Universidade Federal Fluminense

Entropy Rigidity for Surface Anosov Diffeomorphisms

Vítor Gomes Ferreira

Niterói

Setembro 2025

Entropy Rigidity for Surface Anosov Diffeomorphisms

Vítor Gomes Ferreira

Dissertação submetida ao Programa de Pós-Graduação em Matemática da Universidade Federal Fluminense como requisito parcial para a obtenção do grau de Mestre em Matemática.

Orientador:
Prof. Bruno Santiago (UFF)

Niterói
Setembro 2025

Ficha catalográfica automática - SDC/BIME
Gerada com informações fornecidas pelo autor

F383e Ferreira, Vítor Gomes
Entropy Rigidity for Surface Anosov Diffeomorphisms / Vítor Gomes Ferreira. - 2025.
88 f.

Orientador: Bruno Rodrigues Santiago.
Dissertação (mestrado)-Universidade Federal Fluminense, Instituto de Matemática e Estatística, Niterói, 2025.

1. Dinâmica Hiperbólica. 2. Teoria Ergódica. 3. Rigidez. 4. Classificação. 5. Produção intelectual. I. Santiago, Bruno Rodrigues, orientador. II. Universidade Federal Fluminense. Instituto de Matemática e Estatística. III. Título.

CDD - XXX

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

Dissertação de Mestrado da Universidade Federal Fluminense

por

Vítor Gomes Ferreira

apresentada ao Programa de Pós-Graduação em Matemática como requesito parcial para a obtenção do grau de

Mestre em Matemática

Título da tese:

Entropy Rigidity for Surface Anosov Diffeomorphisms

Defendida publicamente em 04 de setembro de 2025.

Diante da banca examinadora composta por:

Bruno Santiago	UFF	Orientador
Jiagang Yang	UFF	Examinador
Sébastien Alvarez	UDELAR	Examinador
Ali Tahzibi	USP	Examinador

DECLARAÇÃO DE CIÊNCIA E CONCORDÂNCIA DO ORIENTADOR

Autor da Dissertação: Vítor Gomes Ferreira
Data da defesa: 04/09/2025
Orientador: Bruno Santiago

Para os devidos fins, declaro **estar ciente** do conteúdo desta **versão corrigida** elaborada em atenção às sugestões dos membros da banca examinadora na sessão de defesa do trabalho, manifestando-me **favoravelmente** ao seu encaminhamento e publicação no **Repositório Institucional da UFF**.

Niterói, 30/09/2025.

gov.br Documento assinado digitalmente
BRUNO RODRIGUES SANTIAGO
Data: 30/09/2025 13:40:57-0300
Verifique em https://validar.iti.gov.br

Nome do Orientador

AGRADECIMENTOS

In first place, I thank God for giving me the capacity to be where I am. I am also very grateful to my parents, for they always encouraged me.

Throughout my path I have met many wonderful teachers who guided me here. Special thanks to Mitchael Martelo, Juliana Coelho, Ralph Teixeira and Pablo Guarino who, with a few spot-on advice, made a great impact on me.

I was also very lucky to have met Caio Caetano in my college freshman year. It was his figure that made me dive into this amazing world of math.

I thank each member of the examining board: Jiagang Yang, Sébastien Alvarez, Ali Tahzibi and Bruno Santiago for accepting this task. Special thanks for Bruno Santiago, for he is a brilliant advisor. His enthusiasm is contagious, which fits perfectly with his attention with didactics. Also, his patience and guidance were key for my development.

And, of most importance, I thank Keila Nunes. For I love her and she is the greatest source of happiness of my days.

At last, I thank CAPES for the financial support via a master's scholarship of number 88887.897113/2023-00.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

Quando as medidas físicas (SRB) são de máxima entropia?

Quando dois sistemas são conjugados de forma suave?

O objetivo deste texto é apresentar um resultado elegante (clássico) de rigidez para difeomorfismos de Anosov em dimensão dois que conecta essas duas questões: precisamente, nós mostramos que a medida SRB de f tem máxima entropia se e só se f for conjugado a um mapa linear de forma suave ao longo da folheação instável. A prova clássica desse resultado (veja cor. 20.4.5 de [KH95]) depende do formalismo termodinâmico: os estados de equilíbrio associados a potenciais diferentes são mutualmente singulares, a menos que os potenciais sejam co-homólogos. Outra maneira de se provar é por meio de uma rigidez dos expoentes de Lyapunov: se o push-forward da medida SRB pela conjugação for uma medida com o mesmo expoente de Lyapunov, então a conjugação tem que ser suave ao longo de folhas instáveis. No texto, iremos apresentar uma demonstração geométrica autocontida desse fato, baseada em um estudo detalhado das medidas condicionais em folhas instáveis da SRB (leafwise measures) e das medidas de máxima entropia (família de Margulis).

Palavras-chave: Sistemas Dinâmicos; Teoria Ergódica; Rigidez; Entropia.

ABSTRACT

When physical SRB measures are of maximal entropy?
When two given systems are smoothly conjugated?

The goal of this text is to outline a beautiful (classical) rigidity result for Anosov systems in dimension two that connects these two questions: Precisely, we show that the SRB measure of f maximizes the entropy if and only if f is smoothly conjugated to its linearization along unstable leaves. The classical proof of this result (see cor. 20.4.5 of [KH95]) relies on thermodynamical formalism: the equilibrium states associated with different potentials must differ, unless the potentials are co-homologous. Another approach can be made via Lyapunov exponent rigidity: if the push-forward of the SRB measure by the conjugacy is a measure with the same Lyapunov exponent, then the conjugacy must be smooth along the leaves. In the text, we shall present a self-contained geometrical proof based on a detailed study of conditional measures along unstable leaves for the SRB (leafwise measures) and for the MME (Margulis family).

Keywords: Dynamical Systems; Ergodic Theory; Rigidity; Entropy.

Contents

CHAPTER 1

Introduction

In this text we will study the dynamics of Anosov diffeomorphisms which are a kind of system described by a simple axiom: its derivative preserves a pair of complementary directions, contracting one and expanding the other. One of the major questions about them is whether you can classify up to conjugacies all of them, i.e. whether you can find an explicit collection of Anosov diffeomorphisms whose all others are conjugated to it. For low dimensions this question was positively answered up to the topological level by Franks-Newhouse (see theorem. 2.3.3).

Since the topological classification is already done, it incites us to explore the class of smooth conjugacies. However, it doesn't take long to notice that this type of conjugacy is very sensible: suppose f and g are two Anosov systems conjugated by a smooth diffeomorphism h, that is $h \circ f = g \circ h$. Then, if you take any periodic point p for f with period n, you can differentiate the conjugacy to obtain

$$dh \cdot d_p f^n = d_{h(p)} g^n \cdot dh$$

Thus, the linear maps $d_p f^n$ and $d_{h(p)} g^n$ are also conjugated, which implies that they have the same spectrum (we say that they have the same periodic data). It is a notorious fact that Anosov diffeomorphisms in low dimension always have a dense set of periodic points; thus this condition on the derivatives must hold in a dense set. This condition is, of course, very fragile: any small perturbation, as smooth as you want, is capable of changing the spectrum at a point, which breaks the equality. Hence, there is no hope of classifying them up to smooth conjugacies.

There is, however, a peculiar phenomena that arises in this situation. A smooth conjugation is so rigid that it, in some sense, can characterize excessively specific properties. I.e. requiring a system to satisfy a very restrictive property may cause all the systems that satisfy it to be smoothly conjugated. This is nothing but a mantra, however, a fruitful one.

An example of such phenomena already appeared in the discursion given above: not just smooth conjugacy implies equality of periodic data, but equality of periodic data also implies smooth conjugacy! (see appendix A).

This equality of periodic data is explicitly gross and it makes sense to imply rigidity. What we will show is something more subtle: Start with a conservative linear Anosov diffeomorphism. For this system, its invariant volume measure coincides with its maximal entropy measure. If you consider a conservative perturbation of this system, it still preserves the volume, however its measure of maximal entropy may not be equal it. What happens if you require it to still be the measure of maximal entropy?

We will see that this condition implies rigidity. In fact, we'll not be restricted to the conservative case: When the system is not conservative, there is a type of measure (the SRB measure) that naturally takes its place. The main theorem that we will prove is the following

Theorem 1.0.1 (Main Theorem)**.** *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is an Anosov diffeomorphism of class C^2 such that its SRB measure μ is of maximal entropy, then f is conjugated to a linear toral automorphism and the conjugacy h is $C^{1+\alpha}$ when restricted to any unstable leaf.*

And as a corollary we will obtain that

Corollary 1.0.2 (The Conservative Case)**.** *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is a conservative Anosov diffeomorphism such that its invariant volume measure maximizes the entropy, then it is $C^{1+\alpha}$ conjugated to a linear toral automorphism.*

The classical proof of these facts uses the thermodynamical formalism: the measure of maximum entropy and the SRB measure are both equilibrium states of different potentials (the identically null potential and the logarithm of the unstable jacobian respectively).

When the equilibrium states are equal, the potentials must be co-homologous, which implies that the logarithm of the unstable jacobian is co-homologous to a constant. From it, you can deduce that the conjugacy between f and its linear counterpart, which sends equilibrium states into equilibrium states, must be absolutely continuous along the unstable foliation, and this forces the smoothness.

A more geometrical approach, which better aligns with this work, is with the use of Ledrappier-Young's theory: By the entropy formula, the Lyapunov exponent of f must be equal to the entropy of the SRB measure. Thus, since topological entropy is invariant along a class of conjugation, if the SRB is of maximal entropy we obtain that its Lyapunov exponent coincides with the logarithm of the unstable eigenvalue of the linear model.

This equality involving the Lyapunov exponents in turn implies in the smoothness of the conjugacy (see, for example [SY19]).

In this work we present a complete self-contained proof of Theorem 1.0.1 and corollary 1.0.2.

Our main focus will lie on the geometrical structure of the conditional measures.

We will present the construction of the Leafwise measures for the SRB measure. These are a family of locally finite borelian measures defined on each unstable leaf that coincides (up to normalization) with the conditional measures. The construction we show in chapter 4 closely follows the work done in [Alv+24].

We'll also present the construction of the affine parameters, also known as normal forms, which are a type of non-stationary linearization. With this affine structure, we will see that the leafwise measures can be identified with the usual Lebesgue measure on the real line.

Also, in chapter 5, we will present the Margulis family.

The main difference between the Margulis family and the Leafwise family is in how it is renormalized under the dynamics: the Margulis family scales with the topological entropy while the Leafwise family scales with the unstable jacobian.

When the SRB measure is of maximal entropy, this allows us to re-obtain Ledrappier-Young's entropy formula by a direct computation.

We then proceed to present a self-contained proof, for our particular context, that the equality between the Lyapunov exponents of the SRB measure with those of the Lebesgue measure for a linear map implies that the conjugacy is Lipschitz continuous and, furthermore, that it is $C^{1+\alpha}$.

It is important to emphasize that there are many similar, but different, ways to achieve the same conclusion that we obtain here. However, in this work, the main purpose is to exhaustively explore the following heuristics: in low dimensions, the preservation of local asymptotic quantities by the conjugation implies in rigidity.

This same heuristics is also present in appendix A, where we show that the coincidence of periodic data, as discoursed before, also implies in smoothness for the conjugacy.

CHAPTER **2**

Anosov Diffeomorphisms

In this chapter I will present some fine properties of Anosov diffeomorphisms. Those form a class of systems that are of great importance to the study of dynamical systems. Their dynamics is very rich and are an important example of a structurally stable system.

In the first two sections we define them and establish some basic results about them. In the third section we discourse about their stability and classification.

2.1 Invariant Manifolds

Definition 2.1.1. A diffeomorphism $f : M \to M$ on a Riemann manifold M is said to be Anosov if there are constants $0 < C$, $\lambda > 1$ and an invariant decomposition $TM = E^s \oplus E^u$ of the tangent bundle such that for all $p \in M$

$$|df^n v| \leq C\lambda^{-n}|v| \quad , \quad \text{for all } v \in E^s(p)$$
$$|df^{-n} u| \leq C\lambda^{-n}|v| \quad , \quad \text{for all } u \in E^u(p)$$

The distributions E^s and E^u are called stable and unstable distributions respectively. The constant λ is a rough estimate on the rate of contraction in the stable direction and expansion in the unstable direction. The constant C is to account for a correction in the firsts iterates. Often times one ignores this constant by making it equal to 1. This can be done if you (as we) is not interested in the fine properties of a Riemann manifold, because it is always possible to redefine the Riemann metric so that it happens:

Proposition 2.1.2. *If $f : M \to M$ is Anosov and M is compact, then there is a Riemann metric $|\cdot|_f$ on M such that the constant C in the definition 2.1.1 is 1.*

Using this adapted norm, an Anosov diffeomorphism is an instantaneous contraction on the stable direction E^s and expansion in the unstable direction E^u, i.e. if $S \subseteq M$ is any sufficiently small submanifold of M tangent to the stable (resp. unstable) distribution E^s (resp. E^u) of a point $p \in M$, then $d(f(p), f(q)) < \tilde{\lambda^{-1}} d(p, q)$ (resp. $d(f(p), f(q)) > \tilde{\lambda^{-1}} d(p, q)$), where $\tilde{\lambda}$ is as close to λ as long as you make $q \to p$ in S.

This is nothing but saying that a map can be well approximated by its own derivative. However, for Anosov diffeomorphisms we require these stable and unstable directions to be defined everywhere and even more they are invariant. If we could integrate these distributions, the resulting manifolds would be invariant and have contraction\expansion properties. One may ask themselves if those invariant manifolds actually capture all the asymptotical behavior of a point or not. With this question I mean the following

Definition 2.1.3. Given a map $f : M \to M$ and a point $p \in M$, the stable and unstable manifold of p are the sets

$$W^s(p) \overset{\text{def.}}{=} \{q \in M | d(f^n q, f^n p) \to 0 \text{ as } n \to +\infty\}$$

$$W^s(p) \overset{\text{def.}}{=} \{q \in M | d(f^n q, f^n p) \to 0 \text{ as } n \to -\infty\}$$

respectively.

The sets $W^s(p)$ and $W^u(p)$ capture all asymptotical behavior at p in the sense that every point whose orbit follows the trajectory of p is at one or another. At first, the name 'manifold' above is only formal, for the sets $W^\sigma(p)$ are not necessarily manifolds. But only at first, because those sets are actually immersed manifolds. It follows from the well know Stable Manifold Theorem:

Theorem 2.1.4. *If $f : M \to M$ is an Anosov Diffeomorphism of class C^r, $r \geq 1$ with invariant decomposition $TM = E^s \oplus E^u$, then there exists an $\varepsilon_0 > 0$ such that for every $0 < \varepsilon < \varepsilon_0$, the set*

$$W^s_\varepsilon(p) \overset{\text{def.}}{=} \{q \in M | d(f^n p, f^n q) < \varepsilon, \forall n \geq 0\}$$

is a C^r embedded disk that is and tangent to E^s and of the same dimension of it. Moreover, the map $p \to W^s_\varepsilon(p)$ is continuous in the C^r topology.

It is clear from its definition that if ε satisfies the Theorem above then

$$W^s(p) = \bigcup_{n \in \mathbb{N}} f^{-n}(W^s_\varepsilon(f^n p))$$

so that $W^s(p)$ is an immersed manifold. To avoid the use of unnecessary parameters, we may write this local stable manifold $W^s_\varepsilon(p)$ as $W^s_{loc}(p)$ without expliciting ε. Also, if $\tilde{W}^s_{loc}(p)$ is a local stable manifold for f^{-1}, we may as well define the local unstable manifold of f as $W^u_{loc}(p) = \tilde{W}^s_{loc}(p)$. We obtain $W^u(p)$ in the same manner.

Remark 2.1.5. In particular, since these manifolds are tangent to the distributions E^σ, the convergence in definition 2.1.3 is exponentially fast.

This Theorem is classical, and the general proof can be easily found in many textbooks; see [Wen16]. Since we are interested in surface diffeomorphisms I will present the particular case for a fixed point in $\dim M = 2$:

Theorem 2.1.6. *If $f : M \to M$ is C^r and $p \in M$ is a fixed point such that the eigen values of $d_p f$ are $\lambda_1 < 1 < \lambda_2$ then there exists a C^1 embedded curve $W^s_{loc}(p)$ satisfying*

1. **Invariance** : $f(W^s_{loc}(p)) \subseteq W^s_{loc}(p)$.
2. **Convergence** : *For all $q \in W^s_{loc}(p)$ we have that $f^n(q)$ converges exponentially fast to p.*
3. **Uniqueness** : *If $q \in M$ is so that $f^n(q)$ converges exponentially fast to p then there exists $N \in \mathbb{N}$ such that $f^N(q) \in W^s_{loc}(p)$.*
4. **Tangency** : *$W^s_{loc}(p)$ is tangent to the eigen space of $d_p f$ associated to λ_1.*

For the proof we will follow section 3.1 of Rafael Potrie's lecture notes [Pot16]. The proof is based on the following steps:

1. We construct a good local representation of f.

2. Considering a certain space of curves, we find one that is invariant by this local representation.

3. We show that this curve restricted to the origin is the one satisfying the conclusions of the Theorem.

Local Expression

Consider a local chart $\Phi : U \to \mathbb{R}^2$ around p with $\Phi(p) = 0$. Since p is a fixed point and f is a diffeomorphism, we have that $f^{-1}(U)$ is a neighborhood of p. Let $V \overset{\text{def.}}{=} U \cap f^{-1}(U)$ and notice that

$$\hat{f} \overset{\text{def.}}{=} \Phi \circ f \circ \Phi^{-1}(V) \to \mathbb{R}^2$$

is well defined. The benefit of this expression is that in this way, $d_{\vec{0}}\hat{f}$ shares the same eigenvalues $0 < \lambda_1 < 1 < \lambda_2$ of $d_p f$. Also, up to linear change of coordinates, we may suppose that the eigenspaces of $d_{\vec{0}}\hat{f}$ are the $\hat{x}$ and $\hat{y}$ axes. In particular

$$d_{\vec{0}}\hat{f} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

By Taylor's Theorem, we can write

$$\hat{f}(x, y) = (\lambda_1 x + \alpha(x, y), \lambda_2 y + \beta(x, y))$$

Where α and β are of class C^1, $\alpha(\vec{0}) = \beta(\vec{0}) = 0$ and $\nabla\alpha(\vec{0}) = \nabla\beta(\vec{0}) = \vec{0}$.

By continuity, given $\varepsilon > 0$, we can take a $\delta > 0$ such that both $\|\alpha|_{B(0,\delta)}\|_{C^1}$ and $\|\beta|_{B(0,\delta)}\|_{C^1} < \varepsilon$ are bounded by ε. Let us now consider a smooth bump function η satisfying

· $\eta(x, y) = 1$, if $\|(x, y)\| < \frac{\delta}{2}$

· $\eta(x, y) = 0$, if $\|(x, y)\| > \delta$

· $\|\nabla\eta(x, y)\| < \frac{4}{\delta}$

Using this function, we can extend $\hat{f}$ to a map in $\mathbb{R}^2$ by defining

$$\overline{f} \overset{\text{def.}}{=} \eta\hat{f} + (1 - \eta)d_{\vec{0}}\hat{f} : \mathbb{R}^2 \to \mathbb{R}^2$$

Near the origin $\overline{f}$ coincides with $\hat{f}$, and far from the origin $\overline{f}$ coincides with $d_{\vec{0}}\hat{f}$. However, ideally, $\hat{f}$ and $d_{\vec{0}}\hat{f}$ are very similar near the origin. Hence, $\overline{f}$ turns out to be very close to the linear map $d_{\vec{0}}\hat{f}$ everywhere. That is, $\overline{f} = d_{\vec{0}}\hat{f} + \vec{r}$ where $\|r\|_{C^1}$ is very small. Explicitly, in coordinates:

$$\overline{f}(x, y) = (\lambda_1 x + \eta(x, y)\alpha(x, y), \lambda_2 y + \eta(x, y)\beta(x, y))$$

Lets write $\overline{\alpha} = \eta\alpha$ and $\overline{\beta} = \eta\beta$.

Claim 2.1.6.1. Given $\overline{\varepsilon} > 0$ it is possible to choose $\varepsilon > 0$ such that

$$\begin{cases} |\overline{\alpha}(p_1) - \overline{\alpha}(p_2)| < \overline{\varepsilon} \min\{\delta, \|p_1 - p_2\|\} \\ |\overline{\beta}(p_1) - \overline{\beta}(p_2)| < \overline{\varepsilon} \min\{\delta, \|p_1 - p_2\|\} \end{cases}$$

Proof. Let $p_1, p_2 \in \mathbb{R}^2$. Notice that $|\eta(p_i)| \leq 1$ and that if $\|p_i\| \leq \delta$ we have

$$|\alpha(p_i)| < \|\alpha|_{B(0,\delta)}\|_{C^1}\|p_i\| < \varepsilon\delta$$

And if $\|p_i\| \geq \delta$ then

$$|\eta(p_i)\alpha(p_i)| = 0 < \varepsilon\delta$$

Thus

$$\begin{aligned} |\overline{\alpha}(p_1) - \overline{\alpha}(p_2)| &\leq |\overline{\alpha}(p_1)| + |\overline{\alpha}(p_2)| \\ &= |\eta(p_1)\alpha(p_1)| + |\eta(p_2)\alpha(p_2)| \\ &= |\eta(p_1)||\alpha(p_1)| + |\eta(p_2)||\alpha(p_2)| \\ &< \varepsilon\delta + \varepsilon\delta \\ &= 2\varepsilon\delta \end{aligned}$$

In another way, if $\|p_1\| \geq 0$ and $\|p_2\| \geq 0$ we have

$$|\overline{\alpha}(p_1) - \overline{\alpha}(p_2)| = 0$$

But if atleast one $p_i < \delta$ (say $p_1 < \delta$), we have

$$\begin{aligned} |\overline{\alpha}(p_1) - \overline{\alpha}(p_2)| &= |\eta(p_1)\alpha(p_1) - \eta(p_2)\alpha(p_2)| \\ &\leq |\eta(p_1)\alpha(p_1) - \eta(p_1)\alpha(p_2)| + |\eta(p_1)\alpha(p_2) - \eta(p_2)\alpha(p_2)| \\ &= |\eta(p_1)||\alpha(p_1) - \alpha(p_2)| + |\alpha(p_2)||\eta(p_1) - \eta(p_2)| \\ &< \varepsilon\|p_1 - p_2\| + \varepsilon\delta\tfrac{4}{\delta}\|p_1 - p_2\| \\ &= 5\varepsilon\|p_1 - p_2\| \end{aligned}$$

The same results holds for β. Finally, taking $\varepsilon < \frac{\overline{\varepsilon}}{5}$ we finish the claim. □

In particular, this claim says that $\|\overline{\alpha}\|_{C^1} < \overline{\varepsilon}$ and $\|\overline{\beta}\|_{C^1} < \overline{\varepsilon}$. Thus, in fact

$$\overline{f} = d_{\vec{0}}\hat{f} + \vec{r}$$

where $\|\vec{r}\|_{C^1}$ is as small as we want. By the inverse function theorem, $\overline{f}$ is a diffeomorphism. Since $\overline{f}$ is just $d_{\vec{0}}\hat{f}$ plus a very small perturbation, we would like to say that $\overline{f}^{-1}$ is $[d_{\vec{0}}\hat{f}]^{-1}$ plus something tiny.

Thankfully to this next technical lemma, we can actually say it

Lemma 2.1.7. *Let $A : E \to E$ be a continuous linear isomorphism and $\Delta : E \to E$ be of class C^1. Then, given $\varepsilon_0 > 0$ there exists a $\delta_0 > 0$ such that is $\|\Delta\|_{C^1} < \delta_0$ then $A + \Delta$ is invertible and*

$$(A + \Delta)^{-1} = A^{-1} + \varepsilon$$

where $\varepsilon : E \to E$ and $\|\varepsilon\|_{C^1} < \varepsilon_0$.

Proof. The fact that for some $\delta_0 > 0$ the map $A + \Delta$ is invertible with C^1 inverse follows directly from the inverse function theorem. Given that we know that, the only way that ε could be defined is by

$$\varepsilon \overset{\text{def.}}{=} (A + \Delta)^{-1} - A^{-1}$$

All we have to do is to verify that for any choice of $\varepsilon_0 > 0$ we can make $\|\varepsilon\|_{C^1} < \varepsilon_0$.

Let $x \in E$ and put $y \stackrel{\text{def.}}{=} (A+\Delta)(x)$. We have

$$\begin{aligned} \|d_y\epsilon\| &= \|d_y(A+\Delta)^{-1} - d_yA^{-1}\| \\ &= \|\, [d_x(A+\Delta)]^{-1} - [d_xA]^{-1}\, \| \\ &= \|\operatorname{Inv}(A+d_x\Delta) - \operatorname{Inv}(A)\| \end{aligned}$$

Since the operator $\operatorname{Inv} : \mathcal{L}(E) \to \mathcal{L}(\mathcal{E})$ is a continuous function, there exists a $\delta_1 > 0$ such that $\|(A+d_x\Delta) - A\| < \delta_1$ implies

$$\|\operatorname{Inv}(A+d_x\Delta) - \operatorname{Inv}(A)\| < \varepsilon_0$$

Taking $\delta < \delta_1$ we obtain

$$\|d_y\epsilon\| < \varepsilon_0\,,\ \forall y \in E$$

Also, by continuity of A^{-1}, there exists some $\delta_2 > 0$ such that $\|y' - y\| < \delta_2$ implies that

$$\|A^{-1}y' - A^{-1}y\| < \varepsilon_0$$

For any $y \in E$, we can find $x \in E$ such that $y = (A+\Delta)(x)$. For this x let $y' \stackrel{\text{def.}}{=} Ax$ and notice that $\|y' - y\| = \|\Delta(x)\| < \delta_0$. Thus if we choose $\delta_0 < \delta_2$ we have $\|A^{-1}y' - A^{-1}y\| < \varepsilon_0$, and hence

$$\|\epsilon(y)\| = \|(A+\Delta)^{-1}(y) - A^{-1}y\| = \|A^{-1}y' - A^{-1}y\| < \varepsilon_0$$

Thus $\|\varepsilon\|_{C^1}$ as desired. ☐

Corollary 2.1.8. *Given $\overline{\varepsilon} > 0$ we can, if necessary, diminish $\varepsilon > 0$ so that we can write*

$$\overline{f}^{-1}(x,y) = (\lambda_1^{-1}x + \theta(x,y), \lambda_2^{-1}y + \vartheta(x,y))$$

where $\|\theta\|_{C^1} < \overline{\varepsilon}$ and $\|\vartheta\|_{C^1} < \overline{\varepsilon}$.

This $\overline{f}$ is our desired local expression. We will now go to the next step in our proof, which is to search for a $\overline{f}$ invariant curve.

Invariant Curve

We will look for an invariant curve among graphs of Lipschitz functions, i.e. lets consider the set

$$\operatorname{Lip}_1 \stackrel{\text{def.}}{=} \{\tilde{\varphi} : \mathbb{R} \to \mathbb{R} | \varphi(0) = 0 \text{ e } Lip\,\varphi \leq 1\}$$

We endow this set with the following norm:

$$\begin{aligned} \|\cdot\| : \operatorname{Lip}_1 &\to \mathbb{R} \\ \varphi &\mapsto \|\varphi\| \stackrel{\text{def.}}{=} \sup_{t\neq 0} \tfrac{|\varphi|}{|t|} \end{aligned}$$

Claim 2.1.8.1. $(\operatorname{Lip}_1, \|\cdot\|)$ is a Banach space.

Proof. It is clear that $\|\cdot\|$ defines a norm, we must only check that it is complete.

Let $(\varphi_n)_{n\in\mathbb{N}} \subseteq \operatorname{Lip}_1$ be a Cauchy sequence. In particular

$$\frac{|\varphi_n(t) - \varphi_m(t)|}{|t|} \leq \|\varphi_n - \varphi_m\| \to 0 \text{ when } n,m \to +\infty$$

Thus $(\varphi_n(t))_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is Cauchy for every $t \in \mathbb{R}$. Since $\mathbb{R}$ is complete, this sequence converges.

Define $\varphi(t) = \lim_{n\to+\infty} \varphi_n(t)$. By definition of being a Cauchy sequence, given $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that for all $m, n > N$ we have

$$\|\varphi_n - \varphi_m\| < \varepsilon$$

Fixing m and taking the limit as $n \to +\infty$ we obtain that

$$\|\varphi - \varphi_m\| < \varepsilon$$

Thus, it seams that $\varphi_n \to \varphi$. It only remais to show that $\varphi \in \mathrm{Lip}_1$. For it, just notice that for this same $m > N$ that we fixed, we have

$$\begin{aligned} |\varphi(t) - \varphi(s)| &\leq |\varphi(t) - \varphi_m(t)| + |\varphi_m(t) - \varphi_m(s)| + |\varphi_m(s) - \varphi(s)| \\ &\leq 2\varepsilon + Lip\,\varphi_m|t - s| \\ &\leq 2\varepsilon + |t - s| \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have

$$|\varphi(t) - \varphi(s)| \leq |t - s|$$

so that $\varphi \in \mathrm{Lip}_1$ and the claim is proven. $\square$

The graph $graph\,\varphi$ of a function $\varphi \in \mathrm{Lip}_1$ is invariant by $\overline{f}$ if $\overline{f}(graph\,\varphi) = graph\,\varphi$, or equivalently, if $graph\,\varphi = \overline{f}^{-1}(graph\,\varphi)$.

To proceed we need the following claim

Claim 2.1.8.2. If $\overline{\varepsilon} > 0$ is small enough, it holds that for all $\varphi \in \mathrm{Lip}_1$ we have that $\overline{f}^{-1}(graph\,\varphi)$ is the graph of a function.

Proof. A set $A \subseteq \mathbb{R}^2$ is the graph of a function in $\mathbb{R}$ if for all $x \in \mathbb{R}$ there is one and only one $y \in \mathbb{R}$ such that $(x, y) \in A$. Equivalently, let $\pi : \mathbb{R}^2 \to \mathbb{R}$ be the projection in the first coordinate; a set $A \subseteq \mathbb{R}^2$ is the graph of a function in $\mathbb{R}$ if the restriction to A of the projection is a bijection.

We have that

$$\begin{aligned} \overline{f}^{-1}(graph\,\varphi) &= \{\overline{f}^{-1}(t, \varphi(t)) | t \in \mathbb{R}\} \\ &= \{(\lambda_1^{-1}t + \vartheta(t, \varphi(t)), \lambda_2^{-1}\varphi(t) + \vartheta(t, \varphi(t)) | t \in \mathbb{R}\} \end{aligned}$$

Thus using t as a parameter of $\overline{f}^{-1}(graph\,\varphi)$, we can write $pi|_{\overline{f}^{-1}(graph\,\varphi)}(t) \equiv \lambda_1^{-1}t + \theta(t, \varphi(t))$. Notice that

$$\begin{aligned} \|\theta(t, \varphi(t)) - \theta(s, \varphi(s))\| &\leq \overline{\varepsilon}\|(t, \varphi(t)) - (s, \varphi(s))\| \\ &= \overline{\varepsilon}\sqrt{(t-s)^2 + (\varphi(t) - \varphi(s))^2} \\ &\leq \overline{\varepsilon}\sqrt{(t-s)^2 + (t-s)^2} \\ &= \sqrt{2}\overline{\varepsilon}|t - s| \end{aligned}$$

Thus $\mathrm{Lip}(\theta(\cdot, \varphi(\cdot))) \leq \sqrt{2}\overline{\varepsilon}$. Hence, by the (Lipschitz) Inverse Function Theorem, if $\overline{\varepsilon}$ is sufficiently small, $\pi|_{\overline{f}^{-1}(graph\,\varphi)}$ is a lipeomorphism and this concludes the claim. $\square$

Since the graph of a function completely characterizes the function, this claim says that for every $\varphi \in \mathrm{Lip}_1$ there is one and only one function whose graph is $\overline{f}^{-1}(graph\,\varphi)$. Thus we have a well defined function $\overline{f^*}$ that takes a map $\varphi \in \mathrm{Lip}_1$ and returns the unique map $\overline{f}^{-1}\varphi$ such that

$$graph\,\overline{f^*\varphi} = \overline{f}^{-1}(graph\,\varphi) \tag{*}$$

Now, the task to find a function whose graph is invariant to $\overline{f}$ has been converted into the task to find a fixed point for the map $\overline{f^*}$. To do it we must first understand the functions $\overline{f^*\varphi}$. Let's start by noticing that, since graphs are parameterized by a single real parameter, the equation (*) implies that for every $t \in \mathbb{R}$ there exists a $N_\varphi(t) \in \mathbb{R}$ such that

$$(t, \overline{f^*\varphi}(t)) = \overline{f}^{-1}(N_\varphi(t), \varphi(N_\varphi(t))) \tag{1}$$

Claim 2.1.8.3. Given $\varphi \in \mathrm{Lip}_1$, the map $N_\varphi : \mathbb{R} \to \mathbb{R}$ given by the formula above is well defined and a bijection.

Proof. If $\tilde{N}_\varphi(t) \in \mathbb{R}$ also satisfies (1), them

$$\begin{cases} (t, \overline{f^*\varphi}(t)) = \overline{f}^{-1}(N_\varphi(t), \varphi(N_\varphi(t))) \\ (t, \overline{f^*\varphi}(t)) = \overline{f}^{-1}(\tilde{N}_\varphi(t), \varphi(\tilde{N}_\varphi(t))) \end{cases}$$

Thus

$$\overline{f}^{-1}(N_\varphi(t), \varphi(N_\varphi(t))) = \overline{f}^{-1}(\tilde{N}_\varphi(t), \varphi(\tilde{N}_\varphi(t)))$$

Since $\overline{f}$ is a bijection, we have that

$$(N_\varphi(t), \varphi(N_\varphi(t))) = (\tilde{N}_\varphi(t), \varphi(\tilde{N}_\varphi(t))) \iff N_\varphi(t) = \tilde{N}_\varphi(t)$$

hence N_φ is well defined. If $t \neq s$ then

$$(t, \overline{f^*\varphi}(t)) \neq (s, \overline{f^*\varphi}(s))$$

Thus

$$\overline{f}^{-1}(N_\varphi(t), \varphi(N_\varphi(t))) \neq \overline{f}^{-1}(N_\varphi(s), \varphi(N_\varphi(s)))$$

And again, since $\overline{f}$ is a bijection, we have

$$\begin{aligned} (N_\varphi(t), \varphi(N_\varphi(t))) \neq (N_\varphi(s), \varphi(N_\varphi(s))) \quad &\iff N_\varphi(t) \neq N_\varphi(s) \text{ ou } \varphi(N_\varphi(t)) \neq \varphi(N_\varphi(s)) \\ &\implies N_\varphi(t) \neq N_\varphi(s) \end{aligned}$$

so that N_φ is injective. To see that N_φ is surjective, let $t \in \mathbb{R}$. By equation (*), there must be a $K(t) \in \mathbb{R}$ such that

$$(K(t), \overline{f^*\varphi}(K(t))) = \overline{f}^{-1}(t, \varphi(t))$$

Thus, by definition of N_φ, we have that $N_\varphi(K(t)) = t$. Hence N_φ is also surjective, and the claim is proved.

□

Back to equation (1), opening the first coordinate shows that

$$t = \lambda_1^{-1} N_\varphi(t) + \theta(N_\varphi(t), \varphi(N_\varphi(t)))$$

By the claim, N_φ is a bijection, so we can write $t = N_\varphi^{-1}(s)$. With this the expression above becomes

$$N_\varphi^{-1}(s) = \lambda_1^{-1} s + \theta(s, \varphi(s))$$

Notice that this is the same expression that we obtained in claim 2.1.8.2, where we saw that this is a lipeomorphism. In particular, the (Lipchitz) Inverse Function Theorem gives that $\mathrm{Lip}(N_\varphi) \leq \frac{1}{\lambda_1^{-1} - \mathrm{Lip}(\theta)}$, i.e.

$$|N_\varphi(t) - N_\varphi(s)| \leq \frac{|t-s|}{\lambda_1^{-1} - \sqrt{2}\overline{\varepsilon}}$$

Now that we have understood a little about the function N_φ, we can go back to equation (2) and open the second coordinate

$$\overline{f^*\varphi}(t) = \lambda_2^{-1} \varphi(N_\varphi(t)) + \vartheta(N_\varphi(t), \varphi(N_\varphi(t)))$$

Claim 2.1.8.4. If $\overline{\varepsilon}$ is small enough, then given $\varphi \in \mathrm{Lip}_1$ we have that $\overline{f^*\varphi} \in \mathrm{Lip}_1$.

Proof. We have that

$$\begin{aligned}
|\overline{f^*\varphi}(t) - \overline{f^*\varphi}(s)| &= |\lambda_2^{-1}\varphi(N_\varphi(t)) + \vartheta(N_\varphi(t), \varphi(N_\varphi(t))) - \lambda_2^{-1}\varphi(N_\varphi(s)) - \vartheta(N_\varphi(s), \varphi(N_\varphi(s)))| \\
&\leq \lambda_2^{-1}|\varphi(N_\varphi(t) - \varphi(N_\varphi(s)| - |\vartheta(N_\varphi(t), \varphi(N_\varphi(t))) - \vartheta(N_\varphi(s), \varphi(N_\varphi(s)))| \\
&\leq \lambda_2^{-1}|N_\varphi(t) - N_\varphi(s)| - \sqrt{2}\overline{\varepsilon}|N_\varphi(t) - N_\varphi(s)| \\
&\leq \frac{\lambda_2^{-1} - \sqrt{2}\overline{\varepsilon}}{\lambda_1^{-1} - \sqrt{2}\overline{\varepsilon}}|t-s|
\end{aligned}$$

Since $\lambda_2^{-1}\lambda_1 < 1$, in the limit where $\overline{\varepsilon \to 0}$ we have that $\frac{\lambda_2^{-1} - \sqrt{2}\overline{\varepsilon}}{\lambda_1^{-1} - \sqrt{2}\overline{\varepsilon}} \to \lambda_2^{-1}\lambda_1 < 1$. Thus, for $\overline{\varepsilon}$ small, we get

$$|\overline{f^*\varphi}(t) - \overline{f^*\varphi}(s)| \leq |t-s|$$

Also, $(0,0) \in graph\, \varphi$, Thus

$$(0,0) = \overline{f}^{-1}(0,0) \in \overline{f}^{-1}(graph\, \varphi) = graph\, \overline{f^*\varphi}$$

and $\overline{f^*\varphi}(0) = 0$. With all that we conclude that $\overline{f^*\varphi} \in \mathrm{Lip}_1$ and finish the claim. □

Thanks to this claim, we now have a well defined map

$$\begin{aligned}
\overline{f^*} : \mathrm{Lip}_1 &\to \mathrm{Lip}_1 \\
\varphi &\mapsto \overline{f^*\varphi}
\end{aligned}$$

It would be great if this map $\overline{f^*}$ were a contraction. In that way we could use Banach's Fixed Point Theorem to find an unique fixed point, i.e. a unique invariant curve $\varphi^* \in \mathrm{Lip}_1$, which is exactly what we want. In fact, it does hold that $\overline{f^*}$ is a contraction, and to see it we will need some computations. Before them, we need the following lemma:

Lemma 2.1.9. *Let (X,d) be a metric space, $\xi \in \mathbb{R}_{>0}$, $0 < \varepsilon < \lambda$ and suppose that $(f_x : \mathbb{R} \to \mathbb{R})_{x \in X}$ is a family of homeomorphisms satisfying*

$$\begin{cases} |f_x(t) - f_y(t)| \leq \xi d(x,y)|t| & (1) \\ |f_x(t) - \lambda t| \leq \varepsilon |t| & (2) \end{cases}$$

Then

$$|f_x^{-1}(s) - f_y^{-1}(s)| \leq \frac{\lambda^{-1}}{(1-\lambda^{-1}\varepsilon)(\lambda-\varepsilon)}\xi d(x,y)|s|$$

Proof. By equation (2) we have that

$$f_x(t) = \lambda t + r_x(t) \tag{i}$$

where $Lip\, r_x \leq \varepsilon$. Using (2) we obtain

$$|f_x(t) - f_y(t)| = |r_x(t) - r_y(t)| \leq \xi d(x,y)|t|$$

Since f_x is an homeomorphism we can write $t = f_x^{-1}(s)$ for some s. Using this in (i) we get

$$f_x^{-1}(s) = \lambda^{-1}s - \lambda^{-1}r_x(f_x^{-1}(s))$$

Thus

$$\begin{aligned}
|f_x^{-1}(s) - f_y^{-1}(s)| &= \lambda^{-1}|r_x(f_x^{-1}(s)) - r_y(f_y^{-1}(s))| \\
&\leq \lambda^{-1}|r_x(f_x^{-1}(s)) - r_y(f_x^{-1}(s))| + \lambda^{-1}|r_y(f_x^{-1}(s)) - r_y(f_y^{-1}(s))| \\
&\leq \lambda^{-1}\xi d(x,y)\|f_x^{-1}(s)\| + \lambda^{-1}Lip\, r_y|f_x^{-1}(s) - f_y^{-1}(s)| \\
&\leq \lambda^{-1}\xi d(x,y)\|f_x^{-1}(s)\| + \lambda^{-1}\varepsilon|f_x^{-1}(s) - f_y^{-1}(s)|
\end{aligned}$$

Rearranging the terms with $|f_x^{-1}(s) - f_y^{-1}(s)|$ we obtain

$$|f_x^{-1}(s) - f_y^{-1}(s)| \leq \frac{\lambda^{-1}}{1-\lambda^{-1}\varepsilon}\xi d(x,y)|f_x^{-1}(s)|$$

More over, by the Lipschitz Inverse Function Theorem, $\mathrm{Lip}(f_x^{-1}) \leq \frac{1}{\lambda - Lip\, r_x} \leq \frac{1}{\lambda-\varepsilon}$. Hence

$$|f_x^{-1}(s) - f_y^{-1}(s)| \leq \frac{\lambda^{-1}}{(1-\lambda^{-1}\varepsilon)(\lambda-\varepsilon)}\xi d(x,y)|s|$$

as desired. □

Corollary 2.1.10. *If $\overline{\varepsilon} > 0$ is small enough, then*

$$|N_\varphi(t) - N_{\tilde{\varphi}}(t)| \leq \frac{\lambda_1\overline{\varepsilon}}{(1-\lambda_1\sqrt{2}\overline{\varepsilon})(\lambda_1^{-1} - \sqrt{2}\overline{\varepsilon})}d(\varphi,\tilde{\varphi})|t|$$

Proof. If necessary, diminish $\overline{\varepsilon} > 0$ so that $\sqrt{2}\overline{\varepsilon} < \lambda_1^{-1}$. Apply the lemma above for $X = \mathrm{Lip}_1$, $\xi = \overline{\varepsilon}$, $\lambda = \lambda_1^{-1}$, $\varepsilon = \sqrt{2}\overline{\varepsilon}$ and the family $(N_\varphi^{-1} : \mathbb{R} \to \mathbb{R})_{\varphi \in \mathrm{Lip}_1}$. □

With this result, we can now show that $\overline{f^*}$ is indeed a contraction:

Claim 2.1.10.1. If $\overline{\varepsilon} > 0$ is small enough, then there exists a $\gamma \in (0,1)$ such that $\mathrm{Lip}(\overline{f^*}) \leq \gamma$.

Proof. We have that

$$\begin{aligned}
|\overline{f^*\varphi}(t) - \overline{f^*\tilde{\varphi}}(t)| &= |\lambda_2^{-1}\varphi(N_\varphi(t)) + \vartheta(N_\varphi(t), \varphi(N_\varphi(t))) - \lambda_2^{-1}\tilde{\varphi}(N_{\tilde{\varphi}}(s)) - \vartheta(N_{\tilde{\varphi}}(s), \tilde{\varphi}(N_{\tilde{\varphi}}(s)))| \\
&\leq \lambda_2^{-1}|\varphi(N_\varphi(t)) - \tilde{\varphi}(N_\varphi(t))| + \lambda_2^{-1}|\tilde{\varphi}(N_\varphi(t)) - \tilde{\varphi}(N_{\tilde{\varphi}}(t))| \\
&\quad + |\vartheta(N_\varphi(t), \varphi(N_\varphi(t))) - \vartheta(N_\varphi(t), \tilde{\varphi}(N_\varphi(t)))| \\
&\quad + |\vartheta(N_\varphi(t), \tilde{\varphi}(N_\varphi(t))) - \vartheta(N_{\tilde{\varphi}}(t), \tilde{\varphi}(N_{\tilde{\varphi}}(t)))|
\end{aligned}$$

Where, by definition

$$|\varphi(N_\varphi(t)) - \tilde{\varphi}(N_\varphi(t))| \leq d(\varphi,\tilde{\varphi})|N_\varphi(t)|$$

Also, since $|N_\varphi(t) - N_\varphi(s)| \leq \frac{|t-s|}{\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}}}$, we have that

$$|\varphi(N_\varphi(t)) - \tilde{\varphi}(N_\varphi(t))| \leq d(\varphi, \tilde{\varphi}) \frac{1}{\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}}} |t|$$

Thus using that $\mathrm{Lip}(\tilde{\varphi}) \leq 1$ and corollary 2.1.10 we see that

$$|\tilde{\varphi}(N_\varphi(t)) - \tilde{\varphi}(N_{\tilde{\varphi}}(t))| \leq \frac{\lambda_1 \bar{\varepsilon}}{(1 - \lambda_1 \sqrt{2\bar{\varepsilon}})(\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}})} d(\varphi, \tilde{\varphi})|t|$$

Simillarly, by corollary 2.1.8 and the same arguments above, we also see that

$$|\vartheta(N_\varphi(t), \varphi(N_\varphi(t))) - \vartheta(N_\varphi(t), \tilde{\varphi}(N_\varphi(t)))| \leq \bar{\varepsilon} d(\varphi, \tilde{\varphi}) \frac{1}{\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}}} |t|$$

and

$$|\vartheta(N_\varphi(t), \tilde{\varphi}(N_\varphi(t))) - \vartheta(N_{\tilde{\varphi}}(t), \tilde{\varphi}(N_{\tilde{\varphi}}(t)))| \leq \sqrt{2}\bar{\varepsilon} \frac{\lambda_1 \bar{\varepsilon}}{(1 - \lambda_1 \sqrt{2\bar{\varepsilon}})(\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}})} d(\varphi, \tilde{\varphi})|t|$$

Thus, gathering all those estimates above, we obtain

$$\begin{aligned} |\overline{f^*\varphi}(t) - \overline{f^*\tilde{\varphi}}(t)| \quad &\leq \Big(\tfrac{\lambda_2^{-1}}{\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}}} + \tfrac{\lambda_2^{-1}\lambda_1 \bar{\varepsilon}}{(1 - \lambda_1 \sqrt{2\bar{\varepsilon}})(\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}})} + \bar{\varepsilon} \tfrac{1}{\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}}} \\ &\qquad + \sqrt{2}\bar{\varepsilon} \tfrac{\lambda_1 \bar{\varepsilon}}{(1 - \lambda_1 \sqrt{2\bar{\varepsilon}})(\lambda_1^{-1} - \sqrt{2\bar{\varepsilon}})}) \Big) d(\varphi, \tilde{\varphi})|t| \end{aligned}$$

In the limit where $\bar{\varepsilon} \to 0$, the coefficient above goes to $\lambda_2^{-1}\lambda_1 < 1$. Hence we can take a $\gamma \in (\lambda_2^{-1}\lambda_1, 1)$, which for $\bar{\varepsilon}$ small, satisfies

$$|\overline{f^*\varphi}(t) - \overline{f^*\tilde{\varphi}}(t)| \leq \gamma d(\varphi, \tilde{\varphi})|t|$$

For $t \neq 0$ this means that

$$\frac{|\overline{f^*\varphi}(t) - \overline{f^*\tilde{\varphi}}(t)|}{|t|} \leq \gamma \|\varphi - \tilde{\varphi}\|$$

Taking the supremum over all $t \neq 0$ we conclude that

$$\|\overline{f^*\varphi} - \overline{f^*\tilde{\varphi}}\| \leq \gamma \|\varphi - \tilde{\varphi}\|$$

Hence $\mathrm{Lip}(\overline{f^*}) \leq \gamma$ as desired. □

Finally, by this claim, $\overline{f^*}$ is a contraction. Hence, by Banach's Fixed Point Theorem, we have an unique invariant map $\varphi^* \in \mathrm{Lip}_1$ whose graph is an invariant curve.

We define $W^s_{loc}(p)$ in the chart Φ by the restriction of this curve to the ball $B_{\frac{\delta}{2}}(0)$. I.e.

$$W^s_{loc}(p) \overset{\text{def.}}{=} \Phi^{-1}\left(graph\, \varphi^* \cap B_{\frac{\delta}{2}}(0) \right)$$

We will now prove that this curve satisfies the conclusions of the Theorem.

Convergence and Invariance

Consider the dynamics restricted to this graph

$$\begin{aligned} \rho := \pi|_{graph\,\varphi^*} \circ \overline{f} \circ \pi|^{-1}_{graph\,\varphi^*} : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \lambda_1 t + \overline{\alpha}(t, \varphi^*(t)) \end{aligned}$$

and notice that the following diagram commutes

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\pi|^{-1}_{graph\,\varphi^*}} & graph\,\varphi^* \\ \rho\downarrow & & \downarrow\overline{f} \\ \mathbb{R} & \xleftarrow[\pi|_{W^s_{loc}}]{} & graph\,\varphi^* \end{array}$$

For $p \in W^s_{loc}$, let $x = \Phi(p)$. Since $x \in graph\,\varphi^*$, we can write $x = \pi|^{-1}_{graph\,\varphi^*}(t)$ for some t. With this we have

$$\rho^n(t) = \pi(\overline{f}^n(p_1))$$

And in particular

$$\overline{f}^n(p_1) = (\rho^n(t), \varphi^*(\rho^n(t)))$$

However, notice that

$$\begin{aligned} |\rho(t) - \rho(s)| &= |\lambda_1 t + \overline{\alpha}(t, \varphi^*(t)) - \lambda_1 s - \overline{\alpha}(s, \varphi^*(s))| \\ &\leq (\lambda_1 + \sqrt{2}\overline{\varepsilon})|t - s| \end{aligned}$$

Thus, for $\overline{\varepsilon} < \frac{1-\lambda_1}{\sqrt{2}}$, we have that ρ is a contraction. Hence, since $\rho(0) = 0$, it holds that $\rho^n(t) \to 0$ exponentially fast. Also, since $\varphi^*(0) = 0$ and φ^* is Lipschitz, it follows that $\overline{f}^n(x) \to 0$ exponentially fast. That is, $f^n(p_0) \to p$ exponentially fast. In particular, $\overline{f}|_{W^s_{loc}}$ is a contraction, so $f(W^s_{loc}) \subseteq W^s_{loc}$.

Uniqueness

Suppose that $q \in M$ is such that $f^n(q) \to p$ exponentially fast. In particular, $f^n(q)$ is eventually in the domain of our local expression. Thus, for $n \in \mathbb{N}$ big, we can construct a sequence of points $(x_n, y_n) = \Phi(f^n(q)) \in \mathbb{R}^2$ such that $(x_n, y_n) \to (0, 0)$ exponentially fast. Now, to prove that $f^n(q)$ is eventually in $W^s_{loc}(p)$ is to prove that (x_n, y_n) is eventually in $graph(\varphi^*)$.

This sequence satisfies $(x_{n+1}, y_{n+1}) = \Phi \circ f \circ \Phi^{-1}(x_n, y_n)$. Hence, since (x_n, y_n) is eventually in $B_{\frac{\delta}{2}}(0)$ (where $\overline{f}$ coincides with $\Phi \circ f \circ \Phi^{-1}$), we have $(x_{n+1}, y_{n+1}) = \overline{f}(x_n, y_n)$ for all n big enough.

Claim 2.1.10.2. If we choose $\overline{\varepsilon} > 0$ small enough, then $|y_{n+1}| \leq |y_n|$ only if $|y_n| < |x_n|$.

Proof. Suppose that for some $n \in \mathbb{N}$ we have $|y_n| \geq |x_n|$. In particular, we have $\|(x_n, y_n)\| \leq \sqrt{2}|y_n|$. Thus, by the definition of the sequence and claim 2.1.6.1, we have

$$|y_{n+1}| = |\lambda_2 y_n + \overline{\beta}(x_n, y_n)| \geq \lambda_2|y_n| - |\overline{\beta}(x_n, y_n)| \geq \lambda_2|y_n| - \sqrt{2}\overline{\varepsilon}|y_n| = (\lambda_2 - \sqrt{2}\overline{\varepsilon})|y_n|$$

Since $\lambda_2 > 1$, we can chose $\overline{\varepsilon} > 0$ small enough so that $\lambda_2 - \sqrt{2}\overline{\varepsilon} > 1$ which proves the claim. □

Since we know that $(x_n, y_n) \to 0$, we also know that $y_n \to 0$. Thus, there must be some n's such that $y_{n+1} < y_n$. By the claim above, for these n's, we must have $y_n < x_n$. Actually, this must hold for every n:

Claim 2.1.10.3. For all $n \in \mathbb{N}$, it holds that $|y_n| < |x_n|$.

Proof. Suppose by contradiction that there is a $n_0 \in \mathbb{N}$ such that $|y_{n_0}| \geq |x_{n_0}|$. Then, as before, we have $\|(x_{n_0}, y_{n_0})\| \leq \sqrt{2}|x_{n_0}|$. Thus, by the definition of the sequence and claim 2.1.6.1

$$|y_{n_0+1}| \geq (\lambda_2 - \sqrt{2}\overline{\varepsilon})|y_{n_0}|$$

and

$$|x_{n_0+1}| = |\lambda_1 x_1 + \overline{\alpha}(x_{n_0}, y_{n_0})| \leq (\lambda_1 + \sqrt{2}\overline{\varepsilon})|y_{n_0}|$$

Thus

$$\left|\frac{y_{n_0+1}}{x_{n_0+1}}\right| \geq \frac{\lambda_2 - \sqrt{2}\overline{\varepsilon}}{\lambda_1 + \sqrt{2}\overline{\varepsilon}}$$

Since $0 < \lambda_1 < \lambda_2$, we can take $\overline{\varepsilon} > 0$ small enough so that the term above is greater than 1. In particular $|y_{n_0+1}| \geq |x_{n_0+1}|$ and by induction $|y_n| \geq |x_n|$ for every $n \geq n_0$. However, by claim 2.1.6.1, this means that y_n is increasing for $n \geq n_0$. This contradicts the fact that $y_n \to 0$ and concludes the claim. □

Now, consider the subset $\text{Lip}_1(q)$ of Lip_1 given by

$$\text{Lip}_1(q) = \{\varphi \in \text{Lip}_1 \, | \varphi(x_n) = y_n, \, \forall n \in \mathbb{N}\}$$

Claim 2.1.10.4. $\text{Lip}_1(q)$ is a non empty closed $\overline{f^*}$-invariant subset of Lip_1.

Proof. It is nonempty because it contains the map φ_0 defined by

$$\varphi_0(x_n) = y_n \quad \text{and} \quad \varphi_0 \text{ is linear between each } x_n$$

which is in Lip_1 by claim 2.1.10.3. To show that it is closed, let $\{\varphi_k\}_{k\in\mathbb{N}} \subseteq \text{Lip}_1(q)$ be a sequence such that $\varphi_k \to \varphi \in \text{Lip}_1$. We have that

$$\varphi(x_n) = \lim_k \varphi_k(x_n) = \lim_k y_n = y_n$$

Thus $\varphi \in \text{Lip}_1(q)$ and $\text{Lip}_1(q)$ is in fact closed. To show that it is $\overline{f^*}$ invariant, let $\varphi \in \text{Lip}_1$. We already know that $\overline{f^*\varphi} \in \text{Lip}_1$, thus we must only show that $\overline{f^*\varphi}(x_n) = y_n$. Recall that by definition of $\overline{f^*}$ we have

$$\overline{f^{-1}}(graph(\varphi)) = graph(\overline{f^*\varphi})$$

Also, by definition of the sequence, we have $(x_n, y_n) = \overline{f^{-1}}(x_{n+1}, y_{n+1})$. Since $\varphi \in \text{Lip}_1(q)$ we have $(x_{n+1}, y_{n_1}) \in graph(\varphi)$, thus

$$(x_n, y_n) \in \overline{f^{-1}}(graph(\varphi)) = graph(\overline{f^*\varphi})$$

I.e. $\overline{f^*\varphi}(x_n) = y_n$ and the claim is proven. □

From this and claim 2.1.10.1, we can use Banach's Fixed Point Theorem to find a $\overline{f^*}$ invariant map φ_0^*. By unicity we must have $\varphi_0^* = \varphi^*$, which means that $\varphi^*(x_n) = y_n$. This is what we wanted.

Tangency and Smoothness

To obtain the smoothness, notice that since $\varphi*$ is Lipschitz, the accumulation points of

$$\frac{\varphi^*(x) - \varphi^*(x_0)}{x - x_0} \quad \text{as } x \to x_0$$

is a subset of $[-1, 1]$. In particular, if you define a cone at $\vec{x} \in \mathbb{R}^2$ of width $\delta > 0$ and diameter $\eta > 0$ tangent to $\hat{x}$ as the set

$$\mathcal{C}(\vec{x}, \delta, \eta) \stackrel{\text{def.}}{=} \{(x, y) \in B_\eta(\vec{x}) | |y| \leq \delta |x|\}$$

It follows that every piece of the graph of $\varphi*$ is contained in a cone of width 1. In particular, the derivative of $\overline{f^{-1}}$ contracts the width of these cones. Hence, since the graph is invariant, for every point $\vec{x} \in graph(\varphi^*)$ you can obtain a tangent line $L \subseteq \bigcap_n \overline{f^{-n}}(\mathcal{C}(\overline{f}^n(\vec{x}), 1, \eta_n))$. With a similar argument, you show that those eigenspaces must vary continuously; otherwise these cones wouldn't degenerate.

At $\vec{0}$ the derivative is diagonal; thus the tangent line must be parallel to the x axis. Hence, $W^s_{loc}(p)$ is tangent to the eigenspace of $d_p f$ associated to the contracting eigenvalue λ_1. And the last item of the theorem is proved. □

2.2 Local Product Structure

A crucial consequence of the existence of those stable and unstable manifolds is that, locally, you can use them as coordinates. This is called the local product structure. To be more precise:

Definition 2.2.1 (Product Neighborhood)**.** A system $f : M \to M$ is said to have the local product structure if for every $p \in M$ there is an open neighborhood U of p parameterized by a continuous map

$$\Phi : W^s_{loc}(p) \times W^u_{loc}(p) \to U$$

such that $\Phi(q_1, q_2) \in W^u(q_1) \cap W^s(q_2)$. Any such neighborhood is called a product neighborhood (or foliated box).

To obtain this, notice that the continuity of df and the uniform contraction of E^s implies that this distribution is continuous:

Lemma 2.2.2. *The distribution E^s is Hölder continuous.*

Proof. See Theorem 19.1.6 of [KH95]. □

The same holds for E^u. Since E^s and E^u are transverse at each point, the angle between them is positive. Thus, by their continuity, it is locally bounded from 0. Hence, for small neighborhoods, the manifolds $W^s_{loc}(p)$ and $W^u_{loc}(p)$ form a collection of smooth uniformly transverse curves; in particular they intersect each other at one and at most one point:

Proposition 2.2.3. *Given $p \in M$ there exists $\varepsilon_0 > 0$ such that for every $\varepsilon < \varepsilon_0$ there is a $\delta > 0$ such that for every neighborhood U of p with diameter less than δ, it holds that $W^s_\varepsilon(q_1) \cap W^u_\varepsilon(q_2)$ consists of a single point for every two points $q_1, q_2 \in U$.*

With this proposition we easily obtain the local product structure:

Theorem 2.2.4. *If $f : M \to M$ is Anosov, then it as the local product structure.*

Proof. Take $p \in M$ and let $\varepsilon, \delta > 0$ be as in the proposition above. We have that $W^s_\delta(p)$ and $W^u_\delta(p)$ are within a δ neighborhood of p. Hence, the map Φ_δ that takes $(q_1, q_2) \in W^s_\delta(p) \times W^u_\delta(p)$ and sends it to the unique element of $W^u_\varepsilon(q_1) \cap W^s_\varepsilon(q_2)$ is well defined continuous map. By unicity, if $q = \Phi_\delta(q_1, q_2)$, then q_1 is the unique element in $W^u_\varepsilon(q) \cap W^s_\varepsilon(p)$ and q_2 is the unique element in $W^s_\varepsilon(q) \cap W^u_\varepsilon(p)$. Thus, for any $\delta' < \delta$, Φ_δ restricted to the closure of $W^s_{\delta'}(p) \times W^u_{\delta'}(p)$ is a homeomorphism. In particular, we obtain a parametrization $\Phi_{\delta'} : W^s_{\delta'}(p) \times W^u_{\delta'}(p)$ for any $\delta' < \delta$. □

A fundamental use of this structure is to obtain Holonomy maps:

Definition 2.2.5. A stable Holonomy map from $p \in M$ to $q \in M$ is any continuous map of the form

$$\begin{array}{rcl} \mathcal{H}_{p\to q} : W^u_{loc}(p) & \to & W^u_{loc}(q) \\ r & \mapsto & \mathcal{H}_{p\to q}(r) \in W^s(r) \bigcap W^u(q) \end{array}$$

such that $\mathcal{H}_{p\to q}(p) = q$

These maps glue pieces of unstable manifolds of a point to those of another by carrying them along stable manifolds. Their regularity is of great importance for the study of the dynamics; see, for example [Gu23]. For us, they will be an essential tool for the construction of the Margulis family in chapter 5.

Using the local product structure we can easily show their existence at short distances: Given a product neighborhood U with parametrization Φ of a point $p \in M$, you can define an Holonomy $\mathcal{H}_{p\to q}$ for any $q \in U \cap W^s_{loc}(p)$ by

$$r \in W^u_{loc}(p) \mapsto \mathcal{H}_{p\to q}(r) \overset{\text{def.}}{=} \Phi(q, r) \in W^u_{loc}(q)$$

by lemma 2.2.2, this map is Hölder.

In fact, in our context of surface Anosov diffeomorphisms, they are actually $C^{1+\alpha}$. We prove it in 4.2.7 using that we can construct a very special kind of parameter for the unstable manifolds (the affine structures 4.2.1).

Their existence is not limited to small neighborhoods; in chapter 5 we'll show that for every open segment of unstable leaf we can map uniformly small sets everywhere inside it. The usefulness of this procedure is that the set we have just mapped inside the other is connected by segments of stable manifolds (which will have uniform length). Thus, as we iterate, this small set gets exponentially closer to the big one. This allows us, in some sense, to pass information from one to another (see proposition 5.1.4).

2.3 Topological Classification on Dimension Two

We have now defined the type of systems that we will work on and some properties that come with them. The question that we ask now is: How many Anosov maps exist?

This question must be made carefully. If you pay attention to the definition of Anosov diffeomorphism, it is clearly open in the C^1 topology. Hence, if there is one Anosov map, then there is an entire infinite-dimensional open set of them. It is not hard to find some starting examples.

The simplest one is an hyperbolic matrix in $\mathbb{R}^n$. I.e. if $A \in GL(n, \mathbb{R})$ is a $n \times n$ real invertible matrix with eigenvalues $|\lambda_1| \leq \cdots \leq \lambda_k < 1 < \lambda_{k+1} \leq \cdots \leq \lambda_n$, then $A : \mathbb{R}^n \to \mathbb{R}^n$ is an Anosov diffeomorphism. However, that's not the best example, because some of their properties we want for them require the space to be compact. We can make a better one by making its domain more interesting: The n-torus is the quotient space $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, where Z^n acts in $\mathbb{R}^n$ by translations. If the coefficients of a $n \times n$ matrix A are all integers, it preservers the the 'grid' $\mathbb{Z}^n$. Thus, the class $[A(p)] \in \mathbb{T}^n$ does not depend on the representative of $[p] \in \mathbb{T}^n$. In particular, we have another example

Example 2.3.1 (Linear Anosov Diffeomorphisms)**.** *If $A \in SL(n, \mathbb{Z})$ is hyperbolic, then $f_A : \mathbb{T}^n \to \mathbb{T}^n$ given by $f_A([p]) \stackrel{\text{def.}}{=} [A(p)]$ is an Anosov diffeomorphism in $\mathbb{T}^n$.*

Taking small perturbations of these examples we obtain an infinity of Anosov diffeomorphisms. The question now is whether we can topologically classify them. By this we mean

Definition 2.3.2. Two continuous maps $f : M \to M$ and $g : N \to N$ are said to be topologically conjugated if there exists an homeomorphism $h : M \to N$ such that $h \circ f = g \circ h$.

Topological conjugacy is an equivalence relation. To topologically classify all Anosov maps, we mean finding all classes of topological conjugacy. This result is partially given when the map is of codimension one, i.e. when either E^s or E^u is one-dimensional:

Theorem 2.3.3. *If $f : M \to M$ is a codimension one Anosov diffeomorphism, then f is topologically conjugated to a linear toral automorphism.*

This is a well-known result, initially proved under some hypothesis by Franks in his thesis and later stated as here by Newhouse [New70]. Curiously, that simple example 2.3.1 turned out to be very general. In particular, in dimensions 2 and 3, every Anosov diffeomorphism must be of codimension one. Hence, for these low dimensions, topologically, there are only linear toral Anosov diffeomorphisms, and the classification is complete.

For higher dimensions, the topological problem is still open for now. In particular, it is not even known which kind of manifolds supports Anosov diffeomorphisms. It is a conjecture that a closed manifold that supports it is homeomorphic to an infranilmanifold, which is, in some sense, a generalization of the torus.

Anyway. Since we are interested in surface Anosov diffeomorphisms $f : M \to M$, we will, without loss of generality, assume that $M = \mathbb{T}^2$ and that he is conjugated to linear Toral automorphism f_A for some hyperbolic matrix $A \in SL(2, \mathbb{Z})$.

A consequence of this classification is that, for our case Holonomies are globally defined:

Corollary 2.3.4. *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is an Anosov diffeomorphism, then for every two unstable leaves $W^u(p)$ and $W^u(q)$ there is a bijective Holonomy map $\mathcal{H} : W^u(p) \to W^u(q)$.*

Proof. Let $h : \mathbb{T}^2 \to \mathbb{T}^2$ be the conjugacy between f and its linearization f_A. Let $e^s \in \mathbb{R}^2$ be the eigenvector of A associated to its contracting eigenvalue. Since the unstable and stable leaves of f_A are (projections of) lines, it follows that the holonomies of f_A between any two points $q_1 \in \mathbb{T}^2$ and $q_2 \in W^s_{f_A}(q_1)$ are given by

$$\begin{array}{rccc} \mathcal{H}_{q_1 \to q_2} : & W^u_{f_A}(q_1) & \to & W^u_{f_A}(q_2) \\ & q & \mapsto & q + te^s \end{array}$$

where $t \in \mathbb{R}$ is given by the distance of q_1 to q_2 in $W^s_{f_A}(q_1)$. Since h is a homeomorphism and $\mathbb{T}^2$ is compact, both h and h^{-1} is uniformly continuous; hence, they send Cauchy sequences in Cauchy sequences. If you recall the definitions of the unstable and stable manifolds (see definition 2.1.3), it means that $h(W_f^\sigma(p)) = W_{f_A}^\sigma(h(p))$ for $\sigma = s, u$.

Thus, for any two unstable leaves $W_f^u(p)$, $W_f^u(q)$ and any $\hat{p} \in W_f^s(p) \cap W_f^u(q)$, the induced map $\mathcal{H}_{p\to\hat{p}} \stackrel{\text{def.}}{=} h^{-1} \circ \mathcal{H}_{h(p)\to h(\hat{p})} \circ h$ is a globally defined Holonomy map for f. By Theorem 2.3.5 ahead, we can always find such $\hat{p} \in W_f^s(p) \cap W_f^u(q)$, which concludes the proof. □

This is of great use because, as said before, Holonomies will be very important for us. In this sense, another consequence of this classification is that, in our case, unstable and stable leaves are dense:

Theorem 2.3.5. *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is an Anosov diffeomorphism, then all its stable and unstable manifolds are dense in $\mathbb{T}^2$.*

Proof. We know that $W_f^\sigma(p) = h^{-1}(W_{f_A}^\sigma(h(p))$ (see the proof of the corollary above). Since h^{-1} is continuous, it suffices to show that the stable and unstable manifolds of f_A are dense. However, the condition that $A \in SL(A, \mathbb{Z})$ with eigenvalues $\lambda_1 < 1 < \lambda_2$ implies that the eigenvectors e^s and e^u of A have irrational angles (there is no non-trivial solution for $e^\sigma \cdot \vec{z} \in \mathbb{Z}$ for $\vec{z} \in \mathbb{Z}$ and $\sigma = s, u$). This, in turn, implies that the projection of the subspaces spanned by e^s and e^u are dense in $\mathbb{T}^2$. □

This result will be used later on in chapter 5 to prove the existence of uniformly bounded Holonomies between points far away (see proposition 5.1.4).

The C^1 Regularity of the Foliations

This subsection enters here as a remark: A foliation $\mathscr{F}$ of a n-dimensional manifold M is a collection of disjoint k-dimensional (k fixed) immersed connected manifolds such that for every $p \in M$ there is $\mathscr{F}_p \in \mathscr{F}$ with $p \in \mathscr{F}_p$.

An atlas $\mathcal{A}$ for a foliation $\mathscr{F}$ is a collection of homeomorphisms $\{\Phi_p : D^k \times D^{n-k} \to M\}$ such that $\Phi(D^k \times \{y\}) \subseteq \mathscr{F}_{\Phi(0,y)}$.

Definition 2.3.6. A foliation $\mathscr{F}$ is said to be of class C^r, $r \geq 1$, if it admits an atlas made of C^r diffeomorphisms.

By Theorem 2.1.4 (and the remarks after it) we have a well defined stable foliation $W^s = \{W^s(p)\}_{p\in M}$ that is continuous. However, this regularity is not optimal for us. There is a weaker notion of regularity for foliations:

Definition 2.3.7. A foliation $\mathscr{F}$ is said to be weak-C^r, $r \geq 1$, if every leaf $\mathscr{F}_p \in \mathscr{F}$ is a C^r immersed manifold and every local holonomy map between its leaves is of class C^r.

It is immediate that C^r regularity implies in weak-C^r regularity. But, it is not true, in general, that weak-C^r regularity implies C^r regularity. However, in Theorem 6.1 of [PSW00] they prove that $C^{r+\alpha}$ and weak-$C^{r+\alpha}$ are actually equivalent if the holonomies are uniform. Theorem 2.1.4 already says that the leaves are C^r, and in Proposition 4.2.7 we give a proof that the holonomies are of class $C^{1+\alpha}$ (with uniformly bounded Jacobian). Thus

Theorem 2.3.8. *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is an Anosov diffeomorphism of class C^2, then its stable and unstable foliations are $C^{1+\alpha}$.*

The extra regularity that we obtain here is a low-dimensional phenomena: to obtain the regularity of the stable foliation, we need the regularity of the stable holonomies, which we'll prove using the affine parameters (see prop. 4.2.1). Those require the unstable leaves to be one-dimensional. Since the argument is symmetric, to obtain the regularity for the unstable foliation (by our methods), we also need the stable leaves to be one-dimensional.

CHAPTER **3**

The Ergodic Theory of Anosov Diffeomorphisms

In this chapter I will present a brief survey on Ergodic Theory. This is a major field of dynamics and is of fundamental importance for the understanding of what follows in the following chapters.

Ergodic Theory is the study of dynamics under the view of measures. It's principal concepts are those of invariant measures and ergodicity. Its roots were set not so long ago (less than 200 years) when, to establish a wonderful result about the energy of a gas, Boltzmann supposed that the state of a system would equally float around all accessible states so that the time averages of any measurable property of the system would equal its average under all states.

This so-called 'Ergodic Hypothesis' is not true in all generality; however, it drew the attention of many. Over time this hypothesis unfolded a grand scope of applicability and culminated in a formal statement known as Birkhoff's Ergodic Theorem, which characterizes exactly when the hypothesis is true.

3.1 The Ergodic Theorem

Birkhoff's Theorem

For this section, $(X, \mathcal{B})$ is a measurable space. I.e. X is a set, and $\mathcal{B}$ is a sigma algebra, and $f : X \to X$ is a measurable function. The following starts the theory:

Definition 3.1.1. A measure μ in $(X, \mathcal{B})$ is said to be f invariant if for every measurable set $A \in \mathcal{B}$ we have that $\mu(A) = \mu(f^{-1}A)$.

As always, let's interpret f as some physical process: the dynamics f takes a state $x \in X$ in the present and returns its evolution $f(x)$ one unit of time in the future. With this in mind, the meaning of an invariant measure is that the chance $\mu(A)$ of finding a state in a configuration A is the same chance $\mu(T^{-1}A)$ of finding a state $x \in f^{-1}(A)$ that in a unit of time will be in A.

The pair (f, μ) is said to be a measurable system. Some examples are:

Example 3.1.2. *The pair* (T, Leb) *where* $T : \mathbb{R}^n \to \mathbb{R}^n$ *is a diffeomorphism with* $JacT = 1$ *and* Leb *is the Lebesgue measure in* $\mathbb{R}^n$*. It follows directly from the change of variables formula:*

$$\mathrm{Leb}(A) = \int_A \mathrm{d}\,\mathrm{Leb} = \int_{T^{-1}A} JacT \mathrm{d}\,\mathrm{Leb} = \int_{T^{-1}A} \mathrm{d}\,\mathrm{Leb} = \mathrm{Leb}(T^{-1}A)$$

Example 3.1.3. *The pair (T, Leb) where $\mathbb{S}^1$ is the circle identified as $\mathbb{R}$ modulo 1 and $T : \mathbb{S}^1 \to \mathbb{S}^1$ is given by $T(x) = 2x \mod 1$. It may seem strange that T preserves Lebesgue, because T of a (small) set has twice the length of the set. However, notice that the definition of invariant measure requires that* $\mathrm{Leb}(A) = \mathrm{Leb}(T^{-1}A)$ *and not that* $\mathrm{Leb}(A) = \mathrm{Leb}(TA)$*. It happens that for every interval $I \subseteq \mathbb{S}^1$, $T^{-1}(I)$ consists of two intervals with half the size of I; consequently, T in fact preserves the measure* Leb.

Example 3.1.4. *The pair $(T, \delta_{p,n})$ where $p \in \mathrm{Per}_n(T)$ is a periodic point and $\delta_{p,n}$ is the sum of diracs along the orbit of p: $\delta_{p,n} \overset{\text{def.}}{=} \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i p}$.*

Example 3.1.4 is also a reflection of a much more interesting behavior: If you consider $\delta = \lim_n \delta_{p,n}$, then it follows directly from the periodicity of p that $\delta = \delta_{p,n}$. But this convergence is not limited to the orbit of p; it also happens that if q is in the stable manifold of p then $\lim_n \delta_{q,n}$ also exists and is equal to this same δ. All those points are said to be in the basin of δ (see definition 3.3.6) and in section 3.3 we study a particular type of measure that have a lot of points in their basin (see Theorem 3.3.7).

Now, we have the most basic definition of the theory and some examples. The first connection of this concept with the dynamics comes in the form of Poincaré's Recurrence Theorem. It says that an invariant measure only 'sees' recurrent points, i.e. for almost every $x \in X$ there is a sequence of iterates $f^{n_1}(x), f^{n_2}(x), \ldots$ such that $f^{n_k}(x) \overset{k\to+\infty}{\longrightarrow} x$. Its proof is simple, so we give it here:

Theorem 3.1.5. *If (f, μ) is a measurable system and μ is a finite Borel measure in a second countable metric space X, then μ-almost every point is recurrent.*

Proof. We first begin by proving the following claim:

Claim 3.1.5.1. For all measurable $A \subseteq X$ with $\mu(A) > 0$ it holds that

$$\mu(\{x \in A | f^n(x) \notin A, \ \forall n \geq 1\}) = 0$$

In fact, suppose this claim was false. Let B be the set above. Since μ is f-invariant, all the sets $B, f^{-1}(B), f^{-2}(B), \ldots$ have the same positive measure. Hence, since μ is finite, they can't be all disjoint. Thus, there are some $n > m \geq 0$ such that $f^{-m}(B) \cap f^{-n}(B) \neq \emptyset$. Consequently

$$f^n(f^{-m}(B) \cap f^{-n}(B)) \subseteq f^{n-m}(B) \cap B \neq \emptyset$$

Take a y in $f^{n-m}(B) \cap B$. This point satisfies $y \in B$ and $f^{n-m}(y) \in B$. But $B \subseteq A$, hence $y \in A$ and $f^k(y) \in A$, where $k \overset{\text{def.}}{=} n - m > 0$. This contradicts the definition of y being in B and proves the claim.

This claim asserts that almost every point in a set returns to itself. To pass from this to recurrence, let's use that X is second countable to take a dense subset $\{x_n\}_{n\in\mathbb{N}} \subseteq X$. The balls $B_{x_n,m} \overset{\text{def.}}{=} B_{\frac{1}{m}}(x_n)$ for $m \in \mathbb{N}$ fixed form a cover of X. By the claim, each $B_{x_n,m}$ has a subset $A_{x_n,m}$ of full measure such that every point in $A_{x_n,m}$ returns to $B_{x_n,m}$. Define A_m by

$$A_m \overset{\text{def.}}{=} \bigcup_{n\in\mathbb{N}} B_{x_n,m}$$

and

$$A \overset{\text{def.}}{=} \bigcap_{m\in\mathbb{N}} A_m$$

Each A_m has full measure; thus, A as full measure. I claim that every $x \in A$ is recurrent. In fact, let V be a neighborhood of x. For a sufficiently small $\varepsilon > 0$, we have that $B_\varepsilon(x) \subseteq V$. Let $m \in \mathbb{N}$ be big enough so that $\frac{1}{m} < \frac{\varepsilon}{2}$, and find a $n \in \mathbb{N}$ such that $d(x_n, x) < \frac{1}{m}$. The ball $B_{x_n,m}$ contains x and is contained in V. By definition of $x \in A$, we have that $x \in A_{x_n,m}$, hence there is a $k > 0$ such that $f^k(x) \in B_{x_n,m} \subseteq V$. Since V was an arbitrary neighborhood of x, we conclude that x is recurrent and the Theorem is proven. □

This Theorem gives a first heuristic of the theory: an invariant measure only sees non-wandering points. It is in favor of the initial intuition that states of a system do not go away. However, there are some qualitative flaws about this theorem. That is, we don't know how much time the points take to return close to themselves. Also, we don't know the behavior of the points along their orbits; we only know that they come back.

Lets try to formalize what we want. Let $C^0(X)$ denote the set of all continuous functions from X to $\mathbb{R}$. Such a continuous function $\varphi \in C^0(X)$ takes a state $x \in X$ of our system and returns a real value $\varphi(x)$ that may be understood as some property of our system that we can observe. For example, if X was an ensemble of gases, some familiar observables would be its pressure P, its volume V, and its temperature T[1]. Given a state $x \in X$, the time average of φ at x is simply its mean value along the orbit of x:

$$\hat{\varphi}(x) = \lim_{n \to +\infty} \frac{\varphi(x) + \varphi(fx) + \cdots + \varphi(f^{n-1}x)}{n} = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(f^i x)$$

when this limit exists. Now, for the space mean, we need to first take a measure μ and then compute $\int \varphi \mathrm{d}\mu$. Also, for this to be a mean, μ should be normalized, i.e. we require that μ is a probability ($\mu(X) = 1$). The ergodic hypothesis is then

$$\text{“}\hat{\varphi}(x) = \int \varphi \mathrm{d}\mu\text{”}$$

At first, there is a drastic ambiguity between these values: The time mean depends on the starting point and the space mean depends on the chosen measure. The major hint relating these values is that for invariant measures $f^{-1} \operatorname{supp} \mu = \operatorname{supp} \mu$ and $\hat{\varphi}(fx) = \hat{\varphi}(x)$, i.e., invariant measures can only 'see' an invariant set and the time mean is invariant under f.

After all this discusion, Birkhoff's Ergodic Theorem enters with the following assertion which is almost what we wanted

Theorem 3.1.6 (Birkhoff's Ergodic Theorem)**.** *Suppose μ is a f-invariant probability measure. Then, for every $\varphi \in L^1(\mu)$, its time mean $\hat{\varphi}$ exists for μ-almost every point in X and*

$$\int \hat{\varphi} \mathrm{d}\mu = \int \varphi \mathrm{d}\mu$$

This theorem is excellent, because it says that you can first take the time mean and then integrate it for the same value. Also, in some sense, it says that an invariant measure is uniform along orbits, thus, it seems to be a combination of measures that only sees 'one orbit' each. Lets try to define this class of measures:

Definition 3.1.7 (Ergodic Measure)**.** An invariant probability measure μ is said to be ergodic if for every measurable subset $A \subseteq X$ that is invariant (in the sense that $f^{-1}A = A$) we have $\mu(A) \in \{0, 1\}$.

[1] Somehow, in many cases, those are all the relevant observables.

If μ is an ergodic measure, we say that (f, μ) is an ergodic system. The support of an ergodic measure is an invariant set and there is no smaller invariant set $A \subseteq \operatorname{supp} \mu$ with intermediate value, i.e. such that $0 < \mu(A) < 1$. In fact, you can prove that $\operatorname{supp} \mu$ will always be the closure of some orbit. This is one way to define an ergodic measure, but, if you may prefer, there are many equivalent definitions for it. Some of interest are the following:

Lemma 3.1.8. *For a f-invariant measure μ, the following is equivalent:*

(1) μ is ergodic.

(2) For every $\varphi \in L^1(\mu)$ invariant (i.e. such that $\varphi(fx) = \varphi(x)$) we have that $\varphi(x)$ is constant for μ-almost every $x \in X$.

(3) For $A \subseteq X$ measurable we have that the mean sojourn time of a point x in A

$$\tau_A(x) \stackrel{\text{def.}}{=} \lim_n \frac{Card\left\{0 \leq j \leq n-1 | f^j(x) \in A\right\}}{n}$$

exists for μ-almost every point and is equal to the measure of A.

Item (3) makes it explicit how an ergodic measure behaves with the dynamics: For a given region $A \subseteq X$, almost every state float around the space spending an amount of time in A proportional to its measure. Item (2) is clearly a strong property, after all it is something that works for every $\varphi \in L^1(\mu)$. If you take in account that the time average $\hat{\varphi}$ of a $\varphi \in L^1(\mu)$ is a f-invariant function and that for X compact we have $C^0(X) \subseteq L^1(\mu)$, we can extract the following case from Birkhoff's Ergodic Theorem

Corollary 3.1.9. *If X is compact and μ is an ergodic measure, then for every observable $\varphi \in C^0(X)$, its time average $\hat{\varphi}(x)$ exists for μ-almost every $x \in X$ and moreover, for these points, $\hat{\varphi}(x) = \int \varphi \mathrm{d}\mu$.*

This was our initial goal. The answer we got is that the Ergodic Hypothesis is true when the measure is Ergodic. Stating it like that seems like we just made an ad hoc: Ergodic measure is a measure satisfying the Ergodic Hypothesis. This is, in fact, what we did. However, the surprising point that we haven't touched yet is that those definitions are very natural. Actually, it happens that every map $f : X \to X$ is ergodic! (For some measure).

Existence of Ergodic Measures

First, let $\mathcal{M}(X)$ be the set of all finite measures in X. It is a vector space thus we can endow it with the weak* topology. Let $\mathcal{P}(X) \subseteq \mathcal{M}(X)$ denote the set of all probability measures in X.

Theorem 3.1.10. *The set $\mathcal{P}(X)$ of all probability measures in X is compact in the weak* topology.*

Lets define the push forward $f_*\mu$ of μ by the measure $f_*\mu(A) \stackrel{\text{def.}}{=} \mu(f^{-1}A)$. With this, an invariant measure (see definition 3.1.1) is precisely a fixed point of the map f_*.

Lemma 3.1.11. *If $f : X \to X$ is continuous, then the map $f_* : \mathcal{M}(X) \to \mathcal{M}(X)$ is continuous in the weak* topology.*

The restriction $f_*|_{\mathcal{P}(X)} : \mathcal{P}(M) \to \mathcal{P}(X)$ is well defined and the set $\mathcal{P}(X)$ is compact and convex, hence, by Tychonoff's fixed point theorem, there exists a $\mu \in \mathcal{P}(X)$ such that $f_*\mu = \mu$. Let $\mathcal{P}_f(X)$ denote the set off all f-invariant measures. What we have shown is that

Theorem 3.1.12. *If $f : X \to X$ is continuous, then there exists an invariant probability measure $\mu \in \mathcal{P}_f(X)$*

Now we know that there exists invariant measures, but we also want ergodic measures. For it, le us recall that a measure ν is said to absolutely continuous with respect to another measure μ if for every measurable set $A \subseteq X$ such that $\mu(A) = 0$ we have $\nu(A) = 0$. To denote it we write $\nu \ll \mu$.

Lemma 3.1.13. *If μ_1 and μ_2 are invariant measures and μ is ergodic, then $\mu_1 = \mu_2$*

Proof. Let φ be a bounded measurable function, then, since μ_2 is ergodic, we have

$$\hat{\varphi}(x) = \int \varphi \mathrm{d}\mu_1$$

for μ_2-almost every point. Since $\mu_1 \ll \mu_2$, it follows that

$$\hat{\varphi}(x) = \int \varphi \mathrm{d}\mu_2$$

is also constant for μ_1-almost every point. Hence

$$\int \varphi \mathrm{d}\mu_1 = \int \varphi \mathrm{d}\mu_2$$

and by the arbitrarily of φ we obtain $\mu_1 = \mu_2$. □

As said before, $\mathcal{P}_f(X)$ is convex. Using the lemma above it follows that the set of ergodic measures are the extrema of $\mathcal{P}_f(M)$:

Lemma 3.1.14. *An invariant measure $\mu \in \mathcal{P}_f(X)$ is ergodic if and only if there is no $t \in (0, 1)$ and $\mu_1, \mu_2 \in \mathcal{P}_f(X)$ with $\mu_1 \neq \mu_2$ such that $\mu = t\mu_1 + (1 - t)\mu_2$.*

Proof. If μ is not ergodic, we can find an invariant set $B \subseteq X$ with intermediate measure $\mu(B) \in (0, 1)$. In particular $\mu(B)$ and $\mu(B^c)$ are both non zero and for every measurable $A \subseteq X$ we can write

$$\mu(A) = \frac{\mu(A \cap B)}{\mu(B)} + \frac{\mu(A \cap B^c}{\mu(B^c)}$$

where both measures $\mu_1(A) \overset{\text{def.}}{=} \frac{\mu(A \cap B)}{\mu(B)}$ and $\frac{\mu(A \cap B^c)}{\mu(B^c)}$ are invariant probability measures. Hence, we have shown that if we cannot write μ like in the statement, μ is ergodic.

Now, suppose that μ is ergodic and that $\mu(A) = t\mu_1(A) + (1 - t)\mu_2(A)$ where $t \in (0, 1)$ and $\mu_1, \mu_2 \in \mathcal{P}_f(X)$. If $\mu(A) = 0$ we must have $\mu_1(A) = 0$ and $\mu_2(A) = 0$. Hence $\mu_1 \ll \mu$ and $\mu_2 \ll \mu$. By the preceding lemma $\mu_1 = \mu$ and $\mu_2 = \mu$. In particular, we showed that $\mu_1 = \mu_2$ and in fact, for μ ergodic, it cannot be written as in the statement. □

By Krein-Milman's Theorem, the set $\mathcal{P}_f(X)$ is the closed convex hull of the set $\mathcal{P}_{erg}(f)$ of all ergodic measures in X with respect to f. In particular, we obtain

Theorem 3.1.15. *If $f : X \to X$ is continuous, then there exists an ergodic measure $\mu \in \mathcal{P}_{erg}(f)$.*

Ergodic Decomposition

Right before definition 3.1.7 I said that invariant measures $seems$ to be combinations of ergodic measures. It is true, and to understand it we have to talk about disintegrations.

The simplest idea if disintegration is the following: Let μ be a measure in X and $B \subseteq X$ be a measurable subset with $\mu(B) \in (0,1)$. Then, for every $A \subseteq X$ measurable, we can write

$$\mu(A) = \mu(A \cap B) + \mu(A \cap B^c) = \frac{\mu(A \cap B)}{\mu(B)}\mu(B) + \frac{\mu(A \cap B^c)}{\mu(B^c)}\mu(B^c)$$

Defining $\mu_x^B(A) \stackrel{\text{def.}}{=} \frac{\mu(A \cap B^c)}{\mu(B)}$ for $x \in B$ and $\mu_x^B(A) \stackrel{\text{def.}}{=} \frac{\mu(A \cap B^c)}{\mu(B^c)}$ for $x \in B^c$ we obtain a family of probability measures $\{\mu_x^B\}_{x \in X}$ such that $\mu_x^B = \mu_y^B$ if x, y are both in B or both in B^c satisfying

$$\mu(A) = \int \mu_x^B(A) \mathrm{d}\mu$$

for every $A \subseteq X$ measurable. This family of measure is called the disintegration of μ with respect to the partition $\xi = \{B, B^c\}$ of X. This procedure can be generalized has follows:

Definition 3.1.16. An atmost countable partition of X is a collection ξ of atmost countably many pairwise disjoint measurable subsets $B \subseteq X$ such that $X = \cup_{B \in \xi} B$.

Given a partition ξ of X and some point $x \in X$, there is only one $B \in \xi$ such that $x \in B$. We denote this B by $\xi(x)$ and we call it the atom of ξ at x. Also, given two partitions ξ_1 and ξ_2 we denote their common refinement $\xi_1 \vee \xi_2$ by the partitions whose atoms are the intersection of their atoms:

$$\xi_1 \vee \xi_2 = \{B_1 \cap B_2 | B_1 \in \xi_1 \text{ and } B_2 \in \xi_2\}$$

Also, if for every $x \in X$ the atom $\xi_1(x)$ is contained in the atom $\xi_2(x)$ we say that ξ_1 is thinner than ξ_2 and we denote it by $\xi_2 \prec \xi_1$.

The simple example I gave before can very easily be extended to a general atmost countable partition ξ instead of the simplest possible choice of $\xi = \{B, B^c\}$. What is not so simple, but very useful, is that we can extend it to a class of very more general partitions:

Definition 3.1.17 (Measurable Partition)**.** A collection ξ of subsets of X is said to be a measurable partition if there is a sequence $\xi_1, \xi_2, \ldots$ of atmost countable partitions of X such that $\xi = \xi_1 \vee \xi_2 \vee \cdots \stackrel{\text{def.}}{=} \vee_{i=1}^{+\infty} \xi_i$.

Measurable partitions are way more general than countable partitions, e.g.

Example 3.1.18. *If X is a separable metric space, then the point partition $\xi = \{\{x\}\}_{x \in X}$ of X will be a measurable partition. To see it, let $\{U_i\}_{i \in \mathbb{N}}$ be a countable basis for the topology of X. Since X is a metric space, it is Hausdorff, thus we have $\{x\} = \cap_{x \in U_i} U_i$ for all $x \in X$. If we define $\xi_i \stackrel{\text{def.}}{=} \{U_i, X \setminus U_i\}$, them ξ_i is a finite partition of X for each $i \in \mathbb{N}$ and $\xi \stackrel{\text{def.}}{=} \vee_{i=1}^{+\infty} \xi_i$ is a measurable partition satisfying $\xi(x) = \cap_i \xi_i(x) \subseteq \cap_{x \in U_i} \xi_i(x) = \cap_{x \in U_i} U_i = \{x\}$ as desired.*

Example 3.1.19. *The partition of $\mathbb{R}^2$ in vertical lines $\{\{x\} \times \mathbb{R}\}_{x \in \mathbb{R}}$ is a measurable partition. To see it consider the intervals $I_{n,i} = [\frac{i}{n}, \frac{i+1}{n})$ for $n \in \mathbb{N}$ and $i \in \mathbb{Z}$. It is clear that each $\xi_n \stackrel{\text{def.}}{=} \{I_{n,i} \times \mathbb{R}\}_{i \in \mathbb{Z}}$ is a countable partition of $\mathbb{R}^2$ whose atoms are vertical columns of width $2/n$. Thus $\xi \stackrel{\text{def.}}{=} \vee_{i=1}^{+\infty} \xi_i$ is a measurable partition and its atoms are vertical lines. The same is true for the partition of $\mathbb{R}^n$ in k-dimensional parallel planes.*

Example 3.1.20. *The partition associated with a foliated box B (see definition 2.2.1) is the partition ξ whose atoms for every point of B are segments of unstable leaves and just $M \setminus B$ for points outside B. This partition is measurable because you are assuming that B is parametrized by a map in $\mathbb{R}^n$ that sends $\dim E^u$-dimensional planes in segments of unstable leaves, so by a local argument, the last example applies here.*

For both these examples, there are uncountably many atoms in the final partitions; thus, it is not clear how we could disintegrate a measure in it (atleast, our previous method does not work anymore). Rokhlin's Theorem comes with great news:

Theorem 3.1.21 (Rokhlin's Disintegration)**.** *If ξ is a measurable partition of X and μ is a measure in X, then there exists a family of probability measures $\{\mu_x^\xi\}_{x\in X}$ such that*

(1) For almost every $x \in X$ their support is their atom: $\mu_x^\xi(\xi(x)) = 1$.

(2) For $y \in \xi(x)$ the components of the disintegration agree: $\mu_x^\xi = \mu_y^\xi$.

(3) μ is a convex combination of them: For every $A \subseteq X$ measurable we have $\mu(A) = \int \mu_x^\xi(A)\mathrm{d}\mu(x)$

Moreover, this family is almost unique in the sense that if $\{\nu_x^\xi\}_{x\in X}$ is another family satisfying these conditions, then $\nu_x^\xi = \mu_x^\xi$ for μ-almost every $x \in X$.

Now we can disintegrate measures in measurable partitions. This will be the main tool used in section 4.3 of chapter 4. For now, we conclude this subsection stating the following

Theorem 3.1.22. *If $f : X \to X$ is continuous, X is a complete separable metric space and μ is a f-invariant probability measure, then there is a suitable choice of measurable partition ξ of X such that μ-almost every component μ_x of its disintegration $\{\mu_x^\xi\}_{x\in X}$ is ergodic.*

3.2 Entropy and Pressure

In the section above, we defined what ergodic measures are and showed that they exist. However, the proof of their existence was not constructive; thus, even though they behave very well with the dynamics within their support, we don't know how much they are actually measuring.

To be more clear, consider as an example a toral automorphism $f_A : \mathbb{T}^n \to \mathbb{T}^n$. It's easy to see that it has a lot of periodic points; in particular, the origin $\vec{0} \in \mathbb{T}^n$ is a fixed point. Hence the Dirac $\delta_{\vec{0}}$ at $\vec{0}$ is ergodic (an invariant set either contains $\vec{0}$ or not). But, in the other side, the Lebesgue measure Leb in $\mathbb{T}^n$ is also ergodic:

Proposition 3.2.1. *For a linear Anosov toral automorphism $f_A : \mathbb{T}^n \to \mathbb{T}^n$, the Lebesgue measure $\mathrm{Leb}_{\mathbb{T}^n}$ is ergodic.*

Proof. Let $\varphi \in L^2(\mathrm{Leb}_{\mathbb{T}^n})$ and write

$$\varphi(\vec{x}) = \sum_{\vec{m}\in\mathbb{Z}^n} \varphi_{\vec{m}} e^{2\pi i \vec{m}\cdot\vec{x}}$$

where the terms $\varphi_{\vec{m}}$ are the Fourier components of φ. Composing with f_A, we have

$$\varphi \circ f_A(\vec{x}) = \sum_{\vec{m}\in\mathbb{Z}^n} \varphi_{\vec{m}} e^{2\pi i \vec{m}\cdot \vec{Ax}} = \sum_{\vec{m}\in\mathbb{Z}^n} \varphi_{\vec{m}} e^{2\pi i A^{\dagger}\vec{m}\cdot\vec{x}} = \sum_{\vec{m}\in\mathbb{Z}^n} \varphi_{\vec{m}} e^{2\pi i A^{\dagger}\vec{m}\cdot\vec{x}}$$

If we suppose that φ is invariant, we have $\varphi \circ f = \varphi$, so their Fourier series must agree, and we have $\varphi_{\vec{m}} = \varphi_{A^{\dagger}\cdot\vec{m}}$. By induction, if $k \in \mathbb{N}$, we must have $\varphi_{\vec{m}} = \varphi_{(A^{\dagger})^k\cdot\vec{m}}$. Since A is hyperbolic, if $\vec{m} \neq \vec{0}$, we have that $\|(A^{\dagger})^k\vec{m}\| \to +\infty$ as $|k| \to +\infty$. Using that $\varphi \in L^2$, we know that $\varphi_{\vec{m}} \to 0$ as $\|\vec{m}\| \to 0$; thus, for $\vec{m} \neq \vec{0}$ we have

$$\varphi_{\vec{m}} = \lim_{n\to+\infty} (A^{\dagger})^k\vec{m} = 0$$

and we conclude that

$$\varphi(\vec{x}) = \varphi_{\vec{0}}$$

is constant. By lemma 3.1.8, we prove the proposition. □

Even though they are both ergodic, the reason for them being so is very different. The Dirac delta $\delta_{\vec{0}}$ is ergodic simply because its support is so tiny that there is no difficulty in behaving well with the dynamics. However, the Lebesgue measure $\mathrm{Leb}_{\mathbb{T}^n}$ has full support; thus, the fact that it is ergodic is not trivial at all, and to know it actually says a lot about the dynamics in the entire manifold.

The contrast above highlights a little problem that we must face now: sometimes there are too many ergodic measures!

In fact, for many systems f, the set of ergodic measures $\mathcal{P}_{erg}(f)$ is dense in the (usually very big) set of invariant measures $\mathcal{P}_f(X)$. In any case, we must be more specific with what we are measuring. In this section we define the entropy of a system and its generalization, the pressure. Those are numbers that, in some sense, measure the complexity of the dynamics. We then define a measure-theoretic analogue, and we relate them using the variational principle. This allows us to define a measure of maximal entropy, which are those that capture the entire complexity of the dynamics.

Topological Entropy

We will measure the complexity of the dynamics by counting the number of orbits in X. Of course, (if X is not countable), there are infinitely many orbits, so we need to be very precise in what we mean by ‘counting orbits’.

Imagine that X represents a set of states of a physical system that you can watch. You, as an observer, cannot fully trust in your eyes, which are imperfect. Hence, if two states are too close to each other, you cannot distinguish them. To express this, given an imprecision $\varepsilon > 0$, we say that a subset E of X is ε-separated if the distance of any pair of points $x, y \in E$ is greater than ε.

Also, there is a dynamic occurring, so the states are evolving with time. Since you are not immortal, nor would you have the patience to wait forever, you can only count finitely many iterates of a state. If you were tracking the iterates of some state, and suddenly you see it splitting in two, then you know that you were actually looking at two states. Hence, even though two states may initially be very close, if they eventually get apart from each other at some moment, you can spot it. To express this, given a time $n \in \mathbb{N}$, we introduce the time n metric d_n by

$$d_n(x,y) = \max\{d(f^i x, f^i y) | 0 \leq i < n\}$$

for every $x, y \in X$. And, for a subset E of X, we say that E is (n, ε)-separated if E is ε-separated for the metric d_n. With this notation, the number $s(n, \varepsilon)$ of orbits that we can distinguish at time n is given by

$$s(n, \varepsilon) = \max\{Card(E)|E \text{ is } (n, \varepsilon)\text{-separated}\} \tag{3.1}$$

As you let the time $n \in \mathbb{N}$ pass, every two states have more and more chance to separate more (i.e. the metric d_n is increasing with n). Hence, the number $s(n, \varepsilon)$ of distinguished orbits increases with n. If it increases slowly, the system is simple. There are many possible rates of growth that we could consider; for reference see [CP24] Theorem 1.2. We are interested in (very) chaotic dynamics; hence, we will care about the exponential rate of growth of $s(n, \varepsilon)^2$:

$$r(\varepsilon) \stackrel{\text{def.}}{=} \limsup_n \frac{1}{n} \log(r(n, \varepsilon))$$

Similarly, if we increase the precision, $r(\varepsilon)$ grows (i.e. if $\varepsilon_1 < \varepsilon_2$, then $r(\varepsilon_2) \leq r(\varepsilon_1)$. Hence, since we want the finest measurement possible, we should take the limit as $\varepsilon \to 0$. This limit will be the entropy:

Definition 3.2.2 (Topological Entropy)**.** The topological entropy of a continuous map $f : X \to X$ defined on a compact metric space X is the number

$$\mathrm{h_{top}}(f) \stackrel{\text{def.}}{=} \lim_{\varepsilon \to 0} s(\varepsilon) = \lim_{\varepsilon \to 0} \limsup_n \frac{1}{n} \log(s(n, \varepsilon))$$

when this limit is finite or (the symbol) $+\infty$ when it is not.

The most important property of entropy is that it is a topological invariant:

Theorem 3.2.3. *If $f_1 : X_1 \to X_1$ and $f_2 : X_2 \to X_2$ are continuous maps on compact metric spaces and $\Phi : X_1 \to X_2$ is a homeomorphism satisfying $\Phi \circ f_1 = f_2 \circ \Phi$, then $\mathrm{h_{top}}(f_1) = \mathrm{h_{top}}(f_2)$.*

In particular, the entropy $\mathrm{h_{top}}(f)$ only depends on the topology of X and not on the metric d that we used to compute it. Talking about computing it, that's very hard from the definition; thus, in general, one must find different equivalent methods to do it. In particular, to ease the exposition, I omitted the non-compact case and completely ignored the definition via (n, ε)-spanning sets that is generally given together with ours. For further details, see [Wal00] chapter 7.

Measure Theoretic Entropy

Now, we will once again try to measure the complexity of our system, but this time we will try to guess its behavior beforehand. In the subsection above, we sought orbits without caring about their trajectory. To take their position into account, we'll partition our space X into smaller pieces and distribute probabilities on them.

A guess here will be represented as a choice of a probability measure $\mu \in \mathcal{P}(X)$. Given a finite partition $\xi = \{A_1, \ldots, A_n\}$ of X, the numbers $\mu(A_i)$ may be seen as (what we suppose to be) the likelihood of looking at our system and finding it in a state $x \in A_i$. Let's represent the 'uncertainty' of this

[2] That's not the only reason to choose the exponential rate. It will be more clear when we relate the topological entropy with the the measure theoretic entropy. To define the latter, we require some properties that lead to a unique possible expression (Theorem 3.2.4). This expression is naturally related to the topological entropy defined here using the exponential rate of growth (Variational Principle3.2.10).

guess by a number $H_\mu(\xi)$. We identify the entropy of a partition as this value. What kind of properties should this function $H_\mu(\cdot)$ satisfy?

First, the uncertainty of a guess only depends on the probability I gave to each set and not on the set itself, i.e. if $\xi = \{A_1, \ldots, A_n\}$, then $H_\mu(\xi) = H(\mu(A_1), \ldots, \mu(A_n))$. Similarly, this uncertainty H should not depend on the order in which I told you the sets; after all, for $\xi = \{A_1, A_2, A_3, \ldots, A_n\}$ we have $\xi = \{A_2, A_1, A_3, \ldots, A_n\}$ (the set is not ordered), then for any permutation σ of n letters we should have $H(p_1, \ldots, p_n) = H(p_{\sigma(1)}, \ldots, p_{\sigma(n)})$. If an outcome is impossible, it doesn't matter, so that $H(p_1, \ldots, p_n, 0) = H(p_1, \ldots, p_n)$. Moreover, if we are so certain about an outcome that we choose a set with probability 1 while every other with probability 0, there will be no uncertainty in this guess to that $H(0, \ldots, 1, \ldots, 0) = 0$. In the completely opposite case, if we don't know the slightest and take a partition $\xi = \{A_1, \ldots, A_n\}$ with all sets with the same probability $1/n$, it must be the most uncertain, so that $H(1/n, \ldots, 1/n)$ is a maximum.

With the above requirements, we are measuring the entropy of partition, static in time. We must not forget to include the dynamics in the discussion. Let's say we guessed $\xi = \{A_1, \ldots, A_n\}$ and we want to measure it two times. At first, you verify the state x of your system and find it in the configuration $A \in \xi$. Right after that, you'll see in which set A_j will be $f(x)$. However, since you know it was in A, you won't just suppose that its chance of being in A_j is $\mu(A_j)$; you can now refine your measure by restricting it to A, i.e. the chance of being in A_j knowing it was in A will be

$$\mu_A(A_j) \overset{\text{def.}}{=} \frac{\mu(A_j \cap A)}{\mu(A)}$$

if $\mu(A) \neq 0$, or zero if $\mu(A) = 0$. To know that $x \in A_i$ and $f(x) \in A_j$ is to know that $x \in A_i \cap f^{-1}(A_j)$, in other words, is to know in which element of $\xi \vee f^{-1}\xi$ is x. Hence, what we are requiring above is that

$$H_\mu(\xi \vee f^{-1}\xi) = H_\mu(\xi) + \sum_{A \in \xi} \mu(A) H_{\mu_A}(\xi)$$

Thankfully, all this discussion has a reward. We required many properties of H_μ, and, in the end, there does not just exist a function satisfying it, but it is unique!

Theorem 3.2.4. *If $H(p_1, \ldots, p_n)$ is a function defined for every $n \in \mathbb{N}$ and collection of real numbers $p_i \geq 0$ satisfying $\sum_i p_i = 1$ such that*

1. *$H(p_{\sigma(1)}, \ldots, p_{\sigma(n)}) = H(p_1, \ldots, p_n)$ for any permutation σ of n letters.*
2. *For every two finite partitions ξ, ζ of X and $\mu \in \mathcal{P}(X)$ it holds that*

$$H_\mu(\xi \vee \zeta) = H_\mu(\xi) + \sum_{A \in \xi} \mu(A) H_{\mu_A}(\zeta).$$

3. *For n fixed, $H(p_1, \ldots, p_n)$ has its maximum at $p_1 = \cdots = p_n = 1/n$.*
4. *$H(p_1, \ldots, p_n) \geq 0$ with equality if and only if some p_i is 1.*
5. *$H(p_1, \ldots, p_n)$ is continuous for n fixed.*
6. *$H(p_1, \ldots, p_n, 0) = H(p_1, \ldots, p_n)$*

Then there exists a $\lambda > 0$ such that $H(p_1, \ldots, p_n) = -\lambda \sum_{p_i \neq 0} p_i \log(p_i)$.

Proof. See [Khi57], page 9. □

This lemma says that there is a canonical way of measuring the entropy of a partition:

Definition 3.2.5. For a finite partition ξ of X its entropy $H_\mu(\xi)$ with respect to a measure $\mu \in \mathcal{P}(X)$ is given by

$$H_\mu(\xi) \stackrel{\text{def.}}{=} -\sum_{\substack{A\in\xi \\ \mu(A)\neq 0}} \mu(A)\log(\mu(A)).$$

Now, similarly as before, we will now define the entropy of our map f with respect to a partition ξ and measure μ as the growth rate of this number when you keep measuring it indeterminately

$$H_\mu(f,\xi) \stackrel{\text{def.}}{=} \limsup_n \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} f^{-i}\xi\right)$$

And finally, the entropy of f with respect to this measure is the supremum over all partitions of the number above:

Definition 3.2.6 (Measure Theoretic Entropy)**.** The metric entropy of a measurable map $f : X \to X$ defined on a measurable space $(X, \mathscr{B})$ with respect to a measure $\mu \in \mathcal{P}(X)$ is the number

$$\mathrm{h}_\mu(f) \stackrel{\text{def.}}{=} \sup_{\substack{\xi \text{ is a} \\ \text{finite partition}}} H_\mu(f,\xi) = -\sup_{\substack{\xi \text{ is a} \\ \text{finite partition}}} \limsup_n \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} f^{-i}\xi\right)$$

when this limit is finite or (the symbol) $+\infty$ when it is not.

Just like for topological entropy, this number is an invariant:

Theorem 3.2.7. *If $f_1 : (X_1,\mu_1) \to (X_1,\mu_1)$ and $f_2 : (X_2,\mu_2) \to (X_2,\mu_2)$ are measure-preserving maps and $\Phi : (X_1,\mu_1) \to (X_2,\mu_2)$ is an invertible measure-preserving map such that $\Phi \circ f_1 = f_2 \circ \Phi$, then $\mathrm{h}_{\mu_1}(f_1) = \mathrm{h}_{\mu_2}(f_2)$.*

Variational Principle

Those two methods of measuring entropy have their differencee: the topological entropy $\mathrm{h}_{\mathrm{top}}(f)$ is uniquely defined for a given map f, while its metric entropy $\mathrm{h}_\mu(f)$ depends on the measure $\mu \in \mathcal{P}_f(X)$. Since there is so much more freedom available for the metric entropy, before we relate them, we may extend the notion of topological entropy to a more general one.

Just like we did for the measure theoretic case, we will count states caring about the place in X that they are. Recall the equation (1) that defines the amount $s(n,\varepsilon)$ of distinguishable orbits of length n and precision ε. Notice that it can be written as

$$s(n,\varepsilon) = \sup\left\{\sum_{x\in E} 1 | E \text{ is } (n,\varepsilon)\text{-separated}\right\}$$

Introduce some observable $\varphi \in C^0(X)$. Instead of summing 1 for each point x in E, we will now consider the output of the observable φ by summing e^φ along its orbit (see the next subsection for the reason why)

$$s(n,\varepsilon,\varphi) \stackrel{\text{def.}}{=} \sup\left\{\sum_{x\in E} e^{\sum_{i=0}^{n-1}\varphi(f^i x)} | E \text{ is } (n,\varepsilon)\text{-separated}\right\}$$

Now, just like before, we'll take the exponential rate of growth of this number and make the limit as ε goes to zero. This limit is called the pressure of f with respect to φ:

Definition 3.2.8 (Pressure)**.** The pressure of a continuous map $f : X \to X$ with respect to an observable $\varphi \in C^0(X)$ is the number

$$P(f, \varphi) \stackrel{\text{def.}}{=} \lim_{\varepsilon \to 0} \limsup_{n} \frac{1}{n} \sup \left\{ \sum_{x \in E} e^{\sum_{i=0}^{n-1} \varphi(f^i x)} | E \text{ is } (n, \varepsilon)\text{-separated} \right\}$$

Remark 3.2.9. The topological entropy is the pressure of f with respect to the constant observable: $P(f, 0) = \mathrm{h}_{\mathrm{top}}(f)$.

Given this way more general definition, the Variational Principle comes as a connection between $\mathrm{h}_\mu(f)$ and $P(f, \cdot)$:

Theorem 3.2.10 (Variational Priciple)**.** *For a continuous map $f : X \to X$ of a compact metric space X and $\varphi \in C^0(X)$, we have*

$$P(f, \varphi) = \sup \left\{ \mathrm{h}_\mu(f) + \int f \mathrm{d}\mu \middle| \mu \in \mathcal{P}_f(X) \right\}$$

Now we may define our special kind of measures:

Definition 3.2.11. A measure $\mu \in \mathcal{P}_f(X)$ is said to be an equilibrium state for an observable $\varphi \in C^0(X)$ if it attains the supremum in Theorem 3.2.10, i.e. if

$$P(f, \varphi) = \mathrm{h}_\mu(f) + \int f \mathrm{d}\mu$$

In particular, μ is said to be a measure of maximal entropy if it is an equilibrium state for the potential $\varpi \equiv 0$.

In chapter 5 we show that every Hölder continuous observable admits an equilibrium state. Moreover, this equilibrium state is a very well-behaved measure.

The Physical Meaning of Counting Orbits

Throughout this section, I tried to motivate the definitions of entropy by saying we are 'counting orbits'. However, when defining both metric entropy and Pressure, there were some arbitrary choices we made. In the latter, the meaning of 'counting orbits' was affected by the choice of measure, and in the former, it was influenced by the exponential of a potential. In this subsection I give a brief background on Thermodynamics and statistical mechanics, which is where lies the physical concepts of these definitions.

The starting point is energy: every physical system has an associated internal energy U. This energy U satisfies the following axiom:

Axiom 3.2.12 (First Law of Thermodynamics)**.** *For a quasi-stationary transformation, the energy U of a system satisfies the following differential equation:*

$$\mathrm{d}U = \mathrm{d}Q - \mathrm{d}W$$

where Q is the heat absorbed by the system and W the work done by the system.

This axiom introduces the idea of Heat and Work; however, it does not precisely define how they are nor how to compute them. This is very similar to the well-known Newton's third law, -$F = m\ddot{r}$- which makes reference to forces without explicitly showing their expression.

In practice, these quantities will depend in how we choose to model our system. The work W is supposed to represent any purely mechanical way of transferring energy, while the heat Q is to account for the thermal ways to do it (in other words, to account for what we can't explain). For example, the work to compress a balloon of helium is $\mathrm{d}W = -P\mathrm{d}V$, where P is its pressure and V the volume displaced.

Another word that this axiom introduces is the 'quasi-stationary transformation'. This comes from the following axiom, which in general is implicitly assumed within the many axioms of thermodynamics:

Axiom 3.2.13. *For every thermodynamical system, there exists a finite number of observable quantities that completely describe their equilibrium state.*

A thermodynamical system will always go to an equilibrium state, i.e. a state that does not undergo any transformations as long as it is isolated. The axiom above says that this final state is uniquely defined as long as you know a finite amount of properties of your system. The main example of the theory is the ideal gas: an ideal gas is a collection of non-interacting free particles. This system can be described by three observables: its pressure P, its volume V, and its temperature T.

It was using this system that the French military Sadi Carnot showed that the most efficient possible thermodynamical cycle (a repeating process that extract the most heat using the minimum work) depends only on the initial and final temperatures. Later Rudolf Clausius showed that in an invertible cycle, the heat absorbed $\mathrm{d}Q_1$ and the heat emitted $\mathrm{d}Q_2$ satisy

$$\frac{\mathrm{d}Q_1}{T_1} = \frac{\mathrm{d}Q_2}{T_2}$$

while in a non-invertible cycle $\frac{\mathrm{d}Q_1}{T_1} < \frac{\mathrm{d}Q_2}{T_2}$. This led him to define the entropy (from Greek entropêe: change) of the system as $S \overset{\text{def.}}{=} Q/T$. This quantity seams to never decrease (experimentally), so the following axiom was declared

Axiom 3.2.14. *The entropy S of a system is non-decreasing:* $\mathrm{d}S \geq 0$. *Moreover,* $\mathrm{d}S > 0$ *if and only if the system suffers a non-invertible transformation.*

For us, this is enough thermodynamics. Now, it is time to connect this independent 'heat' theory with mechanics. The person that did this was Boltzmann and his idea was the following:

Suppose that you can distinguish every possible state of your system (up to a microscopic resolution) and let Ω be the collection of all these micro-states. If you have an observable $\varphi : \Omega \to \mathbb{R}$, the set of all states in Ω such that φ is equal to some fixed value $a \in \mathbb{R}$ is said to be a macro-state. What is the most probable macro-state of your system?

Well, if you don't know anything about how the system evolves but you know Ω, you can do an educated guess. Let $\Omega(\varphi = a) \subseteq \Omega$ be the subset of Ω consisting of all micro-states of your system such that φ is equal to a ($\Omega(\varphi = a) = \varphi^{-1}(a)$). With this definition, the cardinality of $\Omega(\varphi = a)$ is the number of states where $\varphi = 0$. In particular, a good guess is that the most probable value of φ is the one that maximizes the number of available states in $\Omega(\varphi = a)$.

We now found something interesting. This quantity, $card(\Omega(\varphi(\cdot)))$ (as a function of the state), is being maximized. This, which is a defining property of the entropy, led Boltzmann to the following claim:

Axiom 3.2.15 (The Fundamental Law of Statistical Mechanics)**.** *The entropy of a system is given by* $S = k_B \log(card(\Omega))$.

This k_B is just a constant of proportionality (Boltzmann's constant), and the logarithm appears here for the same reasons as in Theorem 3.2.4. The entropy is a function of the values of the observables, i.e. $S = S(\varphi_1 = a_1, \ldots, \varphi_n = a_n)$. This axiom comes with a natural distribution of probability

Definition 3.2.16. The probability of an observable $\varphi : \Omega \to \mathbb{R}$ to assume a value $a \in \mathbb{R}$ is given by

$$P(\varphi = a) = \frac{card(\Omega(\varphi = a))}{card(\Omega)}$$

This definition is self-justifying by the phrase 'The most probable state is the most probable state'.

Now, let us connect it with what we have done in this section. First, let's see why we added the exponential of the potential in the definition of equilibrium states.

Suppose you have a system whose the work done by it is given by $\mathrm{d}W = \varphi \mathrm{d}\eta$, where φ and η are some observables. If it undergoes a transformation while we maintain its energy U constant, the first law of thermodynamics says that

$$0 = T\mathrm{d}S - \varphi \mathrm{d}\eta$$

Thus, up to order 1, we have

$$S = constant + \frac{\varphi}{T}\eta$$

Using the fundamental law of statistical mechanics, we obtain that $P = ce^{\frac{S}{k_B}}$, where c is a constant. Hence $P = Ce^{\frac{\varphi}{k_B T}\eta}$. In a laboratory, you may be able to fix the temperature T and this property η. Then, the probability of a state is given by $P = c_0 e^{\varphi}$. Up to a constant, this is our definition in 3.2.8.

Now, let's understand the connection of this with our definition of metric entropy.

This connection follows from an issue that I hid from you: all the construction above assumes that you can count each possible state in Ω one by one. And moreover, it assumes that $card(\Omega)$ is finite. This, however, is not the case.

There isn't a unique way of defining Ω, as an example: imagine you have a system made of two identical particles. At the beginning, one particle is at your left and the other at your right. This is a configuration of your system. But now, suppose you swap them: the particle in the right goes to the left, and the particle in the left goes to the right. Is this a new configuration? Or is it just the same state?

This question has no definite answer: if you are modeling the system with classical mechanics, then it is a new state. If you are modeling it with quantum mechanics, then no, it is just the same state. This is just an example, but there are many other cases where this happens. Another problem is that Ω may have infinitely many states; in fact, in classical mechanics it will generally be a submanifold of some $\mathbb{R}^n$. For these cases, there is no natural way to count the states; you may either discretize it or use some volume form to measure their size. These different choices are said to define different ensembles for your system, and this freedom is what accounts for the different possible measures in the definition of metric entropy.

3.3 Smooth Ergodic Theory

For this section, M is a Riemann manifold and Leb is the Lebesgue measure in M.

Smooth Ergodic Theory is the branch of ergodic theory that tries to explore invariant smooth measures, i.e. measures absolutely continuous with respect to the volume of your space. However, since absolutely continuous invariant measures are not always available, we need a more suitable class of measures that, in some sense, are still relatable to them:

Definition 3.3.1. A measure $\mu \in \mathcal{P}_f(M)$ is said to be a SRB measure if for every foliated box B (see 2.2.1), the components in B of its disintegration $\{\mu_p^{\xi_B}\}_{p\in M}$ with respect to the partition associated to B (see example 3.1.20) are absolutely continuous with respect to the volume in the leaves. I.e. for all $p \in B$

$$\mu_p^{\xi_B} << \mathrm{Leb}_{W^u(p)}$$

What this definition says is that SRB are measures that are 'smooth along unstable manifolds'. This class of measures is good because they come with many fine properties. Also, they are good because they exist:

Theorem 3.3.2. *If $f : M \to M$ is an Anosov diffeomorphism of class C^2, then there exists a SRB measure μ for f.*

Sketch of the proof. Take any point $p \in M$ and let $D \subseteq W^u_{loc}(p)$ be a small open disk around p contained in its local unstable manifold. If Leb_D is the Lebesgue measure induced in D, let

$$\mu_{n,D} \stackrel{\text{def.}}{=} \frac{1}{n}\sum_{i=0}^{n-1} \frac{f_*^j\,\mathrm{Leb}_D}{f_*^j\,\mathrm{Leb}_D(D)}$$

where f_* is the pushforward by f of a measure. Each of those $\mu_{n,D}$ has absolutely continuous components in the unstable direction; however, are not invariant. With a suitable distortion argument, one can show that any weak* accumulation point μ of the sequence $\{\mu_{D,n}\}_{n\in\mathbb{N}}$ is still absolutely continuous in the unstable direction. Hence, since any accumulation point of this sequence is invariant, μ is a SRB measure. (see chapter 11 of [BDV06]).

The fact that those measures are very regular in the unstable leaves will allow us to obtain precise quantitative estimates for their disintegrations (see. 4.3.17). To see it, instead of using the generic expansion rate λ defined in 2.1.1 we will need a more precise one:

Definition 3.3.3. For $p \in \mathbb{T}^2$, let $\lambda_p^\sigma \stackrel{\text{def.}}{=} |d_pf|_{E^\sigma(p)}|$, and for $n \in \mathbb{Z}$, also define $\lambda_p^\sigma(n) \stackrel{\text{def.}}{=} |d_pf^n|_{E^\sigma(p)}|$, where $\sigma = s, u$.

These pointwise expansion rates λ_p^σ give way more detail than the constant λ. In particular, you can recover λ by setting $\lambda \stackrel{\text{def.}}{=} \min\{\inf_p \lambda_p^u, \sup_p(\lambda_p^s)^{-1}\}$.

In fact, these values are so precise that they are sufficient to characterize SRB measures:

Theorem 3.3.4 (Ledrappier-Young's formula)**.** *Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be an Anosov diffeomorphism; then, for every invariant measure $\mu \in \mathcal{P}_f(M)$, you have*

$$\mathrm{h}_\mu(f) \leq \int \lambda_p^u \mathrm{d}\mu(p)$$

And the equality holds if and only if μ is SRB.

This quantity on the right is called the mean Lyapunov exponent of f:

Definition 3.3.5. The Lyapunov exponent of f at a point $p \in M$ is the following limit

$$\lambda_f^u(p) \overset{\text{def.}}{=} \lim_{n\to+\infty} \frac{1}{n} \log \lambda_p^u(n)$$

Which, by Birkhoff's ergodic theorem, exists for almost every point. The mean Lyapunov exponent of f with respect to μ is

$$\lambda_f^u(\mu) \overset{\text{def.}}{=} \int \lambda_f^u(p) \mathrm{d}\mu$$

Both these quantities will be useful for us. In particular, we'll show a particular case of this formula for our context.

Another property of SRB measures, which we will not use but is worth mentioning, is that, in our context, they are physical measures. This means the following:

Definition 3.3.6. The Basin of Attraction of a measure $\mu \in \mathcal{P}_f(M)$ is the set (see example 3.1.4)

$$\mathbb{B}(\mu) = \{p \in M | \delta_{p,n} \overset{w^*}{\to} \mu\}$$

The measure μ is said to be physical if $\mathrm{Leb}(\mathbb{B}(\mu)) > 0$.

A point p being in $\mathbb{B}(\mu)$ means that the measure perfectly describes its orbit, i.e. the time averages at p converge to the measure μ. Hence, since Leb is our reference measure, the condition $\mathrm{Leb}(\mathbb{B}(\mu)) > 0$ means that a physical measure is a measure that describes the orbits of a significant amount of points. Do not confuse with $\mu(\mathbb{B}(\mu)) > 0$. We require the Lebesgue measure of $\mathbb{B}(\mu)$ to be positive and not the μ measure of it. In fact, for ergodic measures, Birkhoff's Ergodic Theorem gives that $\mu(\mathbb{B}(\mu)) = 1$.

If the diffeomorphism is slightly regular, SRB measures coincide with Physical measures

Theorem 3.3.7. *If $f : M \to M$ is an Anosov diffeomorphism of class C^2, then the SRB is physical.*

Proof. Let $\{\mu_p^\xi\}_{p\in M}$ be a disintegration of μ with respect to a partition subordinated to the unstable foliation. Since

$$1 = \mu(\mathbb{B}(\mu)) = \int \mu_p^\xi(\mathbb{B}(\mu))\mathrm{d}\mu(p)$$

it follows that for at least some $p \in M$, we must have $\mu_p^\xi(\mathbb{B}(\mu)) > 0$. Thus, since $\mu_p^\xi << \mathrm{Leb}_p^u$, we obtain a set $I \overset{\text{def.}}{=} W_{loc}^u(p) \cap \mathbb{B}(\mu)$ with $\mathrm{Leb}_p^u(I) > 0$. It is clear from the definition that every point in the unstable manifold of a point in $\mathbb{B}(\mu)$ will also be in $\mathbb{B}(\mu)$. Hence $\bigcup_{q\in I} W_{loc}^s(q) \subseteq \mathbb{B}(\mu)$. Since the foliation is C^1 and the stable and unstable leaves are transversal, Fubini's Theorem gives that

$$\mathrm{Leb}(\mathbb{B}(\mu)) \geq \mathrm{Leb}\left(\bigcup_{q\in I} W_{loc}^s(q)\right) > 0.$$

And the theorem is proven. □

CHAPTER 4

Leafwise Measures

In the previous chapter we introduced equilibrium states and SRB measures, which are both the measures of most interest for us. Here and forwards, we start developing some very special properties of them that will be of great use for us.

In this chapter we focus on the SRB measure. Specifically, we show that associated with them, there exists a family of locally finite measures defined on entire unstable leaves, which, when restricted to a foliated box and normalized, gives you the disintegration of your SRB.

These measures will not be probabilities. However, in turn, we'll see that they behave very well with the dynamics. So well that in a special parameterization of the leaves (the affine parameters 4.2.1), they are the Lebesgue measure (times 0.5).

4.1 Subordinated Partitions

We want to obtain a more global definition for an SRB measure. In the last chapter (definition 3.3.1) we defined them as measures such that the components of their decomposition in foliated boxes were smooth. A silly manner to make this definition not so local is to ask them to be smooth along entire unstable manifolds. This would save us of some trouble, however, unfortunately, in most cases we cannot ask this. The problem is not like they wouldn't be smooth, but that we can't even decompose a measure along entire unstable manifolds.

In lemma 3.2.4 of [Bro+19] it was proven that the partition of our space into unstable manifolds is measurable if and only if our system has zero entropy. But our interest in this text lies in Anosov diffeomorphisms, which have non-zero entropy. Fortunately, we can avoid this problem. Instead of searching for a partition made of entire unstable manifolds, we can look for one such that each atom is a piece of unstable manifold.

Definition 4.1.1 (Subordinated Partition)**.** A measurable partition ξ is said to be subordinated to the unstable foliation W^u with respect to a measure μ if

(i) For μ-a.e.$p \in \mathbb{T}^2$ there is a number $r(p) > 0$ such that $W^u_{r(p)}(p) \subseteq \xi(p)$.

(ii) There exists an $R > 0$, such that for all $p \in \mathbb{T}^2$ we have $\xi(p) \subseteq W^u_R(p)$.

(iii) $\bigvee_{n=0}^{+\infty} f^{-n}(\xi)$ is the point partition of M.

(iv) ξ is increasing: $\xi \prec f^{-1}(\xi)$.

The main concern for this section is with their existence which, thanks to F. Ledrappier and J.-M. Strelcyn, we know that it is granted:

Theorem 4.1.2. *If $f : \mathbb{T}^2 \to \mathbb{T}^2$ is Anosov and $\mu \in \mathcal{P}_f(\mathbb{T}^2)$, then there exists a partition ξ subordinated to the unstable foliation.*

Since this partition is measurable, we can use Rokhlin's Theorem to disintegrate μ with respect to it. The usefulness that we take from this object is that the measure μ is SRB if and only if its components from this disintegration are absolutely continuous with respect to the leaves volume[1].

The construction that I present here follows appendix D of [Bro+19].

An Ansatz for ξ

At first sight, a subordinated partition may seem like a lot of foliated boxes around the entire manifold. And that is almost what they are. If you restrict yourself to a small neighborhood of a point, conditions (i) and (ii) basically say the subordinated partition looks like a foliated box, but with open segments of unstable leaves of non-constant length. Conditions (iii) and (iv) are just what you would expect from the contraction of unstable leaves by f^{-1}.

For each $p \in \mathbb{T}^2$, consider a foliated ball $B_{R(p)}(p)$ (see 2.2.1) of radius $R(p)$ to be defined around it. Since $\mathbb{T}^2$ is compact, we can choose finitely many of these open balls $B_1, B_2, \dots, B_\kappa$ that cover it. Each B_i defines a measurable partition ξ_i whose atoms are unstable leaves at B_i and the entire complement B_i^c outside B_i (see example 3.1.20). Consider their mutual refinement $\xi_0 \overset{\text{def.}}{=} \bigvee_{n=1}^{k} \xi_i$.

This partition ξ_0 already looks like a subordinated partition. In fact, if we choose $R(p) < R$ for a fixed R, conditions (i), (ii) and (iii) already holds:

(i) If a point p is not in the boundary of any B_i, then each $\xi_i(p)$ contains an open interval of p in $W^u(p)$. Thus, there is a $r(p) > 0$ such that $W^u_{r(p)}(p) \subseteq \xi_0(p)$. Also, the boundary of each B_i can be chosen to have zero measure;[2] thus, this $r(p)$ exists for almost every $p \in \mathbb{T}^2$.

(ii) Every point $p \in \mathbb{T}^2$ is in some B_i, hence $\xi_0(p) \subseteq W^u_R(p)$.

(iii) Since this R is the same for every point, it follows that

$$f^{-n}\xi_0(f^n p) \subseteq f^{-n}W^u_R(f^n p) \subseteq W^u_{\lambda^{-n}R}(p) \overset{n\to+\infty}{\longrightarrow} \{p\}$$

Thus $\bigvee_{n=0}^{+\infty} f^{-n}\xi_0$ is the point partition of $\mathbb{T}^2$.

There is, however, no certainty about the fourth condition. Instead of conditions (i),(ii) and (iii) which require either a static or long-term behavior, condition (iv) is a first-time condition. I.e. it requires that the first (backwards) iterate of an atom to already be contained in another. This is, of course, very restrictive. However, is essential for the construction of Leafwise measures on section 4.3. To obtain this property,

[1] This should seem intuitive, but by no means it is obvious. At least not until section 4.3, where we prove the Superposition Property. There we see that nested partitions have proportional conditional measures. Hence the equivalence.

[2] This follows because there are uncountably many possible radius which give rise to uncountably many disjoint shells. Since a measure is additive, it cannot be positive at each one of those shells.

there is a clever trick: If we want every atom to already fall on another, then, just force it to happen! Define

$$\xi \overset{\text{def.}}{=} \bigvee_{n=0}^{+\infty} f^n \xi_0$$

This partition ξ is thinner than ξ_0, hence it also satisfies (ii) and (iii). Now we also made it satisfy (iv), because

$$f^{-1}\xi = f^{-1}\left(\bigvee_{n=0}^{+\infty} f^n \xi_0\right) = \bigvee_{n=0}^{+\infty} f^{n-1}\xi_0 = f^{-1}\xi_0 \vee \bigvee_{n=0}^{+\infty} f^n \xi_0 = f^{-1}\xi_0 \vee \xi \succ \xi$$

But this came with a price. Property (i) is not necessarily true anymore. The problem is, we iterated so much that we lost track of the boundaries we had to avoid before. The heart of the proof is that it can be fixed by a suitable choice of the previously stated radius $R(p)$.

A Borel-Cantelli Argument

For every $p \in \mathbb{T}^2$, let $0 < R_0(p) < R$ be so that, for any choice of radius $0 < R(p) < R_0(p)$, the foliated ball $B_{R(p)}(p)$ is well defined. When defining ξ_0, we used a simple cardinality argument to achieve a $R(p)$ such that $\mu(\partial B_{R(p)}(p)) = 0$. But now, since we are iterating, we don't just want the measure of $\partial B_{R(p)}(p)$ to be zero, we want that points remains far from it while we iterate them. Lemma 4.1.4 is a refinement of this argument and says that we can always find a such $R(p)$ that keeps points uniformly away from $\partial B_{R(p)}(p)$. But before we prove it, we need a technical lemma which is a particular case of Besicovitch's Covering Theorem that we prove here for completeness:

Lemma 4.1.3. *If Y is a non-empty bounded subset of $\mathbb{R}$ then, for any $\alpha > 0$ we can cover Y with finitely many intervals $I_r(\alpha) \overset{\text{def.}}{=} (r-\alpha, r+\alpha)$ with $r \in Y$ and such that no point is in more than two of them.*

Proof. Let I be a compact interval in $\mathbb{R}$ that contains Y. Being compact, we can cover I by finitely many intervals $I_{r_1}(\alpha/10), \dots, I_{r_l}(\alpha/10)$. For each $i = 1, \dots, l$ take a $r_i' \in I_{r_i} \cap Y$ if this intersection is non empty. For all those with non empty intersection we have $I_{r_i}(\alpha/10) \subseteq I_{r_i'}(\alpha)$. Thus

$$Y \subseteq \bigcup_{i=1}^{l} I_{r_i}(\alpha/10) \subseteq \bigcup_{I_{r_i} \cap Y \neq \emptyset} I_{r_i}(\alpha/10) \subseteq \bigcup_{I_{r_i} \cap Y \neq \emptyset} I_{r_i'}(\alpha)$$

Thus, these intervals $I_{r_i'}(\alpha)$ with $r_i' \in Y$ covers Y. Rearrange these points r_i' so that $r_i' < r_{i+1}'$. If a point $r \in Y$ is in three or more intervals $I_{r_{i_1}'}(\alpha), \dots, I_{r_{i_k}'}(\alpha)$, $i_1 \leq \cdots i_k$. Then the extremal intervals $I_{r_{i_1}'}(\alpha)$ and $I_{r_{i_k}'}(\alpha)$ covers all the other $I_{r_i'}(\alpha)$, $i_1 < i < i_k$. Hence these intervals in the middle are redundant and can be discarded from our cover. Now this point r is in only two intervals of our cover. Each time we do this process we discard at least one of the intervals in our cover, thus, since it is finite, in the worst case we would eventually have only two intervals covering Y and at this point there can't be a point in three or more intervals. So, this process eventually ends leaving us with a finite cover covering at most two times a point, as desired. □

With this little result, we can proceed on our context: Recall the exponent $\lambda > 1$ given in definition 2.1.1.

Lemma 4.1.4. *For $p \in \mathbb{T}^2$, we can find a $R(p) \in [R_0(p)/2, R_0(p)]$ such that for μ-almost every $q \in \mathbb{T}^2$, for all $n \in \mathbb{N}$ sufficiently big we have that $d(f^{-n}(q), \partial B_{R(p)}(p)) > \lambda^{-n}$.*

Proof. For a set $A \subseteq \mathbb{T}^2$, let $B_\delta(A)$ denote the set of points δ close to A. The conclusion of the lemma is equivalent to "almost every point $q \in \mathbb{T}^2$ is only in finitely many of the sets $f^j(B_{\lambda^{-j}}(\partial B_{R(p)}(p)))$". By Borel-Cantelli's lemma it would suffice to have

$$\sum_{j=0}^{+\infty} \mu\left(f^j(B_{\lambda^{-j}}(\partial B_{R(p)}(p)))\right) < +\infty$$

Define an auxiliary measure η on the interval $[0, R(p)]$ by

$$\eta([a,b]) \overset{\text{def.}}{=} \mu\left(\{q \in \mathbb{T}^2 | a \leq d(p,q) \leq b\}\right)$$

This measure measures the boundaries of balls around p. In particular, if for $r \in [0, R_0(p)]$ the intervals $I_r^j \overset{\text{def.}}{=} [r - \lambda^{-j}, r + \lambda^{-j}]$ are in $[0, R_0(p)]$, we have

$$\eta(I_r^j) = \mu\left(B_{\lambda^{-j}}(\partial B_{R(p)}(p))\right)$$

Since μ is f-invariant, this means that

$$\eta(I_r^j) = \mu\left(f^j(B_{\lambda^{-j}}(\partial B_{R(p)}(p)))\right)$$

Thus, our problem is reduced to show that

$$\sum_{j=0}^{+\infty} \eta(I_r^j) < +\infty$$

This series would converge if $\eta(I_r^j) \leq 1/j^2$. Thus, let's track the r's that don't satisfy it. Define

$$Y_j \overset{\text{def.}}{=} \{r \in [0, R_0(p)] | \eta(I_r^j) > 1/j^2\}$$

These sets Y_j are a subset of the compact interval $[0, R_0(p)]$, thus bounded. By lemma 4.1.3 we can find a finite cover of Y_j by intervals $I^j_{r_1^j}, \ldots, I^j_{r^j_{l_j}}$ with $r_i^j \in Y_j$ that covers Y_j intersecting each point no more than twice. Hence

$$\text{Leb}(Y_j) \leq \sum_{i=1}^{l_j} \text{Leb}(I^j_{r_i}) = \sum_{i=1}^{l_j} 2\lambda^{-j} = l_j \cdot 2\lambda^{-j}$$

Also, using the definition of Y_j we have

$$l_j \cdot \frac{1}{j^2} = \sum_{i=1}^{l_j} \frac{1}{j^2} \leq \sum_{i=1}^{l_j} \eta(I^j_{r_i}) \leq 2\eta([0, R_0(p)])$$

Which implies that $l_j \leq 2j^2$. Thus $\text{Leb}(Y_j) \leq 4j^2\lambda^{-j}$ and $\sum_j \text{Leb}(Y_j) < +\infty$. By Borel-Cantelli's lemma, Lebesgue almost every point $r \in [0, R_0(p)]$ is in only finitely many Y_j's. Taking a $R(p) \in [R_0(p)/2, R_0(p)]$ satisfying it we conclude the lemma. □

An immediate consequence of this lemma, is that for this choice of $R(p)$ almost no point falls in $\partial B_{R(p)}(p)$. In particular, if ∂B denotes the union of the boundaries of the finite cover $B_1, \ldots, B_\kappa$, then

Lemma 4.1.5. *For μ-almost every point $q \in \mathbb{T}^2$ we have that $f^n(q) \notin \partial B$ for all $n \in \mathbb{Z}$*

Proof. We want to show that the sets $f^n(\partial B_{R(p)}(p))$ have zero measure. But since μ is f-invariant, it suffices to show it for $n = 0$. In particular, since $\partial B_{R(p)}(p) \subseteq B_{\lambda^{-j}}(\partial B_{R(p)}(p))$, we have

$$\sum_{j=0}^{+\infty} \mu\left(\partial B_{R(p)}(p)\right) \leq \sum_{j=0}^{+\infty} \mu\left(B_{\lambda^{-j}}(\partial B_{R(p)}(p))\right)$$

This last sum was shown to be finite in the previous lemma, and the first sum can only be finite if $\mu\left(\partial B_{R(p)}(p)\right) = 0$.

□

Finally, we have everything ready to prove their existence:

Proof of Theorem 4.1.2

Let ξ be as above, where $R(p)$ was given by lemma 4.1.4 and R was chosen smaller than 1. Conditions (ii), (iii) and (iv) were already verified. We need only to verify condition (i). For it, take $q \in \mathbb{T}^2$ satisfying lemma 4.1.4 and 4.1.5. Let n_0 be the number given by lemma 4.1.4 such that $d(f^{-n}(q), \partial B) > \lambda^{-n}$ for all $n \geq n_0$. By lemma 4.1.5, each atom $\xi_0(f^{-k}q)$, $k \in \mathbb{N}$ contains an open interval of unstable manifold. Thus each iterate $f^k(\xi_0(f^{-k}q))$ also have. Hence, there is a $r(q) > 0$ such that

$$W^u_{r(q)}(q) \subseteq \bigcap_{k=0}^{n_0} f^k(\xi_0(f^{-k}q)) = \left(\bigvee_{k=0}^{n_0} f^k\xi_0\right)(q)$$

I claim that for $n \geq n_0$ we have $(\bigvee_{k=0}^{n} f^k\xi_0)(q) = (\bigvee_{k=0}^{n_0} f^k\xi_0)(q)$. If this claim is right, them

$$\xi(q) = \left(\bigvee_{k=0}^{+\infty} f^k\xi_0\right)(q) = \left(\bigvee_{k=0}^{n_0} f^k\xi_0\right)(q) \supseteq W^u_{r(q)}(q)$$

and the theorem will be proven. To prove the claim, suppose it was not true. Them, there would be a $m > n_0$ such that

$$\bigcap_{k=0}^{m} f^k(\xi_0(f^{-k}q)) \subsetneq \bigcap_{k=0}^{n_0} f^k(\xi_0(f^{-k}q))$$

In particular, this means that for some $m_0 > n_0$ the interval $\bigcap_{k=0}^{n_0} f^k(\xi_0(f^{-k}q))$ is not properly contained in $f^{m_0}(\xi_0(f^{-m_0}q))$. Thus, there is a point q_0 in the boundary of $f^{m_0}(\xi_0(f^{-m_0}q))$ that crosses it. That is, there is a point $q_0 \in f^{m_0}(\partial B) \cap \bigcap_{k=0}^{n_0} f^k(\xi_0(f^{-k}q))$. Since $\bigcap_{k=0}^{n_0} f^k(\xi_0(f^{-k}q))$ is contained in $W^u_R(q)$ we have $d(q, f^{m_0}(\partial B)) < R < 1$. Iterating and using that q and q_0 are in the same unstable leaf we get that $d(f^{-m_0}q, \partial B) < \lambda^{-m_0}$, which contradicts the definition of n_0 and concludes the proof. □

4.2 Affine Structures

In this section we develop once again another tool towards the understanding of the global structure of the unstable and stable foliations. Here we construct the so called normal forms (or Affine Structures) which are a collection of parameterizations of unstable leaves that behave very well under the dynamics. Precisely, we prove the following:

Proposition 4.2.1. *If $f \in \mathrm{Diff}^2_\mu(\mathbb{T}^2)$ is an Anosov diffeomorphism, then there exists an unique continuous family of $C^{1+\alpha}$-diffeomorphisms $H_p : W^u(p) \to E^u(p)$ for $p \in \mathbb{T}^2$ such that:*

(i) H_p *depends* C^0 *on* p *restricted to unstable leaves in the* $C^{1+\alpha}$ *topology.*

(ii) $H_p(p) = 0$ *and* $d_p H_p$ *is the identity*

(iii) $H_q \circ H_p^{-1}$ *for* $q \in W^s(p)$ *is affine*

(iv) $H_{f(p)} \circ f = d_p f \cdot H_p$

These parameters do an incredible job in simplifying many arguments. If you simply interpret $E^u(p)$ as an ordinary copy of $\mathbb{R}$, them these parameters behave just like a linearization of the dynamics. In fact, condition (iv) precisely says that the action of the dynamics under these coordinates is given by its derivative.

There is, of course, a technicality here: The parameter H_p depends on the base point p. This may seem to diminish the effectiveness of this family, however, it is not expected that we would be able to linearize it for an arbitrary f. In fact, it already is a miracle that it is possible to do this non stationary linearization and even though it depends on the base point, items (ii) and (iii) says that they vary very well. Even more, if we apply the dynamics in an unstable neighborhood of a fixed (or periodic) point, you can track the exact length of expansion/contraction of the leaves to be just a multiplication by the unstable eigen value.

As a proof of their usefulness, if their essential role on the next section isn't enough, in the end of this section I show how they can be used to prove the regularity of stable Holonomies. This will be used in the construction of the Margulis family in the next chapter.

Proof of Proposition 4.2.1

The main reference we follow here is [KK06].

To simplify the exposition I will assume the existence of a unitary vector field tangent to the unstable distribution $p \in \mathbb{T}^2 \mapsto v^u(p) \in E^u(p)$. We can identify the unstable direction $E^u(p)$ with $\mathbb{R}$ via the relation $av^u(p) \in E^u(p) \sim a \in \mathbb{R}$. If $\rho : \mathbb{T}^2 \times \mathbb{T}^2 \to \mathbb{R}$ is continuous, then for every $p \in \mathbb{T}^2$ we can define a map $H_p : W^u(p) \to E^u(p)$ by

$$H_p(q) \stackrel{\text{def.}}{=} \int_p^q \rho(p, r)\mathrm{d}r$$

where the integral above is taken in a path connecting p to q contained in $W^u(p)$ and with sign given by the field v^u. Our job now is to find an appropriate ρ which gives us the desired properties of H_p. First, if we had chosen ρ Hölder continuous, we would have that H_p is the integral of a Hölder continuous function, hence $C^{1+\alpha}$. If we also happens to choose ρ positively bounded from bellow, then dH_p is always invertible, and being it one dimensional, we get that H_p would be a diffeomorphism. In fact, by the fundamental theorem of calculus, we have $d_q H_p = \rho(p, q)$, thus if we have $\rho(p, p) = 1$ then $d_p H_p$ is the identity.

All these properties above are very loose and would only give us properties (i) and (ii). Thankfully, property (iv) is way more enlightening. To see it, suppose for a moment that item (iv) holds. Them, for all

$p \in \mathbb{T}^2$ and $q \in W^u(p)$ we would have that

$$\begin{aligned} H_{f(p)} \circ f(q) = d_p f \circ H_p(q) &\iff \int_{f(p)}^{f(q)} \rho(f(p), r)\mathrm{d}r = \lambda_p^u \int_p^q \rho(p, r)\mathrm{d}r \\ &\iff \int_p^q J^u f(r)\rho(f(p), f(r))\mathrm{d}r = \int_p^q \lambda_p^u \rho(p, r)\mathrm{d}r \end{aligned}$$

Since $q \in W^u(p)$ was arbitrary, the integrands must coincide. I.e.

$$J^u f(q)\rho(f(p), f(q)) = \lambda_p^u \rho(p, q)$$

Also, since $W^u(p)$ is one dimensional, we have that $J^u f(q) = \lambda_q^u$. Rearranging and using induction we see that for any $n \in \mathbb{N}$:

$$\rho(f(p), f(q)) = \frac{\lambda_p^u}{\lambda_q^u} \cdot \rho(p, q) = \frac{\lambda_p^u}{\lambda_q^u} \cdot \frac{\lambda_{f^{-1}p}^u}{\lambda_{f^{-1}q}^u} \cdot \rho(f^{-1}p, f^{-1}q) = \cdots = \prod_{i=0}^{n} \frac{\lambda_{f^{-i}p}^u}{\lambda_{f^{-i}q}^u} \cdot \rho(f^{-n}p, f^{-n}q)$$

Since $q \in W^u(p)$, we have that $d(f^{-n}p, f^{-n}q) \to 0$. Hence, since we choose $\rho(r, r) = 1$ for any $r \in \mathbb{T}^2$, it makes sense to suppose that $\rho(f^{-n}p, f^{-n}q) \to 1$. In this case, the formula above gives an explicit expression for our candidate ρ:

$$\rho(p, q) = \prod_{i=1}^{+\infty} \frac{\lambda_{f^{-i}p}^u}{\lambda_{f^{-i}q}^u}, \qquad \text{for } q \in W^u(p)$$

All we have to do now is to prove that it is well defined and satisfies the previous hypothesis.

Lemma 4.2.2 (Distortion Lemma)**.** *For $p \in \mathbb{T}^2$, the function $q \in W^u(p) \mapsto \rho_p(q) \overset{\text{def.}}{=} \rho(p, q) \in \mathbb{R}$ where $\rho(p, q)$ is defined as above is a well defined Hölder continuous function.*

Proof. Fix a $p \in \mathbb{T}^2$. Define a sequence of functions $\rho_n : W^u(p) \to \mathbb{R}$ by

$$\rho_n(q) \overset{\text{def.}}{=} \prod_{i=1}^{n-1} \frac{\lambda_{f^{-i}p}^u}{\lambda_{f^{-i}q}^u}$$

These ρ_n are well defined because $\lambda_r^u > 0$. Also, since f is C^2 and E^u is Hölder continuous, we have that λ_r^u is Hölder continuous. The manifold $\mathbb{T}^2$ is compact and $\lambda_r^u > 0$ is continuous on it, thus it is bounded away from zero and ρ_n is also Hölder continuous. Notice that

$$\log(\rho_n(q)) = \sum_{i=1}^{n-1} \log(\lambda_{f^{-i}p}^u) - \log(\lambda_{f^{-i}q}^u)$$

Hence, using that the logarithm is Lipschitz when restricted to an interval bounded away from zero, we see that

$$\begin{aligned} |\log(\rho_n(q))| \le \sum_{i=1}^{n-1} \left|\log(\lambda_{f^{-i}p}^u) - \log(\lambda_{f^{-i}q}^u)\right| &\le \sum_{i=1}^{n-1} C \left|\lambda_{f^{-i}p}^u - \lambda_{f^{-i}q}^u\right| \\ &\le \sum_{i=1}^{n-1} CC' d(f^{-i}p, f^{-i}q)^\alpha \end{aligned}$$

The series above converges when $n \to +\infty$ because p and q are exponentially asymptotic. In particular, for $m > n$:

$$|\log(\rho_m(q)) - \log(\rho_n(q))| \leq \sum_{i=n}^{m-1} CC' d(f^{-i}p, f^{-i}q)^\alpha$$

So $\log(\rho_n(q))$ is Cauchy and it converges to a function $g(x)$. Thus the sequence ρ_n converges to the function $\rho_p(q) = e^{g(x)}$ which is Hölder with the same exponent of the ρ_n's. □

These ρ_p are what we'll use to define our H_p in place of that hypothetical ρ of before. The lemma above says it is well defined and by construction H_p already satisfies property (iv). It is also clear from their definition that $\rho_p(p) = 1$ and $H_p(p) = 0$, thus property (ii) is also done. Now, for item (i), notice that

Lemma 4.2.3. *For $p', q \in W^u(p)$ we have $\rho_p(q) = \rho_p(p')\rho_{p'}(q)$.*

Proof. It follows from a direct computation:

$$\rho_p(q) = \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-i}p}}{\lambda^u_{f^{-i}q}} = \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-i}p}}{\lambda^u_{f^{-i}p'}} \frac{\lambda^u_{f^{-i}p'}}{\lambda^u_{f^{-i}q}} = \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-i}p}}{\lambda^u_{f^{-i}p'}} \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-i}p'}}{\lambda^u_{f^{-i}q}} = \rho_p(p')\rho_{p'}(q)$$

□

Thus $H_p(q) = \rho_p(p')H_{p'}(q)$ and $dH_p = \rho_p(p')dH_{p'}$. Hence H_p depends continuously on p restricted to unstable leaves in the $C^{1+\alpha}$ topology. Finally, it remains to show property (iii):

Lemma 4.2.4. *For $q \in W^u(p)$, the map $H_q \circ H_p^{-1}$ is affine.*

Proof. To show that a map is affine is equivalent to show that its derivative is constant. In fact, let $r \in \mathbb{R}$ and put $p' \stackrel{\text{def.}}{=} H_p^{-1}(r)$. With that, we have

$$d_r(H_q \circ H_p^{-1}) = d_{p'}H_q \cdot d_r H_p^{-1} = d_{p'}H_q \cdot [d_{p'}H_p]^{-1} = \frac{\rho_q(p')}{\rho_p(p')}$$

Taking $q = p'$ in Lemma 4.2.3, we see that $1/\rho_p(p') = \rho_{p'}(p)$. Thus, by the same Lemma, the expression above is equals to $\rho_q(p')\rho_{p'}(p) = \rho_q(p)$ which does not depends on r. □

Here, we finished the proof of the existence of this family. The ρ_p's we defined looks like they are uniquely defined (up to zero measure), thus the family $\{H_p\}_{p\in\mathbb{T}^2}$ seems to be unique. However, we made some assumptions before we had the expression for the ρ_p's. Thus the unicity is not obvious and we must properly check it:

Lemma 4.2.5. *This family $\{H_p\}_{p\in\mathbb{T}^2}$ is unique.*

Proof. Suppose that $\{\tilde{H}_p\}_{p\in\mathbb{T}^2}$ is another family satisfying the conclusions of Proposition 4.2.1. For each $p \in \mathbb{T}^2$, define a transition map $G_p : \mathbb{R} \to \mathbb{R}$ by $G(t) \stackrel{\text{def.}}{=} H_p \circ \tilde{H}_p^{-1}(t)$. It follows from property (iv) that for any $n \in \mathbb{Z}$ we have that

$$H_p = df^{-n} \cdot H_{f^n(p)} \circ f^n \quad \text{and} \quad \tilde{H}_p = df^{-n} \cdot \tilde{H}_{f^n(p)} \circ f^n$$

Hence

$$G_p(t) = df^n \cdot G_{f^{-n}p} \circ df^{-n} \cdot t = \frac{G_{f^n p}(\lambda^u_{f^n p}(-n)t)}{\lambda^u_{f^n p}(-n)}$$

By property (ii) we have $G_p(0) = 0$ and $G_p'(0) = 1$, thus since $\lambda^u_{f^n p}(-n)t \to 0$ as $n \to +\infty$, we obtain that

$$G_p(t) = \frac{G_{f^n p}(\lambda^u_{f^n p}(-n)t) - G_{f^{-n} p}(0)}{\lambda^u_{f^n p}(-n) - 0} = \lim_{n\to\infty} t\frac{G_{f^n p}(\lambda^u_{f^n p}(-n)t) - G_{f^{-n} p}(0)}{\lambda^u_{f^n p}(-n)t - 0} = t$$

Thus G_p is the identity and $H_p = \tilde{H}_p$. □

Regularity of Holonomies

In this subsection we use these affine parameters to increase the regularity of our Holonomy maps.

Recall that in 2.2, we saw that Holonomies are, locally, translations along the stable direction: If you have an holonomy between p and q, then take a thin (but maybe long) foliated neighborhood containing p and q, what the holonomy does is to take the segment of unstable leaf at p at move it along the stable lines connecting it to an unstable segment at q (in the local chart this is precisely a translation. See 2.2.1). This means that holonomies are unique, in the sense that if they have the same base point then they agree. This, in turn, reveal another feature that make holonomies enjoyable; They commute with the dynamics:

Lemma 4.2.6. *If $\mathcal{H}_{p\to q}$ and $\mathcal{H}_{f^n p\to f^n q}$ are Holonomies from p to q and from $f^n p$ to $f^n q$ respectively, then $f^n \circ \mathcal{H}_{p\to q} = \mathcal{H}_{f^n p\to f^n q} \circ f^n$ when booth sides are well defined.*

Proof. Define the map $\mathcal{H}'_{p\to q}(r) = f^{-n} \circ \mathcal{H}_{f^n p\to f^n q} \circ f^n(r)$. We have $\mathcal{H}'_{p\to q}(p) = q$ and

$$\mathcal{H}'_{p\to q}(r) \in f^{-n}(W^s(f^n r) \cap W^u(f^n q)) = W^s(r) \cap W^u(q)$$

Thus $\mathcal{H}'_{p\to q}$ is an holonomy from p to q. By unicity $\mathcal{H}'_{p\to q} = \mathcal{H}_{p\to q}$ in their respective domains and the lemma is proven. □

Now, as said, Holonomies are translations of peaces of unstable manifolds (which are as smooth as f) along stable manifolds (which are as smooth as f). However, we only required them to be continuous. It may seem plausible that they also are as smooth as f. Sadly, this isn't always the case. Even thought the leaves are smooth, the foliation as a whole is only Hölder. Different sections of leaves are moving away at different rates, and this may hamper the smoothness. Haply, in our case, they do are smooth, for, they are linear in these affine charts we just defined:

Proposition 4.2.7. *If $\mathcal{H}_{p\to q}$ is an holonomy and we define the map $\mathscr{H} \overset{\text{def.}}{=} H_q \circ \mathcal{H}_{p\to q} \circ H_p^{-1}$, where H_p and H_q are the affine parameters at p and q respectively. Then, this map is $C^{1+\alpha}$.*

Proof. We already know that the maps H_p and H_q are $C^{1+\alpha}$, thus the proposition is all about $\mathcal{H}_{p\to q}$ being $C^{1+\alpha}$. Notice that the base point of an holonomy map is only for intuition convenience, i.e. if $r \in \mathrm{Dom}(\mathcal{H}_{p\to q})$ is another point on the domain of $\mathcal{H}_{p\to q}$, then, the map $\tilde{\mathcal{H}}_{r\to \mathcal{H}_{p\to q}(r)}(z) \overset{\text{def.}}{=} \mathcal{H}_{p\to q}(z)$ is an actual holonomy between r and $\mathcal{H}_{p\to q}(r)$. Thus, if we assumed that every holonomy is differentiable at its base point, we obtain that

$$d_r\mathcal{H}_{p\to q} = d_r\mathcal{H}_{r\to \mathcal{H}_{p\to q}}$$

so that the holonomies are actually differentiable everywhere. Since $H_p^{-1}(0) = p$, it means that, to prove the proposition, we need only show that $\mathscr{H}$ is differentiable at 0 and obtain a value that is Hölder on p and q.

For $n \in \mathbb{N}$ and $t \in \mathbb{R}$, lemma 4.2.6 says that

$$\mathscr{H}(t) = \lambda^u_{f^n q}(-n) H_{f^n q} \circ \mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t)$$

Since $H_{f^n q}$ is C^1, and $\mathscr{H}(0) = 0$, by the mean value theorem we have

$$|\mathscr{H}(t)| = |\mathscr{H}(t) - \mathscr{H}(0)| \leq \lambda^u_{f^n q}(-n) \|dH_{f^n q}\|_{n,t} \cdot |J(n,t)|$$

where

$$|J(n,t)| \overset{\text{def.}}{=} d^u(\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n q)$$

is the length of the interval of unstable manifold J connecting the point $\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t)$ to $f^n(q) = \mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(0)$ and $\|dH_{f^n q}\|_{n,t} \overset{\text{def.}}{=} \sup_{h \in J} \|d_h H_{f^n q}\|$ is the maximum of the norm of the derivative of $H_{f^n q}$ along this interval. Actually, since $\|d_h H_{f^n q}\|$ is Hölder in h, with Hölder constants independent of the base point and $\|d_{f^n q} H_{f^n q}\| = 1$, we have that

$$\|dH_{f^n q}\|_{n,t} \leq 1 + \varepsilon_1(|J(n,t)|)$$

where $\varepsilon_1(|J(n,t)|) \to 0$ as $|J(n,t)| \to 0$. Now, we need to control the term $|J(n,t)|$. Notice that

$$\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t) \to f^n(q) \quad \text{as} \quad t \to 0$$

Thus, given $n \in \mathbb{N}$, we can find a $t_n > 0$, such that for $t < t_n$, the points $\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t)$ and $f^n(q)$ are in a same product neighborhood. Since the unstable manifold is C^1, in this neighborhood, the distances along unstable leaves d^u and the distance in the entire manifold d are equivalent. In particular, there is a $C > 0$ such that,

$$d^u(\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n q) \leq C d(\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n q)$$

by compactness, this $C > 0$ can be chosen uniform. Now, for $t \leq t_n$ we can use the triangle inequality to obtain that

$$\begin{aligned}|J(n,t)| \leq C\Big(& d(\mathcal{H}_{f^n p \to f^n q} \circ H^{-1}_{f^n p}(\lambda^u_p(n) t), H^{-1}_{f^n p}(\lambda^u_p(n) t)) \\ & + d(H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n p) \\ & + d(f^n p, f^n q)\Big)\end{aligned}$$

Let $\varepsilon_2(n)$ be the sum of the first and last terms above. Since they are the distance between two points in the same stable manifold, they go to 0 exponentially fast in n, i.e. $\varepsilon_2(n) \overset{\exp}{\to} 0$ as $n \to +\infty$. For the term in the middle, we have

$$\begin{aligned} d(H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n p) &\leq d^u(H^{-1}_{f^n p}(\lambda^u_p(n) t), f^n p) \\ &= d^u(H^{-1}_{f^n p}(\lambda^u_p(n) t), H^{-1}_{f^n p}(0)) \\ &\leq \sup_{|s| \leq \lambda^u_p(n) t} \|d_s H^{-1}_{f^n p}\| \cdot |\lambda^u_p(n) t| \end{aligned}$$

Similarly to before, we can write

$$\sup_{|s| \leq \lambda^u_p(n) t} \|d_s H^{-1}_{f^n p}\| \leq 1 + \varepsilon_3(\lambda^u_p(n)|t|)$$

where $\varepsilon_3(\lambda_p^u(n)t) \to 0$ as $\lambda_p^u(n)t \to 0$. Gathering all those estimates we conclude that,

$$|\mathscr{H}(t)| \leq C\lambda^u_{f^n q}(-n)\left(1+\varepsilon_1(|J(n,t)|)\right)\left(\varepsilon_2(n)+\left[1+\varepsilon_3(\lambda_p^u(n)|t|)\right]\lambda_p^u(n)|t|\right)$$

Dividing by $|t|$ and making the limit $n \to +\infty$, $t \to 0$ while also making $\lambda_p^u(n)|t| \to 0$ we obtain that

$$\limsup_{t\to 0} \frac{|\mathscr{H}(t)|}{|t|} \leq C \lim_n \frac{\lambda_p^u(n)}{\lambda_q^u(n)} = C\rho_p^s(q)$$

where, analogously to lemma 4.2.2, the limit above converges to a Hölder function $\rho_p^s(q)$. Using infimums instead of the supremums above, we obtain the same estimate but for the $\liminf$. Thus concluding the proposition. □

4.3 Leafwise Measures

In the last sections we made some general results, now we fix μ as the SRB measure of f and use those results to construct the Leafwise measures, which are a family of measures defined on W^u which locally gives a disintegration of μ in product neighborhoods that scales with the unstable Lyapunov exponent λ^u of f.

The outline of the proof is that we start with a subordinated partition and we iterate it. These iterates form a sequence of partitions whose atoms gets bigger and bigger. By proving a superposition principle, we show that the disintegrations of our measure in these partitions converges (in some sense) to measures in entire unstable leaves. Using the affine parameters we constructed, we can obtain that their pullback is precisely 0.5 times the Lebesgue measure in $\mathbb{R}$. With that, using the linearity of the dynamics under normal forms, we can return this estimate to the manifolds and conclude the Theorem.

The superposition Property

Let $\xi_0 < W^u$ be a subordinated partition. Since atoms of ξ_0 are intervals of unstable leaves, the dynamics dilates each $\xi_0(p)$ so that by iterating ξ_0 we can define a family of wider and wider subordinated partitions:

Definition 4.3.1. For $n \in \mathbb{N}$ we set $\xi_n \overset{\text{def.}}{=} f^n\xi_0$

In fact, by property (iv) of definition 4.1.1 we have that $\xi_{n+1} \prec \xi_n$. Also, each one of these partitions is measurable, so by Rokhlin's Theorem, we can disintegrate μ with respect to ξ_n to obtain a family of conditional measures $\{\mu_p^{\xi_n}\}_{p\in\mathbb{T}^2}$. Their domain, $\xi_n(p)$, grows with n in the following sense:

Lemma 4.3.2. *For μ-almost every point $p \in \mathbb{T}^2$ and $R > 0$, $\exists n_0 \in \mathbb{N}$ such that*

$$W_R^u(p) \subseteq \xi_n(p)$$

for all $n > n_0$.

Proof. For $\varepsilon > 0$ define the set

$$A(\varepsilon) \overset{\text{def.}}{=} \{p \in \mathbb{T}^2 | W_\varepsilon^u(p) \subseteq \xi(p)\}$$

If $p \in \mathbb{T}^2$ is such that $f^{-n}(p) \in A(\frac{1}{k})$, then, by the uniform expansion of f we have

$$f^n W^u_{\frac{1}{k}}(f^{-n}(p)) \subseteq f^n(\xi(f^{-n}(p))) \implies W^u_{\frac{\lambda^n}{k}}(p) \subseteq \xi_n(p)$$

where $\lambda > 1$ is as in definition 2.1.1. Hence, if $n > \log_{\lambda^u}(kR)$ we have $W^u_R(p) \subseteq \xi_n(p)$ and the lemma is true for this p. Thus, it suffices to prove that the set

$$\bigcup_{\substack{k\in\mathbb{N} \\ n>\log_\lambda(kR)}} f^n A(\tfrac{1}{k}) \tag{$*$}$$

has full measure. For it notice that, by item (i) of 4.1.1, the set $\bigcup_{k\in\mathbb{N}} A(\frac{1}{k})$ has full measure. So, by the f-invariance of μ, so does $(*)$. □

Using this lemma, we see that for a given $R > 0$, almost every measure $\mu_n^{\xi_n}$ with n sufficiently big is well defined in $W^u_R(p)$. We will show that, modulo a normalization, they all agree on W^u_R. For it, use the fact that the family $(\mu_p^{\xi_n})_{p\in\mathbb{T}^2}$ is measurable to define an auxiliary measure $\eta^{\xi_m}_{p,n}$ on $\xi_m(p)$ by:

Definition 4.3.3. If $m > n$ and $A \subseteq \xi_m(p)$, *put*

$$\eta^{\xi_m}_{p,n}(A) = \int_{\xi_m(p)} \left[\int_{\xi_n(q)} \mathbb{1}_A \mathrm{d}\mu_q^{\xi_n}\right] \mathrm{d}\mu_p^{\xi_m}(q)$$

Remark 4.3.4. It will be convenient for later to notice that if $A \subseteq \xi_n(p)$, the expression above greatly simplifies. This happens because if $A \subseteq \xi_n(q)$ then the inner integral is 0 for all $q \notin \xi_n(p)$ and constant for $q \in \xi_n(p)$ (item 2 of 3.1.21). Thus it can be rewritten as

$$\eta^{\xi_m}_{p,n}(A) = \int_{\xi_m(p)} \left[\mathbb{1}_{\xi_n(p)}(q) \int_{\xi_n(q)} \mathbb{1}_A \mathrm{d}\mu_q^{\xi_n}\right] \mathrm{d}\mu_p^{\xi_m}(q) = \mu_p^{\xi_n}(A)\mu_p^{\xi_m}(\xi_n(p))$$

Lemma 4.3.5. *For μ-almost every $p \in \mathbb{T}^2$ we have*

$$\eta^{\xi_m}_{p,n} = \mu_p^{\xi_m}$$

Proof. Let $\varphi : \mathbb{T}^2 \to \mathbb{R}$ be integrable. Then, using that both $\{\mu_p^{\xi_m}\}_{p\in\mathbb{T}^2}$ and $\{\mu_p^{\xi_n}\}_{p\in\mathbb{T}^2}$ are disintegrations of μ we have:

$$\begin{aligned}
\int \left[\int_{\xi_m(p)} \varphi \mathrm{d}\eta^{\xi_m}_{p,n}\right] \mathrm{d}\mu(p) &= \int \left[\int_{\xi_m(p)} \left[\int_{\xi_n(q)} \varphi \mathrm{d}\mu_q^{\xi_n}\right] \mathrm{d}\mu_p^{\xi_m}(q)\right] \mathrm{d}\mu(p) \\
&= \int \left[\int_{\xi_n(p)} \varphi \mathrm{d}\mu_q^{\xi_n}\right] \mathrm{d}\mu(p) \\
&= \int \varphi \mathrm{d}\mu
\end{aligned}$$

Thus $\{\eta^{\xi_m}_{p,n}\}_{p\in\mathbb{T}^2}$ is a disintegration of μ in ξ_m. The lemma follows by the a.e. unicity of disintegrations. □

The appearance of $\eta_{p,n}^{\xi_m}$ combined with this lemma may suggest that $\{\mu_p^{\xi_n}\}_{p\in\xi_m(p)}$ is a disintegration of $\mu_p^{\xi_m}$ in the partition $\{\xi_n(p)\}_{p\in\xi_m(p)}$ of $\xi_m(p)$. This suggestion is true: By item (i) of definition 4.1.1, almost every atom $\xi_n(p)$ contains an open interval, thus, since $\xi_m(p)$ is second countable, it contains at most enumerably many atoms of ξ_n. Hence, the partition of $\xi_m(p)$ by atoms of ξ_n is measurable and the lemma above precisely describes the disintegration of $\mu_p^{\xi_m}$ as $\{\mu_p^{\xi_n}\}_{p\in\xi_m(p)}$.

Lemma 4.3.6. *For μ-almost every $p\in\mathbb{T}^2$ we have $\mu_p^{\xi_m}(\xi_n(p))>0$ for $m>n$.*

Proof. Define $Y \overset{\text{def.}}{=} \{p\in\mathbb{T}^2|\mu_p^{\xi_m}(\xi_n(p))=0\}$. We want to show that $\mu(Y)=0$. For it, notice that

$$\mu(Y)=\int\chi_Y\mathrm{d}\mu=\int\mu_p^{\xi_m}(Y)\mathrm{d}\mu$$

By the remark above, we can write $\xi_m(p)=\cup_{p\in\xi_m(p)}\xi_n(p)$ where this union is atmost countable. Then

$$\mu_p^{\xi_m}(Y)=\sum_{\substack{p\in\xi_m(p)\\ 1-\text{per atom}}}\mu_p^{\xi_m}(\xi_n(p)\cap Y)$$

It suffices to show that the right side is 0. If $\xi_n(p)\cap Y=\emptyset$ we are done. If this is not empty, let $q\in\xi_n(p)\cap Y$. By definition we have $\mu_q^{\xi_m}(\xi_n(p))=0$. However, $\xi_n(p)\subseteq\xi_m(p)$, so $q\in\xi_m(p)$ and by item (ii) of Theorem 3.1.21 we have $\mu_p^{\xi_m}(\xi_n(p))=\mu_q^{\xi_m}(\xi_n(p))$ and consequently $\mu_p^{\xi_m}(\xi_n(p)\cap Y)=0$ as wanted. □

With this lemma we can normalize these measures as follows: Consider the collection of normal forms $\{H_p: W^u(p)\to\mathbb{R}\}_{p\in\mathbb{T}^2}$ and notice that since they are C^1 there is a $C>0$ such that $W_C^u(p)\subseteq H_p^{-1}([-1,1])$ for every $p\in\mathbb{T}^2$. Notice that, by the proof of Theorem 4.1.2, we can choose $0<R<C$ in the definition 4.1.1 so that $\xi(p)\subseteq H_p^{-1}([-1,1])$ for almost every p. Then by the lemma above $\mu_p^{\xi_m}(H_p^{-1}([-1,1]))>0$ and we can set

Definition 4.3.7.

$$\hat{\mu}_p^{\xi_m}\overset{\text{def.}}{=}\frac{\mu_p^{\xi_m}}{H_{p\,*}\mu_p^{\xi_m}([-1,1])}$$

With this definition we have the following stability property:

Proposition 4.3.8. *For μ-almost every $p\in\mathbb{T}^2$, $\forall K\subseteq W^u(p)$ compact, there is a $n_0\in\mathbb{N}$ such that for all $m,n>n_0$ we have*

$$\hat{\mu}_p^{\xi_m}(K)=\hat{\mu}_p^{\xi_n}(K)$$

Proof. Let $p\in\mathbb{T}^2$ be generic. Since $K\subseteq W^u(p)$ is compact, there is a $R>0$ such that $K\subseteq W_R^u(p)$. We can use lemma 4.3.2 to obtain a $n_0\in\mathbb{N}$ such that for all $n_0<n<m$ we have $K\subseteq\xi_n(p)\subseteq\xi_m(p)$. Using lemma 4.3.5 and the remark above it, we have

$$\mu_p^{\xi_m}(K)=\mu_p^{\xi_n}(K)\mu_p^{\xi_m}(\xi_n(p))$$

By lemma 4.3.6 $\mu_p^{\xi_m}(\xi_n(p))>0$ generically, thus

$$\mu_p^{\xi_n}(K)=\frac{\mu_p^{\xi_m}(K)}{\mu_p^{\xi_m}(\xi_n(p))} \tag{1}$$

The only condition to obtain this equality was that K was bounded in W^u, in particular, this is also the case of $H_p^{-1}([-1,1])$, hence, applying everything above to it we obtain

$$H_{p*}\mu_p^{\xi_n}([-1,1]) = \frac{H_{p*}\mu_p^{\xi_m}([-1,1])}{\mu_p^{\xi_m}(\xi_n(p))} \tag{2}$$

Dividing (1) by (2), the term $\mu_p^{\xi_m}(\xi_n(p))$ cancels and we obtain $\hat{\mu}_p^{\xi_m}(K) = \hat{\mu}_p^{\xi_n}(K)$. ☐

One simple corollary, obtained in the proof above, which is interesting by itself to have a name:

Corollary 4.3.9. *(Superposition Property)* $\mu_p^{\xi_n} = \frac{\mu_p^{\xi_m}|_{\xi_n(p)}}{\mu_p^{\xi_m}(\xi_n(p))}$ *for μ-a.e. $p \in \mathbb{T}^2$ if $n < m$.*

By the above proposition, we now see that the family of measures $\{\hat{\mu}_p^{\xi_n}\}_{p\in\mathbb{T}^2}$ are stationary, in the sense that they (almost everywhere) eventually agree. With that, and the fact that the domain of each $\hat{\mu}_p^{\xi_n}$ converges to the entire unstable leaf $W^u(p)$, we can define "pre"-measure $\hat{\mu}_p$ on $W^u(p)$ by setting:

Definition 4.3.10. $\hat{\mu}_p(K) \overset{\text{def.}}{=} \lim_{n\to+\infty} \hat{\mu}_p^{\xi_n}(K)$, for $K \subseteq W^u(p)$ compact.

Here, we are in a standard setting of measure theory: There is a pre-measure defined on all compacts of a Borel Space. So, by Carathéodory's Theorem (see Theorem 1.14 of [Fol99]) we can extend $\overline{\mu}_p$ to a Borel measure μ_p in $W^u(p)$ such that they agree in compacts, i.e. $\hat{\mu}_p(K) = \mu_p(K)$. This family $\{\mu_p\}_{p\in\mathbb{T}^2}$ is the leafwise measures. To see in which sense they give a local disintegration of μ we first need a general lemma:

Lemma 4.3.11. *For μ-a.e. $p \in \mathbb{T}^2$ and $n \in \mathbb{N}$ we have*

$$\mu_p^{\xi_n} = f_*^n \mu_{f^{-n}(p)}^{\xi}$$

Proof. It follows from a direct computation and unicity: Let $\varphi : \mathbb{T}^2 \to \mathbb{R}$ be integrable, then, by the invariance of μ and the fact that $\{\mu_p^{\xi_n}\}_{p\in\mathbb{T}^2}$ is a disintegration of μ, we have

$$\begin{aligned}
\int \varphi \mathrm{d}\mu = \int \varphi \circ f^n \mathrm{d}\mu &= \int \left[\int_{\xi(p)} \varphi \circ f^n \mathrm{d}\mu_p^{\xi} \right] \mathrm{d}\mu \\
&= \int \left[\int_{\xi(f^{-n}(p))} \varphi \circ f^n \mathrm{d}\mu_{f^{-n}(p)}^{\xi} \right] \mathrm{d}\mu \\
&= \int \left[\int_{f^n(\xi(f^{-n}(p)))} \varphi \mathrm{d} f_*^n \mu_{f^{-n}(p)}^{\xi} \right] \mathrm{d}\mu \\
&= \int \left[\int_{\xi_n} \varphi \mathrm{d} f_*^n \mu_{f^{-n}(p)}^{\xi} \right] \mathrm{d}\mu
\end{aligned}$$

Hence, $\{f_*^n \mu_{f^{-n}(p)}\}_{p\in\mathbb{T}^2}$ is a disintegration of μ with respect to ξ_n. By unicity, $f_*^n \mu_{f^{-n}(p)} = \mu_p^{\xi_n}$ for μ-a.e. $p \in \mathbb{T}^2$. ☐

With this result, we can show the following universal property of this family:

Lemma 4.3.12. *There is an almost everywhere defined family $\{\mu_p\}_{p\in\mathbb{T}^2}$ of measures in unstable leaves such that for all partition ξ subordinated to W^u we have*

$$\mu_p^\xi = \frac{\mu_p|_{\xi(p)}}{\mu_p(\xi(p))}$$

Proof. Let $\{\mu_p\}_{p\in\mathbb{T}^2}$ be the family described after Lemma 4.3.10. If ξ is thinner than some ξ_n used in the above construction, i.e. if $\xi_n \prec \xi$ then the claimed property follows from the superposition property (Cor. 4.3.9) and the definition of μ_p:

$$\mu_p^\xi = \frac{\mu_p^{\xi_n}|_{\xi(p)}}{\mu_p^{\xi_n}(\xi(p))} = \frac{H_{p\,*}\mu_p^{\xi_n}([-1,1])}{H_{p\,*}\mu_p^{\xi_n}([-1,1])}\frac{\hat{\mu}_p^{\xi_n}|_{\xi(p)}}{\hat{\mu}_p^{\xi_n}(\xi(p))} = \frac{\mu_p|_{\xi(p)}}{\mu_p(\xi(p))}$$

There is, however, no guarantee that we can find such a $n\in\mathbb{N}$. To overcome this problem, lets use that the inferior bound $r(p)>0$ for the atoms of ξ_0 given in def. 4.1.2 is measurable. Thus, we can find an increasing family of subsets $K_1 \subseteq K_2 \subseteq \cdots K_n \subseteq \cdots \subseteq \mathbb{T}^2$ such that $\mu(K_i)\to 1$ and for each $i\in\mathbb{N}$ there is a $\varepsilon_i>0$ such that $r(p)>\varepsilon_i$.

Since atoms of ξ are uniformly bounded by some constant R, let $n_i\in\mathbb{N}$ be big enough so that $\lambda^{-n}R<\varepsilon_i$. With that choice, for every $p\in K_i$ the atom $f^{-n}\xi(p)$ of $f^{-n}\xi$ is contained in the atom $\xi_0(p)$ of ξ_0. Hence, if we restrict $f^{-n}\xi$ and ξ_0 to partitions $f^{-n}\xi|_{K_i}$ and $\xi_0|_{K_i}$ of K_i, the superposition property gives us that

$$\mu_p^{f^{-n}\xi|_{K_i}} = \frac{\mu_p^{\xi_0|_{K_i}}|_{f^{-n}\xi|_{K_i}}}{\mu_p^{\xi_0|_{K_i}}(f^{-n}\xi|_{K_i}(p))}$$

for almost every $p\in K_i$. By lemma 4.3.11, we have

$$\mu_p^{f^{-n}\xi|_{K_i}} = f_*^{-n}\mu_{f^n p}^{\xi|_{K_i}}$$

and

$$\mu_p^{\xi_0|_{K_i}} = f_*^{-n}\mu_{f^n p}^{\xi_n|_{K_i}}$$

Thus

$$f_*^{-n}\mu_{f^n p}^{\xi|_{K_i}} = \frac{f_*^{-n}\mu_{f^n p}^{\xi_n|_{K_i}}|_{f^{-n}\xi|_{K_i}}}{f_*^{-n}\mu_{f^n p}^{\xi_n|_{K_i}}(f^{-n}\xi|_{K_i}(p))}$$

Since $f^{-n}\xi|_{K_i}(p)) = f^{-n}(\xi|_{k_i}(f^n p))$ and $f_*^{-n}\mu(A)=\mu(f^n A)$, this denominator on the right is just $\mu_{f^n p}^{\xi_n|_{K_i}}(\xi|_{K_i}(f^n p))$. Hence, applying f_*^n, we obtain that

$$\mu_p^{\xi|_{K_i}} = \frac{\mu_p^{\xi_n|_{K_i}}|_{\xi|_{K_i}}}{\mu_p^{\xi_n|_{K_i}}(\xi|_{K_i}(p))} = \frac{\mu_p|_{K_i\cap\xi(p)}}{\mu_p(K_i\cap\xi(p))}$$

for every $p\in f^n K_i$. Since $\bigcup_i K_i$ covers almost every point, the above equality actually holds for the entire atom of ξ almost everywhere and the lemma is proven. □

Scaling Property

Until now, all we used was the properties of a subordinated partition ξ and the fact that μ is f-invariant. We now use the fact it is SRB to verify that these leaf-wise measures $\{\mu_p\}_{p\in\mathbb{T}^2}$ constructed above have a great scaling property under the dynamics.

For it, recall definition 4.3.1 and notice that there is nothing bad by setting $\xi_{-1} \stackrel{\text{def.}}{=} f^{-1}\xi$. This, once again, defines a subordinated partition verifying $\xi \prec \xi_{-1}$ and we can obtain a decomposition $\{\mu_p^{\xi_{-1}}\}_{p\in\mathbb{T}^2}$ of μ with respect to that. Comparing that with the disintegration given by Lemma 4.3.11 allow us to obtain an exclusive property of SRB measures: It's densities are dynamically determined.

For it, recall that by definition of being SRB we have

$$\mu_p^\xi = \rho\,\mathrm{Leb}^u$$

for some measurable function $\rho : \mathbb{T}^2 \to \mathbb{R}$. Thus, we have the following:

Lemma 4.3.13. *For μ-a.e. $p \in \mathbb{T}^2$ and μ_p^ξ-a.e. $q \in \xi(p)$ we have*

$$\rho(q) = \rho(f^{-1}(q))(\lambda^u_{f^{-1}(q)})^{-1}(\mu_p^\xi(\xi_{-1}(p)))^{-1}$$

Proof. Let $A \subseteq \mathbb{T}^2$ be measurable. Since $\xi \prec \xi_{-1}$, the superposition property 4.3.9 says that for μ-a.e. $p \in \mathbb{T}^2$ we can write

$$\mu_p^{\xi_{-1}}(A) = \frac{\mu_p^\xi(A \cap \xi_{-1}(p))}{\mu_p^\xi(\xi_{-1}(p))} = (\mu_p^\xi(\xi_{-1}(p)))^{-1}\int_{A\cap\xi_{-1}(p)} \rho(z)\mathrm{d}\,\mathrm{Leb}^u(z)$$

Also, by lemma 4.3.11

$$\begin{aligned}\mu_p^{\xi_{-1}}(A) = f_*^{-1}\mu^\xi_{f(p)}(A) = \mu^\xi_{f(p)}(f(A)) &= \int_{f(A)\cap\xi(f(p))} \mathrm{d}\mu^\xi_{f(p)}\\ &= \int_{f(A)\cap\xi(f(p))} \rho(z)\mathrm{d}\,\mathrm{Leb}^u(z)\\ &= \int_{A\cap f^{-1}(\xi(f(p)))} \rho\circ f(z)\;\lambda^u_z \mathrm{d}\,\mathrm{Leb}^u(z)\\ &= \int_{A\cap\xi_{-1}(p)} \rho\circ f(z)\;\lambda^u_z \mathrm{d}\,\mathrm{Leb}^u(z)\end{aligned}$$

Thus, comparing those expressions:

$$(\mu_p^\xi(\xi_{-1}(p)))^{-1}\int_{A\cap\xi_{-1}(p)} \rho\mathrm{d}\,\mathrm{Leb}^u = \int_{A\cap\xi_{-1}(p)} \rho\circ f(z)\;\lambda^u_z \mathrm{d}\,\mathrm{Leb}^u(z)$$

Since A was arbitrary, we must have

$$(\mu_p^\xi(\xi_{-1}(p)))^{-1}\rho(z) = \rho\circ f(z)\;\lambda^u_z$$

for μ_p^ξ-a.e. $z \in \xi_{-1}(p)$. Finally, since $f(\xi_{-1}(p)) = \xi(f(p))$, every $q \in \xi(f(p))$ can be written as $q = f(r)$ for some $r \in \xi_{-1}(p)$. Using it, and rearranging the expression above, we get

$$\rho(q) = \rho(f^{-1}(q))(\lambda^u_{f^{-1}(q)})^{-1}(\mu_p^\xi(\xi_{-1}(p)))^{-1}$$

for $\mu^\xi_{f(p)}$-a.e. $q \in \xi(f(p))$. By changing $f(p)$ to p we have proven the lemma. □

That's a curious property, but the term $(\mu_p^\xi(\xi_{-1}(p)))^{-1}$ makes it a little clumsy to say that it 'dynamically defines' the density. Well, with a little rearrangement we can get rid of this term and obtain an expression that is so more visually appealing that I will state it as a lemma on it's own:

Lemma 4.3.14. *For μ-a.e. $p \in \mathbb{T}^2$ and μ_p^ξ-a.e. $q \in \xi(p)$ we have*

$$\frac{\rho(q)}{\rho(p)} = \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-l}(p)}}{\lambda^u_{f^{-l}(q)}}$$

Proof. By lemma 4.3.13 we have

$$\frac{\rho(q)}{\rho(p)} = \frac{\lambda^u_{f^{-1}(p)}\rho(f^{-1}(q))}{\lambda^u_{f^{-1}(q)}\rho(f^{-1}(p))} = \frac{\lambda^u_{f^{-2}(p)}\lambda^u_{f^{-1}(p)}\rho(f^{-2}(q))}{\lambda^u_{f^{-2}(q)}\lambda^u_{f^{-1}(q)}\rho(f^{-2}(p))} = \cdots = \prod_{i=1}^{n} \frac{\lambda^u_{f^{-l}(p)}}{\lambda^u_{f^{-l}(q)}} \frac{\rho(f^{-n}(q))}{\rho(f^{-n}(p))}$$

By lemma 4.2.2, the limit of this product converges. Hence

$$\frac{\rho(q)}{\rho(p)} = \prod_{i=1}^{+\infty} \frac{\lambda^u_{f^{-l}(p)}}{\lambda^u_{f^{-l}(q)}} \cdot \lim_{n\to+\infty} \frac{\rho(f^{-n}(q))}{\rho(f^{-n}(p))}$$

Since $q \in W^u(p)$ and almost every p is recurrent, we get a set of full measure such that this limit is 1. Proving the lemma. □

This lemma says that if you know the value of ρ at a point p, then you know its value at every point of $\xi(p)$. If you compare this infinity product with the Jacobian ρ_p of the normal forms $\{H_p\}_{p\in\mathbb{T}^2}$ introduced in subsection 4.2 you will see that they are the same thing! It is not a coincidence, of course this was made to be like that. With this, we have that $\rho(q) = C d_q H_p$ where C is constant on atoms. That very convenient expression invites us to explore the behavior of μ_p^ξ in these affine coordinate charts. For it, define the following

Definition 4.3.15. For $p \in \mathbb{T}^2$ and $I \subseteq \mathbb{R}$, put $\overline{\mu}_p(I) \stackrel{\text{def.}}{=} (H_p)_*\mu_p(I)$

Those $\overline{\mu}_p$ are a collection of measures in $\mathbb{R}$ and they reveal the following amazing regularity of μ_p

Lemma 4.3.16. *For μ-a.e. $p \in \mathbb{T}^2$ we have $\overline{\mu}_p = 0.5\,\mathrm{Leb}_{\mathbb{R}}$.*

Proof. Let $I \subseteq \mathbb{R}$ be compact and find $n \in \mathbb{N}$ big enough so that $H_p^{-1}(I) \subseteq \xi_n(p)$ then, by lemma 4.3.12, for all $J \subseteq I$ we have $\mu_p(H_p^{-1}(J)) = C_0\mu_p^{\xi_n}(H_p^{-1}(J))$ where $C_0 = \mu_p(\xi_n(p))$ is constant. Then

$$\begin{aligned} \overline{\mu}_p(J) = \int_J \mathrm{d}\overline{\mu}_p = \int_J \mathrm{d}(H_p)_*\mu_p &= \int_{H_p^{-1}(J)} \mathrm{d}\mu_p \\ &= C_0 \int_{H_p^{-1}(J)} \mathrm{d}\mu_p^{\xi_n} \\ &= C_0 \int_{H_p^{-1}(J)} \rho(z)\mathrm{d}\,\mathrm{Leb}^u(z) \\ &= C_0 \int_J d_z H_p^{-1}(z)\rho\circ H_p^{-1}(z)\mathrm{d}\,\mathrm{Leb}_{\mathbb{R}}(z) \end{aligned}$$

By the remark above definition 4.3.15, we have $\rho \circ H_p^{-1}(z) = C d_{H_p^{-1}(z)} H_p$. Also, by the inverse function theorem we have $d_z H_p^{-1} = (d_{H_p^{-1}(z)} H_p)^{-1}$, thus

$$d_z H_p^{-1}(z)\rho\circ H_p^{-1}(z) = C$$

Hence

$$\overline{\mu}_p(J) = C_0 C \int_J \mathrm{d}\,\mathrm{Leb}_{\mathbb{R}}(z)$$

Since $J \subseteq I$ was arbitrary, the measures coincides minus a constant in I. And since I was an arbitrary compact in $\mathbb{R}$, the measures coincides in $\mathbb{R}$ minus a constant. To determine the constant, just notice that by definitions 4.3.10 and 4.3.7, for $n \in \mathbb{N}$ sufficiently big we have

$$\overline{\mu}_p(H_p([-1,1]) = \frac{\mu_p^{\xi_n}(H_p([-1,1])}{\mu_p^{\xi_n}(H_p([-1,1])} = 1$$

Thus

$$1 = C_0 C \operatorname{Leb}_{\mathbb{R}}(H_p([-1,1]) = 2C_0 C$$

so that $C_0 C = 0.5$ and $\overline{\mu}_p = 0.5 \operatorname{Leb}_{\mathbb{R}}$. □

The fact the leaf-wise measures are exactly (0.5)Lebesgue under the normal forms reparameterization is really strong. In particular, since in normal forms f acts on a set by multiplying it by it's derivative, we can recover this property to the manifold:

Lemma 4.3.17. *For μ-a.e. $p \in \mathbb{T}^2$ and $n \in \mathbb{N}$ we have $f_*^n \mu_p = \lambda^u_{f^n p}(-n)\mu_{f^n(p)}$*

Proof. It follows from a direct calculation: Let $A \subseteq \mathbb{T}^2$ and define $B \overset{\text{def.}}{=} H_{f^n(p)}(A) \subseteq \mathbb{R}$. Then, by the item (iv) of Prop. 4.2.1 we have

$$\begin{aligned} f_*^n \mu_p(A) &= \mu_p(f^{-n} \circ H^{-1}_{f^n(p)}(B)) &&= \mu_p(H_p^{-1}\lambda^u_{f^n p}(-n)B) &&= \overline{\mu}_p(\lambda^u_{f^n p}(-n)B) \\ &= 0.5 \operatorname{Leb}_{\mathbb{R}}(\lambda^u_{f^n p}(-n)B) &&= \lambda^u_{f^n p}(-n) 0.5 \operatorname{Leb}_{\mathbb{R}}(B) &&= \lambda^u_{f^n p}(-n)\overline{\mu}_p(B) \\ &= \lambda^u_{f^n p}(-n)\mu_p(H_p^{-1}(B)) &&= \lambda^u_{f^n p}(-n)\mu_p(A). \end{aligned}$$

□

CHAPTER **5**

The Margulis Family

In the previous chapter we focused on the SRB measure and constructed a family of measures defined on entire unstable leaves that behaved very well under the dynamics. Now it's the turn of the measure of maximal entropy ν. Here we construct a family of measures defined on W^u and W^s which locally gives a disintegration of ν in product neighborhoods that scales with the topological entropy h_{top} of f.

This family is called the Margulis Family and their applicability in fact holds for equilibrium states in general and not just the maximal entropy measure. Since the effort to construct them for the general case is similar to our case, we present general one. The outline of the proof is to define a linear operator of functions on leaves and use the Riez Representation Theorem to obtain a measure on each leaf. For these measures to behave well we will construct them as to be a fixed point of a good action.

The main reference for this chapter is [Alv13]

5.1 A fixed point for an action

When we write $A \subseteq W^\sigma$ we mean $A \subseteq W^\sigma(p)$ for some $p \in M$, $\sigma = s, u$. Also, for $A, B \subseteq W^u$ we denote an Holonomy from A to B by $\mathcal{H}_{A\to B}$, i.e. we are just dropping the base point of the holonomy and denoting it by its domain and codomain ($\mathcal{H}_{A\to B}(A) \subseteq B$). In particular, if you interpret a point as a singleton we have $\mathcal{H}_{\{p\}\to\{q\}} = \mathcal{H}_{p\to q}$ and this notation agree with the definition 2.2.5.

Definition 5.1.1. Given $A, B \subseteq W^u$, we say A and B are ε-equivalent if there is a well defined surjective Holonomy map $\mathcal{H}_{A\to B} : A \to B$ such that $d(\mathcal{H}(x), x) < \varepsilon$, $\forall x \in B$.

Consider the set of continuous functions over unstable leaves with compact support

$$C_c(W^u) \stackrel{\text{def.}}{=} \{\varphi : M \to \mathbb{R} |\, \operatorname{supp}(\varphi) \subseteq W^u \text{ is compact and } \varphi|_{\operatorname{supp}(\varphi)} \text{ is continuous}\}$$

Definition 5.1.2. We say $\varphi_1, \varphi_2 \in C_c(W^u)$ are ε-equivalent if $\operatorname{supp}(\varphi_1)$ and $\operatorname{supp}(\varphi_2)$ are ε-equivalent via an Holonomy $\mathcal{H}$ and $\varphi_1 = \varphi_2 \circ \mathcal{H}$.

It is a curious feature of transitive systems that for any open $A \subseteq W^u$, all sufficiently small unstable balls are ε-equivalent to A. We see it in Proposition 5.1.4, just after a little definition.

Definition 5.1.3. A disk of radius ε and center in $p \in M$ is the set

$$D_\varepsilon(p) \stackrel{\text{def.}}{=} \bigcup_{q\in W^s_\varepsilon(p)} W^u_\varepsilon(q) \stackrel{\Phi}{\cong} (-1,1)^2 \subseteq \mathbb{R}^2$$

when well defined via a foliated chart $\Phi : D_\varepsilon(p) \to (-1,1)^2$. A disk can be endowed with an analogue of the sum metric in $\mathbb{R}^n$: $d^+(p,q) \stackrel{\text{def.}}{=} d^s(p,r) + d^u(r,q)$, where $r \in W^u_\varepsilon(q) \cap W^s_\varepsilon(p)$.

Proposition 5.1.4. *For any $A \subseteq W^u$ open, there exists $r, \varepsilon > 0$ such that $\forall p \in M$, $B^u_r(p)$ is ε-equivalent to a subset of A.*

Proof. Let $p \in M$ and $q \in A$. Find $r > 0$ such that there are well defined product neighborhoods $D_{2r}(p)$ and $D_{2r}(q)$. By theorem 2.3.5, $W^s(p)$ is dense. Thus there exists $\varepsilon > 0$ such that $W^s_\varepsilon(p) \bigcap D_{2r}(q) \neq \emptyset$. In fact, since W^s and W^u are continuously transverse, ε can be chosen so that for a small enough r, $W^s_\varepsilon(z)$ intersects $W^u_{2r}(q)$ in a single point for each $z \in W^u_{2r}(p)$. Thus there is a well defined map

$$\mathcal{H} : D_{2r}(p) \to D_{2r}(q)$$

which in restriction to any unstable plaque $W^u \cap D_{2r}(p)$ gives an Holonomy map. If needed, diminish r so that $D_{2r}(q)$ is above A. Now, if $z \in D_r(p)$ we have $B^u_r(z) \subseteq D_{2r}(p)$ thus $\mathcal{H}|_{B^u_r(z)}$ is an Holonomy and

$$d(\mathcal{H}|_{B^u_r(z)}(x), x) < \varepsilon \quad , \forall x \in B^u_r(z)$$

so it establish an ε-equivalence between $B^u_r(z)$ and a subset of A for every $z \in D_r(p)$. To finish, cover M by finitely many such neighborhoods $D_{r_i}(p_i)$, take r to be the minimum of such $r_i's$ and ε the maximum of each associated ε_i.

□

Let $\psi : M \to \mathbb{R}$ be a Hölder observable. We define densities k^ψ_n by

$$k^\psi_n(p) \stackrel{\text{def.}}{=} \exp\left(\sum_{l=0}^{n-1} \psi \circ f^{-l}(p)\right)$$

Thoose satisfies the following cocycle relation: $\forall n_1, n_2 \in \mathbb{N}$ and $p \in M$

$$k^\psi_{n_1+n_2}(p) = k^\psi_{n_2}(p) k^\psi_{n_1}(f^{-n_2}(p)) \tag{5.1}$$

We will be interested in the action on $C_c(W^u)$ of the functionals

$$L^\psi_n(\varphi) \stackrel{\text{def.}}{=} \int k^\psi_n \varphi \circ f^{-n} \mathrm{d}\,\mathrm{Leb}^u \tag{5.2}$$

As we explore it, we will see that they behave very well as the time n advances. In fact, fixing one observable, it is possible to compare all the others with its value. To achieve those comparisons we need many more lemmas and some very numericals as the one below

Lemma 5.1.5. *If $n \in \mathbb{N}$ and $(a_l)^n_{l=1}, (b_l)^n_{l=1} \subseteq \mathbb{R}$ are finite sequences with $b_l > 0$ then*

$$\frac{\sum_{l=1}^{n} a_l}{\sum_{l=1}^{n} b_l} \leq \max_{1 \leq l \leq n} \left(\frac{a_l}{b_l} \right)$$

Proof. Let $S = \frac{\sum_{l=1}^{n} a_l}{\sum_{l=1}^{n} b_l}$ and suppose the Lemma is false. Then, for all $0 \leq l \leq n$ we have

$$\frac{a_l}{b_l} < S \quad \Longrightarrow \quad a_l < b_l S$$

Thus

$$S = \frac{\sum_{l=1}^{n} a_n}{\sum_{l=1}^{n} b_n} < \frac{\sum_{l=1}^{n} b_l S}{\sum_{l=1}^{n} b_l} = S$$

a contradiction.

□

Lemma 5.1.6. *If M and N are manifolds endowed with the Lebesgue measure* Leb*, $g : M \to N$ is a local diffeomorphism and $\varphi_1 : M \to \mathbb{R}$, $\varphi_2 : N \to \mathbb{R}_{>0}$ are continuous and integrable, we have*

$$\frac{\int_M \varphi_1 \mathrm{d}\,\mathrm{Leb}}{\int_{g(M)} \varphi_2 \mathrm{d}\,\mathrm{Leb}} \leq \sup_{p \in M} \left(\frac{\varphi_1(p)}{\varphi_2 \circ g(p)\ Jac\, g(p)} \right)$$

Proof. By the Change of Variables Theorem

$$\int_{g(M)} \varphi_2 \mathrm{d}\,\mathrm{Leb} = \int_M \varphi_2 \circ g\ Jac\, g\ \ \mathrm{d}\,\mathrm{Leb}$$

Since φ_1 and φ_2 are continuous and integrable, their Lebesgue integral can be taken as a limit of Riemann sums. In particular, if we take a partition $P_n = \{X_{1,n}, \ldots, X_{n,n}\}$ such that $\mathrm{Leb}(X_{i,n}) = \mathrm{Leb}(M)/n$ and $x_{i,n} \in X_{i,n}$ then using lemma 5.1.5

$$\frac{\displaystyle\int_M \varphi_1 \mathrm{d}\,\mathrm{Leb}}{\displaystyle\int_{g(M)} \varphi_2 \mathrm{d}\,\mathrm{Leb}} = \lim_{n\to+\infty} \frac{\displaystyle\sum_{l=1}^{n} \varphi_1(x_{l,n})\,\mathrm{Leb}(X_{l,n})}{\displaystyle\sum_{l=1}^{n} \varphi_2(g(x_{l,n}))\ Jac\,g(x_{l,n})\,\mathrm{Leb}(X_{l,n})}$$
$$= \lim_{n\to+\infty} \frac{\displaystyle\sum_{l=1}^{n} \varphi_1(x_{l,n})}{\displaystyle\sum_{l=1}^{n} \varphi_2(g(x_{l,n}))\ Jac\,g(x_{l,n})}$$
$$\leq \lim_{n\to+\infty} \max_{1\leq l\leq n} \frac{\varphi_1(x_{l,n})}{\varphi_2(g(x_{l,n}))\ Jac\,g(x_{l,n})}$$
$$\leq \sup_{p\in M} \left(\frac{\varphi_1(p)}{\varphi_2(g(p))\ Jac\,g(p)} \right)$$

.

□

Lemma 5.1.7. *If $A \subseteq W^u$ is open with compact closure. Then, $\exists C = C(A) > 0$ such that $\forall p \in M$ and $n \in \mathbb{N}$:*

$$\int_{f^n(B^u_{r(A)}(p))} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u \leq C \int_{f^n(A)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u$$

Proof. Let $p \in M$ and $B = \mathcal{H}_{p\to A}(B^u_r(p))$ be the subset of A ε-equivalent to $B^u_r(p)$ given by the Proposition 5.1.4 .By lemma 5.1.6 we have

$$\frac{\displaystyle\int_{f^n(B^u_{r(A)}(p))} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u}{\displaystyle\int_{f^n(B)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u} \leq \sup_{q\in B^u_r(p)} \left(\frac{k_n^{\psi}(f^n(q))}{k_n^{\psi}(f^n \circ \mathcal{H}_{p\to A}(q))} \right) \sup_{q\in B^u_r(p)} \left(\frac{1}{Jac\mathcal{H}_{f^n(p)\to f^n(B)}(q)} \right)$$

Denote $\hat{q} \stackrel{\text{def.}}{=} \mathcal{H}_{p\to A}(q)$ and notice that

$$\log\left(\frac{k_n^{\psi}(f^n(q))}{k_n^{\psi}(f^n(\hat{q}))}\right) = \sum_{l=0}^{n-1} \psi(f^{n-l}(q)) - \psi(f^{n-l}(\hat{q})) \leq \sum_{l=1}^{n} \mathrm{Höl}(\psi) d(f^l(q), f^l(\hat{q}))^{\alpha}$$
$$\leq \mathrm{Höl}(\psi)\varepsilon^{\alpha} \sum_{l=1}^{n} (\lambda^s)^{\alpha l}$$
$$\leq \mathrm{Höl}(\psi)\varepsilon^{\alpha} \sum_{l=1}^{+\infty} (\lambda^s)^{\alpha l}$$

which is bounded independently of p.

Also, Proposition 4.2.7 gives us that $Jac\mathcal{H}_{f^n(p)\to f^n(B)}(q)$ is uniformly close to 1, hence the second term is also bounded. Finally, $f^n(B) \subseteq f^n(A)$ and k_n^{ψ} is positive, so

$$\int_{f^n(B)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u \leq \int_{f^n(A)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u$$

Gathering all those estimates and applying to the first equation, finishes the demonstration.

□

Lemma 5.1.8. *Let $\varphi_0 \in C_c(W^u)$ be positive and non-null. Then, for all $\varphi \in C_c(W^u)$ non-negative, there exists a $C = C(\varphi) > 0$ such that $\forall n \in \mathbb{N}$:*

$$L_n^{\psi}(\varphi) \leq C L_n^{\psi}(\varphi_0)$$

Proof. Let $\varepsilon > 0$ be small enough so that $A \overset{\text{def.}}{=} \{x \in M | \varphi_0(x) > \varepsilon\}$ is non-empty. In particular, A is open with compact closure, so we can obtain a constant $r(A) > 0$ given by Proposition 2.2.4.

Let $K \subseteq W^u$ be a compact such that $\operatorname{supp}(\varphi) \subseteq K$ and cover K by a finite number N of u-balls of radius $r(A)$. By lemma 5.1.7 we have:

$$\begin{aligned}\int k_n^{\psi} \varphi \circ f^{-n} \mathrm{d}\,\mathrm{Leb}^u \leq \|\varphi\|_\infty \int_{f^n(K)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u \quad &\leq \|\varphi\|_\infty N C(A) \int_{f^n(A)} k_n^{\psi} \mathrm{d}\,\mathrm{Leb}^u \\ &\leq \frac{\|\varphi\|_\infty N C(A)}{\varepsilon} \int k_n^{\psi} \varphi_0 \circ f^{-n} \mathrm{d}\,\mathrm{Leb}^u \\ &= C(\varphi) L_n^{\psi}(\varphi_0)\ , \quad C(\varphi) \overset{\text{def.}}{=} \frac{\|\varphi\|_\infty N C(A)}{\varepsilon}\end{aligned}$$

□

Now, consider the set $\mathcal{L}$ of functions L on $C_c(W^u)$ that restrained to each unstable leaf is linear:

$$\mathcal{L} \overset{\text{def.}}{=} \{L : C_c(W^u) \to \mathbb{R} |\ \forall p \in M\ ,\ \ L|_{C_c(W^u(p))} \text{ is linear}\}$$

$\mathcal{L}$ is a vector space and can be endowed with the product topology via the identification:

$$L \in \mathcal{L} \mapsto (L(\varphi))_{\varphi \in C_c(W^u)} \in \prod_{\varphi \in C_c(W^u)} \mathbb{R}_\varphi \tag{5.3}$$

The dynamics naturally acts in $\mathcal{L}$ by the action F^{ψ} given by

$$F_n^{\psi} L(\varphi) \overset{\text{def.}}{=} L(k_n^{\psi} \varphi \circ f^{-n})$$

We fix a $\varphi_0 \in C_c(W^u)$ positive with $\varphi_0 \geq 1$ for an open $A \subseteq W^u$ and set a renormalized action $\hat{F}^{\psi}$ by

$$\hat{F}_n^{\psi} L \overset{\text{def.}}{=} \frac{F_n^{\psi} L}{F_n^{\psi} L(\varphi_0)}$$

Lemma 5.1.9. *The action of F_n^{ψ} is continuous.*

Proof. An open in the base of the product topology is given by taking a finite number of intervals $I_1, \ldots, I_N$ and functions $\varphi_1, \ldots, \varphi_N$ in $C_c(W^u)$ to obtain a set of the form:

$$U \overset{\text{def.}}{=} \{L \in \mathcal{L} |\ L(\varphi_i) \in I_i\ i = 1, \ldots, N\}$$

The action simply gives

$$(F_n^\psi)^{-1}U = \{L \in \mathcal{L}|\ L(k_n^\psi \varphi_i \circ f^{-n}) \in I_i\ i = 1, \dots, N\}$$

which is clearly open, since $k_n^\psi \varphi_i \circ f^{-n} \in C_c(W^u)$.

□

The cocycle relations in (5.1) implies F^ψ and $\hat{F}^\psi$ satisfies the following properties

$$F_{n_1}^\psi F_{n_2}^\psi L(\varphi) = F_{n_1}^\psi L(k_{n_2}^\psi \varphi \circ f^{-n_2}) = L(k_{n_2+n_1}^\psi \varphi \circ f^{-(n_2+n_1)}) = F_{n_1+n_2}^\psi L(\varphi)$$

$$\hat{F}_{n_1}^\psi \hat{F}_{n_2}^\psi L = \hat{F}_{n_1}^\psi \frac{F_{n_2}^\psi L}{F_{n_2}^\psi L(\varphi_0)} = \frac{F_{n_1}^\psi \frac{F_{n_2}^\psi L}{F_{n_2}^\psi L(\varphi_0)}}{F_{n_1}^\psi \frac{F_{n_2}^\psi L(\varphi_0)}{F_{n_2}^\psi L(\varphi_0)}} = \frac{F_{n_1}^\psi F_{n_2}^\psi L}{F_{n_1}^\psi F_{n_2}^\psi L(\varphi_0)} = \frac{F_{n_1+n_2}^\psi L}{F_{n_1+n_2}^\psi L(\varphi_0)} = \hat{F}_{n_1+n_2}^\psi L$$

and acting in the previously defined functionals (5.2) we have

$$F_{n_1}^\psi L_{n_2}^\psi(\varphi) = L_{n_2}^\psi(k_{n_1}\varphi \circ f^{-n_1}) = \int k_{n_1+n_2}^\psi \varphi \circ f^{-(n_1+n_2)} = L_{n_1+n_2}^\psi(\varphi)$$

$$\hat{F}_{n_1}^\psi \left(\frac{L_{n_2}^\psi}{L_{n_2}^\psi(\varphi_0)}\right) = \frac{F_{n_1}^\psi \left(\frac{L_{n_2}^\psi}{L_{n_2}^\psi(\varphi_0)}\right)}{F_{n_1}^\psi \left(\frac{L_{n_2}^\psi(\varphi_0)}{L_{n_2}^\psi(\varphi_0)}\right)} = \frac{F_{n_1}^\psi L_{n_2}^\psi}{F_{n_1}^\psi L_{n_2}^\psi(\varphi_0)} = \frac{L_{n_1+n_2}^\psi}{L_{n_1+n_2}^\psi(\varphi_0)}$$

The first two properties shows that F^ψ and $\hat{F}^\psi$ are in fact actions. And, the last property means that $\hat{F}^\psi$ preservers the space of functionals of the form $\frac{L_n^\psi}{L_n^\psi(\varphi_0)}$. We denote the closure of all the convex combinations of thoose functionals by $\mathscr{X}_0$:

$$\mathscr{X}_0 \overset{\text{def.}}{=} \overline{\text{convhull}}\left(\left\{\frac{L_n^\psi}{L_n^\psi(\varphi_0)} | n \geq 0\right\}\right)$$

Lemma 5.1.10. *All $L \in \mathscr{X}_0$ are positive, i.e. if $\varphi \in C_c(W^u)$ and $\varphi \geq 0$ then $L(\varphi) \geq 0$*

Proof. Clearly each L_n^ψ is positive and this property is maintained by either convex combinations and limits.

□

Lemma 5.1.11. *$\mathscr{X}_0$ is compact*

Proof. Suppose $\varphi \in C_c(W^u)$ is non-negative. Then by lemma 5.1.8 $\exists C(\varphi) > 0$ such that $\frac{L_n^\psi}{L_n^\psi(\varphi_0)} \leq C(\varphi)\ \forall n \in \mathbb{N}$. This estimate is kept by convex combinations and limts, thus, it holds for every $L \in \mathscr{X}_0$. Also, by taking positive and negative parts and by linearity, if $\varphi \in C_c(W^u)$ is arbitrary, the same holds. I.e. $\exists C = C(\varphi) > 0$ such that $\forall L \in \mathscr{X}_0$.

$$-C(\varphi) \leq L(\varphi) \leq C(\varphi)$$

By the identification 5.3, it means that $\mathscr{X}_0$ is contained in $\prod_{\varphi \in C_c(W^u)}[-C(\varphi), C(\varphi)]$, which by Tychonoff's Theorem is compact. Since $\mathscr{X}_0$ is closed, it means that $\mathscr{X}_0$ is compact.

□

For $N \in \mathbb{N}$ we can define the subset of $\mathscr{X}_0$ given by

$$\mathscr{X}_N \stackrel{\text{def.}}{=} \overline{\text{convhull}} \left(\left\{ \frac{L_n^\psi}{L_n^\psi(\varphi_0)} | n \geq N \right\} \right)$$

and their intersection

$$\mathscr{X}_\infty \stackrel{\text{def.}}{=} \bigcap_{N \geq 0} \mathscr{X}_N$$

which by Lemma 5.1.11, being the intersection of decreasing compact sets, is non-empty.

Lemma 5.1.12. $\exists m \in \mathscr{X}_\infty$ *such that* $\hat{F}_n^\psi m = m \;\; \forall n \in \mathbb{N}$.

Proof. As we saw, $\hat{F}_n^\psi$ is continuous and $\hat{F}_n^\psi \mathscr{X}_N \subseteq \mathscr{X}_N$. Thus, $\hat{F}_n^\psi \mathscr{X}_\infty \subseteq \mathscr{X}_\infty$. Also, $\mathscr{X}_\infty$ is compact and convex. Thus, by Tychonoff's fixed point theorem, $\exists m \in \mathscr{X}_\infty$ such that $\hat{F}_1^\psi m = m$. Since $\hat{F}^\psi$ is an action, m is fixed for all $n \in \mathbb{N}$.

□

5.2 Estimates on the fixed point

This fixed point satisfies some very special properties. To see it, given $p \in M$, for $q \in W^s(p)$ and $r \in W^u(p)$ define

$$k_\psi^s(p, q) \stackrel{\text{def.}}{=} \exp \left(\sum_{l=1}^{+\infty} \psi \circ f^l(q) - \psi \circ f^l(p) \right)$$

$$k_\psi^u(p, r) \stackrel{\text{def.}}{=} \exp \left(\sum_{l=1}^{+\infty} \psi \circ f^{-l}(r) - \psi \circ f^{-l}(p) \right)$$

Since ψ is Hölder and stable and unstable leaves are contracting for f and f^{-1} respectively; the sums above converges. With that definition we have:

Lemma 5.2.1. *If* $\varphi \in C_c(W^u)$, $p \in \text{supp}(\varphi)$ *and* $\mathcal{H}_{p \to p'}^s : K \to K'$ *is a stable Holonomy between* p *and* $p' \in W^s(p)$ *then*

$$m(\varphi \circ \mathcal{H}_{p' \to p}^s) = m(k_\psi^s(\cdot, \mathcal{H}_{p \to p'}^s(\cdot))\varphi)$$

Proof. We'll show that

$$m_N \left(\frac{\varphi \circ \mathcal{H}_{p' \to p}^s}{k_\psi^s(\mathcal{H}_{p' \to p}^s(\cdot), \cdot)} \right) - m_N(\varphi) \stackrel{N \to +\infty}{\longrightarrow} 0$$

independently of $m_N \in \mathscr{X}_\infty$. This shows that

$$m_\infty \left(\frac{\varphi \circ \mathcal{H}_{p' \to p}^s}{k_\psi^s(\mathcal{H}_{p' \to p}^s(\cdot), \cdot)} \right) = m_\infty(\varphi)$$

for any $m_\infty \in \mathscr{X}_\infty$. So, using the linearity of m_∞ and arbitrariness of $K \subseteq \text{supp}(\varphi)$, it concludes the Lemma.

Let $N \geq 0$ and define

$$\phi \stackrel{\text{def.}}{=} \frac{\varphi \circ \mathcal{H}^s_{p' \to p}}{k^s_\psi(\mathcal{H}^s_{p' \to p}(\cdot), \cdot)}$$

Changing variables we have

$$\begin{aligned} L^\psi_n(\phi) &= \int_{f^N(K')} \frac{k^\psi_N(q')\varphi \circ \mathcal{H}^s_{p' \to p}(f^{-n}(q))}{k^s_\psi(\mathcal{H}^s_{p' \to p}(f^{-N}(q'), f^{-N}(q'))} \mathrm{d}\,\mathrm{Leb}^u(q') \\ &= \int_{f^N(K)} \frac{k^\psi_N(\mathcal{H}^s_{f^N(p) \to f^N(p')}(q))}{k^s_\psi(f^{-N}(q), \mathcal{H}^s_{p \to p'}(f^{-N}(q))} \varphi \circ f^{-N}(q) Jac\mathcal{H}^s_{f^N(p) \to f^N(p')}(q) \mathrm{d}\,\mathrm{Leb}^u(q) \end{aligned}$$

For $q \in f^N(K)$ write $q' = \mathcal{H}^s_{f^N(p) \to f^N(p')}(q)$ and define

$$R_N(q) \stackrel{\text{def.}}{=} \frac{k^\psi_N(q')}{k^\psi_N(q) k^s_\psi(f^{-N}(q), f^{-N}(q'))} Jac\mathcal{H}^s_{f^N(p) \to f^N(p')}(q)$$

With this we see that

$$L^\psi_n(\phi) = \int_{f^N(K)} R_N(q) k^\psi_N(q) \varphi \circ f^{-n}(q) \mathrm{d}\,\mathrm{Leb}^u(q)$$

Thus

$$L^\psi_N(\phi) - L^\psi_N(\varphi) = \int_{f^N(K)} (R_N(q) - 1) k^\psi_N(q) \varphi \circ f^{-N}(q) \mathrm{d}\,\mathrm{Leb}^u(q)$$

And

$$\left| \frac{L^\psi_N(\phi)}{L^\psi_N(\varphi_0)} - \frac{L^\psi_N(\varphi)}{L^\psi_N(\varphi_0)} \right| \leq \sup_{q \in f^N(K)} |R_N(q) - 1| \frac{L^\psi_N(\varphi)}{L^\psi_N(\varphi_0)} \tag{i}$$

By Lemma 5.1.8, $\frac{L^\psi_N(\varphi)}{L^\psi_N(\varphi_0)}$ is bounded independently of $N \geq 0$. Also, by Proposition 4.2.7 and the fact that $p' \in W^s(p)$ we have $Jac\mathcal{H}^s_{f^N(p) \to f^N(p')} \to 1$. Finally

$$\frac{k^\psi_N(q')}{k^\psi_N(q) k^s_\psi(f^{-N}(q), f^{-N}(q'))} = \exp\left(\sum_{l=N}^{+\infty} \psi \circ f^l(f^{-N}(q)) - \psi \circ f^l(f^{-N}(q')) \right)$$

which goes to zero since ψ is Hölder, $d(f^{-N}(q), f^{-N}(q')) \leq d(K, K')$ is bounded and $q' \in W^s(q)$. Thus $R_N(q)$ goes to 1 and so (i) goes to 0.

As any $m_N \in \mathscr{X}_N$ can be written as the limit of convex combinations of $\frac{L^\psi_N}{L^\psi_N(\varphi_0)}$, all the above estimates holds for m_N. I.e. $|m_N(\phi) - m_N(\varphi)| \to 0$. Thus $m_\infty(\phi) = m_\infty(\varphi)$ as we wanted.

□

Proposition 5.2.2. *There exists $P_\psi \in \mathbb{R}$ such that $\forall \varphi \in C_c(W^u)$ and $\forall n \in \mathbb{N}$:*

$$m(\varphi \circ f^n) = e^{-NP_\psi} m(k^\psi_n \varphi)$$

Proof. Let $n \in \mathbb{N}$ and write $a_n = m(k_n^\psi \varphi_0 \circ f^{-n})$. Then, by 5.1 and the fact that m is a fixed point, we have for $m, n \in \mathbb{N}$

$$
\begin{aligned}
a_{m+n} = m(k_m^\psi k_n^\psi \circ f^{-m} \varphi_0 \circ f^{-n} \circ f^{-m}) &= m(k_m^\psi \varphi_0 \circ f^{-m}) \hat{F}_m^\psi m(k_n^\psi \varphi_0 \circ f^{-n}) \\
&= m(k_m^\psi \varphi_0 \circ f^{-m}) m(k_n^\psi \varphi_0 \circ f^{-n}) \\
&= a_m a_n
\end{aligned}
$$

thus $\exists P_\psi \in \mathbb{R}$ such that $a_n = a_0 e^{nP_\psi}$. Since $a_0 = m(\varphi_0) = 1$, we have $a_n = e^{nP_\psi}$.
Again, with more computations we get

$$
m(\varphi \circ f^n) = \hat{F}_n^\psi m(\varphi \circ f^n) = \frac{m(k_n^\psi \varphi)}{m(k_n^\psi \varphi_0 \circ f^{-n})} = e^{-nP_\psi} m(k_n^\psi \varphi)
$$

□

Now, each leaf in W^u is a locally compact Hausdorff space and in restriction to it, m is a positive linear functional. Thus, by the Riez Representation Theorem, we obtain a measure $\nu_{\psi,p}^u$ in each leaf. In fact, by re-constructing everything above for f^{-1}, we also obtain a family of measures $\nu_{\psi,p}^s$ in the stable leaves. Precisely:

Theorem 5.2.3. *There exists two family of measures $(\nu_{\psi,p}^u)_{p \in M}$ and $(\nu_{\psi,p}^s)_{p \in M}$ in the unstables and stable leaves respectively, such that*

a) *(Leafwise) If $q \in W^u(p)$ then $\nu_{\psi,q}^u = \nu_{\psi,p}^u$. And if $r \in W^s(p)$ then $\nu_{\psi,r}^s = \nu_{\psi,p}^s$.*

b) *(Holonomy deformation) For $p' \in W^s(p)$ and q in the domain of a stable Holonomy $\mathcal{H}_{p \to p'}^s$:*

$$
\frac{\mathrm{d}\mathcal{H}_{p' \to p}^s * \nu_{\psi,p'}^u}{\mathrm{d}\nu_{\psi,p}^u}(q) = k_\psi^s(q, \mathcal{H}_{p \to p'}^s(q))
$$

For $p'' \in W^u(p)$ and r in the domain of an unstable Holonomy $\mathcal{H}_{p \to p''}^u$:

$$
\frac{\mathrm{d}\mathcal{H}_{p'' \to p}^u * \nu_{\psi,p''}^s}{\mathrm{d}\nu_{\psi,p}^s}(r) = k_\psi^u(r, \mathcal{H}_{p \to p''}^u(r))
$$

c) *(Dynamical deformation) There exists real numbers P_ψ, P'_ψ such that $\forall p \in M$ and $n \in \mathbb{N}$ we have:*

$$
\frac{\mathrm{d}f^n * \nu_{\psi,f^{-n}(p)}^u}{\mathrm{d}\nu_{\psi,p}^u}(p) = \exp\left(\sum_{l=0}^{n-1} (\psi \circ f^{-n}(p) - P_\psi)\right)
$$

and

$$
\frac{\mathrm{d}f^n * \nu_{\psi,f^{-n}(p)}^s}{\mathrm{d}\nu_{\psi,p}^s}(p) = \exp\left(-\sum_{l=0}^{n-1} (\psi \circ f^{-n}(p) - P'_\psi)\right)
$$

Proof. As said, the existence is just a consequence of the Riez representation Theorem. And to obtain properties (b) and (c), apply Lemma 5.2.1 and Proposition 5.2.2.

□

Property (b) shows that both measures varies continuously along the domain of the other. Thus we can integrate one with respect to the other to obtain a measure in M.

Lemma 5.2.4. *There exists a probability measure ν_ψ in M that is locally given by the integration of $\rho^u_{\psi,p}\nu^u_{\psi,p}$ by ν^s_{ψ,p_0}. Where $\rho^u_{\psi,p}(q) \overset{\text{def.}}{=} k^u_\psi(p,q)$.*

Proof. The local existence of this measure is given by the continuity with respect to the holonomies as discoursed in the remark above. The measure is finite because $\nu^u_{\psi,p}$ and ν^s_{ψ,p_0} are locally finite and $\mathbb{T}^2$ is compact. Thus, up to a constant multiplication, we can say it is a probability.

What remains to prove is that these local expressions are coherent. I.e. the value measured does not depend in which stable leaves you choose to integrate $\rho^u_{\psi,p}\nu^u_{\psi,p}$. But it is no problem, since the Holonomy deforms ν^s_{ψ,p_0} as

$$\frac{\rho^u_{\psi,p_0}}{\rho^u_{\psi,p_0''}} = k^u_\psi(p_0,p_0'') = \frac{\mathrm{d}\mathcal{H}^u_{p_0''\to p_0} * \nu^s_{\psi,p_0''}}{\mathrm{d}\nu^s_{\psi,p_0}}(p_0)$$

Thus the collage is well defined.

□

5.3 Passing to the measure

This measure ν_ψ is an equilibrium state. To see it we need to pass these estimates on the fixed point to it. Lets start by seeing that it is atleast invariant:

Lemma 5.3.1. *We have $P_\psi = P'_\psi$, thus ν_ψ is f-invariant.*

Proof. By integrating $A \subseteq M$ and using how the dynamics deforms the measure (property (c)) we see that $\nu_\psi(f^n(A)) = e^{n(P_\psi - P'_\psi)}\nu_\psi(A)$. Putting $A = M$ we obtain $P_\psi = P'_\psi$. And thus, ν_ψ is invariant.

□

To confirm that ν_ψ is an equilibrium state is basically to show that P_ψ is the pressure of f with respect to ψ. For it we will need some computation. But before it, we need two more lemmas.

Lemma 5.3.2. *There exists $\varepsilon_0 > 0$ and $C > 1$ such that for all $p \in \mathbb{T}^2$ and $\varepsilon < \varepsilon_0$*

$$B_{C^{-1}\varepsilon}(p) \subseteq D_\varepsilon(p) \subseteq B_{C\varepsilon}(p)$$

Proof. Since distances in leaves are surely atleast bigger that in the entire manifold, i.e. $d(p,q) \le d^+(p,q)$ (when defined). We have

$$D_\varepsilon(p) \subseteq B_{2\varepsilon}(p)$$

Thus it remais to show that $\exists C > 2$ such that $d^+(p,q) \le C d(p,q)$.

Let $p \in \mathbb{T}^2$ and $\varepsilon_0 > 0$ be so that $D_{3\varepsilon_0}(p)$ is well defined. Put $U = D_{\varepsilon_0}(p)$, and notice that $\forall q \in U$ we have a well defined disk $D_{\varepsilon_0}(q) \subseteq K \overset{\text{def.}}{=} \overline{D_{2\varepsilon_0}(p)}$. Consider a foliated chart

$$\Phi : D_{3\varepsilon_0}(p) \to (-1,1)^2 \subseteq \mathbb{R}^2$$

and set $M \stackrel{\text{def.}}{=} \sup_{x\in K}\{Jac\Phi(x), Jac\Phi^{-1}(x)\}$. If $p, q \in U$ and $r = (p, q)$ we have

$$
\begin{aligned}
d^+(p,q) = \int_{p\to r} \mathrm{d}l^s + \int_{r\to q} \mathrm{d}l^u \quad &= \int_{\Phi(p)}^{\Phi(r)} Jac\Phi \mathrm{d}l + \int_{\Phi(r)}^{\Phi(q)} Jac\Phi \mathrm{d}l \\
&\le M[d(\Phi(p),\Phi(r)) + d(\Phi(r),\Phi(q))] \\
&= M d^+_{\mathbb{R}^2}(\Phi(p),\Phi(q))
\end{aligned}
$$

But, in $\mathbb{R}^2$, the metrics $d_{\mathbb{R}^2}$ and $d^+_{\mathbb{R}^2}$ are strongly equivalent, thus $\exists C_0 > 0$ such that $d^+_{\mathbb{R}^2} \le C_0 d_{\mathbb{R}^2}$ and we obtain

$$d^+(p,q) \le M d^+_{\mathbb{R}^2}(\Phi(p),\Phi(q)) \le MC d_{\mathbb{R}^2}(\Phi(p),\Phi(q)) \le M^2 C_0 d(p,q)$$

Setting $C = \max\{2, MC_0\}$ concludes the lemma for $q \in U$. To extend it to the entire torus, cover $\mathbb{T}^2$ by finite such U's, take the minimum of all ε_0's and the maximum of all C's.

□

Lemma 5.3.3. $\forall \varepsilon > 0$, $\exists C_\varepsilon > 1$ *such that* $\forall p \in \mathbb{T}^2$, $C_\varepsilon^{-1} < \nu_\psi(B_\varepsilon(p)) < C_\varepsilon$.

Proof. Since ν_ψ is a probability, the upper bound is given by $C_\varepsilon > 1$. Also, we may just show the lemma for sufficiently small ε. Let ε_0 and C be the constants given by lemma 5.3.2. For $0 < \varepsilon < \varepsilon_0$, let $0 < \delta < C^{-1}\varepsilon/2$ be small enough so that $D_\delta(p)$ is well defined for all $p \in \mathbb{T}^2$. By compacity, cover $\mathbb{T}^2$ by finite many disks $D_\delta(p_i)$, $i = 1, \dots, k$. Notice that, with those choices, for every $p \in \mathbb{T}^2$, there is an $i \in \{1, \dots, l\}$ such that $D_\delta(p_i) \subseteq B_\varepsilon(p)$. Hence, it suffices to see that $\nu_\psi(D_\delta(p_i)) > C_\varepsilon^{-1}$ for some $C_\varepsilon > 1$. For it, let χ_i, $i = 1, \dots, k$ be continuous bump functions satisfying

$$\chi_i|_{D_{\delta/2}(p_i)} \equiv 1 \text{ and } \chi_i|_{D^c_\delta(p_i)} \equiv 0$$

For every $q \in W^s_{\delta/2}(p_i)$, $\chi_i^q \stackrel{\text{def.}}{=} \chi_i|_{W^u(q_i)} \in C_c(W^u)$ is positive and non-null. Thus, we are under the hypothesis of lemma 5.1.8 and there exists $C^u_q > 0$ such that, for all $n \in \mathbb{N}$,

$$C^u_q \le \frac{L_n(\chi_i^q)}{L_n(\varphi_0)}$$

Since the measures λ^u_q are the limits of convex combinations of such quotients, we must have $\lambda_q(\chi_i^q) > C^u_q > 0$. Similarly, $\lambda^s_{p_i}(\chi_i|_{W^s_{\delta/2}(p_i)}) > 0$. Now, the measure ν_ψ, in a product disc, is given by the product measure of λ^u and λ^s, which we have just shown are strictly positive at the subsets $D_{\delta/2}(p_i)$ of $D_\delta(p_i)$. Thus $\nu_\psi(D_\delta(p_i)) > 0$ for $i = 1, \dots, k$ and taking C_ε^{-1} as the minimum of those values, the lemma is proven.

□

We recall that the time n metric by f is given by

$$d_n(p,q) = \max_{0\le l\le n-1} d(f^l(p), f^l(q))$$

and denote by $B_{\varepsilon,n}(p)$ the time n ball of radius ε at p. Restricting this metric to paths contained in leaves, we can similarly to def 5.1.3 define $D_{\varepsilon,n}(p) = \bigcup_{q\in W^s_{\varepsilon,n}(p)} W^u_{\varepsilon,n}(q)$ as the time n disk of radius ε at p, where $W^\sigma_{\varepsilon,n}(p)$ is the ε neighborhood of p in its W^σ leaf with respect to this metric.

Lemma 5.3.4. *There exist* $\varepsilon_0 > 0$ *such that for* $0 < \varepsilon < \varepsilon_0$, $\exists E_\varepsilon > 0$ *such that* $\forall p \in \mathbb{T}^2$ *and* $n \geq 0$ *we have*

$$E_\varepsilon^{-1} \leq \frac{\nu_\psi(B_{\varepsilon,n}(p))}{\exp\left(\sum_{l=1}^{n} \psi \circ f^l(p) - P_\psi\right)} \leq E_\varepsilon$$

Proof. By lemma 5.3.2, it suffices to prove the lemma by replacing $B_\varepsilon(p)$ with $D_\varepsilon(p)$. Using a metric adapted to f as in Prop. 2.1.2, the instantaneous contraction of W^s and dilatation of W^u gives us $W^s_{\varepsilon,n}(p) = W^s_\varepsilon(p)$ and $W^u_{\varepsilon,n}(p) = f^{-n}(W^u_\varepsilon(f^n(p)))$. Also, for $q \in W^u_\varepsilon(p)$ we have

$$\left|\log\left(\frac{\exp\left(\sum_{l=1}^{n} \psi \circ f^l(q)\right)}{\exp\left(\sum_{l=1}^{n} \psi \circ f^l(p)\right)}\right)\right| \leq \sum_{l=1}^{n} |\psi(f^l(q)) - \psi(f^l(p))| \leq \varepsilon^\alpha H\ddot{o}l(\psi) \sum_{l=1}^{+\infty} \lambda^{l\alpha} < C$$

Thus $\exp\left(\sum_{l=1}^{n} \psi \circ f^l(q)\right) \leq e^C \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p)\right)$. Swapping p and q we get

$$e^{-C} \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p)\right) \leq \exp\left(\sum_{l=1}^{n} \psi \circ f^l(q)\right) \leq e^C \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p)\right)$$

Let $E_\varepsilon = C_\varepsilon e^C$, where C_ε is the constant given by lemma 5.3.3. Property c) of Theorem 5.2.3 then gives

$$E_\varepsilon^{-1} \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p) - P_\psi\right) \leq \nu_\psi(D_{\varepsilon,n}(p)) \leq E_\varepsilon \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p) - P_\psi\right)$$

as desired.

□

We say that a set E is (n, ε)-separated if $d_n(p, q) \geq \varepsilon$ for all $p, q \in E$. We remember that the pressure of an observable ψ with respect to f is given by the beautiful expression below

$$P(\psi) = \lim_{\varepsilon\to 0} \limsup_{n\to+\infty} \frac{1}{n} \sup\left\{\log\left(\sum_{p\in E} \exp\left(\sum_{l=0}^{n-1} \psi \circ f^l(p)\right)\right) \mid E \text{ is } (n, \varepsilon) \text{ separated}\right\}$$

Lemma 5.3.5. P_ψ *is the pressure* $P(\psi)$.

Proof. Let E be a maximal (n, ε)-separated set. In particular, being maximal implies $\mathbb{T}^2 = \bigcup_{p\in E} B_{\varepsilon,n}(p)$, thus by the precedent lemma

$$1 \leq \sum_{p\in E} \nu_\psi(B_{\varepsilon,n}(p)) \leq E_\varepsilon \sum_{p\in E} \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p) - P_\psi\right)$$

so that $P_\psi \leq P(\psi \circ f)$. Now, E being (n, ε)-separated implies that the time n balls $B_{\varepsilon/2,n}(p)$ are all mutually disjoint. Thus lemma 5.3.4 gives

$$1 \geq \sum_{p\in E} \nu_\psi(B_{\varepsilon/2,n}(p)) \geq E_{\varepsilon/2}^{-1} \sum_{p\in E} \exp\left(\sum_{l=1}^{n} \psi \circ f^l(p) - P_\psi\right)$$

so that $P(\psi \circ f) \leq P_\psi$, hence $P_\psi = P(\psi \circ f)$. By a theorem of Walters (Teo. 9.7 (vii))[Wal00] $P(\psi \circ f) = P(\psi)$. So $P_\psi = P(\psi)$.

□

Lemma 5.3.6. *The measure ν_ψ is a Gibbs state for ψ with respect to f.*

Proof. Using Birkov's Ergodic Theorem, lemma 5.3.4 readily gives

$$\lim_{\varepsilon \to 0} \limsup_{n\to\infty} \frac{-\log \nu_\psi(B_{\varepsilon,n}(p))}{n} \leq P_\psi - \int \psi \mathrm{d}\nu_\psi \leq \lim_{\varepsilon \to 0} \liminf_{n\to\infty} \frac{-\log \nu_\psi(B_{\varepsilon,n}(p))}{n}$$

So the limits above exists and are equal. A theorem by Brin & Katok [BK83] states that this limit is the topological entropy $h_{top}(f)$ of f. Thus ν_ψ is a measure maximizing the Variational Principle and the lemma is proven.

□

CHAPTER **6**

Lyapunov Exponents Rigidity

Now, we have everything ready to prove the Main Theorem and its corollary. We will start by proving lemma 6.1.1 which, if you recall that we are assuming $\mathrm{h}_{\mathrm{top}}(f) = \mathrm{h}_{\mu}(f)$, is Ledrappier-Young's formula in out context. Then, after some technical lemmas, we use it to prove lemma 6.1.5 and show that the conjugacy is Lipschitz along the unstable foliation.

To pass from Lipschitz continuity to $C^{1+\alpha}$ regularity will only require absolutely continuity from h.

In the second section, we will prove the conservative case (corollary 1.0.2) which, will be reduced to an application of Journé's lemma.

6.1 Proof of The Main Theorem

Let $f \in \mathrm{Diff}^2_{\mu}(\mathbb{T}^2)$, μ be the SRB measure for f and suppose that $h_\mu(f) = h_{top}(f)$. Since μ is both a SRB and a measure of maximal entropy, we have two associated family of measures: The leaf-wise measures $\{\mu_p\}_{p\in\mathbb{T}^2}$ and the Margulis family $\{\nu^u_p\}_{p\in\mathbb{T}^2}$ associated to the zero potential $\psi = 0$. Let $\xi < W^u$ be a subordinated partition. By the superposition property (Cor. 4.3.9) their mutual normalization must agree, i.e.

$$\frac{\mu_p}{\mu_p(\xi(p))} = \frac{\nu^u_p}{\nu^u_p(\xi(p))}$$

This leads to the following equality:

Lemma 6.1.1. $\lambda^u_f(\mu) = h_{top}(f)$

Proof. By the Poincaré Recurrence Theorem, we can find $p \in \mathbb{T}^2$ recurrent. For each $n \in \mathbb{N}$ let $A_n \overset{\text{def.}}{=} H^{-1}_{f^n(p)}([-1,1])$. By Lemma 4.3.17 and Theo. 5.2.3 item (c) we have

$$\mu_p(f^{-n}(A_n)) = \lambda^u_{f^n p}(-n)\mu_{f^n(p)}(A_n) = \lambda^u_{f^n p}(-n)0.5\,\mathrm{Leb}_R([-1,1]) = \lambda^u_{f^n p}(-n)$$

and

$$\nu^u_p(f^{-n}(A_n)) = e^{-nh_{top}(f)}\nu^u_{f^n(p)}(A_n)$$

Thus, by the remark above and a little rearrangement, we have

$$\lambda^u_{f^n p}(-n) = \frac{\mu_p(\xi(p))}{\nu^u_p(\xi(p))}\nu^u_{f^n(p)}(A_n)e^{-nh_{top}(f)} \tag{$*$}$$

Since p is recurrent we can take a subsequence $(f^{n_k}(p))_{k\in\mathbb{N}}$ of $(f^n(p))_{n\in\mathbb{N}}$ such that $f^{n_k}(p) \to p$. Thus, since A_n are of uniformly bounded length, all A_{n_k} lies on a compact set of $\mathbb{T}^2$. Hence, because the measures ν_p^u are locally finite and varies continuously (Theorem 5.2.3 item (b)), we conclude that $\nu_{f^n(p)}(A_n)$ is bounded. Applying the logarithm, dividing by n_k and taking the limit in the expression above we have

$$\lim_{k\to+\infty}\frac{1}{n_k}\log\lambda^u_{f^np}(-n_k) = -h_{top}(f)$$

Using that $\lambda^u_{f^np}(-n_k) = (\lambda^u_p(n_k))^{-1}$ we obtain exactly what we wanted. □

Remark 6.1.2. By our hypothesis, the above lemma can be written as $\lambda^u_f(\mu) = h_\mu(f)$. This in turn, is Ledrappier-Young's Theorem, which is true in a much wider setting. Here, we recovered this result in a particular case by exploring pure geometrical properties of the system.

Let f_A and $h : \mathbb{T}^2 \to \mathbb{T}^2$ be as in the Franks-Newhouse Theorem 2.3.3. We already have that h is Hölder, and we want to promote it to C^1 in the unstable direction. For it, we first will promote it to Lipschitz. And for that, we will need to measure some sets:

Definition 6.1.3. For $J \subseteq W^u(p)$ measurable, we denote by $|J|_p$ it's length in normal forms at p and $|J|$ it's length in W^u. I.e. $|J|_p \overset{\text{def.}}{=} \mathrm{Leb}_{\mathbb{R}}(H_p(J))$ and $|J| \overset{\text{def.}}{=} \mathrm{Leb}^u(J)$. Also, if $q, r \in W^u(p)$ we write $[q, r]$ to denote the shortest interval in $W^u(p)$ connecting q and r.

In particular, for $q, r \in W^u(p)$, we have $|[q, r]| = d^u(q, r)$. And since normal forms are C^1 along unstable leaves and their Jacobian is 1 at the base point (Prop. 4.2.1 item (ii)), they are locally Lipschitz. This means that if $\varepsilon > 0$, there exists a $C > 1$ such that for any set $J \subseteq W^u_\varepsilon(p)$ we have

$$\frac{1}{C}|J| \le |J|_p \le C|J|$$

This is interesting, for the dynamics is really well behaved under normal forms, in the sense that $|f^n(J)|_{f^n(p)} = \lambda^u_p(n)|J|_p$. And fortunately, even though the equivalence between lengths is only local, the following lemma says that, somehow, to promote a Hölder function to Lipschitz is a local matter.

Lemma 6.1.4. *If for a $p \in \mathbb{T}^2$ there exists a $\delta_0 > 0$ and a $C > 0$ such that for any interval $J \subseteq W^u(p)$ such that $|J| < \delta_0$ we have*

$$|h(J)| \le C|J|$$

Then $h|_{W^u(p)}$ is Lipschitz.

Proof. Suppose that the conclusion is false. Then there exists sequences $(q_n)_{n\in\mathbb{N}}, (r_n)_{n\in\mathbb{N}} \subseteq W^u(p)$ such that

$$d^u(h(q_n), h(r_n)) \ge n d^u(q_n, r_n)$$

I claim that $d^u(q_n, r_n) \to 0$ as $n \to +\infty$. In fact, since h is Hölder along unstable leaves, we have a $C_0 > 0$ and $\alpha \in (0, 1)$ such that

$$d^u(h(q_n), h(r_n)) \le C_0(d^u(q_n, r_n))^\alpha$$

But, if the claim was false, there would be a $\delta > 0$ such that $d^u(h(q_n), h(r_n)) \ge \delta$. However, the function $t \in \mathbb{R}_{>\delta} \mapsto t^\alpha \in \mathbb{R}_{>0}$ is C^1 bounded. Thus $\exists C_1 > 0$ such that

$$t^\alpha < C_1 t$$

Hence,

$$d^u(h(q_n), h(r_n)) \leq C_0 C_1 d^u(q_n, r_n)$$

This contradicts the definition of these sequences and cannot happen. The claim is proved. Now, let $J_n = [q_n, r_n]$. We have $|J_n| \to 0$ but

$$|h(J_n)| = |h(q_n), h(r_n)| = d(h(q_n), h(r_n)) \geq n d^u(q_n, r_n) = n|J_n|$$

Which is a contradiction to the hypothesis and the lemma is proven. □

To obtain the hypothesis of the lemma above, recall that the topological entropy $h_{top}(f_A)$ of the linear model f_A is simply given by the logarithm $\log(\lambda^u(A))$ of A greatest eigen value $\lambda^u(A)$. Since topological entropy is a conjugacy invariant, we have $h_{top}(f_A) = h_{top}(f)$ and in particular $h_{top}(f) = \log(\lambda^u(A))$. Together with this, we use equation $(*)$ of Lemma 6.1.1 to prove the following

Lemma 6.1.5. *There exists constants $\delta_0 > 0$ and a $C > 0$ such that for μ-a.e. $p \in \mathbb{T}^2$ for all $J \subseteq W^u(p)$ with $|J| < \delta_0$ we have*

$$|h(J)| \leq C|J|$$

Proof. Consider the intervals $J_n(p) \stackrel{\text{def.}}{=} f^{-n}(J_0(p))$ where $J_0(p) = H_p^{-1}([-1, 1])$. Since all J_0's are the image of the same set $[-1, 1]$ by H_p^{-1}, their length is bounded by an uniform constant. Furthermore, since the dynamics contracts unstable intervals in the past we have that $|J_n(p)| \to 0$ uniformly on p.

In particular, since they are all mutually bounded, $|J_n(p)|$ is proportional to $|J_n(p)|_x$. Thus it suffices to show the lemma for $|\cdot|_p$ instead of $|\cdot|$. In fact, all we need to show is that $|h(J_n(p))|_p / |J_n(p)|_p$ is bounded. For convenience, will drop the base point of $J_n(p)$ for the rest of the proof.

Using how the dynamics acts linearly under normal forms we have

$$|J_n|_p = |f^{-n}(J_0)|_p = \lambda^u_{f^n p}(-n)|J_0|_p = 2\lambda^u_{f^n p}(-n)$$

and by the commutativity of h with the dynamics,

$$|h(J_n)| = |h \circ f^{-n}(J_n)| = |f_A^{-n} \circ h(J_0)| = (\lambda^u(A))^{-n}|h(J_0)|$$

Since h is Hölder and the J_0's are bounded, we have $h(J_0) \leq 2C_0$ for some constant. Also, using the remark above the lemma, $(\lambda^u(A))^{-n} = e^{-nh_{top}(f)}$. Thus, the above equation becomes

$$|h(J_n)| \leq 2C_0 e^{-nh_{top}(f)}$$

Dividing both terms we get

$$\frac{|h(J_n)|}{|J_n|} \leq C_0 \frac{e^{-nh_{top}(f)}}{\lambda^u_{f^n p}(-n)}$$

By the equation $(*)$ of Lemma 6.1.1, we have

$$\frac{|h(J_n)|}{|J_n|} \leq C_0 \frac{\nu_p^u(\xi(p))}{\mu_p(\xi(x))} (\nu^u_{f^n(p)}(A_n))^{-1}$$

where $A_n = J_0(f^n p)$. Similarly to what was done there, we may suppose p is recurrent. So, A_n lies in a compact product neighborhood of p, and since the A_n's are all uniformly bounded bellow and ν_p^u varies continuously, $\nu^u_{f^n(p)}(A_n)$ is also bounded below. So the above expression is bounded and the lemma is proven. □

An immediate consequence of the two lemmas above is

Corollary 6.1.6. *For every $p \in \mathbb{T}^2$, $h|_{W^u(p)}$ is Lipschitz.*

Now that h is Lipschitz, we can finally show that it is C^1 in the unstable direction and with that we finish the proof of the Theorem:

Lemma 6.1.7. *For every $p \in \mathbb{T}^2$, $h|_{W^u(p)}$ is C^1.*

Proof. Since $h|_{W^u(p)}$ is Lipschitz, it is absolutely continuous. Thus, the pushback $f^* \operatorname{Leb}$, is also absolute continuous on unstable leaves. In particular, since h is a conjugation and Leb is f_A invariant, $h^* \operatorname{Leb}$ is f invariant. Thus, $h^* \operatorname{Leb}$ is the SRB measure for f. Hence, we have $h^* \operatorname{Leb} = \mu$. Consider a subordinated partition $\xi < W^u$. For $p \in \mathbb{T}^2$ and $q \in W^u(p)$ we have

$$\operatorname{Leb}^u_{h(p)}([h(p), h(q)]) = h^* \operatorname{Leb}^u_p([p, q]) = \mu^\xi_p([p, q]) = \int_{[p,q]} \rho \mathrm{d} \operatorname{Leb}^u$$

If we take smooth arc-length parameterizations of $W^u(p)$ and $W^u(h(p))$ identifying them with $\mathbb{R}$ we have

$$h(q) - h(p) = \int_p^q \rho \mathrm{d} \operatorname{Leb}_{\mathbb{R}}$$

Thus, $h|_{W^u(p)}$ is the integral of a continuous function, hence is C^1. □

6.2 The Conservative Case

For the conservative case, the result above implies that the conjugation is in fact C^1. To see it, let $f \in \operatorname{Diff}_{\operatorname{Leb}_{\mathbb{T}^2}}(\mathbb{T}^2)$ and suppose that $\mathrm{h}_{\mathrm{top}}(f) = \mathrm{h}_{\operatorname{Leb}_{\mathbb{T}^2}}(f)$.

The Lebesgue measure $\operatorname{Leb}_{\mathbb{T}^2}$ is, of course, the SRB measure for f. Thus, by the main theorem 1.0.1, the conjugacy $h : M \to M$ between f and its linear counterpart f_A is C^1 along unstable leaves. However, $\operatorname{Leb}_{\mathbb{T}^2}$ is also invariant for f^{-1}, so it is the SRB measure of f^{-1} as well. Since f is invertible, we have $\mathrm{h}_{\mathrm{top}}(f) = \mathrm{h}_{\mathrm{top}}(f^{-1})$ and $\mathrm{h}_{\operatorname{Leb}_{\mathbb{T}^2}}(f) = \mathrm{h}_{\operatorname{Leb}_{\mathbb{T}^2}}(f^{-1})$, hence $\mathrm{h}_{\mathrm{top}}(f^{-1}) = \mathrm{h}_{\operatorname{Leb}_{\mathbb{T}^2}}(f^{-1})$ and it also satisfies the hypothesis of the main theorem. Thus, the conjugacy h between f^{-1} and its linear counterpart f_A^{-1} is C^1 along unstable leaves of f^{-1}.

The unstable leaves of f^{-1} are the stable leaves of f. Hence we have obtained that h is C^1 when restricted to either the unstable or stable manifold. Since these manifolds form two continuous transverse foliations, any real function that is C^1 along them will be C^1 in the entire manifold:

Lemma 6.2.1. *If an observable $\varphi \in C^0(M)$ is C^1 when restricted to the leaves of two continuously transverse foliations W^u and W^s, then φ is C^1*

Proof. Since differentiability is a local matter, we may take coordinates around a point and treat φ as a function in $C^0(\mathbb{R}^n)$. Similarly, we may suppose that W^s and W^u have a global product structure. Since φ is C^1 along W^u, there is map $L^u : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R})$ that for every $x \in \mathbb{R}^n$ associates the unique linear transformation $L^u_x : T_x W^u \to \mathbb{R}$ that satisfies $\varphi(y) - \varphi(x) = L^u_x(y - x) + R^u(y - x)$, where $R^u(y - x)/|y - x|$ goes to 0 as $y \to x$ as long as $y - x \in T_x W^u$. The same for L^s. These maps are continuous along their respective foliations.

Now, I will define $L:\mathbb{R}^n \to L(\mathbb{R}^n,\mathbb{R})$ as the map that associates for ever $x\in\mathbb{R}^n$ the linear map $L_x = L_x^u \oplus L_x^s : T_xW^u \oplus T_xW^s$. I claim that this map L is continuous and is the derivative of φ at each point. To prove it, take $x, y \in \mathbb{R}^n$ and let z be the unique element in $W^u(x)\cap W^s(y)$. We have

$$\begin{aligned}
\varphi(y)-\varphi(x) &= \varphi(y)-\varphi(z)+\varphi(z)-\varphi(x)\\
&= L_z^s(y-z)+R^s(y-z)+L_x^u(z-x)+R^u(z-x)
\end{aligned}$$

As $y\to x$, we have that $z\to x$ along a leaf of W^s. Hence $L_z^s \to L_x^s$, i.e. $L_z^s = L_x^s + \varepsilon_z$ where the operator norm of ε_z goes to 0 as $z\to x$. Thus

$$\begin{aligned}
\varphi(y)-\varphi(x) &= L_x^s(y-z)+L_x^u(z-x)+R^s(y-z)+R^u(z-x)+\varepsilon_z(y-z)\\
&= L_x(y-x)+R^s(y-z)+R^u(z-x)+\varepsilon_z(y-z)
\end{aligned}$$

All terms other than $L_x(y-x)$ above are of a higher order than $O(|y-x|)$, thus L_x is in fact the derivative of φ at x. In particular, for $\tilde{z}\in W^u(y)\cap W^s(x)$ we also have

$$\varphi(y)-\varphi(x) = L_y(y-x)+R^s(x-\tilde{z})+R^u(\tilde{z}-y)+\varepsilon_{\tilde{z}}(x-\tilde{z})$$

so that $L_y - L_x = R(y-x)$ where $R(y-x)\to 0$ as $y\to 0$, that is, L is continuous. □

To use this lemma for h, we remember that a function $h: M\to M$ is C^1 if and only if $\varphi\circ h : M\to\mathbb{R}$ is C^1 for every $\varphi\in C^1(M)$. Hence, since h being C^1 restricted to stable and unstable leaves implies that $\varphi\circ h$ also is, the lemma above gives that

Corollary 6.2.2. *The conjugacy h is C^1.*

Which proves corollary 1.0.2 and concludes the text.

APPENDIX **A**

Periodic Data Rigidity

Two systems $f, g \in \mathrm{Diff}(M)$ conjugated by a homeomorphism $h : M \to M$ are said to have the same periodic data if for every point $p \in \mathrm{Per}_n(f)$ you have

$$spec(d_p f^n) = spec(d_{h(p)} g^n)$$

In this appendix we prove the following

Theorem A.0.1. *Let $f, g : \mathbb{T}^2 \to \mathbb{T}^2$ be two Anosov diffeomorphisms of class C^2 with the same periodic data and let $h : \mathbb{T}^2 \to \mathbb{T}^2$ be a conjugation between f and g. Then h is $C^{1+\alpha}$.*

This theorem can be generalized for dimension 3. However its proof requires way more effort than for the two dimensional case and it also requires more hyphothesis: you must either require f and g to be C^1 close to a linear automorphism or that one of them admits a partially hyperbolic invariant splitting $T\mathbb{T}^3 = E^s \oplus E^{wu} \oplus E^u$ and the conjugacy h is homotopic to a linear automorphism (see [GG08]). In fact, just like our main theorem, this is a low dimensional phenomena: there exists counterexamples for dimension $d \geq 4$.

To prove this theorem we will use the affine parameters 4.2.1, some properties of holonomies (as lemma 4.2.6) and the density of the foliations (see theorem 2.3.5).

Proof of Theorem A.0.1

Lets first prove that $h|_{W^u}$ is Lipschitz.

Claim A.0.1.1. For all $p \in \mathbb{T}^2$, $h|_{W^u(p)}$ is Lipschitz.

Proof. Let $p \in M$ and take an interval $I \subseteq W^u_f(p)$ in the unstable leaf of p. Write is image by h as $\hat{I} \overset{\text{def.}}{=} h(I)$ and $\hat{p} = f(p)$. Since h is a conjugacy between f and g we have $\hat{I} \subseteq W^u_g(\hat{p})$.

In dimension 2, we know that the unstable leaves are one-dimensional. Thus, the length of an interval is nothing more than the distance between its end points. With this observation, the task to show that $h_{W^u(p)}$ is Lipschitz becomes the task to bound the ratio $|\hat{I}|/|I|$ from above. Also, since the normal forms are an uniform family of diffeomorphisms, we can measure lengths under their linearized coordinates, i.e. we only need to bound $|\hat{I}|_p/|I|_p$.

Also, since the conjugacy is Hölder, we may assume that $|I|_p < 1$ (see lemma 6.1.4).

Let $q \in \mathrm{Per}_k(f)$ be a periodic point for f. Since the stable leaves of f are dense (see Theorem 2.3.5), we can find a stable holonomy map $\mathcal{H}_{I\to J} : I \to J \subseteq W^u_f(q)$ where $J \overset{\text{def.}}{=} \mathcal{H}_{I\to J}(I)$. In particular, since holonomies commutes with the dynamics, it induces an Holonomy $\hat{\mathcal{H}}_{\hat{I}\to\hat{J}} : \hat{I} \to \hat{J} \subseteq W^u_g(\hat{q})$ where $\hat{J} \stackrel{\text{def.}}{=} \hat{\mathcal{H}}_{\hat{I}\to\hat{J}}(\hat{I})$ and $\hat{q} \stackrel{\text{def.}}{=} h(q)$.

Since the holonomies are uniformly C^1, we can find a $C > 1$ such that

$$C|I|_p^{-1} \leq |J|_q \leq C|I|_p \quad \text{and} \quad C|\hat{I}|_{\hat{p}}^{-1} \leq |\hat{J}|_{\hat{q}} \leq C|\hat{I}_{\hat{p}}|$$

Thus

$$\frac{|\hat{I}|_{\hat{p}}}{|I|_p} \leq C^2 \frac{|\hat{J}|_{\hat{q}}}{|J|_q}$$

Thus it suffices to bound $|\hat{J}|_{\hat{q}}/|J|_q$. For it, define $J_{nk} \subseteq W^u_f(q)$ by $J_{nk} \stackrel{\text{def.}}{=} f^{nk}(J)$ and let $\hat{J}_{nk} \stackrel{\text{def.}}{=} h(J_{nk})$. Take $n_0 \in \mathbb{N}$ big enough so that $|J_{n_0 k}|_q > 1$. Since h is Hölder, this implies that $|\hat{J}_{n_0 k}|_{\hat{q}}/|J_{n_0 k}|_q < C_0$ where C_0 is the Hölder constant of h. However, by item (iv) of prop. 4.2.1, we have

$$|J_{n_0 k}|_q = |d_q f^k|^{n_0} |J|$$

Also, since h commutes with the dynamics, we have $\hat{J}_{n_0 k} = g^{n_0 k}(\hat{J})$. Thus, using the same item

$$|\hat{J}_{n_0 k}|_q = |d_{\hat{q}} g^k|^{n_0} |\hat{J}|$$

But, since they have the same periodic data, we have $|d_q f^k|^{n_0} = |d_{\hat{q}} g^k|^{n_0}$, thus

$$\frac{|\hat{J}|_q}{|J|_q} = \frac{|\hat{J}_{n_0 k}|_{\hat{q}}}{|J_{n_0 k}|_q} < C_0$$

which concludes the claim. □

To pass from $h|_{W^u}$ Lipschitz to it being $C^{1+\alpha}$ apply lemma 6.1.7. Also, since f and g have the same periodic data, it follows that f^{-1} and g^{-1} also have the same periodic data. Thus, since h is also a conjugacy between f^{-1} and g^{-1}, all arguments above are symmetrical and can be used to prove that $h|_{W^s}$ is also C^1. By lemma 6.2.1 (Journé's lemma), it follows that h is $C^{1+\alpha}$ and the theorem is proven. □

Bibliography

[Alv+24] S. Alvarez et al. “Rigidity of *U*-Gibbs measures near conservative Anosov diffeomorphisms on $\mathbb{T}^3$”. In: *Journal of the European Mathematical Society* (2024). DOI: https://doi.org/10.4171/jems/1517. URL: https://ems.press/journals/jems/articles/14298144.

[Alv13] S. Alvarez. “Mesures de Gibbs et mesures harmoniques pour les feuilletages aux feuilles courbées négativement”. Theses. Université de Bourgogne, Dec. 2013. URL: https://theses.hal.science/tel-00958080.

[BDV06] C. Bonatti, L. Díaz, and M. Viana. *Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective*. Encyclopaedia of Mathematical Sciences. Springer Berlin Heidelberg, 2006. ISBN: 9783540268444. URL: https://books.google.com.na/books?id=FS57Tv3YHwYC.

[BK83] M. Brin and A. Katok. “On local entropy”. In: *Geometric Dynamics*. Ed. by J. Palis. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983, pp. 30–38. ISBN: 978-3-540-40969-4.

[Bro+19] A. W. Brown et al. *Entropy, Lyapunov exponents, and rigidity of group actions*. 2019. arXiv: 1809.09192 [math.DS]. URL: https://arxiv.org/abs/1809.09192.

[CP24] J. Correa and H. de Paula. *Flexibility of generalized entropy for wandering dynamics*. 2024. arXiv: 2303.14780 [math.DS]. URL: https://arxiv.org/abs/2303.14780.

[Fol99] G. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999. ISBN: 9780471317166. URL: https://books.google.com.br/books?id=N8jVDwAAQBAJ.

[GG08] A. Gogolev and M. Guysinsky. *C^1-differentiable conjugacy of Anosov diffeomorphisms on three dimensional torus*. 2008. DOI: 10.3934/dcds.2008.22.183. URL: https://www.aimsciences.org/article/id/9ddf1e22-e38f-4b67-9315-74903e250b3d.

[Gu23] R. Gu. “Smooth Stable Foliations of Anosov Diffeomorphisms”. In: *arXiv preprint arXiv:2310.19088* (2023).

[KH95] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1995.

[Khi57] A. I. Khinchin. *Mathematical Foundations of Information Theory*. New York, Dover Publications, 1957.

[KK06] B. Kalinin and A. Katok. "Measure rigidity beyond uniform hyperbolicity: invariant measures for Cartan actions on tori". In: *arXiv preprint math/0602176* (2006).

[New70] S. E. Newhouse. "On Codimension One Anosov Diffeomorphisms". In: *American Journal of Mathematics* 92.3 (1970), pp. 761–770. ISSN: 00029327, 10806377. URL: http://www.jstor.org/stable/2373372 (visited on 07/23/2025).

[Pot16] R. Potrie. *Introduction to Non-Uniform and Partial Hyperbolicity*. Publicaciones matemáticas del Uruguay v. 16. Centro de Matemática de la Universidad de la República, 2016, pp. 127–167. URL: https://pmu.uy/pmu16/pmu16-0127.pdf.

[PSW00] C. Pugh, M. Shub, and A. Wilkinson. "Holder foliations (vol 86, pg 517, 1997)". In: *Duke Mathematical Journal* 105 (Oct. 2000), pp. 105–106.

[SY19] R. Saghin and J. Yang. "Lyapunov exponents and rigidity of Anosov automorphisms and skew products". In: *Advances in Mathematics* 355 (2019), p. 106764. ISSN: 0001-8708. DOI: https://doi.org/10.1016/j.aim.2019.106764. URL: https://www.sciencedirect.com/science/article/pii/S0001870819303822.

[Wal00] P. Walters. *An introduction to ergodic theory*. Vol. 79. Springer Science & Business Media, 2000.

[Wen16] L. Wen. *Differentiable Dynamical Systems*. Graduate Studies in Mathematics. American Mathematical Society, 2016. ISBN: 9781470427993. URL: https://books.google.com.br/books?id=swG1DAAAQBAJ.