mathbb{R}^{d_u + d_s} \ : \ w = \xi\}$. Denote by $J_3 H(x, \xi, y)$ the Jacobian of this holonomy.

(3.6.4) $J_3 H(x, \xi, y) = J_3 H(x, \zeta, y) \, J_3 H(\zeta, \xi, H(x, \zeta, y))$ by equation (3.6.1).

(3.6.5) The maps $y \mapsto H(x, \xi, y)$ and $y \mapsto J_3 H(x, \xi, y)$ are τ'-Hölder (see (A.2) and [33]).

(3.6.6) There is $C > 1$ such that $C \geq J_3 H(x, \xi, y) \geq C^{-1}$ for each x, ξ, y (see [33]).

(3.6.7) There is $C > 0$ such that $\sup_y |1 - J_3 H(x, \xi, y)| \leq C \, |x - \xi|$ for all x, ξ (Lemma A.3).

The instrumental change of variables $\Psi : B_{3r}(0) \times B_{3r}(0) \subset \mathbb{R}^d \to U$ is defined by

$$\Psi(x, y) := (x, H(0, x, y)) \ . \tag{3.6.8}$$

Note that equation (3.6.1) implies that

$$\Psi^{-1}(x, y) = (x, H(x, 0, y)) \ . \tag{3.6.9}$$

Figure 7 shows how Ψ is related to neighbourhoods of $(0, 0)$ of various sizes. Other basic properties are described in the next Lemma.

---

[38] This can be achieved since $|H(0, \xi, 0)| \leq C|\xi|^2$, by the $\mathcal{C}^2$ smoothness of the unstable manifolds, and by the τ'-Hölder continuity of the foliation, see the appendix.

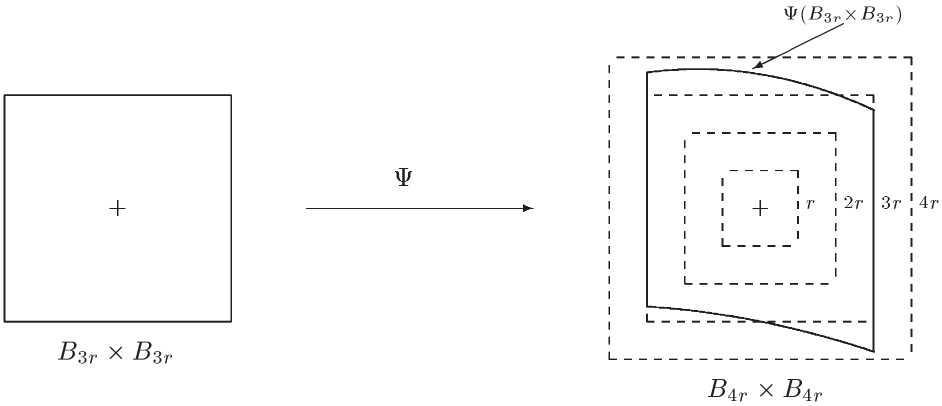

FIGURE 7. Neighbourhoods related to the Ψ coordinate change

Lemma 3.6.3. *The change of coordinates Ψ has the following properties:*

a) In the new coordinates the unstable foliation consists of parallel subspaces.

b) Ψ is absolutely continuous and its Jacobian $J\Psi(x,y) = J_3H(0,x,y)$.

c) $\sup_y |J\Psi(x,y) - J\Psi(\xi,y)| \le C\,|x-\xi|$.

d) For each $(x,\xi,y) \in B_{3r}(0)\times B_{3r}(0)\times B_{3r}(0)$, $E^u(\xi,H(x,\xi,y)) = \{(w, D_2H(x,\xi,y)w) \,:\, w \in \mathbb{R}^{d_u}\}$.[39]

Proof. a) is obvious, b) follows from Fubini's theorem, because Ψ is a skew product over the identity, c) is a consequence of (3.6.4) and (3.6.7), and d) is again obvious, since, as already remarked, $\xi \mapsto (\xi, H(x,\xi,y))$ parametrises an unstable manifold. $\square$

Next we need an auxiliary computation.

Lemma 3.6.4. *Let $f \in \mathcal{C}^0(\mathbb{R}^{d_u+d_s})$ with $\operatorname{supp}(f) \subset B_r(0)$ be such that $\xi \mapsto f(\xi,v)$ is of class $\mathcal{C}^1$ for all v. Then, for all $(u,v) \in B_r(0)$,*

$$\frac{1}{Z_{2r}}\int_{B_r(0)} f(\xi,v)d\xi = f(u,v) + \sum_{i=1}^{d_u} \frac{1}{Z_{2r}}\int_{B_r(0)} D_i f(\xi,v)\Theta_i(\xi,u)d\xi,$$

where Z_{2r} denotes the d_u-dimensional volume of $B_{2r}(0)$ and

$$\Theta_i(\xi,u) := \frac{(2r)^{d_u} - \|\xi-u\|^{d_u}}{d_u\|\xi-u\|^{d_u}}(\xi_i - u_i).$$

Proof. The result follows by using polar coordinates around the point (u,v) and observing that $f = 0$ outside the ball $B_r(0)$:

$$\begin{aligned}
\int_{B_r(0)} f(\xi,v)\,d\xi &= \int_{S^{d_u}} \sigma(d\theta)\int_0^{2r} d\rho\,\rho^{d_u-1} f(\rho\theta+u,v)\\
&= \int_{S^{d_u}} \sigma(d\theta)\int_0^{2r} d\rho\,\rho^{d_u-1}\left(f(u,v) + \int_0^\rho Df_{|t\theta+u,v}\cdot\theta\,dt\right)\\
&= Z_{2r}\,f(u,v) + \sum_{i=1}^{d_u}\int_{S^{d_u}}\sigma(d\theta)\int_0^{2r} dt\,t^{d_u-1}D_if(t\theta+u,v)\theta_i\frac{(2r)^{d_u}-t^{d_u}}{d_u\,t^{d_u-1}}\\
&= Z_{2r}\,f(u,v) + \sum_{i=1}^{d_u}\int_{B_r(0)} D_if(\xi,v)(\xi_i-u_i)\frac{(2r)^{d_u}-\|\xi-u\|^{d_u}}{d_u\,\|\xi-u\|^{d_u}}
\end{aligned}$$

where $\sigma(d\theta)$ is the measure on the $d_u - 1$ dimensional spherical surface S^{d_u}. $\square$

Lemma 3.6.4 allows to write the integral of f against a test function in a way that makes the proof of Lemma 3.6.2 particularly easy.

[39]Here, and in the following, by D_2 we mean the derivative with respect to the second variable.

Proof of Lemma 3.6.2. Let $\{\psi_k\}$ be a smooth partition of unity such that each $V_k := \operatorname{supp}\psi_k$ is contained in an open set U_k where $\{U_k\}$ satisfy the smallness requirements previously discussed for the generic local neighbourhood U. In addition U_k and V_k are chosen such that, if Φ_k denotes the chart associated to U_k so that $\Phi_k U_k = B_{4r}(0) \times B_{4r}(0)$, then $\Phi_k(V_k) \subset B_r(0) \times B_r(0)$.

Let us start by considering any $(U,V) \in \{(U_k, V_k)\}$ and consider $f \in \mathcal{C}^1$, $\operatorname{supp} f \subset V$.

Applying the coordinate change from Lemma 3.6.3 to the function $f \circ \Psi$ and using the formula from Lemma 3.6.4 we have

$$
\begin{aligned}
&\int_V f\varphi\, dm \\
&= \int_{B_r(0)} \int_{B_r(0)} du\, dv\; f(\Psi(u,v))\, \varphi(\Psi(u,v))\, J\Psi(u,v) \\
&= \int_{B_r(0)} \int_{B_r(0)} dv\, d\xi\; f(\Psi(\xi,v)) \frac{1}{Z_{2r}} \int_{B_r(0)} du\; \varphi(\Psi(u,v)) J\Psi(u,v) \\
&- \sum_{i=1}^{d_u} \int_{B_r(0)} \int_{B_r(0)} \frac{1}{Z_{2r}} \int_{B_r(0)} du\, dv\, d\xi\; d_{\Psi(\xi,v)} f((\mathbf{e}_i, D_2 H(0,\xi,v)\mathbf{e}_i)) \Theta_i(\xi,u) \varphi(\Psi(u,v)) J\Psi(u,v),
\end{aligned}
$$

where $\mathbf{e}_i$ $(i = 1, \dots, d_u)$ are the canonical unit vectors in $\mathbb{R}^{d_u}$. This suggests to define

$$
\begin{aligned}
\tilde{\phi}^{\varphi} \circ \Psi(\xi,v) &:= J\Psi(\xi,v)^{-1} \frac{1}{Z_{2r}} \int_{B_r(0)} du\; \varphi(\Psi(u,v)) J\Psi(u,v); \\
\tilde{V}_i^{\varphi} \circ \Psi(\xi,v) &:= J\Psi(\xi,v)^{-1} \frac{1}{Z_{2r}} \int_{B_r(0)} du\; \Theta_i(\xi,u) \varphi(\Psi(u,v)) J\Psi(u,v) (\mathbf{e}_i, D_2 H(0,\xi,v)\mathbf{e}_i)\,.
\end{aligned}
$$

Notice that equation (3.6.9) and (3.6.5) imply

$$
J\Psi(\Psi^{-1}(x,y)) = J_3 H(0,x,H(x,0,y)) = (J_3 H(x,0,y))^{-1}
$$

because $H(x,x,y) = y$, and

$$
J\Psi(u, H(x,0,y)) = J_3 H(0,u,H(x,0,y)) = \frac{J_3 H(x,u,y)}{J_3 H(x,0,y)}\,.
$$

Hence, for each $(x,y) \in B_r(0) \times B_r(0)$,

$$
\begin{aligned}
\tilde{\phi}^{\varphi}(x,y) &= \frac{1}{Z_{2r}} \int_{B_r(0)} du\; \varphi(u, H(x,u,y)) J_3 H(x,u,y) \\
\tilde{V}_i^{\varphi}(x,y) &= \frac{1}{Z_{2r}} \int_{B_r(0)} du\; \Theta_i(x,u) \varphi(u, H(x,u,y)) J_3 H(x,u,y) (\mathbf{e}_i, D_2 H(x,x,y)\mathbf{e}_i)
\end{aligned}
\tag{3.6.10}
$$

since $D_2 H(0,x,H(x,0,y)) = D_2 H(x,x,y)$ [40] and

$$
\int_V f\varphi\, dm = \int_V f\tilde{\phi}^{\varphi}\, dm - \sum_i \int_V df(\tilde{V}_i^{\varphi})\, dm\,. \tag{3.6.11}
$$

Let $\tilde{V}^{\varphi} := \sum_i \tilde{V}_i^{\varphi}$.

Lemma 3.6.5. *There exists $\tilde{K} > 0$ such that, for each $\varphi \in \mathcal{D}_\gamma$, $\tilde{\phi}^{\varphi} \in \mathcal{C}^{\gamma\tau'}(V,\mathbb{R})$, $\tilde{V}^{\varphi} \in \mathcal{C}^{\gamma\tau'}(V,\mathbb{R}^d)$ and*

$$
\|\tilde{\phi}^{\varphi}\|_{\mathcal{C}^{\gamma\tau'}} + \|\tilde{V}^{\varphi}\|_{\mathcal{C}^{\gamma\tau'}} \le \tilde{K}.
$$

Lemma 3.6.5 is clearly the local version of what we aim at; we postpone its proof to the end of the section and use it beforehand to conclude the argument.

To go to the global level the first step is to define globally $\tilde{\phi}^{\varphi}$ and $\tilde{V}^{\varphi}$. Let $\chi \in \mathcal{C}^2(\mathcal{M},\mathbb{R})$ be such that $\chi \equiv 1$ on V and $\chi \equiv 0$ outside $B_{2r}(0) \times B_{2r} \subset \Psi(B_{3r}(0) \times B_{3r}(0)) \subset U$. Then $\hat{\phi}^{\varphi} := \chi\tilde{\phi}^{\varphi}$ and $\hat{V}^{\varphi} := \chi\tilde{V}^{\varphi}$ are well defined on all of $\mathcal{M}$, $\hat{K}^{-1}\hat{\phi}^{\varphi} \in \mathcal{D}_\beta$, $\hat{K}^{-1}\hat{V}^{\varphi} \in \mathcal{V}_\beta$ for suitable $\hat{K} > 0$, and since f is supported in V it follows from (3.6.11) that

$$
\int_{\mathcal{M}} f\varphi\, dm = \int_{\mathcal{M}} f\hat{\phi}^{\varphi}\, dm - \int_{\mathcal{M}} df(\hat{V}^{\varphi})\, dm\,. \tag{3.6.12}
$$

[40] Just differentiate the particular instance of (3.6.1), $H(0,\xi,H(x,0,y)) = H(x,\xi,y)$, with respect to ξ.

We can now conclude the argument by going back to our partition of unity. Given $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$ and $\varphi \in \mathcal{D}_\beta$ for each ψ_k we have $\hat{\phi}_k^\varphi, \hat{V}_k^\varphi$ such that

$$\begin{aligned}\int_{\mathcal{M}} f\varphi &= \sum_k \int_{\mathcal{M}} (f\psi_k)\varphi = \sum_k \int_{\mathcal{M}} f\psi_k \hat{\phi}^\varphi - \int_{\mathcal{M}} d(f\psi_k)(\hat{V}^\varphi)\\ &= \sum_k \int_{\mathcal{M}} f(\psi_k\hat{\phi}^\varphi + d\psi_k(\hat{V}_k^\varphi)) - \int_{\mathcal{M}} df(\psi_k \hat{V}_k^\varphi)\\ &= \int_{\mathcal{M}} f\phi^\varphi - \int_{\mathcal{M}} df(V^\varphi),\end{aligned}$$

where

$$\phi^\varphi := \sum_k \psi_k\hat{\phi}^\varphi + d\psi_k(\hat{V}_k^\varphi) \quad \text{and} \quad V^\varphi := \sum_k \psi_k \hat{V}_k^\varphi \ .$$

□

Proof of Lemma 3.6.5. From formula (3.6.10) it is clear that the supremum norms of the above test functions are bounded, more care is needed to control the Hölder norms. First of all, as stated in (3.6.5), $H(x,\xi,y)$ and $J(x,\xi,y)$ are τ'-Hölder in y and the same for $D_\xi H(x,\xi,y)$.[41]

We discuss $\tilde{\phi}^\varphi$ first. The idea is to see how the function varies when moving along the unstable manifold and when varying y. Since

$$\begin{aligned}\tilde{\phi}^\varphi(x', H(x,x',y)) &= \frac{1}{Z_{2r}} \int_{B_r(0)} du\ \varphi(u, H(x',u,H(x,x',y)))J_3H(x',u,H(x,x',y))\\ &= \tilde{\phi}^\varphi(x,y) J_3 H(x',x,y),\end{aligned}$$

it follows

$$|\tilde{\phi}^\varphi(x,y) - \tilde{\phi}^\varphi(x', H(x,x',y))| = |\tilde{\phi}^\varphi(x,y)|\,|1 - J_3H(x',x,y)| \le C|\varphi|_\infty |x - x'| \ ,$$

where we have used (3.6.1) and the fact that the Jacobian of the holonomy is close to the unity when the two transversal manifolds are close, see (3.6.7). Next, given $y, y' \in \mathbb{R}^{d_s}$, we can define a map $G : \mathbb{R}^{d_u} \to \mathbb{R}^{d_u}$ by the stable holonomy H^s between the unstable manifold of (x,y) and of (x,y'). Namely, for each point $(\xi, H(x,\xi,y))$ of the first unstable manifold we define G such that

$$H^s(\xi, H(x,\xi,y)) = (G(\xi), H(x, G(\xi), y')).$$

The reader can look at Figure 8 to have a quick pictorial idea of all the quantities defined so far. Accordingly,

$$\tilde{\phi}^\varphi(x,y') = \frac{1}{Z_{2r}} \int_{G^{-1}(B_r(0))} d\xi\ \varphi(G(\xi), H(x,G(\xi),y')J_3H(x,G(\xi),y')\, JG(\xi).$$

But $|1 - JG|_\infty \le C|y - y'|^{\tau'}$ since the holonomy is close to unity and the distance between the two manifolds is bounded by $|y-y'|^{\tau'}$.[42] By the same reason, since the stable foliation is uniformly transversal to the unstable one, $\|G(\xi) - \xi\| \le C|y-y'|^{\tau'}$. Using (3.6.5), (3.6.4) and (3.6.7) one shows that $|J_3H(x,\xi,y) - J_3H(x,G(\xi),y')| \le C|y-y'|^{\tau'}$. Then it follows, for each $(x,y) \in B_r(0)\times B_r(0)$ with $|y - y'| \le \delta^{\frac{1}{\tau'}}$,

$$\begin{aligned}|\tilde{\phi}^\varphi(x,y) - \tilde{\phi}^\varphi(x,y')| &\le 2C|\varphi|_\infty|y-y'|^{\tau'} + \frac{C}{Z_{2r}}\int_{B_r(0)} d\xi\ |\varphi(\xi, H(x,\xi,y)) - \varphi(H^s(\xi, H(x,\xi,y)))|\\ &\le 2C|\varphi|_\infty |y-y'|^{\tau'} + C\, H^s_\gamma(\varphi)\,|y-y'|^{\tau'\gamma} \le 3C\,|y-y'|^{\gamma\tau'} \ .\end{aligned}$$

This clearly implies that the test function $\tilde{\phi}^\varphi$ is $\tau'\gamma$-Hölder with Hölder norm bounded uniformly in $\varphi \in \mathcal{D}_\gamma$.

The study of the vectors $\tilde{V}_i^\varphi$ is completely similar apart for the presence of the singular kernel Θ_i. To control it, we will use the following result.

[41] This last statement is nothing else than the τ'-Hölder continuity of the unstable foliation.

[42] This time we have an holonomy between unstable manifolds, for this case the claimed estimate on the Jacobian can be found also in [33].

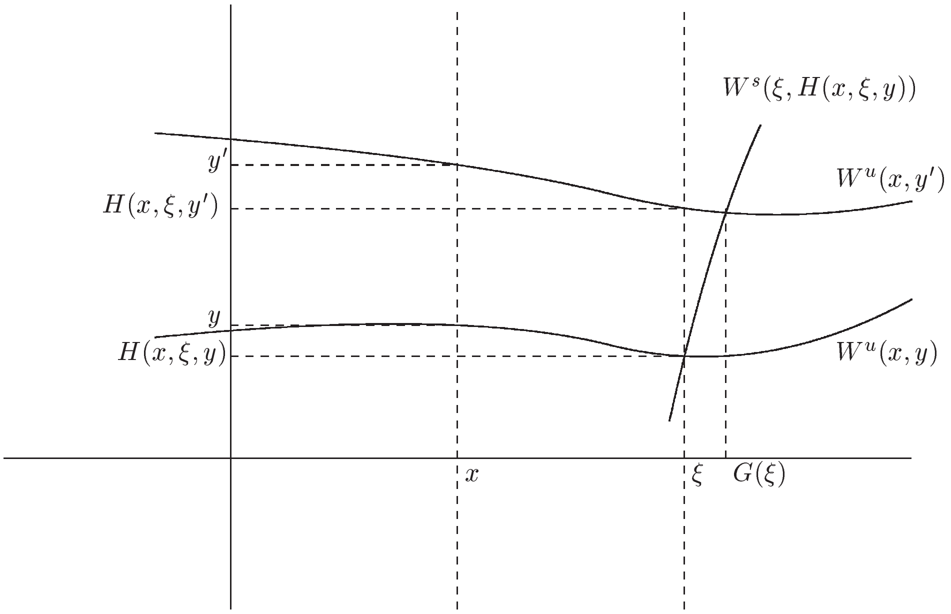

FIGURE 8

Sub-lemma 3.6.6. *There exists a constant $C > 0$ such that, for each $\phi \in L^\infty$,*

$$\frac{1}{Z_{2r}}\left|\int_{B_r(0)} \Theta_i(x,u)\phi(u)\,du - \int_{B_r(0)} \Theta_i(x',u)\phi(u)\,du\right| \leq C|x-x'|\ln|x-x'|^{-1}\,|\phi|_\infty$$
$$\frac{1}{Z_{2r}}\left|\int_{B_r(0)} \Theta_i(x,\xi)\phi(\xi)\,d\xi - \int_{B_r(0)} \Theta_i(x,G(\xi))\phi(\xi)\,d\xi\right| \leq C|y-y'|^{\tau'}\ln|y-y'|^{-1}\,|\phi|_\infty.$$

Proof. An explicit computation yields

$$Z_{2r}^{-1}\,|\Theta_i(x,u)| \leq C\|x-u\|^{-(d_u-1)}$$
$$Z_{2r}^{-1}\,|\Theta_i(x,u) - \Theta_i(x',u')| \leq C\,\frac{\|(x-u)-(x'-u')\|}{\min\{\|x-u\|^{d_u}, \|x'-u'\|^{d_u}\}}$$

for some fixed constant C. The estimate is then done by dividing the integral into two parts. In the first inequality one considers separately the integral in a ball of radius $2\|x-x'\|$ around x and the rest (where $\|x'-u\| \geq \frac{1}{2}\|x-u\|$).

This yields

$$\text{const}\,|\phi|_\infty \int_0^{2\|x-x'\|} d\rho + \text{const}\,|\phi|_\infty \int_{2\|x-x'\|}^{r} \rho^{-1}\,d\rho$$

from which the wanted result follows. The other inequality is done in the same way by separating a ball of radius $a\|y-y'\|^{\tau'} \geq |1-G|_\infty$ around u with a large enough. □

The regularity of $\tilde V_i^\varphi$ follows then trivially from the same arguments used to study $\tilde\phi^\varphi$ and by Sub-Lemma 3.6.6. □

3.7. **Proofs: Unstable seminorm.**

Proof of Lemma 2.1.4. In a sufficiently small neighbourhood U we can use the coordinates Ψ introduced in the proof of Lemma 3.6.2, *i.e.* $\Psi(\xi,y) = (\xi, H(0,\xi,y))$. Recall from Lemma 3.6.3 that $E^u(\Psi(\xi,y)) = \{(w, D_2H(0,\xi,y)w) \;:\; w \in \mathbb{R}^{d_u}\}$. We can then construct a vector field basis of the type requested in the lemma as $V_i \circ \Psi(\xi,y) = \frac{\partial}{\partial\xi}\Psi(\xi,y)\mathbf{e}_i = (\mathbf{e}_i, D_2H(0,\xi,y)\mathbf{e}_i)$. Note that, according to Lemma 3.6.3 and the regularity of the unstable distribution stated in (A.1), the vector fields $C_0^{-1}V_i$ are in $\mathcal{V}_\beta$ for a suitable $C_0 > 0$.

Consider now some $\varphi \in \mathcal{C}^\beta(\mathcal{M}, \mathbb{R}^{d_u})$ with $\operatorname{supp}\varphi \subset U$ and $|\varphi|_\infty \leq 1$, $H^s_\beta(\varphi) \leq 1$ which is $\mathcal{C}^1$ when restricted to any unstable manifold, and let $V_\varphi := \sum_i \varphi_i V_i$. Then $C^{-1}V_\varphi \in \mathcal{V}_\beta$ for some

$C > 0$ that does not depend on φ and, denoting $\tilde{U} = \Psi^{-1}U$, we have

$$\begin{aligned}\int_{\mathcal{M}} df(V_\varphi)\,dm &= \sum_i \int_{\mathcal{M}} \varphi_i\, df(V_i)\,dm \\ &= \sum_i \int_{\tilde{U}} \varphi_i(\Psi(\xi,y))\, d_{\Psi(\xi,y)} f(V_i(\Psi(\xi,y)))\, J\Psi(\xi,y)\,d\xi dy \\ &= \sum_i \int_{\tilde{U}} (\varphi_i \circ \Psi)(\xi,y)\, J\Psi(\xi,y)\, \frac{\partial}{\partial \xi_i}(f\circ\Psi)(\xi,y)\,d\xi dy\end{aligned}$$

In view of (3.6.2) the functions $\varphi_i \circ \Psi(\xi,y)$ are Lipschitz in the variable ξ, and in view of Lemma 3.6.3 also $J\Psi(\xi,y)$ is Lipschitz with respect to ξ. This allows to integrate by parts obtaining

$$\begin{aligned}\int_{\mathcal{M}} df(V_\varphi)\,dm &= -\sum_i \int_{\tilde{U}} \frac{\partial}{\partial \xi_i}\big((\varphi_i\circ\Psi)(\xi,y)J\Psi(\xi,y)\big)\, f(\Psi(\xi,y))\,d\xi dy \\ &= -\sum_i \int_{\mathcal{M}} f\, d\varphi_i(V_i)\,dm + \sum_i \int_{\mathcal{M}} f\cdot \varphi_i A_i\,dm\end{aligned} \tag{3.7.1}$$

where

$$d_{(x,y)}\varphi_i(V_i(x,y)) := \frac{\partial}{\partial \xi_i}(\varphi_i\circ\Psi)(\Psi^{-1}(x,y)) \quad \text{and} \quad A_i := -\left[\frac{\frac{\partial}{\partial \xi_i}J\Psi}{J\Psi}\right]\circ \Psi^{-1}.$$

In particular,

$$\sum_{i=1}^{d_u} d\varphi_i(V_i) = \operatorname{div}^u(\varphi\circ\Psi)\circ\Psi^{-1} := \sum_{i=1}^{d_u} \frac{\partial}{\partial \xi_i}(\varphi_i\circ\Psi)\circ\Psi^{-1}\ . \tag{3.7.2}$$

In view of Lemma A.3, not only $|A_i|_\infty < \infty$, but also $H_s^{\beta'}(A_i) < \infty$. Finally, if $\tilde{V}_i$ is a different basis then there must exist functions $\alpha_{ij} \in \mathcal{C}^{\tau'}(\mathcal{M},\mathbb{R})$, but $\mathcal{C}^1$ when restricted to any unstable manifold, such that $V_i = \sum_j \alpha_{ij}\tilde{V}_j$. Then the test vector field $v^u = \sum_i \varphi_i V_i$ can be written as $v^u = \sum_j \tilde{\varphi}_j \tilde{V}_j$ with $\tilde{\varphi}_j := \sum_i \varphi_i \alpha_{ij}$. Thus

$$\sum_j d\tilde{\varphi}_j(\tilde{V}_j) = \sum_{ij} \alpha_{ij} d\varphi_i(\tilde{V}_j) + \sum_{ij} d\alpha_{ij}(\tilde{V}_j)\varphi_i = \sum_i d\varphi_i(V_j) + \sum_{ij} d\alpha_{ij}(\tilde{V}_j)\varphi_i,$$

which yields the same formula only with different functions A_i. □

Proof of Proposition 2.3.2(d). Let $h = \Pi_j f \in \Pi_j \mathcal{B}$. Without loss of generality we can assume that $f \in \mathcal{C}^1(\mathcal{M},\mathbb{R})$. We must show that the measure μ_h, locally[43] conditioned to the unstable foliation, is absolutely continuous on μ–almost every fibre with respect to the Riemannian measure on the fibre with a density of bounded variation. So let U be a small neighbourhood on which, as in the last proof, we can use the change of coordinates Ψ from (3.6.8), and recall that $\tilde{U} = \Psi^{-1}U$. Since the maps $x \mapsto \Psi(x,y)$ for each fixed y are absolutely continuous with Jacobians close to 1 which are Lipschitz in the variable x (see (3.6.7)), it suffices to prove the same property for the measure $\tilde{\mu}_h := \mu_h \circ \Psi$ and its conditional measures on the subspaces $N_y := \{(x,v) \in \mathbb{R}^{d_u+d_s} : v = y\}$. Hence we must show that there exists a constant $C > 0$ such that for each $\phi \in \mathcal{C}^1(\mathbb{R}^{d_u+d_s},\mathbb{R}^{d_u})$ with $\operatorname{supp}(\phi) \subset \tilde{U}$, $(\xi,y) \mapsto \phi(\xi,y)$,

$$\int_{\tilde{U}} \operatorname{div}^u(\phi)(\xi,y)\,\tilde{\mu}_h(d\xi dy) \le C\cdot \int_{\tilde{U}} |\phi|\,d\tilde{\mu}\ , \tag{3.7.3}$$

where $\tilde{\mu} := \tilde{\mu}_{\Pi_1 1}$. We will shortly see that this estimate implies for $\tilde{\mu}$-almost every N_y

$$\int_{\tilde{U}\cap N_y} \operatorname{div}^u(\phi)(\xi,y)\,(\tilde{\mu}_h)_y(d\xi) \le C\cdot |\phi|_\infty \quad \forall \phi \in \mathcal{C}^1 \tag{3.7.4}$$

[43] Globally, the σ-algebra generated by the unstable foliation is coarser than we intend it to be. For example, in the case of torus automorphisms, it contains only sets of Lebesgue measure 0 or 1.

where $(\tilde{\mu}_h)_y$ denotes the conditional measure of $\tilde{\mu}_h$ on N_y. This implies the announced absolute continuity of $(\tilde{\mu}_h)_y$ w.r.t. Lebesgue measure on N_y with a density of bounded variation.[44] To obtain (3.7.4) from (3.7.3) it is convenient to introduce some natural notations: let $\mathcal{F}^u$ be the σ-algebra generated by the subspaces N_y; $\mu_h^u := \mu_h|_{\mathcal{F}^u}$ the measure restricted to the σ-algebra $\mathcal{F}^u$ (in our case this is simply a marginal of μ_h); $\mathcal{C}_u^1$ be the $\mathcal{C}^1$ functions that are $\mathcal{F}^u$ measurable; and by $\mathbb{E}_\mu$ let us designate the expectation with respect to the measure μ. Clearly, in this coordinate free language, $\int \phi\, d(\tilde{\mu}_h)_y$ corresponds to the conditional expectation $\mathbb{E}_{\tilde{\mu}_h}(\phi \mid \mathcal{F}^u)(N_y)$. Finally, let us designate by $|\cdot|_p^{h,u} := |\cdot|_{L^p(\tilde{\mu}_h^u)}$ the norms in the spaces $L^p(\tilde{\mu}_h^u)$. Then

$$
\begin{aligned}
|\mathbb{E}_{\tilde{\mu}_h}(\operatorname{div}^u \phi \mid \mathcal{F}^u)|_\infty^{h,u} &= \sup_{\substack{\psi \in L^1(\tilde{\mu}_h^u) \\ |\psi|_1^{h,u} \le 1}} \mathbb{E}_{\tilde{\mu}_h^u}(\psi \mathbb{E}_{\tilde{\mu}_h}(\operatorname{div}^u \phi \mid \mathcal{F}^u)) \\
&= \sup_{\substack{\psi \in L^1(\tilde{\mu}_h^u) \\ |\psi|_1^{h,u} \le 1}} \mathbb{E}_{\tilde{\mu}_h}(\psi \operatorname{div}^u \phi) \\
&= \sup_{\substack{\psi \in \mathcal{C}_u^1 \\ |\psi|_1^{h,u} \le 1}} \mathbb{E}_{\tilde{\mu}_h}(\psi \operatorname{div}^u \phi) \\
&= \sup_{\substack{\psi \in \mathcal{C}_u^1 \\ |\psi|_1^{h,u} \le 1}} \mathbb{E}_{\tilde{\mu}_h}(\operatorname{div}^u(\psi\phi)) \\
&\le C \cdot |\phi|_\infty \int_{\tilde{U}} |\psi|\, d\tilde{\mu}\ ,
\end{aligned}
$$

where the third equality is due to the fact that the $\mathcal{F}^u$ measurable functions are simply functions of only one coordinate (hence the obvious denseness of $\mathcal{C}^1$ in the corresponding L^1), and the last inequality follows by (3.7.3).

This proves that for each function $\phi \in \mathcal{C}^1$ there exists a set $A(\phi)$, $\tilde{\mu}(A(\phi)) = 0$, such that if $N_y \notin A(\phi)$, (3.7.4) applies. Since $\mathcal{C}^1$ is separable, (3.7.4) follows by standard approximation arguments.

To conclude we are left with the proof of estimate (3.7.3). Let $\varphi := \phi \circ \Psi^{-1}$. Because of the continuity properties of Ψ the vector field φ is of the type considered in the previous proof, and we define $V_\varphi = \sum_i \varphi_i V_i$ as there. Then $\sum_i d\varphi_i(V_i) \circ \Psi = \operatorname{div}^u(\varphi \circ \Psi) = \operatorname{div}^u(\phi)$.

Since $\mu_h = \mu_{\Pi_j f}$ is defined in terms of $\mathcal{L}^k f$ (see (3.2.1)), we start with the following observation: For $v \in \mathcal{V}_\beta$ recall from Sub-lemma 3.1.1 that $(R_n v)(x) = A^{-2}\lambda_u^n\, d_{T^n x} T^{-n}(v(T^n x)) \in \mathcal{V}_\beta$ for $n = 0, \ldots, N$ where N was chosen such that $(\sigma \cdot \min\{\lambda_u, \lambda_s^\beta\})^N > 9A^2$. Hence, applying identity (3.1.4) repeatedly (for $n = N$), we get

$$
\int_{\mathcal{M}} d(\mathcal{L}^{kN} f)(V_\varphi)\, dm = (A^2\lambda_u^{-N})^k \int_{\mathcal{M}} df(R_N^k V_\varphi)\, dm + \int_{\mathcal{M}} f \cdot \sum_{l=0}^{k-1} (A^2\lambda_u^{-N})^l \Delta_N^{R_N^l V_\varphi} \circ T^{(k-l)N}\, dm
$$

where Δ_N^v denotes the distortion defined in (3.1.3) and where $\left|\int_{\mathcal{M}} df(R_N^k V_\varphi)\, dm\right| \le C \cdot \|f\|_u$ for all $k \ge 0$ because $C^{-1} V_\varphi \in \mathcal{V}_\beta$ and $R_N(\mathcal{V}_\beta) \subset V_\beta$. In the rest of the proof we will, for simplifying the estimate, assume that $N = 1$ and $A = 1$ so that each integer n is trivially a multiple of N and

[44] To see it, denote by ν the measure $(\tilde{\mu}_h)_y$ on $\mathbb{R}^{d_u}$. By taking a convolution with a smooth mollifier j_ε we have that $\nu * j_\varepsilon =: \nu_\varepsilon$ is absolutely continuous with respect to Lebesgue measure dm on $\mathbb{R}^{d_u}$. Moreover, ν_ε converges weakly to ν as $\varepsilon \to 0$. In addition,

$$
\int \operatorname{div}(\phi)\, d\nu_\varepsilon = \int \operatorname{div}(\phi) * j_\varepsilon\, d\nu = \int \operatorname{div}(\phi * j_\varepsilon)\, d\nu \le C|\phi * j_\varepsilon|_\infty \le C|\phi|_\infty.
$$

Hence the densities p_ε of ν_ε with respect to Lebesgue are uniformly bounded in bounded variation norm. Accordingly, they form a relatively compact sequence in $L^1(\mathbb{R}^{d_u})$, [15]. Let $p \in L^1(\mathbb{R}^{d_u})$ be an accumulation point and $\{\varepsilon_k\}$ a sequence such that $\lim_{k\to\infty} \|p_{\varepsilon_k} - p\|_1 = 0$. Then $p \in \mathrm{BV}(\mathbb{R}^{d_u})$ and for each $\phi \in \mathcal{C}^0(\mathbb{R}^{d_u})$

$$
\int \phi p\, dm = \lim_{k\to\infty} \int \phi p_{\varepsilon_k}\, dm = \lim_{k\to\infty} \int \phi\, d\nu_{\varepsilon_k} = \int \phi\, d\nu
$$

that is $p\, dm = d\nu$.

write R instead of R_1.[45] Observing (3.2.1) and (3.7.1), it follows that

$$\begin{aligned}
\int_{\tilde U} \operatorname{div}^u(\phi)(\xi,y)\,\tilde\mu_h(d\xi dy) &= \int_{\mathcal M}\sum_i d\varphi_i(V_i)\,d\mu_h \\
&= \lim_{n\to\infty}\frac1n\sum_{k=0}^{n-1}\lambda_j^{-k}\int_{\mathcal M}\sum_i d\varphi_i(V_i)\,\mathcal L^k f\,dm \\
&= \lim_{n\to\infty}\frac1n\sum_{k=0}^{n-1}\lambda_j^{-k}\left(-\int_{\mathcal M} d(\mathcal L^k f)(V_\varphi)\,dm+\int_{\mathcal M}\mathcal L^k f\sum_i\varphi_i A_i\,dm\right) \\
&= \lim_{n\to\infty}\frac1n\sum_{k=0}^{n-1}\lambda_j^{-k}\left(-\int_{\mathcal M} f\cdot\sum_{l=0}^{k-1}\lambda_u^{-l}\Delta_1^{R^lV_\varphi}\circ T^{k-l}\,dm+\int_{\mathcal M}\mathcal L^k f\sum_i\varphi_i A_i\,dm\right) \\
&= -\sum_{l=0}^{\infty}\left(\frac{\lambda_u^{-1}}{\lambda_j}\right)^l\lim_{n\to\infty}\frac1n\int_{\mathcal M}\sum_{k=l+1}^{n-1}\mathcal L^{k-l}f\cdot\Delta_1^{R^lV_\varphi}\,dm+\lim_{n\to\infty}\frac1n\sum_{k=0}^{n-1}\lambda_j^{-k}\int_{\mathcal M}\mathcal L^k f\sum_i\varphi_iA_i\,dm \\
&= -\int_{\mathcal M}\sum_{l=0}^{\infty}\left(\frac{\lambda_u^{-1}}{\lambda_j}\right)^l\Delta_1^{R^lV_\varphi}\,d\mu_h+\int\sum_i\varphi_iA_i\,d\mu_h
\end{aligned}$$

where, for interchanging the sum and the limit, we made use of the fact that

$$\left|\int_{\mathcal M} f\cdot\lambda_u^{-l}\Delta_1^{R^lV_\varphi}\circ T^{k-l}\,dm\right| \le \sigma^l\,\|f\|_\infty\,|\Delta_1^{R^lV_\varphi}|_\infty \le \sigma^l\,\|f\|_\infty\,|R^lV_\varphi|_\infty \le C\,\sigma^l\,\|f\|_\infty\ .$$

Since $|\Delta_1^{R^lV_\varphi(x)}| \le \text{const}\cdot|V_\varphi(T^lx)| \le \text{const}\cdot|\varphi(T^lx)|$ with a constant that depends only on the Jacobian g of T and on the vector fields V_i, it follows that the first integral is bounded by $\frac{\text{const}}{1-\lambda_u^{-1}}\int|\varphi|\,d\mu \le \text{const}\cdot\int|\phi|\,d\tilde\mu$. (Observe that μ_h has bounded density w.r.t. μ and that μ is T-invariant.) For the second integral the corresponding bound follows from the boundedness of the A_i. This proves estimate (3.7.3) and concludes the proof of Proposition 2.3.2(d). □

Proof of Lemma 2.7.1. As usual it suffices to consider $f\in\mathcal C^1(\mathcal M,\mathbb R)$. By definition there are $g_n\in\mathcal B$ such that

$$\|f-g_n\|_w\to 0 \quad\text{and}\quad \limsup_{n\to\infty}\|g_n\|\le\|f\|^{\sim}\ .$$

Since $\mathcal C^1(\mathcal M,\mathbb R)$ is dense in $\mathcal B$, we may assume that $g_n\in\mathcal C^1(\mathcal M)$.

Our first goal is to estimate $\|f\|_s=\sup_{\varphi\in\mathcal D_\beta}\int f\varphi\,dm$ in terms of $\|g_n\|_s$. To this end we approximate a given $\varphi\in\mathcal D_\beta$ by functions $\tilde\varphi_\rho$, $\frac12\tilde\varphi_\rho\in\mathcal D_\beta$, $\rho>0$, which are such that $H^s_\gamma(\tilde\varphi_\rho)\le K_\rho$ for some $K_\rho>0$ and $|\varphi-\tilde\varphi_\rho|_\infty\le\rho^\beta$.[46] Hence

$$\begin{aligned}
\int f\varphi\,dm &= \int(f-g_n)\,\tilde\varphi_\rho\,dm+\int g_n\,\tilde\varphi_\rho\,dm+\int f\,(\varphi-\tilde\varphi_\rho)\,dm \\
&\le \|f-g_n\|_w\,K_\rho+2\|g_n\|_s+\|f\|_\infty\,\rho^\beta
\end{aligned}$$

so that, letting first $n\to\infty$ and then $\rho\to0$, we see that $\|f\|_s\le\liminf_{n\to\infty}2\|g_n\|_s$.

Next we have to estimate $\|f\|_u=\sup_{v\in\mathcal V_\beta}\int df(v)\,dm$ in terms of $\|g_n\|_u$. We start with $v\in\mathcal V_\beta\cap\mathcal C^{1+\alpha}(\mathcal M)$ so that $C_v^{-1}\operatorname{div}(v)\in\mathcal D_\gamma$ for some $C_v>0$. Then

$$\int df(v)\,dm=-\int(f-g_n)\operatorname{div}(v)\,dm+\int dg_n(v)\,dm\le\|f-g_n\|_w\,C_v+\|g_n\|_u\ ,$$

from which $\int df(v)\,dm\le\liminf_{n\to\infty}\|g_n\|_u$. If we can extend this estimate to all $v\in\mathcal V_\beta$, it follows that $\|f\|_u\le\liminf_{n\to\infty}\|g_n\|_u$ and hence $\|f\|\le\liminf_{n\to\infty}\|g_n\|\le2\|f\|^{\sim}$. As we will have to work in local coordinates we end up with a weaker estimate $\|\cdot\|\le C\|\cdot\|^{\sim}$ where C is a constant determined by a suitable smooth partition of unity, *cf.* Lemma 3.6.1 and the remarks thereafter. Since we are in dimension $d=2$, there is a local $\mathcal C^{1+\alpha}$ change of coordinates that straightens both

[45] If $N>1$ one has to use the decomposition $n=kN+j$, apply the subsequent estimate with R_N instead of R_1, and treat the remaining j iterates separately – as we have done so often in previous sections.

[46] Such $\tilde\varphi_\rho$ are constructed similarly as in Sub-lemma 3.1.3, except that there the construction was carried out for a fixed ρ. Again we can carry out the construction leafwise, which now means that we can deal with each one-dimensional stable fibre $W^s(x)$ separately. We parametrise it by arc-length and define $\tilde\varphi_\rho(x):=(2\rho)^{-1}\int_{W^s_\rho(x)}\varphi(t)\,dt$, where $W^s_\rho(x):=\{y\in W^s(x):d^s(x,y)<\rho\}$. Then $\tilde\varphi_\rho$ is differentiable along each $W^s(x)$, so $H^s_\gamma(\tilde\varphi_\rho)\le K_\rho$ for some $K_\rho>0$, $|\varphi-\tilde\varphi_\rho|_\infty\le(2\rho)^\beta\,H^s_\beta(\varphi)\le(2\rho)^\beta$, and $\frac12\tilde\varphi_\rho\in\mathcal D_\beta$.

foliations simultaneously. Therefore it suffices to show that each vector field $v(\xi,\eta) = \varphi(\xi,\eta)\binom{1}{0}$ with $\varphi \in \mathcal{D}_\beta$ can be approximated in the $L^1_m(\mathbb{T}^2)$-sense by $\mathcal{C}^{1+\alpha}$–vector fields $\tilde\varphi(\xi,\eta)\binom{1}{0}$ with $\tilde\varphi \in \mathcal{D}_\beta$. But this can be done by smoothening φ with convolution kernels in both coordinate directions independently. □

We remark that in the case of linear automorphisms there is no need to pass to local coordinates in this proof. Moreover, instead of $\frac{1}{2}\tilde\varphi \in \mathcal{D}_\beta$ the same proof yields $\tilde\varphi \in \mathcal{D}_\beta$ (in the estimate of $\|f\|_s$). Hence $\|f\| = \|f\|^{\sim}$ for all $f \in \mathcal{B}$ in this case. Furthermore, also automorphisms of higher dimensional tori can be treated in the same way, because the averaging procedures that produce $\tilde\varphi_\rho$ and the $C^{1+\alpha}$-approximations to v can can be easily adapted.

Appendix A. On splitting and foliation regularity

In this appendix we collect, for the reader's convenience, some information on the regularity properties of the invariant foliations in Anosov systems that are used in the paper.

First of all, as already mentioned, for $\mathcal{C}^2$ Anosov diffeomorphisms the invariant distributions (sometimes called *splittings*) are known to be uniformly Hölder continuous. Let us be more precise.

For each invertible linear map L let $\theta(L) := \|L^{-1}\|^{-1}$. We define $\|d^s_xT\| = \|d_xT|_{E^s_x}\|$, $\|d^u_xT\| = \|d_xT|_{E^u_x}\|$, $\theta(d^s_xT) = \theta(d_xT|_{E^s_x})$ and $\theta(d^u_xT) = \theta(d_uT|_{E^u_x})$. Then the following holds ([25], [19])[47]

(A.1)
- If there exists $\tau > 0$ such that, for each $x \in \mathcal{M}$, $\|d^s_xT\|\,\|d^u_xT\|^\tau < \theta(d^u_xT)$, then $E^s \in \mathcal{C}^\tau$.
- If there exists $\tau > 0$ such that, for each $x \in \mathcal{M}$, $\theta(d^s_xT)^\tau\,\theta(d^u_xT) > \|d^s_xT\|$, then $E^u \in \mathcal{C}^\tau$.

Moreover the Hölder continuity is uniform (that is the τ-Hölder norm of the distributions is bounded). The above conditions are often called *τ-pinching* or *bunching* conditions.

Remark A.1. *Note that since the map is Anosov and $\mathcal{C}^2$ we have $\sup\|d^s_xT\| < 1$, $\inf\theta(d^u_xT) > 1$ and $\sup\|d^u_xT\| < \infty$. Hence there always exists $\tau > 0$ for which the bunching conditions are satisfied, from which the announced Hölder continuity follows. Moreover, if the stable foliation is one dimensional then $\|d^s_xT\| = \theta(d^s_xT)$, hence the corresponding bunching condition holds for some $\tau > 1$, that is $E^s \in \mathcal{C}^{1+\alpha}$; the same consideration applies to the unstable foliation.*

The next relevant fact is that the above splittings are integrable. The integral manifolds are the stable and unstable manifolds, respectively. They form invariant continuous foliations. Each leaf of such foliations is as smooth as the map and it is tangent, at each point, to the corresponding distribution, [25]. In addition, for $\mathcal{C}^r$ maps, the $\mathcal{C}^r$ derivatives of such manifolds (viewed as graphs over the corresponding distributions) are uniformly bounded, [21]. In our case this implies that the curvature is uniformly bounded. Finally, the foliations are uniformly transversal and regular. For our purposes we need to detail these last two properties more precisely.

Lemma A.2 ([25], Proposition 6.4.13). *Denote by $W^s_\epsilon(x)$ and $W^u_\epsilon(x)$ the ϵ-ball in the stable and unstable manifolds, respectively, centred at x. There exists and $\epsilon > 0$ such that for any $x, y \in \mathcal{M}$ the intersection $W^u_\epsilon(x) \cap W^s_\epsilon(y)$ consists of at most one point that will be called $[x,y]$. In addition, there exists a $\delta > 0$ such that whenever $d(x,y) < \delta$, then $W^u_\epsilon(x) \cap W^s_\epsilon(y) \neq \emptyset$.*

In general it is also true that $[\cdot,\cdot]$ is continuous where it is defined. The above property is sometimes called *product structure*. Given $x \in \mathcal{M}$, if we consider any smooth chart of its stable and unstable manifolds and, with an harmless abuse of notations, identify them with $\mathbb{R}^{d^s}$ and $\mathbb{R}^{d^u}$, we can define, in a δ–neighbourhood of x, the map $\Gamma(y) = ([y,x],[x,y])$. This is the map we use in section 3.3 and it has the obvious property of transforming, locally, both foliations into a collection of parallel hyperplanes. By the above comments the map Γ is continuous but we need to know much more about its regularity.

In the case in which both distribution are $\mathcal{C}^{1+\alpha}$, it follows by Frobenius' theorem that the map Γ is $\mathcal{C}^{1+\alpha}$ (see Lemma 3.3.1 or section six of [35]). If the splitting is only Hölder the situation is more subtle.

Probably for historical reasons the regularity of the foliations is usually stated in terms of the corresponding *holonomy maps*. Given a foliation $\mathcal{F}$ and two transversal manifolds W_1, W_2, the

[47] In terms of the expansion rates used in the paper we have: $\|d^sT\| \sim \lambda_s$; $\|d^uT\| \sim \mu_u$; $\theta(d^sT) \sim \mu_s$ and $\theta(d^uT) \sim \lambda^u$.

associated holonomy $H_{W_1,W_2} : W_1 \to W_2$ is defined by $\{H_{W_1,W_2}(x)\} := F(x) \cap W_2$, where $F(x)$ is the fibre of the foliation that contains x. We will call *stable holonomy* any holonomy constructed via the stable foliations and *unstable holonomy* holonomies constructed by the unstable foliation. The basic result on holonomies is given by the following [35].

(A.2)
- If there exists $\tau > 0$ such that, for each $x \in \mathcal{M}$, $\|d^s_x T\| \, \|d^u_x T\|^\tau < \theta(d^u_x T)$, then the stable holonomies are uniformly $\mathcal{C}^\tau$.
- If there exists $\tau > 0$ such that, for each $x \in \mathcal{M}$, $\theta(d^s_x T)^\tau \, \theta(d^u_x T) > \|d^s_x T\|$, then the unstable holonomies are uniformly $\mathcal{C}^\tau$.

Thus the same considerations of Remark A.1 apply to the holonomies.

The relation between regularity of holonomies and regularity of the foliation (in the sense that the local foliation charts are regular) is discussed in detail in [35, section 6]. Here we restrict ourselves to what is needed in the present paper.

Consider any $\bar{x} \in \mathcal{M}$ and a sufficiently small neighbourhood U such that Lemma A.2 applies. Next choose a smooth chart and use it to identify all the tangent spaces with the tangent space of $\bar{x}$. We will not distinguish between points on the manifold and corresponding points in the chart since this causes no ambiguity, hence from now on we will identify $\bar{x}$ with 0. We can also assume, without loss of generality that the hyperplanes $\Sigma_x := \{(x, \xi)\}$ are uniformly transversal to the unstable foliation. Then, for each two points $(x, 0), (z, 0) \in U$, we can introduce the unstable holonomies $H(x, z, \cdot) : \Sigma_x \to \Sigma_z$ such that $(z, H(x, z, y))$ is the intersection of the unstable manifold of (x, y) with the hyperplane Σ_z. Now (A.2) asserts that all the maps $H(x, z, \cdot)$ are uniformly τ-Hölder.

This it is not yet enough for our purposes: we need to talk about the Jacobian of the functions Γ and H. Since we use Γ only in the case where the foliations are $\mathcal{C}^{1+\alpha}$, no difficulty arises in this case, in fact the regularity of the holonomies implies that Γ is $\mathcal{C}^{1+\alpha}$ in each variable and then it is $\mathcal{C}^{1+\alpha}$ by [23].[48] The situation for H is in principle less clear since a Hölder function may very well fail to have a Jacobian. Nevertheless, in our case H has a Jacobian and such a Jacobian is itself a Hölder function, [33].

The last information that we need is that the Jacobian of holonomies between close manifolds are close to one, they are differentiable with respect to z and the derivative is a regular test function. This last properties should be obvious to experts, yet we could not locate a clear cut reference applying directly to our case (typically unstable holonomies between stable manifolds are considered). Here is the sketch of a simple proof. Remember the change of coordinates $\Psi(\xi, \eta) = (\xi, H(0, \xi, \eta))$ introduced in section 3.6.

Lemma A.3. *There exists a constant $C > 0$ such that the Jacobian $J(x, z, \cdot)$ of $H(x, z, \cdot)$ satisfies the property*

$$|1 - J(x, z, \cdot)|_\infty \le C|x - z|.$$

In addition there exists $\beta' \in (0, \beta)$ such that, calling $A_i \circ \Psi(z, y) := \partial_{z_i} \ln J(0, z, y)$,

$$H^s_{\beta'}(A_i) < \infty \ .$$

Proof. Considering instead of T a sufficiently high iterate, we may assume that the constant $A = 1$ in (2.1.1), *i.e.* that T expands and contracts strictly along the unstable resp. stable leaves.[49] Let $\Sigma_x = \{(x, y)\}_{y \in B_r(0)}$. As already mentioned $J(x, z, \cdot)$ is just the Jacobian of the holonomy between Σ_x and Σ_z. Calling E_x the tangent space to Σ_x, and setting $y' := H(x, z, y)$, the following formula holds, [33],

$$\begin{aligned} J(x,z,y) &= \prod_{n=1}^{\infty} \frac{\det(D_{\bar{x}_n} T|_{(D_{\bar{x}_0} T^{-n}) E_x})}{\det(D_{\bar{z}_n} T|_{(D_{\bar{z}_0} T^{-n}) E_z})} \\ &= \exp\left(\sum_{n=1}^{\infty} \ln(\det(D_{\bar{x}_n} T|_{(D_{\bar{x}_0} T^{-n}) E_x})) - \ln(\det(D_{\bar{z}_n} T|_{(D_{\bar{z}_0} T^{-n}) E_z})) \right), \end{aligned} \tag{A.3}$$

where, for convenience, we have set $\bar{x}_n := T^{-n}(x, y)$ and $\bar{z}_n := T^{-n}(z, y')$.

Since (x, y) and (z, y') belong to the same unstable manifold it follows that $d(\bar{x}_n, \bar{z}_n) \le \lambda_u^{-n}$. Thus the problem of estimating the size of the Jacobian is reduced to the one of understanding

[48] In fact, the hypothesis of higher regularity is not really necessary since it is true that the Jacobian of Γ always exists and is Hölder, [6].

[49] Note just that all powers of T have the same stable and unstable foliations.

the "distance" between the tangent spaces $(D_{\bar{x}_0}T^{-n})E_x$ and $(D_{\bar{z}_0}T^{-n})E_z$. This requires a little preparation. First of all, as usual in the construction of stable and unstable manifolds, it is convenient to choose appropriate coordinates in the tangent spaces at the points $\bar{x}_n$. We will choose coordinates such that $\{(\xi, 0) : \xi \in \mathbb{R}^{d_u}\}$ is the unstable subspace and $\{(0, \eta) : \eta \in \mathbb{R}^{d_s}\}$ the stable one. Clearly, by using such adapted coordinates, the maps $D_{\bar{x}_n}T$ are all represented by matrices of the form

$$\begin{pmatrix} A_n & 0 \\ 0 & D_n \end{pmatrix}$$

where $\|A_n^{-1}\| \leq \lambda_u^{-1}$ and $\|D_n\| \leq \lambda_s$. Since we want to compare nearby matrices it is convenient to use coordinates related to (x, y) also for the quantities associated to (z, y'). To do so just use the exponential map to extend the above coordinates in the tangent space of $\bar{x}_n$ to coordinates in a neighbourhood of $\bar{x}_n$. Clearly, such neighbourhoods can be chosen all of uniform size, hence if $d(x, z)$ is small enough, $\bar{z}_n$ always belongs to the corresponding neighbourhood. Since the above coordinates are uniformly smooth and the map T is $\mathcal{C}^2$ it follows that the derivatives $D_{\bar{z}_n}T$ are all represented by matrices of the form

$$\begin{pmatrix} A'_n & B'_n \\ C'_n & D'_n \end{pmatrix} := \begin{pmatrix} A_n & 0 \\ 0 & D_n \end{pmatrix} + \Lambda_n \tag{A.4}$$

where $\|\Lambda_n\| \leq c d(\bar{x}_n, \bar{z}_n)$.

Using the above coordinates the spaces $E_x^n(y) := (D_{\bar{x}_0}T^{-n})E_x$ and $E_z^n(y) := (D_{\bar{z}_0}T^{-n})E_z$ can be represented by $d^s \times d^s$ matrices, namely $E_x^n(y) = \{(U_n(y)\eta, \eta) : \eta \in \mathbb{R}^{d_s}\}$ and $E_z^n = \{(U'_n(y)\eta, \eta) : \eta \in \mathbb{R}^{d_s}\}$. The basic estimate on the above subspaces is given in the next lemma the proof of which is postponed to the end of the appendix.

Sub-Lemma A.4. *There exists a constant $c > 0$ such that*

$$\|U_n(y) - U'_n(y)\| \leq c\lambda_u^{-n}|x - z|$$

By this sub-lemma it follows that there exists a constant $c_1 > 0$ such that

$$\left|\ln[\det(D_{\bar{x}_n}T|_{E_x^n(y)})] - \ln[\det(D_{\bar{z}_n}T|_{E_z^n(y)})]\right| \leq c_1\lambda_u^{-n}|x - z|,$$

from which the first statement of Lemma A.3 follows.

In addition, since each term of the sum in (A.3) is clearly $\mathcal{C}^1$, the above estimate on the Lipschitz constant is indeed an estimate on the corresponding derivative. In order to show β'-Hölder continuity of this derivative we note first the formula

$$\begin{aligned} \partial_{z_i} \ln J(x, z, y) &= -\sum_{n=1}^{\infty} \partial_\Lambda \ln(\det(\Lambda))|_{\Lambda = D_{\bar{z}_n}T|_{E_z^n(y)}} \cdot \partial_{z_i}\left(D_{\bar{z}_n}T|_{E_z^n(y)}\right) \\ &= -\sum_{n=1}^{\infty} \partial_\Lambda \ln(\det(\Lambda))|_{\Lambda = D_{\bar{z}_n}T|_{E_z^n(y)}} \cdot \sum_j \Bigg[(\partial_{\xi_j} D_\xi T|_{E_z^n(y)})|_{\xi = \bar{z}_n} \frac{\partial z_j^n}{\partial z_i} \\ &\qquad + \partial_V(D_{\bar{z}_n}T|_{E_V})|_{V = U_n(y)} \cdot \frac{\partial U'_n(y)}{\partial z_i}\Bigg] \end{aligned} \tag{A.5}$$

where $E_V := \{(V\eta, \eta) : \eta \in \mathbb{R}^{d_s}\}$. Since the maps $D_{\bar{z}_n}T|_{E_z^n(y)}$ are all uniformly invertible (because of hyperbolicity), all the derivatives appearing in the above formula are well defined. Varying z the points $\bar{z}_0$ move along the unstable manifold, by definition. It follows that

$$\left|\frac{\partial z_j^n}{\partial z_i}\right| \leq c\lambda_u^{-n}$$

and, together with Sub-Lemma A.4, this implies that the above series converges exponentially fast.

Next, let $x = 0$, $\bar{z} := \Psi(z, y)$, $\bar{\zeta} := \Psi(\xi, \eta)$ and, as before, $y' := H(0, z, y)$, $\eta' := H(0, \xi, \eta)$. Accordingly

$$(|z - \xi| + |y' - \eta'|) \leq c\, d(\bar{z}, \bar{\zeta}) \ .$$

Now each term in (A.5) is clearly Lipschitz with respect to the variables z, y' with Lipschitz constant bounded by μ_s^{-n}. It follows[50]

$$
\begin{aligned}
|\partial_{z_i} \ln J(0,z,y) - \partial_{z_i} \ln J(0,\xi,\eta)| &\leq \sum_{i=1}^{\bar{n}} C\mu_s^{-i}(|y'-\eta'|+|z-\xi|) + C\lambda_u^{-\bar{n}} \\
&\leq c\, d(\bar{z},\bar{\zeta})^{\frac{\ln\lambda_u}{\ln \mu_s^{-1}\lambda_u}} =: c\, d(\bar{z},\bar{\zeta})^{\beta'},
\end{aligned}
$$

where we have chosen $\bar{n}$ appropriately. □

Proof of Sub-Lemma A.4. To start we need to obtain a recursive equation for the U_n. Since

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} U\eta \\ \eta \end{pmatrix} = \begin{pmatrix} (AU+B)\eta \\ (CU+D)\eta \end{pmatrix}
$$

it follows that the space $\{(U\eta,\eta) : \eta \in \mathbb{R}^{d_s}\}$ is sent into the space $\{(\tilde{U}\eta,\eta) : \eta \in \mathbb{R}^{d_s}\}$ with $\tilde{U} := (AU+B)(CU+D)^{-1}$, *i.e.* $U = (A-\tilde{U}C)^{-1}(\tilde{U}D - B)$. Accordingly,

$$
\begin{aligned}
U_n &= A_n^{-1} U_{n-1} D_n \\
U'_n &= (A'_n - U'_{n-1}C'_n)^{-1}(U'_{n-1}D'_n - B'_n)
\end{aligned}
$$

where $U_0 = U'_0 = 0$. From this follows immediately (recall that we assumed $A = 1$)

$$
\|U_n\| \leq \lambda_u^{-n}\lambda_s^n,
$$

and, observing (A.4),

$$
\begin{aligned}
\|U_n - U'_n\| &\leq \|A_n^{-1}(U_{n-1} - U'_{n-1})D_n\| + c\, d(\bar{x}_n, \bar{z}_n) \\
&\leq \lambda_u^{-1}\lambda_s \|U_{n-1} - U'_{n-1}\| + c\, d(\bar{x}_n,\bar{z}_n) \\
&\leq c\sum_{i=0}^{n-1} (\lambda_u^{-1}\lambda_s)^i\, d(\bar{x}_{n-i},\, \bar{z}_{n-i}) \\
&\leq \text{const}\cdot\lambda_u^{-n}|x-z|.
\end{aligned}
$$

□

References

[1] V.I. Bakhtin, *A direct method of constructing an invariant measure on hyperbolic attractor*, Izv. AN SSSR, ser. mat., **56**:5 (1992), 934-957.

[2] V.Baladi, *Positive Transfer Operators and Decay of Correlations*, Advanced Series in Nonlinear Dynamics, **16**, World Scientific, Singapore (2000).

[3] V. Baladi, L.-S. Young, *On the spectra of randomly perturbed expanding maps*, Comm. Math. Phys. **156**:2 (1993), 355-385; **166**:1 (1994), 219–220.

[4] M. Blank, G. Keller, *Random perturbations of chaotic dynamical systems: stability of the spectrum*, Nonlinearity, **11** (1998), 1351-1364.

[5] M. Blank, *Discreteness and continuity in problems of chaotic dynamics*, Amer. Math. Soc., Providence, Rhode Island (1997).

[6] X.Bressaud, C.Liverani, *Anosov diffeomorphisms and coupling*, to appear in Ergodic Theory and Dynamical Systems.

[7] W.M.Boothby, *An Introduction to Differential manifolds and Riemannian Geometry*, Academic Press, New York, San Francisco London (1975).

[8] F.Brini, S.Siboni, G.Turchetti, S.Vaienti, *Decay of correlations for the automorphisms of the torus* $\mathbb{T}^2$, Nonlinearity **10** (1997), 1257-1268.

[9] J.Buzzi, G.Keller, *Zeta functions and Transfer operators for multidimensional piecewise affine and expanding maps,* to appear in Ergod. Th.& Dynam. Sys.

[10] I.P.Cornfeld, S.V.Fomin, Ya.G.Sinai, *Ergodic Theory*, Grundlehren der Mathem. Wiss. 245, Springer-Verlag New York, Heidelberg, Berlin (1982).

[11] M.Dellnitz, O.Junge, *On the approximation of complicated dynamical behavior,* SIAM J. Numerical Analysis **36** (1999), 491-515.

[12] M.Dellnitz, O.Junge, *Set Oriented Numerical Methods for Dynamical Systems,* B. Fiedler, G. Iooss and N. Kopell (eds.): Handbook of Dynamical Systems III: Towards Applications, World Scientific (2001), to appear.

[13] M.P. Do Carmo, *Riemannian Geometry,* Birkhäuser, Boston (1992).

[14] N.Dunford, J.T.Schwartz, *Linear Operators, Part I*, Wiley, New York (1957).

[15] L.E.Evans, R.F.Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, New York, London, Tokyo (1992).

[16] H.Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin, Heidelberg (1969).

[50] The following estimate can be improved, but since we are not able to strengthen it sufficiently to get $\beta' = \tau'$ in the lemma, we content ourselves with what follows.

[17] G.Froyland, *Computer-assisted bounds for the rate of decay of correlations,* Commun. Math. Phys. **189** (1997), 237-257.

[18] G.Froyland, *Extracting dynamical behaviour via Markov models,* Alistair Mees (eds.): "Nonlinear Dynamics and Statistics: Proceedings, Newton Institute, Cambridge, 1998", pp. 283-324, Birkhäuser (2000).

[19] B.Hasselblatt, *Regularity of the Anosov splitting and of horospheric foliations*, Ergod. Th.& Dynam. Sys. **14** (1994), 645-666.

[20] H.Hennion, *Sur un théorème spectral et son application aux noyaux Lipchitziens*, Proceedings of the American Mathematical Society, **118** (1993), 627–634.

[21] M.Hirsch, C.Pugh, M.Shub, *Invariant Manifolds*, Lecture Notes in Math. **583** (1977).

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

[23] J.-L.Journé, *A regularity lemma for functions of several variables*, Rev. Mat. Iberoamericana **4** (1988) 187-193.

[24] O.Junge, *Mengenorientierte Methoden zur numerischen Analyse dynamischer Systeme,* PhD thesis, University of Paderborn, Paderborn (1999).

[25] A.Katok, B.Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, **54**, G.-C.Rota editor, Cambridge University Press, Cambridge (1995).

[26] G.Keller, *Stochastic stability in some chaotic dynamical systems,* Monatsh. Math. **94** (1982), 313-333.

[27] G.Keller, C.Liverani, *Stability of the spectrum for transfer operators*, Annali della Scuola Normale Superiore di Pisa, Scienze Fisiche e Matematiche, (4) **XXVIII** (1999), 141–152.

[28] Yu.Kifer, *Random Perturbations of Dynamical Systems*, Progress in Probability and Statistics 16, Birkhäuser, Boston (1988).

[29] Yu.Kifer, *Computations in dynamical systems via random perturbations,* Discr. Cont. Dynam. Sys. **3** (1997), 457-476.

[30] A.Yu.Kitaev, *Fredholm determinants for hyperbolic diffeomorphisms of finite smoothness*, Nonlinearity **12** (1999), 141-179.

[31] C.Liverani, *Decay of Correlations*, Annals of Mathematics, **142** (1995), 239-301.

[32] C.Liverani, *Rigorous numerical investigation of the statistical properties of piecewise expanding maps–A feasibility study*, to appear in Nonlinearity.

[33] R.Mañe, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, Berlin Heidelberg (1987).

[34] R.D.Nussbaum, *The radius of the essential spectrum,* Duke Math. J. **37** (1970), 473-478.

[35] C.Pugh, M.Shub, A.Wilkinson, *Hölder Foliations*, Duke Math. Journal **86** (1997), 517-546.

[36] W. Rudin, *Real and Complex Analysis*, Mc Graw-Hill series in higher Mathematics, Second Edition (1974).

[37] H.H.Rugh, *Generalized Fredholm determinants and Selberg zeta functions for Axiom A dynamical systems*, Ergod. Th.& Dynam. Sys. **16** (1996), 805-819.

Michael Blank, Institute for Information Transmission Problems, B.Karetnij Per. 19, Moscow, Russia

E-mail address: blank@obs-nice.fr

Gerhard Keller, Mathematisches Institut, Universität Erlangen, Bismarckstr. $1\frac{1}{2}$, 91054 Erlangen, Germany

E-mail address: keller@mi.uni-erlangen.de

Carlangelo Liverani, Dipartimento di Matematica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italy

E-mail address: liverani@mat.uniroma2.it