Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 37 (2020) 727–755

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

Symbolic dynamics for one dimensional maps with nonuniform expansion

Yuri Lima [1]

Departamento de Matemática, Universidade Federal do Ceará (UFC), Campus do Pici, Bloco 914, CEP 60440-900, Fortaleza – CE, Brazil

Received 26 May 2018; received in revised form 4 October 2019; accepted 31 October 2019
Available online 6 November 2019

Abstract

Given a piecewise $C^{1+\beta}$ map of the interval, possibly with critical points and discontinuities, we construct a symbolic model for invariant probability measures with nonuniform expansion that do not approach the critical points and discontinuities exponentially fast almost surely. More specifically, for each $\chi > 0$ we construct a finite-to-one Hölder continuous map from a countable topological Markov shift to the natural extension of the interval map, that codes the lifts of all invariant probability measures as above with Lyapunov exponent greater than χ almost everywhere.

MSC: 37B10; 37D25; 37E05

Keywords: Interval map; Markov partition; Pesin theory; Symbolic dynamics

1. Introduction

The quadratic family $\{f_a\}_{0\leq a\leq 4}$ is the family of one dimensional interval maps $f_a : [0, 1] \to [0, 1]$, $f_a(x) = ax(1 - x)$. Although simple to describe, it exhibits complicated dynamical behaviour: for a set of parameters of positive Lebesgue measure, f_a has an absolutely continuous invariant measure with positive Lyapunov exponents [33,4,5], see also [58]. The idea to prove this is to construct a partition of the interval with good symbolic properties that allows to understand the orbit of the critical point $x = 0.5$, so that for many parameters the critical value $f_a(0.5)$ has positive Lyapunov exponent (this latter property is known as the *Collet-Eckmann condition*). This idea has far reaching applications, see e.g. [42,25,41].

The present works goes in the reverse direction of the above idea: it considers piecewise $C^{1+\beta}$ maps $f : [0, 1] \to [0, 1]$ of the interval with positive Lyapunov exponent and constructs *finite-to-one Hölder continuous* symbolic extensions of the maps. We require f to satisfy the regularity conditions (A1)–(A3), that will be shortly described. These conditions allow f to have both critical points (where the first derivate vanishes) and discontinuities (where the first

E-mail address: yurilima@gmail.com.

[1] The author is supported by Instituto Serrapilheira, grant Serra-1709-20498.

https://doi.org/10.1016/j.anihpc.2019.10.001
0294-1449/

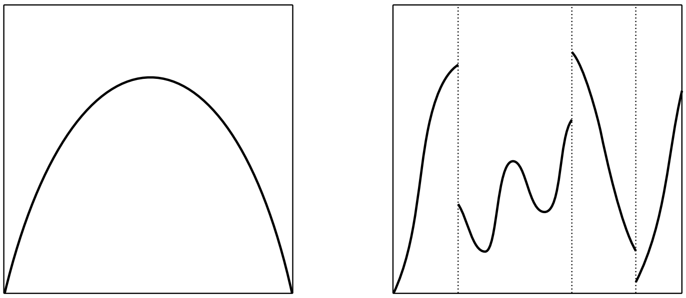

Fig. 1. Examples of maps covered by our results.

derivative can explode), and include the quadratic family, multimodal maps with non-flat critical points, piecewise continuous maps with discontinuities of polynomial type, and combinations of these classes, see Fig. 1 for examples.

Our main result is the construction of a symbolic model for the natural extension $\widehat{f} : \widehat{[0,1]} \to \widehat{[0,1]}$ of f. See Section 1.5 for the definition of the natural extension and its main properties. Let us describe which maps we consider and which measures we are able to code. For the sake of simplicity, we consider maps defined on the interval $M = [0, 0.5]$, since this interval has diameter less than one (and so we do not need to introduce multiplicative constants in the assumptions (A1)–(A3) below). Let $f : M \to M$ be a map with *discontinuity set* $\mathscr{D}$. We assume that f is $C^{1+\beta}$ in the set $M \backslash \mathscr{D}$, for some $\beta \in (0,1)$. Let $\mathscr{C} := \{x \in M \backslash \mathscr{D} : df_x = 0\}$ denote the *critical set* of f.

SINGULAR SET: The *singular set* of f is $\mathscr{S} := \mathscr{C} \cup \mathscr{D}$.

We allow $\mathscr{S}$ to be infinite, e.g. when f is the Gauss map. Let $B(x,r) \subset M$ denote the ball with centre x and radius r. We assume that f satisfies the following properties.

REGULARITY OF f: There exist constants $a, \mathfrak{K} > 1$ s.t. for all $x \in M$ with $x, f(x) \notin \mathscr{S}$ there is $\min\{d(x,\mathscr{S})^a, d(f(x),\mathscr{S})^a\} < \mathfrak{r}(x) < 1$ s.t. for $D_x := B(x, 2\mathfrak{r}(x))$ and $E_x := B(f(x), 2\mathfrak{r}(x))$ the following holds:

(A1) The restriction of f to D_x is a diffeomorphism onto its image; the inverse branch of f taking $f(x)$ to x is a well-defined diffeomorphism from E_x onto its image.
(A2) For all $y \in D_x$ it holds $d(x,\mathscr{S})^a \le |df_y| \le d(x,\mathscr{S})^{-a}$; for all $z \in E_x$ it holds $d(x,\mathscr{S})^a \le |dg_z| \le d(x,\mathscr{S})^{-a}$, where g is the inverse branch of f taking $f(x)$ to x.
(A3) For all $y, z \in D_x$ it holds $|df_y - df_z| \le \mathfrak{K}|y-z|^\beta$; for all $y, z \in E_x$ it holds $|dg_y - dg_z| \le \mathfrak{K}|y-z|^\beta$.

Now we describe the measures that we code. We borrow the notation from [38]. Let μ be an f–invariant probability measure.

f–ADAPTED MEASURE: The measure μ is called *f–adapted* if $\log d(x,\mathscr{S}) \in L^1(\mu)$. A fortiori, $\mu(\mathscr{S}) = 0$.

χ–EXPANDING MEASURE: Given $\chi > 0$, the measure μ is called *χ–expanding* if $\lim_{n\to\infty} \frac{1}{n} \log |df_x^n| > \chi$ for μ–a.e. $x \in M$.

The next theorem is the main result of this paper. Below, $\widehat{f} : \widehat{M} \to \widehat{M}$ is the natural extension of f, and $\widehat{\mu}$ is the lift of μ, see Subsection 1.5 for the definition.

Theorem 1.1. *Assume that f satisfies* (A1)–(A3). *For all $\chi > 0$, there exists a countable topological Markov shift (Σ, σ) and $\pi : \Sigma \to \widehat{M}$ Hölder continuous s.t.:*

(1) $\pi \circ \sigma = \widehat{f} \circ \pi$.
(2) $\pi[\Sigma^\#]$ *has full $\widehat{\mu}$–measure for every f–adapted χ–expanding measure μ.*
(3) *For all $\widehat{x} \in \pi[\Sigma^\#]$, the set $\{\underline{v} \in \Sigma^\# : \pi(\underline{v}) = \widehat{x}\}$ is finite.*

Above, $\Sigma^\#$ is the *recurrent set* of Σ; it carries all σ–invariant probability measures, see Section 1.4. Therefore we are able to code simultaneously all the measures with nonuniform expansion greater than χ almost everywhere that do not approach the singular set exponentially fast.

It is important to make some comments on the assumption of f–adaptedness. By the Birkhoff ergodic theorem, if μ is f–adapted then $\lim_{n\to\infty}\frac{1}{n}\log d(f^n(x),\mathscr{S})=0$ μ–a.e. Ledrappier considered this latter condition for interval maps with critical points [37], where he used the terminology *non-degenerate measure*. Katok and Strelcyn implicitly used that the Lebesgue measure is adapted to billiard maps and then used that $\lim_{n\to\infty}\frac{1}{n}\log d(f^n(x),\mathscr{S})=0$ a.e. [35]. For an invariant measure of a three dimensional flow with positive speed, Lima and Sarig constructed a Poincaré section for which the respective Poincaré return map f satisfies $\lim_{n\to\infty}\frac{1}{n}\log d(f^n(x),\mathscr{S})=0$ almost surely with respect to the induced measure [39], and Lima and Matheus used the assumption of adaptability in their coding of billiard maps [38]. For one dimensional maps satisfying (A1)–(A3), if $\mathscr{D}=\emptyset$ and $\mathscr{C}$ is finite then every ergodic invariant probability measure that is not supported in an attracting periodic orbit satisfies $\lim_{n\to\infty}\frac{1}{n}\log d(f^n(x),\mathscr{S})=0$ a.e. [44], see also [49, Appendix A] for a proof that works under weaker assumptions. It would be interesting to obtain the same conclusion when $\mathscr{S}$ is finite.

Now we make a comparison of our method with the one developed by Hofbauer [30,31], known as *Hofbauer towers*. These towers were first constructed to analyze measures of maximal entropy, and they provide a precise combinatorial description of one dimensional maps. In comparison to Hofbauer's method, our method explores the nonuniform expansion of χ–expanding measures. It constitutes the first implementation, for non-invertible systems, of the recent constructions of Markovian symbolic dynamics for nonuniformly hyperbolic systems [53,39,38,7]. The novelty of the present paper is that, contrary to Hofbauer's method, our construction has the following advantages:

- The extension map π constructed in Theorem 1.1 is *Hölder continuous*.
- While Hofbauer's method only works in very specific higher dimensional cases (see Section 1.2 below), our method is more robust and will be extended, in a forthcoming work, to higher dimensional maps such as complex-valued functions, Viana maps, and general nonuniformly hyperbolic maps.

In addition to these, we emphasize that $\mathscr{S}$ can be infinite.

1.1. Applications

Assume that f satisfies (A1)–(A3) and has finite and positive topological entropy $h\in(0,\infty)$. In this section, we discuss two applications of Theorem 1.1:

- Estimates on the number of periodic points.
- Understanding of equilibrium measures.

The first can be obtained in great generality, but the second is subtler due to the presence of discontinuities and eventual non-finiteness of $\mathscr{S}$. As a matter of fact, the measure of maximal entropy may not exist, even for C^r maps [19,51]. When $\mathscr{D}=\emptyset$ and $\mathscr{C}$ is finite, f is called a multimodal map. There are many results on ergodic and statistical properties of equilibrium measures for one dimensional maps in various contexts, such as uniformly expanding maps [40,56,57,11], piecewise continuous maps [30,31], piecewise monotone maps [28], and multimodal maps [37,22,6, 45,15,16,32].

As mentioned above, Theorem 1.1 allows $\mathscr{S}$ to be infinite. When there is an f–adapted measure of maximal entropy, we obtain exponential estimates on the number of periodic points.

Theorem 1.2. *Assume f satisfies* (A1)–(A3) *with topological entropy $h\in(0,\infty)$. If there is an f–adapted measure of maximal entropy, then $\exists C>0$, $p\geq 1$ s.t. $\#\{x\in[0,1]:f^{pn}(x)=x\}\geq Ce^{hpn}$ for all $n\geq 1$.*

Proof. The set of periodic points of f is in bijection with the set of periodic points of its natural extension $\widehat{f}$. Since (Σ,σ) is a finite extension of $\widehat{f}$, the growth rate of periodic points of $\widehat{f}$ is at least that of σ. Let μ be an ergodic f–adapted measure of maximal entropy for f. Its lift $\widehat{\mu}$ is an ergodic measure of maximal entropy for $\widehat{f}$ satisfying the assumptions of Theorem 1.1. Proceeding as in [53, §13], $\widehat{\mu}$ lifts to an ergodic measure of maximal entropy ν for (Σ,σ). By [26,27], ν is supported on a topologically transitive countable topological Markov shift, and there is $p \geq 1$ s.t. for every vertex v it holds $\#\{\underline{v} \in \Sigma : \sigma^{pn}(\underline{v}) = \underline{v}, v_0 = v\} \asymp e^{hpn}$. Hence, there exists $C > 0$ s.t. $\#\{x \in [0,1] : f^{pn}(x) = x\} \geq Ce^{hpn}$. □

The number p equals the period of the transitive component supporting the measure of maximal entropy. For surface maps, Buzzi gave conditions for which $p = 1$ [17], and we expect his method also applies here.

Now we turn attention to equilibrium measures. Let $(X, \mathscr{A}, T)$ be a probability measure-preserving system, and $\varphi : X \to \mathbb{R}$ a continuous potential. The following definitions are standard.

TOPOLOGICAL PRESSURE: The *topological pressure* of φ is $P_{\rm top}(\varphi) := \sup\{h_\mu(T) + \int \varphi d\mu\}$, where the supremum ranges over all T–invariant probability measures.

EQUILIBRIUM MEASURE: An *equilibrium measure* for φ is a T–invariant probability measure μ s.t. $P_{\rm top}(\varphi) = h_\mu(T) + \int \varphi d\mu$.

Due to the presence of discontinuities, the topological pressure might be infinite (e.g. the Gauss map has infinite topological entropy), and equilibrium measures may not exist. Theorem 1.1 provides countability results for a class of potentials and equilibrium measures: we require φ to be a bounded Hölder continuous potential with finite topological pressure, and consider equilibrium measures that are f–adapted and χ–expanding for some $\chi > 0$. Call a measure *expanding* if it is χ–expanding for some $\chi > 0$. Observe that, when the Ruelle inequality applies, every ergodic measure with positive metric entropy is expanding.

Theorem 1.3. *Assume that f satisfies* (A1)–(A3). *Every equilibrium measure of a bounded Hölder continuous potential with finite topological pressure has at most countably many f–adapted expanding ergodic components. Furthermore, the lift to $\widehat{M}$ of each such ergodic component is Bernoulli up to a period.*

Proof. Let $\varphi : M \to \mathbb{R}$ be bounded, Hölder continuous, with $P_{\rm top}(\varphi) < \infty$. We prove that, for each $\chi > 0$, φ possesses at most countably many f–adapted χ–expanding equilibrium measures. The first part of the theorem follows by taking the union of these measures for $\chi_n = \frac{1}{n}$, $n > 0$.

Fix $\chi > 0$. Let $\vartheta : \widehat{M} \to M$ be the projection into the zeroth coordinate, see Subsection 1.5, and define $\widehat{\varphi} : \widehat{M} \to \mathbb{R}$ by $\widehat{\varphi} = \varphi \circ \vartheta$. Then $\widehat{\varphi}$ is bounded, Hölder continuous and has finite topological pressure $P_{\rm top}(\widehat{\varphi}) = P_{\rm top}(\varphi)$. Furthermore, μ is an ergodic equilibrium measure for φ iff $\widehat{\mu}$ is an ergodic equilibrium measure for $\widehat{\varphi}$. When μ is additionally f–adapted and χ–expanding, we can apply the procedure in [53, §13] and Theorem 1.1 to lift $\widehat{\mu}$ to an ergodic equilibrium measure ν in (Σ, σ). By ergodicity, ν is carried by a topologically transitive countable topological Markov shift. The potential associated to ν is $\Phi = \widehat{\varphi} \circ \pi$, which is bounded, Hölder continuous and has finite topological pressure $P_{\rm top}(\Phi) = P_{\rm top}(\widehat{\varphi}) = P_{\rm top}(\varphi)$. By [13, Thm. 1.1], each topologically transitive countable topological Markov shift carries at most one equilibrium measure for Φ, hence there are at most countably many such ν. This proves the first part of the theorem. By [52], each such ν is Bernoulli up to a period. Since this latter property is preserved by finite-to-one extensions, the same occurs to $\widehat{\mu}$. This concludes the proof of the theorem. □

1.2. Related literature

Symbolic models in dynamics have a longstanding history that can be traced back to the work of Hadamard on closed geodesics of hyperbolic surfaces, see e.g. [36]. The late sixties and early seventies saw a great deal of development of symbolic dynamics for uniformly hyperbolic diffeomorphisms and flows, through the works of Adler & Weiss [2,3], Sinaĭ [54,55], Bowen [8,9], Ratner [47,48]. Below we discuss other relevant contexts.

HOFBAUER TOWERS: Takahashi developed a combinatorial method to construct an isomorphism between a large subset X of the natural extension of β–shifts and countable topological Markov shifts [57]. Hofbauer proved that X carries all measures of positive entropy and hence β–shifts have a unique measure of maximal entropy [29]. Hofbauer later extended his construction to piecewise continuous interval maps [30,31]. The symbolic models obtained by his methods are called *Hofbauer towers*, and they have been extensively used to establish ergodic properties of one dimensional maps.

HIGHER DIMENSIONAL HOFBAUER TOWERS: Buzzi constructed Hofbauer towers for piecewise expanding affine maps in any dimension [18], for perturbations of fibred products of one dimensional maps [20], and for arbitrary piecewise invertible maps whose entropy generated by the boundary of some dynamically relevant partition is *less* than the topological entropy of the map [21]. These Hofbauer towers carry all invariant measures with entropy close enough to the topological entropy of the system. We remark that, contrary to us, Buzzi's conditions make no reference to the nonuniform hyperbolicity of the system.

INDUCING SCHEMES: Many systems, although not hyperbolic, do have sets on which it is possible to define a (not necessarily first) return map on which the map becomes uniformly hyperbolic. This process is known as *inducing*. Indeed, Hofbauer towers can be seen as inducing schemes for which the map becomes uniformly expanding, see [12] for this relation. It is possible to understand ergodic theoretical properties of invariant measures that lift to an inducing scheme, as done for one-dimensional maps [45], higher dimensional ones that do not have full "boundary entropy" [46], and expanding measures [43].

YOCCOZ PUZZLES: Yoccoz constructed Markov structures for quadratic maps of the complex plane, nowadays called *Yoccoz puzzles*, and used them to establish the MLC conjecture for finitely renormalizable parameters as well as a proof of Jakobson's theorem, see [58].

NONUNIFORMLY HYPERBOLIC DIFFEOMORPHISMS: Katok constructed horseshoes of large topological entropy for $C^{1+\beta}$ diffeomorphisms [34]. These horseshoes usually have zero measure for measures of maximal entropy. Sarig constructed a "horseshoe" of full entropy for $C^{1+\beta}$ surface diffeomorphisms [53]: for each $\chi > 0$ there is a countable topological Markov shift that is an extension of the diffeomorphism and codes all χ–hyperbolic measures simultaneously. Ben Ovadia extended the work of Sarig to higher dimension [7].

NONUNIFORMLY HYPERBOLIC THREE-DIMENSIONAL FLOWS: Lima and Sarig constructed symbolic models for nonuniformly hyperbolic three dimensional flows with positive speed [39]. The idea is to build a "good" Poincaré section and construct a Markov partition for the Poincaré return map f.

BILLIARDS: Dynamical billiards are maps with discontinuities. Katok and Strelcyn constructed invariant manifolds for nonuniformly hyperbolic billiard maps [35]. Bunimovich, Chernov, and Sinaĭ constructed countable Markov partitions for two dimensional dispersing billiard maps [14]. Young constructed an inducing scheme for certain two dimensional dispersing billiard maps and used it to prove exponential decay of correlations [59]. All these results are for Liouville measures. Lima and Matheus constructed countable Markov partitions for two dimensional billiard maps and nonuniformly hyperbolic (not necessarily Liouville) measures that are adapted to the billiard map [38].

1.3. Method of proof

The proof of Theorem 1.1 is based on [53], [39] and [38], and follows the steps below:

(1) The derivative cocycle df induces an invertible cocycle $\widehat{df}$, defined on a fibre bundle over the natural extension space $\widehat{M}$, with the same spectrum as df.
(2) If μ is f–adapted and χ–expanding, then $\widehat{\mu}$–a.e. $\widehat{x} \in \widehat{M}$ has a Pesin chart $\Psi_{\widehat{x}} : [-Q_\varepsilon(\widehat{x}), Q_\varepsilon(\widehat{x})] \to M$ s.t. $\lim_{n\to\infty} \frac{1}{n} \log Q_\varepsilon(\widehat{f}^n(\widehat{x})) = 0$. In Pesin charts, the inverse branches of f are uniform contractions.
(3) Introduce the ε–chart $\Psi_{\widehat{x}}^p$ as the restriction of $\Psi_{\widehat{x}}$ to $[-p, p]$. The parameter p gives a definite size for the unstable manifold at $\widehat{x}$, hence different ε–charts $\Psi_{\widehat{x}}^p$, $\Psi_{\widehat{x}}^q$ define unstable manifolds of different sizes.
(4) Construct a countable collection of ε–charts that are dense in the space of all ε–charts, where denseness is defined in terms of finitely many parameters of $\widehat{x}$.

(5) Draw an edge $\Psi_{\widehat{x}}^{p} \leftarrow \Psi_{\widehat{y}}^{q}$ between ε–charts when an inverse branch of f can be represented in these charts by a uniform contraction and the parameter q is as large as possible. Each path of ε–charts defines an element of $\widehat{M}$, and this coding induces a countable cover on a subset of $\widehat{M}$. The requirement that q is as large as possible guarantees that this cover is locally finite.
(6) Apply a refinement procedure to this cover. The resulting partition defines a countable topological Markov shift (Σ, σ) and a coding $\pi : \Sigma \to \widehat{M}$ that satisfy Theorem 1.1.

It is important to stress the importance of having a countable locally finite cover in step (5). If not, its refinement could be an uncountable partition (imagine e.g. the cover of $\mathbb{R}$ by intervals with rational endpoints). When the countable cover is locally finite, i.e. each element of the cover intersects at most finitely many others, then its refinement is again countable. Local finiteness is crucial, and it is the reason to choose q as large as possible.

Contrary to [53,39,38], we find no difficulty on the control of the geometry of M (all exponential maps are identities) neither with the geometry of stable and unstable directions (the stable direction is trivial). Hence the methods we use in steps (2)–(5) are more clear and more easily implemented than those in [53,39,38]. For example, we do not make use of graph transforms. On the other hand, a difficulty for the implementation of steps (2)–(5) is that neither $\widehat{M}$ nor $\widehat{f}$ are smooth objects. This is not a big issue, since what we want is to control the action of f and its inverse branches, and this can be made by controlling the action of $\widehat{f}^{\pm 1}$ in the zeroth coordinate. Working with natural extensions makes step (1) heavier in notation, and step (6) more complicated to implement.

Natural extensions have been previously used to investigate nonuniformly expanding systems. Up to the author's knowledge, the first one to use this approach was Ledrappier, in the context of absolutely continuous invariant measures of interval maps [37]. Other employments of this approach are [24,23].

The methods employed in this article require some familiarity with the articles [53,39,38], and a first reading might be difficulty for those not familiar with the referred literature. Unfortunately, a self-contained exposition would lead to a lengthy manuscript, thus preventing to focus on the novelty of the work.

1.4. Preliminaries

Let $\mathscr{G} = (V, E)$ be an oriented graph, where $V =$ vertex set and $E =$ edge set. We denote edges by $v \to w$, and we assume that V is countable.

Topological Markov shift (TMS): A *topological Markov shift* (TMS) is a pair (Σ, σ) where

$$\Sigma := \{\mathbb{Z}\text{–indexed paths on } \mathscr{G}\} = \left\{\underline{v} = \{v_n\}_{n\in\mathbb{Z}} \in V^{\mathbb{Z}} : v_n \to v_{n+1}, \forall n \in \mathbb{Z}\right\}$$

and $\sigma : \Sigma \to \Sigma$ is the *left shift*, $[\sigma(\underline{v})]_n = v_{n+1}$. The *recurrent set* of Σ is

$$\Sigma^\# := \left\{\underline{v} \in \Sigma : \exists v, w \in V \text{ s.t. } \begin{array}{l} v_n = v \text{ for infinitely many } n > 0 \\ v_n = w \text{ for infinitely many } n < 0 \end{array}\right\}.$$

We endow Σ with the distance $d(\underline{v}, \underline{w}) := \exp[-\min\{|n| : n \in \mathbb{Z} \text{ s.t. } v_n \neq w_n\}]$. This choice is not canonical, and affects the Hölder regularity of π in Theorem 1.1.

Write $a = e^{\pm\varepsilon} b$ when $e^{-\varepsilon} \le \frac{a}{b} \le e^{\varepsilon}$, and $a = \pm b$ when $-|b| \le a \le |b|$. Given an open set $U \subset \mathbb{R}$ and $h : U \to \mathbb{R}$, let $\|h\|_0 := \sup_{x\in U} \|h(x)\|$ denote the C^0 norm of h. For $0 < \beta < 1$, let $\mathrm{Hol}_\beta(h) := \sup \frac{\|h(x)-h(y)\|}{\|x-y\|^\beta}$ where the supremum ranges over distinct elements $x, y \in U$. If h is differentiable, let $\|h\|_1 := \|h\|_0 + \|dh\|_0$ denote its C^1 norm, and $\|h\|_{1+\beta} := \|h\|_1 + \mathrm{Hol}_\beta(dh)$ its $C^{1+\beta}$ norm. For $r > 0$, define $R[r] := [-r, r] \subset \mathbb{R}$, where $\mathbb{R}$ is endowed with the usual euclidean distance.

1.5. Natural extensions

Most of the discussion below is classical, see e.g. [50] or [1, §3.1]. Given a (possibly non-invertible) map $f : M \to M$, let

$$\widehat{M} := \{\widehat{x} = (x_n)_{n\in\mathbb{Z}} : f(x_{n-1}) = x_n, \forall n \in \mathbb{Z}\}.$$

Although $\widehat{M}$ does depend on f, we do not write this dependence. Endow $\widehat{M}$ with the distance $\widehat{d}(\widehat{x},\widehat{y}) := \sup\{2^n d(x_n, y_n) : n \leq 0\}$; then $\widehat{M}$ is a compact metric space. As for TMS, the definition of $\widehat{d}$ is not canonical and reflects the Hölder regularity of π in Theorem 1.1. For each $n \in \mathbb{Z}$, let $\vartheta_n : \widehat{M} \to M$ be the projection into the n–th coordinate, $\vartheta_n[\widehat{x}] = x_n$. Let $\widehat{\mathscr{B}}$ be the sigma-algebra in $\widehat{M}$ generated by $\{\vartheta_n : n \leq 0\}$, i.e. $\widehat{\mathscr{B}}$ is the smallest sigma-algebra that makes all ϑ_n, $n \leq 0$, measurable.

NATURAL EXTENSION OF f: The *natural extension* of f is the map $\widehat{f} : \widehat{M} \to \widehat{M}$ defined by $\widehat{f}(\dots, x_{-1}; x_0, \dots) = (\dots, x_0; f(x_0), \dots)$. It is an invertible map, with inverse $\widehat{f}^{-1}(\dots, x_{-1}; x_0, \dots) = (\dots, x_{-2}; x_{-1}, \dots)$.

Note that $\widehat{f}$ is indeed an extension of f, since $\vartheta_0 \circ \widehat{f} = f \circ \vartheta_0$. It is the smallest invertible extension of f: any other invertible extension of f is an extension of $\widehat{f}$. The benefit of considering the natural extension is that, in addition to having an invertible map explicitly defined, its complexity is the same as that of f: there is a bijection between f–invariant and $\widehat{f}$–invariant probability measures, as follows.

PROJECTION OF A MEASURE: If $\widehat{\mu}$ is an $\widehat{f}$–invariant probability measure, then $\mu = \widehat{\mu} \circ \vartheta_0^{-1}$ is an f–invariant probability measure.

LIFT OF A MEASURE: If μ is an f–invariant probability measure, let $\widehat{\mu}$ be the unique probability measure on $\widehat{M}$ s.t. $\widehat{\mu}[\{\widehat{x} \in \widehat{M} : x_n \in A\}] = \mu[A]$ for all $A \subset M$ Borel and all $n \leq 0$.

It is clear that $\widehat{\mu}$ is $\widehat{f}$–invariant. What is less clear is that the projection and lift procedures above are inverse operations, and that they preserve the Kolmogorov-Sinaĭ entropy, see [50]. Here is one consequence of this fact: μ is an equilibrium measure for a potential $\varphi : M \to \mathbb{R}$ iff $\widehat{\mu}$ is an equilibrium measure for $\varphi \circ \vartheta_0 : \widehat{M} \to \mathbb{R}$. In particular, the topological entropies of f and $\widehat{f}$ coincide, and μ is a measure of maximal entropy for f iff $\widehat{\mu}$ is a measure of maximal entropy for $\widehat{f}$.

Now let $N = \bigsqcup_{x \in M} N_x$ be a vector bundle over M, and let $A : N \to N$ measurable s.t. for every $x \in M$ the restriction $A \restriction_{N_x}$ is a linear isomorphism $A_x : N_x \to N_{f(x)}$. For example, if f is a differentiable endomorphism on a manifold M, then we can take $N = TM$ and $A = df$. The map A defines a (possibly non-invertible) cocycle $(A^{(n)})_{n \geq 0}$ over f by $A_x^{(n)} = A_{f^{n-1}(x)} \cdots A_{f(x)} A_x$ for $x \in M$, $n \geq 0$. There is a way of extending $(A^{(n)})_{n \geq 0}$ to an invertible cocycle over $\widehat{f}$. For $\widehat{x} \in \widehat{M}$, let $N_{\widehat{x}} := N_{\vartheta_0[\widehat{x}]}$ and let $\widehat{N} := \bigsqcup_{\widehat{x} \in \widehat{M}} N_{\widehat{x}}$, a vector bundle over $\widehat{M}$. Define the map $\widehat{A} : \widehat{N} \to \widehat{N}$, $\widehat{A}_{\widehat{x}} := A_{\vartheta_0[\widehat{x}]}$. For $\widehat{x} = (x_n)_{n \in \mathbb{Z}}$, define

$$\widehat{A}_{\widehat{x}}^{(n)} := \begin{cases} A_{x_0}^{(n)} & \text{, if } n \geq 0 \\ A_{x_{-n}}^{-1} \cdots A_{x_{-2}}^{-1} A_{x_{-1}}^{-1} & \text{, if } n \leq 0. \end{cases}$$

By definition, $\widehat{A}_{\widehat{x}}^{(m+n)} = \widehat{A}_{\widehat{f}^n(\widehat{x})}^{(m)} \widehat{A}_{\widehat{x}}^{(n)}$ for all $\widehat{x} \in \widehat{M}$ and all $m, n \in \mathbb{Z}$, hence $(\widehat{A}^{(n)})_{n \in \mathbb{Z}}$ is an invertible cocycle over $\widehat{f}$.

Whenever it is convenient, we will write ϑ to represent ϑ_0.

2. Pesin theory

We define changes of coordinates for which the inverse branches of f become uniformly contracting. Fix $\chi > 0$.

THE SET NUE_χ: It is the set of points $\widehat{x} \in \widehat{M} \backslash \bigcup_{n \in \mathbb{Z}} \widehat{f}^n(\vartheta^{-1}[\mathscr{S}])$ s.t.

$$\lim_{n \to \pm\infty} \frac{1}{n} \log \|\widehat{df}_{\widehat{x}}^{(n)}\| > \chi.$$

PARAMETER $u(\widehat{x})$: For $\widehat{x} \in \mathrm{NUE}_\chi$, define

$$u(\widehat{x}) := \left(\sum_{n \geq 0} e^{2n\chi} |\widehat{df}_{\widehat{x}}^{(-n)}|^2 \right)^{1/2}.$$

It is clear that $1 < u(\widehat{x}) < \infty$. The parameter $u(\widehat{x})$ controls the quality of hyperbolicity (contraction) of the inverse branches of f.

PESIN CHART $\Psi_{\widehat{x}}$: For $\widehat{x}\in\mathrm{NUE}_\chi$, the *Pesin chart* at $\widehat{x}$ is the map $\Psi_{\widehat{x}}:\mathbb{R}\to\mathbb{R}$, $\Psi_{\widehat{x}}(t):=\frac{1}{u(\widehat{x})}t+\vartheta[\widehat{x}]$. It is a diffeomorphism with $\Psi_{\widehat{x}}(0)=\vartheta[\widehat{x}]$ and $d\Psi_{\widehat{x}}\equiv\frac{1}{u(\widehat{x})}$.

Lemma 2.1. *For all $\widehat{x}\in\mathrm{NUE}_\chi$, the composition $F_{\widehat{x}}:=\Psi_{\widehat{f}(\widehat{x})}^{-1}\circ f\circ\Psi_{\widehat{x}}$ is a diffeomorphism from $R[2\mathfrak{r}(x_0)]$ onto its image, and it satisfies $F_{\widehat{x}}(0)=0$ and $|(dF_{\widehat{x}})_0|>e^\chi$.*

Proof. Let $\widehat{x}=(x_n)_{n\in\mathbb{Z}}$. The maps $\Psi_{\widehat{f}(\widehat{x})}^{-1}$, $\Psi_{\widehat{x}}$ are globally defined diffeomorphisms. Observing that $\Psi_{\widehat{x}}(R[2\mathfrak{r}(x_0)])\subset B(x_0,2\mathfrak{r}(x_0))$ and that, by (A1), the restriction of f to $B(x_0,2\mathfrak{r}(x_0))$ is a diffeomorphism onto its image, it follows that $F_{\widehat{x}}$ is a diffeomorphism from $R[2\mathfrak{r}(x_0)]$ onto its image. Also, $F_{\widehat{x}}(0)=(\Psi_{\widehat{f}(\widehat{x})}^{-1}\circ f)(\vartheta[\widehat{x}])=\Psi_{\widehat{f}(\widehat{x})}^{-1}(\vartheta[\widehat{f}(\widehat{x})])=0$. It remains to estimate $|(dF_{\widehat{x}})_0|$. Write $x=\vartheta[\widehat{x}]$ and note that $(dF_{\widehat{x}})_0=df_x\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{x})}$. Since $\widehat{df}_{\widehat{x}}^{(-n)}=\widehat{df}_{\widehat{f}(\widehat{x})}^{(-n-1)}\circ\widehat{df}_{\widehat{x}}=\widehat{df}_{\widehat{f}(\widehat{x})}^{(-n-1)}\circ df_x$ for $n\geq 0$, we have

$$\begin{aligned}u(\widehat{f}(\widehat{x}))^2&=1+\sum_{n\geq 0}e^{2(n+1)\chi}|\widehat{df}_{\widehat{f}(\widehat{x})}^{(-n-1)}|^2=1+e^{2\chi}|df_x|^{-2}\sum_{n\geq 0}e^{2n\chi}|\widehat{df}_{\widehat{x}}^{(-n)}|^2\\&=1+e^{2\chi}|df_x|^{-2}u(\widehat{x})^2,\end{aligned}$$

therefore $|(dF_{\widehat{x}})_0|^2=e^{2\chi}\frac{u(\widehat{f}(\widehat{x}))^2}{u(\widehat{f}(\widehat{x}))^2-1}>e^{2\chi}$. □

Now we control the distance of trajectories of $\widehat{f}$ to the singular set $\mathscr{S}$. For $\widehat{x}\in\widehat{M}$, define $\rho(\widehat{x}):=d(\{\vartheta_{-1}[\widehat{x}],\vartheta_0[\widehat{x}],\vartheta_1[\widehat{x}]\},\mathscr{S})$.

REGULAR SET: The *regular set of f* is

$$\begin{aligned}\mathrm{Reg}&:=\left\{\widehat{x}\in\widehat{M}\backslash\bigcup_{n\in\mathbb{Z}}\widehat{f}^n(\vartheta^{-1}[\mathscr{S}]):\lim_{n\to\pm\infty}\tfrac{1}{|n|}\log\rho(\widehat{f}^n(\widehat{x}))=0\right\}\\&=\left\{\widehat{x}\in\widehat{M}\backslash\bigcup_{n\in\mathbb{Z}}\widehat{f}^n(\vartheta^{-1}[\mathscr{S}]):\lim_{n\to\pm\infty}\tfrac{1}{|n|}\log d(\vartheta_n[\widehat{x}],\mathscr{S})=0\right\}.\end{aligned}$$

THE SET NUE_χ^*: It is the set of $\widehat{x}\in\mathrm{NUE}_\chi$ with the following properties:

(1) $\widehat{x}\in\mathrm{Reg}$.
(2) There exists a sequence $n_k\to-\infty$ s.t. $u(\widehat{f}^{n_k}(\widehat{x}))\to u(\widehat{x})$.
(3) $\lim_{n\to-\infty}\frac{1}{n}\log u(\widehat{f}^n(\widehat{x}))=0$.

The next lemma shows that NUE_χ^* carries the measures that are relevant to us.

Lemma 2.2. *If $\widehat{\mu}=$ lift of f–adapted χ–expanding measure, then $\widehat{\mu}[\mathrm{NUE}_\chi^*]=1$.*

Proof. Let μ be an f–adapted χ–expanding measure, and let $\widehat{\mu}$ be its lift. Let $\widehat{X}_0:=\widehat{M}\backslash\bigcup_{n\in\mathbb{Z}}\widehat{f}^n(\vartheta^{-1}(\mathscr{S}))$. By f–adaptedness and the definition of $\widehat{\mu}$ we have $\widehat{\mu}(\vartheta^{-1}[\mathscr{S}])=\mu(\mathscr{S})=0$, hence $\widehat{\mu}(\widehat{X}_0)=1$. By (A2), $|\log|df_x||\leq a|\log d(x,\mathscr{S})|$ and so f–adaptedness implies that $\log|df_x|\in L^1(\mu)$. Hence $\int|\log|\widehat{df}_{\widehat{x}}||d\widehat{\mu}(\widehat{x})=\int|\log|df_x||d\mu(x)<\infty$, thus proving that $\log|\widehat{df}|\in L^1(\widehat{\mu})$. By the Birkhoff ergodic theorem, $\exists\widehat{X}_1\subset\widehat{M}$ with $\widehat{\mu}(\widehat{X}_1)=1$ s.t. $\lim_{n\to\pm\infty}\frac{1}{n}\log|\widehat{df}_{\widehat{x}}^{(n)}|$ exists for all $\widehat{x}\in\widehat{X}_1$. Now let Y be the set of $x\in M$ s.t. $\lim_{n\to+\infty}\frac{1}{n}\log|df_x^n|>\chi$, and let $\widehat{X}_2=\vartheta^{-1}(Y)$. We also have $\widehat{\mu}(\widehat{X}_2)=1$, therefore $\mu(\widehat{X}_0\cap\widehat{X}_1\cap\widehat{X}_2)=1$. Since $\widehat{df}_{\widehat{x}}^{(n)}=df_{\vartheta[\widehat{x}]}^n$ for $n\geq 0$, we have that $\lim_{n\to\pm\infty}\frac{1}{n}\log|\widehat{df}_{\widehat{x}}^{(n)}|>\chi$ for all $\widehat{x}\in\widehat{X}_1\cap\widehat{X}_2$, therefore $\widehat{X}_0\cap\widehat{X}_1\cap\widehat{X}_2\subset\mathrm{NUE}_\chi$. This implies that $\widehat{\mu}[\mathrm{NUE}_\chi]=1$.

To prove that $\widehat{\mu}[\mathrm{NUE}_\chi^*]=1$, it remains to show that conditions (1)–(3) hold $\widehat{\mu}$–a.e. We have $\int|\log d(\vartheta[\widehat{x}],\mathscr{S})|d\widehat{\mu}(\widehat{x})=\int|\log d(x,\mathscr{S})|d\mu(x)<\infty$. By $\widehat{f}$–invariance, $\log d(\vartheta_{-1}[\widehat{x}],\mathscr{S})$, $\log d(\vartheta_0[\widehat{x}],\mathscr{S})$, $\log d(\vartheta_1[\widehat{x}],\mathscr{S})$ are in $L^1(\mu)$, hence is also their minimum $\log\rho(\widehat{x})$. By the Birkhoff ergodic theorem[2] we get $\widehat{\mu}(\text{Reg})=1$. By the Poincaré recurrence theorem, condition (2) holds $\widehat{\mu}$–a.e.

Finally, we check that condition (3) holds $\widehat{\mu}$–a.e. For $\widehat{x}\in\text{NUE}_\chi$, let $\varphi(\widehat{x}):=\log|(dF_{\widehat{x}})_0|$. It is enough to show that $\varphi\in L^1(\widehat{\mu})$, because of the following:

- By the proof of Lemma 2.1, $\varphi=\psi+[\log u\circ\widehat{f}-\log u]$, where $\psi:=\log|\widehat{df}|\in L^1(\widehat{\mu})$.
- By the Poincaré recurrence theorem, $\liminf_{n\to-\infty}\frac{1}{n}\log u(\widehat{f}^n(\widehat{x}))=0$ $\widehat{\mu}$–a.e., hence $\liminf_{n\to-\infty}\frac{\varphi_n(\widehat{x})}{n}=\liminf_{n\to-\infty}\frac{\psi_n(\widehat{x})}{n}$ $\widehat{\mu}$–a.e., where φ_n,ψ_n denote the Birkhoff sums of φ,ψ with respect to $\widehat{f}$.
- By the Birkhoff ergodic theorem, $\lim_{n\to-\infty}\frac{\varphi_n(\widehat{x})}{n}=\lim_{n\to-\infty}\frac{\psi_n(\widehat{x})}{n}$ $\widehat{\mu}$–a.e., hence $\lim_{n\to-\infty}\frac{1}{n}\log u(\widehat{f}^n(\widehat{x}))=0$ $\widehat{\mu}$–a.e.

We show that $\varphi\in L^1(\widehat{\mu})$. By Lemma 2.1 we have $\varphi>\chi$, hence it is enough to prove that $\int\varphi d\widehat{\mu}<\infty$. Since $2\varphi(\widehat{x})=\log\left(e^{2\chi}\frac{u(\widehat{f}(\widehat{x}))^2}{u(\widehat{f}(\widehat{x}))^2-1}\right)$ and $\widehat{\mu}$ is $\widehat{f}$–invariant, this former claim is equivalent to showing that $\int\log\left(e^{-2\chi}\frac{u(\widehat{x})^2-1}{u(\widehat{x})^2}\right)>-\infty$. By (A2), $u(\widehat{x})^2\geq 1+e^{2\chi}\|\widehat{df}_{\widehat{x}}^{(-1)}\|^2\geq 1+e^{2\chi}d(\vartheta_{-1}[\widehat{x}],\mathscr{S})^{2a}\geq 1+e^{2\chi}\rho(\widehat{x})^{2a}$ thus

$$e^{-2\chi}\frac{u(\widehat{x})^2-1}{u(\widehat{x})^2}=e^{-2\chi}\left(1-\frac{1}{u(\widehat{x})^2}\right)\geq\frac{\rho(\widehat{x})^{2a}}{1+e^{2\chi}\rho(\widehat{x})^{2a}}\geq\frac{\rho(\widehat{x})^{2a}}{1+e^{2\chi}},$$

hence $\int\log\left(e^{-2\chi}\frac{u(\widehat{x})^2-1}{u(\widehat{x})^2}\right)d\widehat{\mu}(\widehat{x})\geq 2a\int\log\rho(\widehat{x})d\widehat{\mu}(\widehat{x})-\log(1+e^{2\chi})>-\infty$. This completes the proof of the lemma. □

2.1. Inverse branches

By (A1), the inverse branch of f that sends $f(x)$ to x is a well-defined diffeomorphism from $E_x=B(f(x),2\mathfrak{r}(x))$ onto its image.

INVERSE BRANCH OF f: Let $g_x:E_x\to g_x(E_x)$ be the inverse branch of f that sends $f(x)$ to x. For $\widehat{x}=(x_n)_{n\in\mathbb{Z}}\in\widehat{M}$, define $g_{\widehat{x}}:=g_{x_{-1}}$.

Note that $g_{\widehat{x}}(x_0)=x_{-1}$. We want to mimic the behaviour of $(dF_{\widehat{x}})_0$ to the inverse branches $g_{\widehat{x}}$. For that, we need to reduce the domains of Pesin charts to intervals so that their images do not intersect the singular set $\mathscr{S}$ and at the same time we can control the variation of df. Given $\varepsilon>0$, let $I_\varepsilon:=\{e^{-\frac{1}{3}\varepsilon n}:n\geq 0\}$.

PARAMETER $Q_\varepsilon(\widehat{x})$: For $\widehat{x}\in\text{NUE}_\chi$, let $Q_\varepsilon(\widehat{x}):=\max\{q\in I_\varepsilon:q\leq\widetilde{Q}_\varepsilon(\widehat{x})\}$, where

$$\widetilde{Q}_\varepsilon(\widehat{x})=\varepsilon^{3/\beta}\min\left\{u(\widehat{x})^{-24/\beta},u(\widehat{f}^{-1}(\widehat{x}))^{-12/\beta}\rho(\widehat{x})^{72a/\beta}\right\}.$$

The term $\varepsilon^{3/\beta}$ will allow to absorb multiplicative constants. The choice of $Q_\varepsilon(\widehat{x})$ guarantees that the inverse of $F_{\widehat{f}^{-1}(\widehat{x})}$ is well-defined in $R[10Q_\varepsilon(\widehat{x})]$ and that it is close to a linear contraction (Theorem 2.4), and it also allows to compare nearby Pesin charts (Proposition 2.5). We have the following trivial bounds:

$$Q_\varepsilon(\widehat{x})\leq\varepsilon^{3/\beta},\ u(\widehat{x})Q_\varepsilon(\widehat{x})^{\beta/24}\leq\varepsilon^{1/8},\ u(\widehat{f}^{-1}(\widehat{x}))Q_\varepsilon(\widehat{x})^{\beta/12}\leq\varepsilon^{1/4},$$
$$\rho(\widehat{x})^{-a}Q_\varepsilon(\widehat{x})^{\beta/72}\leq\varepsilon^{1/24}.$$

Let $\widehat{x}=(x_n)_{n\in\mathbb{Z}}\in\text{NUE}_\chi$. The definition of $Q_\varepsilon(\widehat{x})$ is strong enough to guarantee that $\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])$ is contained in neighbourhoods where (A1)–(A3) hold:

---

[2] Here we are using that if $\varphi:M\to\mathbb{R}$ satisfies $\int|\varphi|d\mu<\infty$ then $\lim_{n\to\pm\infty}\frac{1}{n}\varphi(f^n(x))=0$ μ–a.e. Indeed, by the Birkhoff ergodic theorem $\widetilde{\varphi}(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\varphi(f^i(x))$ exists μ–a.e., hence $\lim_{n\to\infty}\frac{1}{n}\varphi(f^n(x))=\lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=0}^{n}\varphi(f^i(x))-\frac{1}{n}\sum_{i=0}^{n-1}\varphi(f^i(x))\right]=0$ μ–a.e. The same argument works for $n\to-\infty$.

- $\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})]) \subset D_{x_0} \cap E_{x_{-1}}$: we have $\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})]) \subset B(x_0, 10Q_\varepsilon(\widehat{x})) \subset D_{x_0} \cap E_{x_{-1}}$, since $10Q_\varepsilon(\widehat{x}) < 10\varepsilon^{3/\beta}\rho(\widehat{x})^a < \mathfrak{r}(x_{-1}), \mathfrak{r}(x_0)$.
- $g_{\widehat{x}}(\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])) \subset D_{x_{-1}}$: by the previous item, $g_{\widehat{x}}(\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})]))$ is well-defined. By (A2),

$$g_{\widehat{x}}(\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])) \subset g_{\widehat{x}}[B(x_0, 10Q_\varepsilon(\widehat{x}))] \subset B(x_{-1}, 10Q_\varepsilon(\widehat{x})d(x_{-1}, \mathscr{S})^{-a}) \subset D_{x_{-1}},$$

  since $10Q_\varepsilon(\widehat{x})d(x_{-1}, \mathscr{S})^{-a} < 10\varepsilon^{3/\beta}\rho(\widehat{x})^{2a}d(x_{-1}, \mathscr{S})^{-a} < \rho(\widehat{x})^a < \mathfrak{r}(x_{-1})$.
- $\Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})]) \subset g_{\widehat{f}(\widehat{x})}(E_{x_0})$: noting that $g_{\widehat{f}(\widehat{x})} = g_{x_0}$, assumption (A2) implies $g_{x_0}(E_{x_0}) \supset B(x_0, 2\mathfrak{r}(x_0)d(x_0, \mathscr{S})^a) \supset B(x_0, 10Q_\varepsilon(\widehat{x}))$, since $10Q_\varepsilon(\widehat{x}) < 10\varepsilon^{3/\beta}\rho(\widehat{x})^{2a} < 2\mathfrak{r}(x_0)d(x_0, \mathscr{S})^a$. Thus $g_{x_0}(E_{x_0}) \supset \Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])$.

The third item implies the following: if $y \in \Psi_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])$, then y is the unique pre-image of $f(y)$ in $g_{\widehat{f}(\widehat{x})}(E_{x_0})$, and $y = g_{\widehat{f}(\widehat{x})}(f(y))$.

Lemma 2.3 *(Tempering Kernel). If $\widehat{x} \in \mathrm{NUE}^*_\chi$, then*

$$\lim_{n\to-\infty} \tfrac{1}{|n|} \log Q_\varepsilon(\widehat{f}^n(\widehat{x})) = 0.$$

Proof. Clearly $\limsup_{n\to-\infty} \frac{1}{|n|} \log Q_\varepsilon(\widehat{f}^n(\widehat{x})) \le 0$. Conversely, $\widehat{x} \in \mathrm{Reg}$ implies that $\lim_{n\to-\infty} \frac{1}{|n|} \log \rho(\widehat{f}^n(\widehat{x})) = 0$. By condition (3) in the definition of NUE^*_χ, $\lim_{n\to-\infty} \frac{1}{|n|} \log u(\widehat{f}^n(\widehat{x})) = 0$, therefore $\liminf_{n\to-\infty} \frac{1}{|n|} \log Q_\varepsilon(\widehat{f}^n(\widehat{x})) \ge 0$. □

Inverse branches of f in Pesin charts: For $\widehat{x} \in \mathrm{NUE}_\chi$, let $G_{\widehat{x}} : R[2\mathfrak{r}(x_{-1})] \to \mathbb{R}$ be the composition defined by $G_{\widehat{x}} := \Psi^{-1}_{\widehat{f}^{-1}(\widehat{x})} \circ g_{\widehat{x}} \circ \Psi_{\widehat{x}}$.

If $\widehat{x} = (x_n)_{n\in\mathbb{Z}}$ then $G_{\widehat{x}}$ is the representation of $g_{\widehat{x}}$ in Pesin charts. Note that $G_{\widehat{x}}$ is a diffeomorphism from $R[2\mathfrak{r}(x_{-1})]$ onto its image, since $\Psi^{-1}_{\widehat{f}^{-1}(\widehat{x})}, \Psi_{\widehat{x}}$ are globally defined diffeomorphisms and $g_{\widehat{x}}$ is a diffeomorphism from $E_{x_{-1}} \supset \Psi_{\widehat{x}}(R[2\mathfrak{r}(x_{-1})])$ onto its image, by (A1). Also, $G_{\widehat{x}}$ is the inverse of $F_{\widehat{f}^{-1}(\widehat{x})}$ where the compositions are well-defined. The next theorem gives a better understanding of $G_{\widehat{x}}$.

Theorem 2.4. *The following holds for all $\varepsilon > 0$ small enough: If $\widehat{x} \in \mathrm{NUE}_\chi$ then $G_{\widehat{x}}$ is a diffeomorphism from $R[10Q_\varepsilon(\widehat{x})]$ onto its image, and it can be written as $G_{\widehat{x}}(t) = At + h(t)$ where:*

(1) $|A| < e^{-\chi}$.
(2) $h(0) = dh_0 = 0$ *and* $\|h\|_{1+\beta/2} < \varepsilon$, *where the norm is taken in* $R[10Q_\varepsilon(\widehat{x})]$.

In particular, $\|dG_{\widehat{x}}\|_0 < e^{-\chi/2}$.

Proof. Write $\widehat{x} = (x_n)_{n\in\mathbb{Z}}$. Since $10Q_\varepsilon(\widehat{x}) < \rho(\widehat{x})^a < 2\mathfrak{r}(x_{-1})$, it follows from the previous paragraph that the restriction of $G_{\widehat{x}}$ to $R[10Q_\varepsilon(\widehat{x})]$ is a diffeomorphism onto its image. Now we check (1)–(2). Let $A := (dG_{\widehat{x}})_0$ and $h : R[10Q_\varepsilon(\widehat{x})] \to \mathbb{R}$ s.t. $G_{\widehat{x}}(t) = At + h(t)$. By Lemma 2.1, $|A| = |(dF_{\widehat{f}^{-1}(\widehat{x})})_0|^{-1} < e^{-\chi}$. Clearly $h(0) = h'(0) = 0$, so it remains to estimate $\|h\|_{1+\beta/2}$.

Claim: $|(dG_{\widehat{x}})_{t_1} - (dG_{\widehat{x}})_{t_2}| \le \frac{\varepsilon}{3}|t_1 - t_2|^{\beta/2}$ for all $t_1, t_2 \in R[10Q_\varepsilon(\widehat{x})]$.

Before proving the claim, let us show how to conclude (2). If $\varepsilon > 0$ is small enough then $R[10Q_\varepsilon(\widehat{x})] \subset R[1]$. Applying the claim with $t_2 = 0$, we get $|h'(t)| \le \frac{\varepsilon}{3}|t|^{\beta/2} < \frac{\varepsilon}{3}$. By the mean value inequality, $|h(t)| \le \frac{\varepsilon}{3}|t| < \frac{\varepsilon}{3}$, hence $\|h\|_{1+\beta/2} < \varepsilon$.

Proof of the claim. We have $\Psi_{\widehat{x}}(t_1), \Psi_{\widehat{x}}(t_2) \in E_{x_{-1}}$. Using that $\Psi_{\widehat{x}}$ is a contraction and assumption (A3), we get:

$$|(dG_{\widehat{x}})_{t_1} - (dG_{\widehat{x}})_{t_2}| = \frac{u(\widehat{f}^{-1}(\widehat{x}))}{u(\widehat{x})}|g'(\Psi_{\widehat{x}}(t_1)) - g'(\Psi_{\widehat{x}}(t_2))| \le \mathfrak{K}u(\widehat{f}^{-1}(\widehat{x}))|t_1 - t_2|^\beta.$$

Since $|t_1 - t_2| < 20Q_\varepsilon(\widehat{x})$, if $\varepsilon > 0$ is small then:

$$\mathfrak{K}u(\widehat{f}^{-1}(\widehat{x}))|t_1 - t_2|^{\beta/2} < 20\mathfrak{K}u(\widehat{f}^{-1}(\widehat{x}))\varepsilon^{3/2}u(\widehat{f}^{-1}(\widehat{x}))^{-6} < 20\mathfrak{K}\varepsilon^{3/2} < \varepsilon.$$

This completes the proof of the claim. □

If $\varepsilon > 0$ is small enough then $\|dG_{\widehat{x}}\|_0 \le |(dG_{\widehat{x}})_0| + \|dh\|_0 < e^{-\chi} + \varepsilon < e^{-\chi/2}$. □

2.2. The overlap condition

Our next goal is to identify when two Pesin charts $\Psi_{\widehat{x}}, \Psi_{\widehat{y}}$ are close. Even when $\vartheta[\widehat{x}]$ and $\vartheta[\widehat{y}]$ are nearby points of M, the distortions of $\Psi_{\widehat{x}}$ and $\Psi_{\widehat{y}}$ might be very different. The values controlling such distortions are $u(\widehat{x})$ and $u(\widehat{y})$, so we need to compare them. Another requirement for Pesin charts to be close is that their domains of definition have comparable sizes. Because of this, we consider Pesin charts with different domains, which we call *ε–charts*.

ε–CHART: An *ε–chart* $\Psi_{\widehat{x}}^p$ is the restriction $\Psi_{\widehat{x}}\restriction_{[-p,p]}$, where $0 < p \le Q_\varepsilon(\widehat{x})$.

Note that for each $\widehat{x} \in \mathrm{NUE}_\chi$ there are infinitely many ε–charts centred at $\widehat{x}$.

ε–OVERLAP: Two ε–charts $\Psi_{\widehat{x}_1}^{p_1}, \Psi_{\widehat{x}_2}^{p_2}$ are said to *ε–overlap* if $\frac{p_1}{p_2} = e^{\pm\varepsilon}$ and

$$d(\vartheta[\widehat{x}_1], \vartheta[\widehat{x}_2]) + |u(\widehat{x}_1)^{-1} - u(\widehat{x}_2)^{-1}| < (p_1p_2)^4.$$

When this happens, we write $\Psi_{\widehat{x}_1}^{p_1} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_2}^{p_2}$.

Clearly, if $\Psi_{\widehat{x}_1}^{p_1} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_2}^{p_2}$ then $\Psi_{\widehat{x}_1}^{cp_1} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_2}^{cp_2}$ for all $c > 1$ s.t. $cp_i \le Q_\varepsilon(\widehat{x}_i)$. The next proposition shows that ε–overlap is strong enough to guarantee that the Pesin charts are close.

Proposition 2.5. *The following holds for $\varepsilon > 0$ small enough. If $\Psi_{\widehat{x}_1}^{p_1} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_2}^{p_2}$ then:*

(1) CONTROL OF u: $\frac{u(\widehat{x}_1)}{u(\widehat{x}_2)} = e^{\pm(p_1p_2)^3}$.
(2) OVERLAP: $\Psi_{\widehat{x}_i}(R[e^{-2\varepsilon}p_i]) \subset \Psi_{\widehat{x}_j}(R[p_j])$ *for* $i, j = 1, 2$.
(3) CHANGE OF COORDINATES: *For* $i, j = 1, 2$ *it holds* $\|\Psi_{\widehat{x}_i}^{-1} \circ \Psi_{\widehat{x}_j} - \mathrm{Id}\|_2 < \varepsilon(p_1p_2)^3$ *where the norm is taken in* $R[1]$.

Proof. (1) Since $\varepsilon > 0$ is small, it is enough to prove that $\left|\frac{u(\widehat{x}_1)}{u(\widehat{x}_2)} - 1\right| < \varepsilon^{3/\beta}(p_1p_2)^3$. By assumption, $|u(\widehat{x}_1)^{-1} - u(\widehat{x}_2)^{-1}| < (p_1p_2)^4$. Also $u(\widehat{x}_1) < \frac{\varepsilon^{3/\beta}}{Q_\varepsilon(\widehat{x}_1)} < \frac{\varepsilon^{3/\beta}}{p_1p_2}$, therefore

$$\left|\frac{u(\widehat{x}_1)}{u(\widehat{x}_2)} - 1\right| = u(\widehat{x}_1)|u(\widehat{x}_1)^{-1} - u(\widehat{x}_2)^{-1}| < \varepsilon^{3/\beta}(p_1p_2)^3.$$

(2) We prove that $\Psi_{\widehat{x}_1}(R[e^{-2\varepsilon}p_1]) \subset \Psi_{\widehat{x}_2}(R[p_2])$. If $t \in R[e^{-2\varepsilon}p_1]$ then

$$d(\Psi_{\widehat{x}_1}(t), \Psi_{\widehat{x}_2}(t)) \le d(\vartheta[\widehat{x}_1], \vartheta[\widehat{x}_2]) + |u(\widehat{x}_1)^{-1} - u(\widehat{x}_2)^{-1}| < (p_1p_2)^4,$$

therefore $\Psi_{\widehat{x}_1}(t) \in B(\Psi_{\widehat{x}_2}(t), (p_1p_2)^4)$. We have $B(\Psi_{\widehat{x}_2}(t), (p_1p_2)^4) \subset \Psi_{\widehat{x}_2}(B)$ where $B = B(t, u(\widehat{x}_2)(p_1p_2)^4)$. If $t' \in B$ then $|t'| \le |t| + u(\widehat{x}_2)(p_1p_2)^4 \le (e^{-\varepsilon} + \varepsilon^{3/\beta})p_2 < p_2$ for $\varepsilon > 0$ small enough, thus $B \subset R[p_2]$.

(3) By direct calculation,

$$(\Psi_{\widehat{x}_2}^{-1} \circ \Psi_{\widehat{x}_1} - \mathrm{Id})(t) = \left(\frac{u(\widehat{x}_2)}{u(\widehat{x}_1)} - 1\right)t + u(\widehat{x}_2)(\vartheta[\widehat{x}_1] - \vartheta[\widehat{x}_2])$$

is a linear function. By part (1), the C^2 norm taken in $R[1]$ is

$$\|\Psi_{\widehat{x}_2}^{-1} \circ \Psi_{\widehat{x}_1} - \mathrm{Id}\|_2 = \|\Psi_{\widehat{x}_2}^{-1} \circ \Psi_{\widehat{x}_1} - \mathrm{Id}\|_1 \le 2\left|\frac{u(\widehat{x}_2)}{u(\widehat{x}_1)} - 1\right| + u(\widehat{x}_2)d(\vartheta[\widehat{x}_1], \vartheta[\widehat{x}_2])$$
$$< 2\varepsilon^{3/\beta}(p_1p_2)^3 + \varepsilon^{3/\beta}(p_1p_2)^3 < \varepsilon(p_1p_2)^3. \quad \square$$

2.3. The map $G_{\widehat{x},\widehat{y}}$

Let $\widehat{x},\widehat{y}\in\mathrm{NUE}_\chi$, and assume that $\Psi^p_{\widehat{f}^{-1}(\widehat{x})}\overset{\varepsilon}{\approx}\Psi^q_{\widehat{y}}$. We want to change $\Psi_{\widehat{f}^{-1}(\widehat{x})}$ by $\Psi_{\widehat{y}}$ in $G_{\widehat{x}}$ and obtain a result similar to Theorem 2.4.

THE MAP $G_{\widehat{x},\widehat{y}}$: If $\Psi^p_{\widehat{f}^{-1}(\widehat{x})}\overset{\varepsilon}{\approx}\Psi^q_{\widehat{y}}$, let $G_{\widehat{x},\widehat{y}}=\Psi_{\widehat{y}}^{-1}\circ g_{\widehat{x}}\circ\Psi_{\widehat{x}}$ wherever this composition is well-defined.

Note that $G_{\widehat{x},\widehat{y}}$ is the representation of $g_{\widehat{x}}$ in the charts $\Psi_{\widehat{x}}$ and $\Psi_{\widehat{y}}$. Alternatively, by Proposition 2.5, $G_{\widehat{x},\widehat{y}}:=\Psi_{\widehat{y}}^{-1}\circ\Psi_{\widehat{f}^{-1}(\widehat{x})}\circ G_{\widehat{x}}$ is a small perturbation of $G_{\widehat{x}}$. The next result makes this claim more precise.

Theorem 2.6. *The following holds for all $\varepsilon>0$ small enough: If $\widehat{x},\widehat{y}\in\mathrm{NUE}_\chi$ and $\Psi^p_{\widehat{f}^{-1}(\widehat{x})}\overset{\varepsilon}{\approx}\Psi^q_{\widehat{y}}$, then $G_{\widehat{x},\widehat{y}}$ is well-defined in $R[10Q_\varepsilon(\widehat{x})]$ and can be written as $G_{\widehat{x},\widehat{y}}(t)=At+h(t)$ where:*

(1) $|A|<e^{-\chi}$, *cf. Theorem 2.4.*
(2) $|h(0)|<\varepsilon(pq)^3$, $|dh_0|<\varepsilon(pq)^3$, *and* $\mathrm{Hol}_{\beta/3}(dh)<\varepsilon$ *where the norm is taken in* $R[10Q_\varepsilon(\widehat{x})]$.

In particular, $G_{\widehat{x},\widehat{y}}$ contracts at least by a factor of $e^{-\chi/2}$.

Proof. We write $G_{\widehat{x},\widehat{y}}=:H\circ G_{\widehat{x}}$ and see $G_{\widehat{x},\widehat{y}}$ as a small perturbation of $G_{\widehat{x}}$. By Theorem 2.4,

$$G_{\widehat{x}}(0)=0,\ \|dG_{\widehat{x}}\|_0<1,\ |(dG_{\widehat{x}})_{t_1}-(dG_{\widehat{x}})_{t_2}|\le\varepsilon|t_1-t_2|^{\beta/2}\text{ for }t_1,t_2\in R[10Q_\varepsilon(\widehat{x})]$$

where the C^0 norm is taken in $R[10Q_\varepsilon(\widehat{x})]$, and by Proposition 2.5(3) the function H is affine with

$$\|H-\mathrm{Id}\|_0<\varepsilon(pq)^3,\ \|d(H-\mathrm{Id})\|_0<\varepsilon(pq)^3$$

where the C^0 norms are taken in $R[1]$.

It is easy to show that $G_{\widehat{x},\widehat{y}}$ is well-defined in $R[10Q_\varepsilon(\widehat{x})]$: since $G_{\widehat{x}}(R[10Q_\varepsilon(\widehat{x})])\subset B(0,10Q_\varepsilon(\widehat{x}))\subset R[1]$, Proposition 2.5(3) implies that $G_{\widehat{x},\widehat{y}}$ is well-defined. Now we prove (1)–(2). Define $h:=G_{\widehat{x},\widehat{y}}-(dG_{\widehat{x}})_0=H\circ G_{\widehat{x}}-(dG_{\widehat{x}})_0$, where $(dG_{\widehat{x}})_0$ represents the linear functional on $\mathbb{R}$ defined by the derivative. Then $|h(0)|=|H(0)|<\varepsilon(pq)^3$ and $|dh_0|=|dH_0-1||(dG_{\widehat{x}})_0|<\varepsilon(pq)^3$. Finally, if $\varepsilon>0$ is small enough then for all $t_1,t_2\in R[10Q_\varepsilon(\widehat{x})]$ we have

$$\begin{aligned}|dh_{t_1}-dh_{t_2}|&=|dH_{G_{\widehat{x}}(t_1)}(dG_{\widehat{x}})_{t_1}-dH_{G_{\widehat{x}}(t_2)}(dG_{\widehat{x}})_{t_2}|=\|dH\|_0|(dG_{\widehat{x}})_{t_1}-(dG_{\widehat{x}})_{t_2}|\\&\le 2\varepsilon|t_1-t_2|^{\beta/2}<\varepsilon|t_1-t_2|^{\beta/3}.\end{aligned}$$

This completes the proof of (2). In particular, if $\varepsilon>0$ is small enough then $\|dh\|_0\le\varepsilon(pq)^3+\varepsilon(10Q_\varepsilon(\widehat{x}))^{\beta/3}<\varepsilon$ and hence $|G_{\widehat{x},\widehat{y}}(t_1)-G_{\widehat{x},\widehat{y}}(t_2)|\le(|A|+\|dh\|_0)|t_1-t_2|\le(e^{-\chi}+\varepsilon)|t_1-t_2|\le e^{-\chi/2}|t_1-t_2|$ for all $t_1,t_2\in R[10Q_\varepsilon(\widehat{x})]$. □

2.4. ε–generalized pseudo-orbits and the parameter $q_\varepsilon(\widehat{x})$

Now we define when we can pass from one ε–chart to another via the action of $\widehat{f}^{-1}$. We will define two such notions, one weak and one strong. While in [53,39,38] the authors only define one notion (similar to the strong notion presented below), here we also require a weaker one that will be relevant for us in Section 5.

WEAK EDGE $v\overset{\varepsilon}{\dashleftarrow}w$: Given ε–charts $v=\Psi^q_{\widehat{y}}$ and $w=\Psi^p_{\widehat{x}}$, we draw a weak edge from w to v if:

(WE1) OVERLAP: $\Psi^q_{\widehat{f}^{-1}(\widehat{x})}\overset{\varepsilon}{\approx}\Psi^q_{\widehat{y}}$.
(WE2) CONTROL OF PARAMETERS: $p\le e^\varepsilon q$.

When this happens, we write $v \overset{\varepsilon}{\dashleftarrow} w$.

Clearly, if $\Psi^q_{\widehat{y}} \overset{\varepsilon}{\dashleftarrow} \Psi^p_{\widehat{x}}$ then $\Psi^{cq}_{\widehat{y}} \overset{\varepsilon}{\dashleftarrow} \Psi^{cp}_{\widehat{x}}$ for all $c > 1$ s.t. $cq \le Q_\varepsilon(\widehat{y})$ and $cp \le Q_\varepsilon(\widehat{x})$. For $\varepsilon > 0$ small, define $\delta_\varepsilon := e^{-\varepsilon n} \in I_\varepsilon$ where n is the unique positive integer s.t. $e^{-\varepsilon n} < \varepsilon \le e^{-\varepsilon(n-1)}$. In particular, $\delta_\varepsilon < \varepsilon$.

EDGE $v \overset{\varepsilon}{\leftarrow} w$: Given ε–charts $v = \Psi^q_{\widehat{y}}$ and $w = \Psi^p_{\widehat{x}}$, we draw an edge from w to v if the following holds:

(E1) OVERLAP: $\Psi^q_{\widehat{f}^{-1}(\widehat{x})} \overset{\varepsilon}{\approx} \Psi^q_{\widehat{y}}$.
(E2) CONTROL OF PARAMETERS:
  (E2.1) $d(\vartheta_1[\widehat{y}], \vartheta_0[\widehat{x}]) < q$.
  (E2.2) $\frac{u(\widehat{f}(\widehat{y}))}{u(\widehat{x})} = e^{\pm q}$.
  (E2.3) $p = \min\{e^\varepsilon q, \delta_\varepsilon Q_\varepsilon(\widehat{x})\}$.

When this happens, we write $v \overset{\varepsilon}{\leftarrow} w$.

It is not hard to see that condition (E2.1) follows from (E1) and assumption (A2), but for reference purposes we write it separately. The parameters p, q are the sizes of unstable manifolds in the charts, and the greedy recursion in (E2.3) implies that, fixed an unstable manifold at $\widehat{y}$, the unstable manifold at $\widehat{x}$ is as big as possible. This maximality is crucial to prove the inverse theorem (Theorem 4.1).

Remark 2.7. Since f is nonuniformly expanding, our definition of edge is different from those in [53,39,38] in two senses. On one hand, we only need to consider one overlap and one recursive relation. On the other hand, the *lack of symmetry* between f and its inverse requires us to control some parameters separately, as stated in (E2.1) and (E2.2).

Lemma 2.8. *The following holds for all $\varepsilon > 0$ small enough. If $\Psi^q_{\widehat{y}} \overset{\varepsilon}{\dashleftarrow} \Psi^p_{\widehat{x}}$ then:*

(1) $G_{\widehat{x},\widehat{y}}(R[p]) \subset R[q]$.
(2) *If $x \in \Psi_{\widehat{x}}(R[p])$ then $g_{\widehat{x}}(x)$ is the unique $y \in \Psi_{\widehat{y}}(R[q])$ s.t. $f(y) = x$.*

Proof. By Theorem 2.6 and (WE2), $G_{\widehat{x},\widehat{y}}(R[p]) \subset B(G_{\widehat{x},\widehat{y}}(0), e^{-\frac{\chi}{2}}p) \subset R[\varepsilon q^6 + e^{-\frac{\chi}{2}}p] \subset R[q]$, since $\varepsilon q^6 + e^{-\frac{\chi}{2}}p < \varepsilon q + e^{-\frac{\chi}{2}+\varepsilon}q = (\varepsilon + e^{-\frac{\chi}{2}+\varepsilon})q < q$ for $\varepsilon > 0$ small enough. This proves part (1). Now take $x = \Psi_{\widehat{x}}(t) \in \Psi_{\widehat{x}}(R[p])$ and let $y = g_{\widehat{x}}(x) = (\Psi_{\widehat{y}} \circ G_{\widehat{x},\widehat{y}})(t)$. By definition $f(y) = x$, and by part (1) it holds $y \in \Psi_{\widehat{y}}(R[q])$. This proves the existence of y. To prove its uniqueness, note that

$$\Psi_{\widehat{y}}(R[q]) \subset \Psi_{\widehat{f}^{-1}(\widehat{x})}(R[e^{2\varepsilon}q]) \subset \Psi_{\widehat{f}^{-1}(\widehat{x})}(R[10Q_\varepsilon(\widehat{f}^{-1}(\widehat{x}))]) \subset g_{\widehat{x}}(E_{\vartheta_{-1}[\widehat{x}]}),$$

where in the first inclusion we used Proposition 2.5(2) and in the last we used the third item proved before Lemma 2.3. Since $g_{\widehat{x}} : E_{\vartheta_{-1}[\widehat{x}]} \to g_{\widehat{x}}(E_{\vartheta_{-1}[\widehat{x}]})$ is a diffeomorphism, there is at most one $y \in g_{\widehat{x}}(E_{\vartheta_{-1}[\widehat{x}]})$ s.t. $f(y) = x$. □

ε–GENERALIZED PSEUDO-ORBIT (ε–GPO): An *ε–generalized pseudo-orbit* (*ε–gpo*) is a sequence $\underline{v} = \{v_n\}_{n\in\mathbb{Z}}$ of ε–charts s.t. $v_n \overset{\varepsilon}{\leftarrow} v_{n+1}$ for all $n \in \mathbb{Z}$. A *weak ε–gpo* is a sequence $\underline{v} = \{v_n\}_{n\in\mathbb{Z}}$ of ε–charts s.t. $v_n \overset{\varepsilon}{\dashleftarrow} v_{n+1}$ for all $n \in \mathbb{Z}$.

By definition, a necessary condition for drawing a weak edge $\Psi^q_{\widehat{y}} \overset{\varepsilon}{\dashleftarrow} \Psi^p_{\widehat{x}}$ is that $p \le e^\varepsilon q$. We would like to draw an edge $\Psi^{Q_\varepsilon(\widehat{x})}_{\widehat{x}} \overset{\varepsilon}{\dashleftarrow} \Psi^{Q_\varepsilon(\widehat{f}(\widehat{x}))}_{\widehat{f}(\widehat{x})}$, but in general we cannot, because $\frac{Q_\varepsilon(\widehat{f}(\widehat{x}))}{Q_\varepsilon(\widehat{x})}$ might be bigger than e^ε. To bypass this, we introduce the parameter $q_\varepsilon(\widehat{x})$ below.

PARAMETER $q_\varepsilon(\widehat{x})$: For $\widehat{x} \in \mathrm{NUE}^*_\chi$, let $q_\varepsilon(\widehat{x}) := \delta_\varepsilon \min\{e^{\varepsilon|n|}Q_\varepsilon(\widehat{f}^n(\widehat{x})) : n \le 0\}$.

The above minimum is the greedy way of defining values in I_ε smaller than $\varepsilon Q_\varepsilon$ with the required regularity property, as we now show.

Lemma 2.9. *For all $\widehat{x} \in \mathrm{NUE}_\chi^*$, the following holds:*

(1) GOOD DEFINITION: $0 < q_\varepsilon(\widehat{x}) < \varepsilon Q_\varepsilon(\widehat{x})$.
(2) GREEDY ALGORITHM: $q_\varepsilon(\widehat{f}^n(\widehat{x})) = \min\{e^\varepsilon q_\varepsilon(\widehat{f}^{n-1}(\widehat{x})), \delta_\varepsilon Q_\varepsilon(\widehat{f}^n(\widehat{x}))\}$, $\forall n \in \mathbb{Z}$.

Proof. By Lemma 2.3, $\inf\{e^{\varepsilon|n|} Q_\varepsilon(\widehat{f}^n(\widehat{x})) : n \le 0\} > 0$. Since zero is the only accumulation point of I_ε, $q_\varepsilon(\widehat{x})$ is well-defined and positive. It is clear that $q_\varepsilon(\widehat{x}) \le \delta_\varepsilon Q_\varepsilon(\widehat{x}) < \varepsilon Q_\varepsilon(\widehat{x})$, hence (1) is proved. For (2), fix $n \in \mathbb{Z}$ and note that

$$\begin{aligned} q_\varepsilon(f^n(\widehat{x})) &= \delta_\varepsilon \min\{e^{\varepsilon|m|} Q_\varepsilon(\widehat{f}^m(\widehat{f}^n(\widehat{x}))) : m \le 0\} \\ &= \min\{\delta_\varepsilon \min\{e^{\varepsilon|m|} Q_\varepsilon(\widehat{f}^{m+n}(\widehat{x})) : m \le -1\}, \delta_\varepsilon Q_\varepsilon(\widehat{f}^n(\widehat{x}))\} \\ &= \min\{e^\varepsilon \delta_\varepsilon \min\{e^{\varepsilon|m|} Q_\varepsilon(\widehat{f}^m(\widehat{f}^{n-1}(x))) : m \le 0\}, \delta_\varepsilon Q_\varepsilon(\widehat{f}^n(\widehat{x}))\} \\ &= \min\{e^\varepsilon q_\varepsilon(\widehat{f}^{n-1}(\widehat{x})), \delta_\varepsilon Q_\varepsilon(\widehat{f}^n(\widehat{x}))\}. \quad \square \end{aligned}$$

THE SET $\mathrm{NUE}_\chi^\#$: It is the set of $\widehat{x} \in \mathrm{NUE}_\chi^*$ s.t. $\limsup_{n\to+\infty} q_\varepsilon(\widehat{f}^n(\widehat{x})) > 0$ and $\limsup_{n\to-\infty} q_\varepsilon(\widehat{f}^n(\widehat{x})) > 0$.

Note that, while NUE_χ^* is defined by a set of conditions on the past orbit of $\widehat{x}$, the set $\mathrm{NUE}_\chi^\#$ is defined by conditions *both* on the past and on the future. This additional condition is important for the proof of Theorem 5.6.

2.5. Stable and unstable sets of weak ε–gpo's

Call a sequence $\underline{v}^+ = \{v_n\}_{n\ge0}$ a *positive weak ε–gpo* if $v_n \overset{\varepsilon}{\dashleftarrow} v_{n+1}$ for all $n \ge 0$. Similarly, a *negative weak ε–gpo* is a sequence $\underline{v}^- = \{v_n\}_{n\le0}$ s.t. $v_{n-1} \overset{\varepsilon}{\dashleftarrow} v_n$ for all $n \le 0$. Remember $\vartheta_n : \widehat{M} \to M$, $\vartheta_n[(x_k)_{k\in\mathbb{Z}}] = x_n$.

STABLE/UNSTABLE SET OF POSITIVE/NEGATIVE WEAK ε–GPO: The *stable set* of a positive weak ε–gpo $\underline{v}^+ = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\ge0}$ is

$$V^s[\underline{v}^+] := \{\widehat{x} \in \widehat{M} : \vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n]), \forall n \ge 0\}.$$

The *unstable set* of a negative weak ε–gpo $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\le0}$ is

$$V^u[\underline{v}^-] := \{\widehat{x} \in \widehat{M} : \vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n]), \forall n \le 0\}.$$

For a weak ε–gpo $\underline{v} = \{v_n\}_{n\in\mathbb{Z}}$, let $V^s[\underline{v}] := V^s[\{v_n\}_{n\ge0}]$ and $V^u[\underline{v}] := V^u[\{v_n\}_{n\le0}]$.

STABLE/UNSTABLE SETS AT v: Given an ε–chart v, a *stable set at* v is any $V^s[\underline{v}^+]$ where $\underline{v}^+$ is a positive weak ε–gpo with $v_0 = v$. Similarly, an *unstable set at* v is any $V^u[\underline{v}^-]$ where $\underline{v}^-$ is a negative weak ε–gpo with $v_0 = v$.

In the sequel, the notations $\underline{v}^+$, $\{v_n\}_{n\ge0}$ always mean a positive weak ε–gpo, and the notations $\underline{v}^-$, $\{v_n\}_{n\ge0}$ always mean a negative weak ε–gpo. The next lemma gives alternative characterizations of stable and unstable sets.

Lemma 2.10. *The following holds for all $\varepsilon > 0$ small enough.*

(1) *If $\underline{v}^+ = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\ge0}$ is a positive weak ε–gpo then $V^s[\underline{v}^+] = \vartheta^{-1}[x]$, where $x \in M$ is uniquely defined by $f^n(x) \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \ge 0$.*
(2) *If $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\le0}$ is a negative weak ε–gpo then*

$$\begin{aligned} V^u[\underline{v}^-] &= \{\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in \widehat{M} : x_0 \in \Psi_{\widehat{x}_0}(R[p_0]) \text{ and } x_{n-1} = g_{\widehat{x}_n}(x_n), \forall n \le 0\} \\ &= \{(\Psi_{\widehat{x}_n}(t_n))_{n\in\mathbb{Z}} \in \widehat{M} : t_0 \in R[p_0] \text{ and } t_{n-1} = G_{\widehat{x}_n,\widehat{x}_{n-1}}(t_n), \forall n \le 0\}. \end{aligned}$$

In other words, a stable set is the set of all possible pasts of a single $x \in M$, and an unstable set is isomorphic to the interval $\Psi_{\widehat{x}_0}(R[p_0])$, i.e. an element of an unstable set is uniquely determined by its zeroth coordinate.

Proof. Let $\underline{v}^+ = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\geq 0}$ be a positive weak ε–gpo. Firstly we prove that there exists a unique $x \in M$ s.t. $f^n(x) \in \Psi_{\widehat{x}_n}(R[p_n])$, $\forall n \geq 0$. Any such x is defined by a sequence $(t_n)_{n\geq 0}$ of values $t_n \in R[p_n]$ s.t. $f^n(x) = \Psi_{\widehat{x}_n}(t_n)$ and $t_n = G_{\widehat{x}_{n+1},\widehat{x}_n}(t_{n+1})$ for all $n \geq 0$. By Theorem 2.6, each $G_{\widehat{x}_{n+1},\widehat{x}_n}$ contracts at least by a factor $e^{-\chi/2}$, hence t_0 is the intersection of the descending chain of compact intervals $I_n := (G_{\widehat{x}_1,\widehat{x}_0} \circ \cdots \circ G_{\widehat{x}_n,\widehat{x}_{n-1}})(R[p_n])$, $n \geq 0$. By a similar reasoning, t_n is uniquely defined for all $n \geq 0$. By this uniqueness, $t_n = G_{\widehat{x}_{n+1},\widehat{x}_n}(t_{n+1})$ for all $n \geq 0$.

(1) Let $x \in M$ s.t. $f^n(x) \in \Psi_{\widehat{x}_n}(R[p_n])$, $\forall n \geq 0$. Take $\widehat{x} \in V^s[\underline{v}^+]$, and let $x_0 = \vartheta[\widehat{x}]$. Since $f^n(x_0) = \vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \geq 0$, we have $x_0 = x$, thus $V^s[\underline{v}^+] \subset \vartheta^{-1}[x]$. Conversely, $\widehat{y} \in \vartheta^{-1}[x] \Rightarrow \vartheta_n[\widehat{y}] = f^n(x) \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \geq 0$, hence $V^s[\underline{v}^+] \supset \vartheta^{-1}[x]$.

(2) Fix a negative weak ε–gpo $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}$. It is easy to see that the two alternative characterizations of $V^u[\underline{v}^-]$ are equivalent: $x_0 \in \Psi_{\widehat{x}_0}(R[p_0])$ iff $x_0 = \Psi_{\widehat{x}_0}(t_0)$ for some $t_0 \in R[p_0]$; $x_{n-1} = g_{\widehat{x}_n}(x_n)$ iff $x_{n-1} = \Psi_{\widehat{x}_{n-1}}(t_{n-1})$ and $x_n = \Psi_{\widehat{x}_n}(t_n)$ with $t_{n-1} = G_{\widehat{x}_n,\widehat{x}_{n-1}}(t_n)$. Hence it is enough to show the second characterization. Take $\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in V^u[\underline{v}^-]$. Fix $n \leq 0$. By assumption, $x_{n-1} \in \Psi_{\widehat{x}_{n-1}}(R[p_{n-1}])$, $x_n \in \Psi_{\widehat{x}_n}(R[p_n])$ and $f(x_{n-1}) = x_n$. By Lemma 2.8(2), it follows that $x_{n-1} = g_{\widehat{x}_n}(x_n)$.

Conversely, take $\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in \widehat{M}$ s.t. $x_0 \in \Psi_{\widehat{x}_0}(R[p_0])$ and $x_{n-1} = g_{\widehat{x}_n}(x_n)$ for all $n \leq 0$. By Lemma 2.8(1), we have $g_{\widehat{x}_n}(\Psi_{\widehat{x}_n}(R[p_n])) \subset \Psi_{\widehat{x}_{n-1}}(R[p_{n-1}])$ for all $n \in \mathbb{Z}$. Applying this for $n = 0$, we get that $x_{-1} = g_{\widehat{x}_0}(x_0) \in \Psi_{\widehat{x}_{-1}}(R[p_{-1}])$. By induction, it follows that $x_n \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \leq 0$. □

Here are the main properties of stable and unstable sets.

Proposition 2.11. *The following holds for all $\varepsilon > 0$ small enough.*

(1) PRODUCT STRUCTURE: *If V^s / V^u is a stable/unstable set at v then $V^s \cap V^u$ consists of a single element of $\widehat{M}$.*
(2) INVARIANCE:
$$\widehat{f}(V^s[\{v_n\}_{n\geq 0}]) \subset V^s[\{v_n\}_{n\geq 1}] \text{ and } \widehat{f}^{-1}(V^u[\{v_n\}_{n\leq 0}]) \subset V^u[\{v_n\}_{n\leq -1}].$$
(3) HYPERBOLICITY: *If $\widehat{y},\widehat{z} \in V^s[\underline{v}^+]$ then $d(\widehat{f}^n(\widehat{y}), \widehat{f}^n(\widehat{z})) = 2^{-n}d(\widehat{y},\widehat{z})$ for all $n \geq 0$. If $\widehat{y},\widehat{z} \in V^u[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}]$ then for all $n \leq 0$:*
   (a) $d(\widehat{f}^n(\widehat{y}), \widehat{f}^n(\widehat{z})) \leq 2p_0 e^{\frac{\chi}{2}n} d(\widehat{y},\widehat{z})$.
   (b) $|\log \|\widehat{df}_{\widehat{y}}^{(n)}\| - \log \|\widehat{df}_{\widehat{z}}^{(n)}\|| < Q_\varepsilon(\widehat{x}_0)^{\beta/4}$. *In particular,* $\frac{u(\widehat{y})}{u(\widehat{z})} = e^{\pm Q_\varepsilon(\widehat{x}_0)^{\beta/4}}$.
(4) DISJOINTNESS: *Let $v = \Psi_{\widehat{x}}^p$ and $w = \Psi_{\widehat{y}}^q$ with $\widehat{x} = \widehat{y}$. If V^s, W^s are stable sets at v, w then they are either disjoint or coincide. If V^u, W^u are unstable sets at v, w then they are either disjoint or one contains the other.*
(5) *Let $v \overset{\varepsilon}{\leftarrow\!\!-\!\!-} w$. If V^u is an unstable set at v then $\widehat{f}(V^u)$ intersects every stable set at w at a single element.*

Proof. (1) Write $v = \Psi_{\widehat{x}_0}^{p_0}$, $V^s = V^s[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\geq 0}]$, $V^u = V^u[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}]$. By Lemma 2.10(1), $\exists x \in M$ s.t. $V^s = \vartheta^{-1}[x]$. Any element $\widehat{x} = (x_n)_{n\in\mathbb{Z}}$ in $V^s \cap V^u$ satisfies $x_n = f^n(x)$ for all $n \geq 0$, and $x_{n-1} = g_{\widehat{x}_n}(x_n)$ for all $n \leq 0$. These conditions uniquely characterize $\widehat{x}$, hence $V^s \cap V^u$ is a singleton.

(2) If $\underline{v}^+ = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\geq 0}$ is a positive weak ε–gpo then $\widehat{x} \in V^s[\underline{v}^+] \Rightarrow \vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n])$, $\forall n \geq 0 \Rightarrow \vartheta_n[\widehat{f}(\widehat{x})] \in \Psi_{\widehat{x}_{n+1}}(R[p_{n+1}])$, $\forall n \geq 0 \Rightarrow \widehat{f}(\widehat{x}) \in V^s[\{v_n\}_{n\geq 1}]$. By a similar reason, if $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}$ is a negative weak ε–gpo then $\widehat{x} \in V^u[\underline{v}^-] \Rightarrow \vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n])$, $\forall n \leq 0 \Rightarrow \vartheta_n[\widehat{f}^{-1}(\widehat{x})] \in \Psi_{\widehat{x}_{n-1}}(R[p_{n-1}])$, $\forall n \leq 0 \Rightarrow \widehat{f}^{-1}(\widehat{x}) \in V^u[\{v_n\}_{n\leq -1}]$.

(3) Write $V^s[\underline{v}^+] = \vartheta^{-1}(x)$, and let $\widehat{y} = (y_n)_{n\in\mathbb{Z}}, \widehat{z} = (z_n)_{n\in\mathbb{Z}}$ be in $V^s[\underline{v}^+]$. For $n \geq 0$ we have $y_n = z_n = f^n(x)$, thus $d(\widehat{f}^n(\widehat{y}), \widehat{f}^n(\widehat{z})) = \sup\{2^{-k}d(y_{n-k}, z_{n-k}) : k \geq 0\} = \sup\{2^{-k}d(y_{n-k}, z_{n-k}) : k \geq n\} = 2^{-n}d(\widehat{y},\widehat{z})$.

Now let $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}$ be a negative weak ε–gpo, and take $\widehat{y} = (y_n)_{n\in\mathbb{Z}}, \widehat{z} = (z_n)_{n\in\mathbb{Z}} \in V^u[\underline{v}^-]$. By Lemma 2.10(2), for all $n \leq 0$ we can write $y_n = \Psi_{\widehat{x}_n}(t_n)$, $z_n = \Psi_{\widehat{x}_n}(t_n')$, where $t_0, t_0' \in R[p_0]$ and $t_{n-1} = G_{\widehat{x}_n,\widehat{x}_{n-1}}(t_n)$, $t_{n-1}' = G_{\widehat{x}_n,\widehat{x}_{n-1}}(t_n')$. Define $\Delta_n := t_n - t_n'$ for $n \leq 0$. By Theorem 2.6, $|\Delta_{n-1}| \leq e^{-\frac{\chi}{2}}|\Delta_n|$ for all $n \leq 0$, therefore $|\Delta_n| \leq e^{\frac{\chi}{2}n}|\Delta_0| \leq 2p_0 e^{\frac{\chi}{2}n}$ for all $n \leq 0$, and so $d(y_n, z_n) \leq 2p_0 e^{\frac{\chi}{2}n}$ (since $\Psi_{\widehat{x}_n}$ is 1–Lipschitz). We conclude that $d(\widehat{f}^n(\widehat{y}),\widehat{f}^n(\widehat{z}))\leq 2p_0e^{\frac{\chi}{2}n}d(\widehat{y},\widehat{z})$ for all $n\leq 0$. To prove (b), we proceed exactly as in the proof of Proposition 6.2(1)(c) of [38].

(4) Let V^s, W^s be stable sets in v, w respectively. By Lemma 2.10(1), $\exists y,z\in M$ s.t. $V^s=\vartheta^{-1}[y]$ and $W^s=\vartheta^{-1}[z]$, hence either $V^s\cap W^s=\emptyset$ or $V^s=W^s$.

Now let $V^u=V^u[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}]$ and $W^u=V^u[\{\Psi_{\widehat{y}_n}^{q_n}\}_{n\leq 0}]$, with $v=\Psi_{\widehat{x}}^{p}$, $w=\Psi_{\widehat{x}}^{q}$. If $V^u\cap W^u=\emptyset$ then there is nothing to prove, so assume that $V^u\cap W^u\neq\emptyset$. Assuming that $p\leq q$, we will prove that $V^u\subset W^u$ (the other case is identical). In the sequel, "n small enough" means that $n\leq 0$ and $|n|$ is large enough. The following claims hold.

- If n is small enough then $\vartheta_n[V^u]\subset\Psi_{\widehat{x}_n}(R[\frac{1}{2}p_n])$: take $\widehat{x}=(x_n)_{n\in\mathbb{Z}}\in V^u$. For $n\leq 0$ write $x_n=\Psi_{\widehat{x}_n}(t_n)$ with $t_0\in R[p_0]$ and $t_{n-1}=G_{\widehat{x}_n,\widehat{x}_{n-1}}(t_n)$, where $G_{\widehat{x}_n,\widehat{x}_{n-1}}(t)=A_nt+h_n(t)$. It is enough to show that $|t_n|<\frac{1}{2}p_n$ for n small enough. Start noting that $\|dh_n\|_0<2\varepsilon^2$, where the norm is taken in $R[\frac{1}{2}p_n]$. This is a direct consequence of Theorem 2.6: $\|dh_n\|_0\leq|d(h_n)_0|+\varepsilon p_n^{\beta/3}<\varepsilon p_n^6+\varepsilon^2<2\varepsilon^2$. Theorem 2.6 also says that $|A_n|<e^{-\chi}$ and $|h_n(0)|<\varepsilon p_n^6$, therefore if $\varepsilon>0$ is small enough then the following holds for all $n\leq 0$:

$$|t_{n-1}|\leq|A_n||t_n|+|h_n(t_n)|<e^{-\chi}|t_n|+\varepsilon p_n^6+2\varepsilon^2|t_n|<e^{-\frac{\chi}{2}}|t_n|+\varepsilon p_n.$$

  By (WE2), $p_k\leq e^{\varepsilon(k-\ell)}p_\ell$ whenever $\ell\leq k$, hence for all $n\leq 0$ we have

$$\begin{aligned}|t_n|&\leq e^{\frac{\chi}{2}n}|t_0|+\varepsilon(p_{n+1}+e^{-\frac{\chi}{2}}p_{n+2}+\cdots+e^{\frac{\chi}{2}(n+1)}p_0)\\&\leq e^{\frac{\chi}{2}n}p_0+\varepsilon e^{\varepsilon}(p_n+e^{-\frac{\chi}{2}}p_{n+1}+\cdots+e^{\frac{\chi}{2}(n+1)}p_{-1})\\&\leq\left[e^{(\frac{\chi}{2}-\varepsilon)n}+\varepsilon e^{\varepsilon}\sum_{i=n+1}^{0}e^{(\frac{\chi}{2}-\varepsilon)i}\right]p_n<\tfrac{1}{2}p_n,\end{aligned}$$

  since $e^{(\frac{\chi}{2}-\varepsilon)n}<\frac{1}{4}$ for n small enough and $\varepsilon e^{\varepsilon}\sum_{i=n+1}^{0}e^{(\frac{\chi}{2}-\varepsilon)i}<\frac{2\varepsilon}{1-e^{-\chi/4}}<\frac{1}{4}$.
- If n is small enough then $\vartheta_n[W^u]\subset\Psi_{\widehat{x}_n}(R[p_n])$: fix $\widehat{y}=(y_n)_{n\in\mathbb{Z}}\in V^u\cap W^u$, and take $\widehat{z}=(z_n)_{n\in\mathbb{Z}}\in W^u$. In part (3) we proved that $d(y_n,z_n)\leq 2q_0e^{\frac{\chi}{2}n}$ for all $n\leq 0$, thus

$$|\Psi_{\widehat{x}_n}^{-1}(y_n)-\Psi_{\widehat{x}_n}^{-1}(z_n)|\leq 2q_0u(\widehat{x}_n)e^{\frac{\chi}{2}n}\leq 2q_0p_n^{-1}e^{\frac{\chi}{2}n}\leq 2q_0p_0^{-1}e^{(\frac{\chi}{2}-\varepsilon)n},$$

  since $u(\widehat{x}_n)\leq Q_\varepsilon(\widehat{x}_n)^{-1}\leq p_n^{-1}$ and $p_0\leq e^{-\varepsilon n}p_n$. If n is small enough then $\Psi_{\widehat{x}_n}^{-1}(y_n)\in R[\frac{1}{2}p_n]$ and $2q_0p_0^{-1}e^{(\frac{\chi}{2}-\varepsilon)n}<\frac{1}{2}p_0e^{\varepsilon n}\leq\frac{1}{2}p_n$, hence $\Psi_{\widehat{x}_n}^{-1}(z_n)\in R[p_n]$.

Fix $n\leq 0$ s.t. both items above hold for all $N\leq n$. By the definition of unstable sets, we have $\widehat{f}^n(V^u),\widehat{f}^n(W^u)\subset V^u[\{\Psi_{\widehat{x}_k}^{p_k}\}_{k\leq n}]$. By Lemma 2.10(2), and since the inverse branches $g_{\widehat{x}_k}$ send intervals onto intervals, $\exists\alpha,\beta,\alpha',\beta'\in\mathbb{R}$ s.t.:

$$\widehat{f}^n(V^u)=\left\{(x_k)_{k\in\mathbb{Z}}\in\widehat{M}:\begin{array}{l}\text{for } k\leq 0 \text{ we can write } x_k=\Psi_{\widehat{x}_{n+k}}(t_k)\text{ with}\\ t_0\in[\alpha,\beta]\text{ and } t_{k-1}=G_{\widehat{x}_{n+k},\widehat{x}_{n+k-1}}(t_k)\end{array}\right\}$$

$$\widehat{f}^n(W^u)=\left\{(x_k)_{k\in\mathbb{Z}}\in\widehat{M}:\begin{array}{l}\text{for } k\leq 0 \text{ we can write } x_k=\Psi_{\widehat{x}_{n+k}}(t_k')\text{ with}\\ t_0'\in[\alpha',\beta']\text{ and } t_{k-1}'=G_{\widehat{x}_{n+k},\widehat{x}_{n+k-1}}(t_k')\end{array}\right\}.$$

To prove that $V^u\subset W^u$ it is enough to show that $[\alpha,\beta]\subset[\alpha',\beta']$: if this happens then $\widehat{f}^n(V^u)\subset\widehat{f}^n(W^u)$ and thus $V^u\subset W^u$. By contradiction, assume that $[\alpha,\beta]\not\subset[\alpha',\beta']$, then either $\alpha<\alpha'$ and/or $\beta'<\beta$. By symmetry, we may assume $\alpha<\alpha'$. Thus $A'=\Psi_{\widehat{x}_n}(\alpha')$ belongs to the interior of the interval with endpoints $A=\Psi_{\widehat{x}_n}(\alpha)$ and $B=\Psi_{\widehat{x}_n}(\beta)$. Since f is continuous inside the ranges of ε–charts, $f^{-n}(A')$ belongs to the interior of the interval with endpoints $f^{-n}(A)$ and $f^{-n}(B)$. This latter interval is $\Psi_{\widehat{x}}(R[p])$. But $f^{-n}(A')$ is one of the endpoints of $\Psi_{\widehat{x}}(R[q])$, therefore $f^{-n}(A')\in\Psi_{\widehat{x}}(R[p])$ iff $q<p$, which contradicts our assumption. The proof is complete.

(5) Write $v = \Psi_{x_0}^{p_0}$, $w = \Psi_{x_1}^{p_1}$ and $V^u = V^u[\underline{v}^-]$ where $\underline{v}^- = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\le 0}$ is a negative weak ε–gpo. Let $V^s = \vartheta^{-1}[x]$ be a stable set at w. We want to show that $\widehat{f}(V^u) \cap V^s$ consists of a single element. By Lemma 2.10(2),

$$V^u = \{\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in \widehat{M} : x_0 \in \Psi_{\widehat{x}_0}(R[p_0]) \text{ and } x_{n-1} = g_{\widehat{x}_n}(x_n), \forall n \le 0\}$$

hence

$$\widehat{f}(V^u) = \{\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in \widehat{M} : x_{-1} \in \Psi_{\widehat{x}_0}(R[p_0]) \text{ and } x_{n-1} = g_{\widehat{x}_n}(x_n), \forall n \le -1\}.$$

Any $\widehat{x} = (x_n)_{n\in\mathbb{Z}} \in \widehat{f}(V^u) \cap V^s$ must satisfy $x_n = f^n(x)$ for $n \ge 0$ and $x_{n-1} = g_{\widehat{x}_n}(x_n)$ for $n \le -1$, therefore $\widehat{x}$ is uniquely defined by the choice of x_{-1}. By Lemma 2.8(2), there is a unique $x_{-1} \in \Psi_{\widehat{x}_0}(R[p_0])$ s.t. $f(x_{-1}) = x$. □

SHADOWING: A weak ε–gpo $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$ is said to *shadow* a point $\widehat{x} \in \widehat{M}$ if $\vartheta_n[\widehat{x}] \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \in \mathbb{Z}$.

Lemma 2.12. *Every weak ε–gpo shadows a unique element of $\widehat{M}$.*

Proof. Let $\underline{v} = \{v_n\}_{n\in\mathbb{Z}}$ be a weak ε–gpo, and let $V^s = V^s[\{v_n\}_{n\ge 0}]$, $V^u = V^u[\{v_n\}_{n\le 0}]$. By the definition of V^s and V^u, any point $\widehat{x} \in \widehat{M}$ shadowed by $\underline{v}$ belongs to $V^s \cap V^u$. By Proposition 2.11(1), this intersection consists of a single element of $\widehat{M}$. □

3. Coarse graining

In this section, we construct a countable set of ε–charts whose set of (strong) ε–gpo's shadows all relevant orbits of $\widehat{f}$.

Theorem 3.1. *For all $\varepsilon > 0$ sufficiently small, there exists a countable family $\mathscr{A}$ of ε–charts with the following properties:*

(1) DISCRETENESS*: For all $t > 0$, the set $\{\Psi_{\widehat{x}}^p \in \mathscr{A} : p > t\}$ is finite.*
(2) SUFFICIENCY: *If $\widehat{x} \in \mathrm{NUE}_\chi^*$ then there is an ε–gpo $\underline{v} \in \mathscr{A}^{\mathbb{Z}}$ that shadows $\widehat{x}$.*
(3) RELEVANCE: *For all $v \in \mathscr{A}$ there is an ε–gpo $\underline{v} \in \mathscr{A}^{\mathbb{Z}}$ with $v_0 = v$ that shadows a point in NUE_χ^*.*

Parts (1) and (3) are essential to prove the inverse theorem (Theorem 4.1). Part (2) and Lemma 2.2 imply that if μ is f–adapted and χ–expanding then $\widehat{\mu}$–a.e. $\widehat{x} \in \widehat{M}$ is shadowed by an ε–gpo whose vertices belong to $\mathscr{A}$.

Proof. When M is compact and f is a diffeomorphism, the above statement is consequence of Propositions 3.5, 4.5 and Lemmas 4.6, 4.7 of [53]. When M is compact with boundary and f is a local diffeomorphism with bounded derivatives, this is [39, Prop. 4.3]. When f is a surface map with discontinuities and possibly unbounded derivatives, this is [38, Theorem 5.1]. We follow the strategy of [38], adapted to our context.

For $t > 0$, let $M_t = \{x \in M : d(x, \mathscr{S}) \ge t\}$. Since M has finite diameter (we are even assuming it is smaller than one), each M_t is compact. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Fix a countable open cover $\mathscr{P} = \{D_i\}_{i\in\mathbb{N}_0}$ of $M\backslash\mathscr{S}$ s.t.:

- $D_i := D_{z_i} = B(z_i, 2\mathfrak{r}(z_i))$ for some $z_i \in M$.
- For every $t > 0$, $\{D \in \mathscr{P} : D \cap M_t \ne \emptyset\}$ is finite.

Let $X := M^3 \times (0,\infty)^3 \times (0,1]$. For $\widehat{x} \in \mathrm{NUE}_\chi^*$, let $\Gamma(\widehat{x}) = (\underline{\widehat{x}}, \underline{u}, \underline{Q}) \in X$ where:

$$\underline{\widehat{x}} = (\vartheta_{-1}[\widehat{x}], \vartheta_0[\widehat{x}], \vartheta_1[\widehat{x}]),\ \underline{u} = (u(\widehat{f}^{-1}(\widehat{x})), u(\widehat{x}), u(\widehat{f}(\widehat{x}))),\ \underline{Q} = Q_\varepsilon(\widehat{x}).$$

Let $Y = \{\Gamma(\widehat{x}) : \widehat{x} \in \mathrm{NUE}_\chi^*\}$. We want to construct a countable dense subset of Y. Since the maps $\widehat{x} \mapsto u(\widehat{x})$, $Q_\varepsilon(\widehat{x})$ are not necessarily continuous, we apply a precompactness argument. For vectors $\underline{k} = (k_{-1}, k_0, k_1)$, $\underline{\ell} = (\ell_{-1}, \ell_0, \ell_1)$, $\underline{a} = (a_{-1}, a_0, a_1) \in \mathbb{N}_0^3$ and $m \in \mathbb{N}_0$, define

$$Y_{\underline{k},\underline{\ell},\underline{a},m} := \left\{ \Gamma(\widehat{x}) \in Y : \begin{array}{cc} e^{-k_i-1} \le d(\vartheta_i[\widehat{x}], \mathscr{S}) < e^{-k_i}, & -1 \le i \le 1 \\ e^{\ell_i} \le u(\widehat{f^i}(\widehat{x})) < e^{\ell_i+1}, & -1 \le i \le 1 \\ \vartheta_i[\widehat{x}] \in D_{a_i}, & -1 \le i \le 1 \\ e^{-m-1} \le Q_\varepsilon(\widehat{x}) < e^{-m} & \end{array} \right\}.$$

CLAIM 1: $Y = \bigcup_{\substack{\underline{k},\underline{\ell},\underline{a}\in\mathbb{N}_0^3 \\ m\in\mathbb{N}_0}} Y_{\underline{k},\underline{\ell},\underline{a},m}$, and each $Y_{\underline{k},\underline{\ell},\underline{a},m}$ is precompact in X.

Proof of Claim 1. The first statement is clear, so we focus on the second. Fix $\underline{k}, \underline{\ell}, \underline{a} \in \mathbb{N}_0^3$, $m \in \mathbb{N}_0$, and take $\Gamma(\widehat{x}) \in Y_{\underline{k},\underline{\ell},\underline{a},m}$. Then

$$\underline{\widehat{x}} \in M_{e^{-k_{-1}-1}} \times M_{e^{-k_0-1}} \times M_{e^{-k_1-1}},$$

a precompact subset of M^3. For $|i| \le 1$ we have $1 \le u(\widehat{f^i}(\widehat{x})) < e^{\ell_i+1}$, hence $\underline{u}$ belongs to a compact subset of $(0,\infty)^3$. Also $Q_\varepsilon(\widehat{x}) \in [e^{-m-1}, 1]$, therefore $\underline{Q}$ belongs to a compact subinterval of $(0, 1]$. The product of precompact sets is precompact, thus the claim is proved.

Let $j \ge 0$. By Claim 1, there is a finite set $Y_{\underline{k},\underline{\ell},\underline{a},m}(j) \subset Y_{\underline{k},\underline{\ell},\underline{a},m}$ s.t. for every $\Gamma(\widehat{x}) \in Y_{\underline{k},\underline{\ell},\underline{a},m}$ there exists $\Gamma(\widehat{y}) \in Y_{\underline{k},\underline{\ell},\underline{a},m}(j)$ s.t.:

(a) $d(\vartheta_i[\widehat{x}], \vartheta_i[\widehat{y}]) + |u(\widehat{f^i}(\widehat{x}))^{-1} - u(\widehat{f^i}(\widehat{y}))^{-1}| < e^{-8(j+2)}$ for $|i| \le 1$.
(b) $\frac{Q_\varepsilon(\widehat{x})}{Q_\varepsilon(\widehat{y})} = e^{\pm\frac{\varepsilon}{3}}$.

Remind that $I_\varepsilon := \{e^{-\frac{1}{3}\varepsilon n} : n \ge 0\}$.

THE ALPHABET $\mathscr{A}$: Let $\mathscr{A}$ be the countable family of $\Psi_{\widehat{x}}^{p}$ s.t.:

(CG1) $\Gamma(\widehat{x}) \in Y_{\underline{k},\underline{\ell},\underline{a},m}(j)$ for some $(\underline{k}, \underline{\ell}, \underline{a}, m, j) \in \mathbb{N}_0^3 \times \mathbb{N}_0^3 \times \mathbb{N}_0^3 \times \mathbb{N}_0 \times \mathbb{N}_0$.
(CG2) $p \in I_\varepsilon$, $p \le \delta_\varepsilon Q_\varepsilon(\widehat{x})$, and $e^{-j-2} \le p \le e^{-j+2}$.

Proof of discreteness. By the proof of Lemma 2.1 and assumption (A2),

$$u(\widehat{f}(\widehat{x}))^2 = 1 + e^{2\chi}|df_{\vartheta_0[\widehat{x}]}|^{-2}u(\widehat{x})^2 \le 2e^{2\chi}\rho(\widehat{x})^{-2a}u(\widehat{x})^2 < e^{2\chi+2}\rho(\widehat{x})^{-2a}u(\widehat{x})^2$$

hence $u(\widehat{f}(\widehat{x})) < e^{\chi+1}\rho(\widehat{x})^{-a}u(\widehat{x})$. We will use this estimate below.

Fix $0 < t < 1$, and let $\Psi_{\widehat{x}}^{p} \in \mathscr{A}$ with $p > t$. Start noting that $\rho(\widehat{x}) > \rho(\widehat{x})^{2a} > Q_\varepsilon(\widehat{x}) > p > t$. If $\Gamma(\widehat{x}) \in Y_{\underline{k},\underline{\ell},\underline{a},m}(j)$ then:

- Finiteness of $\underline{k}$: for $|i| \le 1$, $e^{-k_i} > d(\vartheta_i[\widehat{x}], \mathscr{S}) \ge \rho(\widehat{x}) > t$, hence $k_i < |\log t|$.
- Finiteness of $\underline{\ell}$: for $i = -1, 0$, $e^{\ell_i} \le u(\widehat{f^i}(\widehat{x})) < Q_\varepsilon(\widehat{x})^{-1} < t^{-1}$, hence $\ell_i < |\log t|$. For $i = 1$, the estimate in the beginning of the proof implies $e^{\ell_1} \le u(\widehat{f}(\widehat{x})) < e^{\chi+1}\rho(\widehat{x})^{-a}u(\widehat{x}) < e^{\chi+1}t^{-(a+1)}$, hence $\ell_1 < \chi + 1 + (a+1)|\log t| =: T_t$, which is bigger than $|\log t|$.
- Finiteness of $\underline{a}$: for $|i| \le 1$, $\vartheta_i[\widehat{x}] \in D_{a_i} \cap M_t$ hence D_{a_i} belongs to the finite set $\{D \in \mathscr{P} : D \cap M_t \neq \emptyset\}$.
- Finiteness of m: we have $e^{-m} > Q_\varepsilon(\widehat{x}) > t$, hence $m < |\log t|$.
- Finiteness of j: $t < p \le e^{-j+2}$, hence $j \le |\log t| + 2$.
- Finiteness of p: $\#\{p \in I_\varepsilon : p > t\} \le \#(I_\varepsilon \cap (t, 1])$ is finite.

The first five items above give that, for $\underline{a} \in \mathbb{N}_0^3$ and $t > 0$,

$$\#\left\{\Gamma(\widehat{x}) : \begin{array}{c} \Psi_{\widehat{x}}^{p} \in \mathscr{A} \text{ s.t. } p > t \text{ and} \\ \vartheta_i[\widehat{x}] \in D_{a_i} \text{ for } |i| \le 1 \end{array}\right\} \le \sum_{j=0}^{\lceil|\log t|\rceil+2} \sum_{k_i,\ell_i,m=0}^{T_t} \#Y_{\underline{k},\underline{\ell},\underline{a},m}(j)$$

is the finite sum of finite terms, hence finite. Together with the last item above and the choice of $\mathscr{P}$, we obtain that

$$\#\left\{\Psi_{\widehat{x}}^p \in \mathscr{A} : p > t\right\} \leq \sum_{j=0}^{\lceil |\log t|\rceil+2} \sum_{k_i,\ell_i,m=0}^{T_t} \#Y_{\underline{k},\underline{\ell},\underline{a},m}(j)$$
$$\times (\#\{D \in \mathscr{P} : D \cap M_t \neq \emptyset\})^3 \times (\#(I_\varepsilon \cap (t,1]))$$

is finite. This proves the discreteness property of $\mathscr{A}$.

Proof of sufficiency. Let $\widehat{x} \in \text{NUE}_\chi^*$. Take $(k_i)_{i\in\mathbb{Z}}, (\ell_i)_{i\in\mathbb{Z}}, (m_i)_{i\in\mathbb{Z}}, (a_i)_{i\in\mathbb{Z}}, (j_i)_{i\in\mathbb{Z}}$ s.t.:

$$d(\vartheta_i[\widehat{x}], \mathscr{S}) \in [e^{-k_i-1}, e^{-k_i}),\ u(\widehat{f}^i(\widehat{x})) \in [e^{\ell_i}, e^{\ell_i+1}),\ Q_\varepsilon(\widehat{f}^i(\widehat{x})) \in [e^{-m_i-1}, e^{-m_i}),$$
$$\vartheta_i[\widehat{x}] \in D_{a_i},\ q_\varepsilon(\widehat{f}^i(\widehat{x})) \in [e^{-j_i-1}, e^{-j_i+1}).$$

For $n \in \mathbb{Z}$, define

$$\underline{k}^{(n)} = (k_{n-1}, k_n, k_{n+1}),\ \underline{\ell}^{(n)} = (\ell_{n-1}, \ell_n, \ell_{n+1}),\ \underline{a}^{(n)} = (a_{n-1}, a_n, a_{n+1}).$$

Then $\Gamma(\widehat{f}^n(\widehat{x})) \in Y_{\underline{k}^{(n)},\underline{\ell}^{(n)},\underline{a}^{(n)},m_n}$. Take $\Gamma(\widehat{x}_n) \in Y_{\underline{k}^{(n)},\underline{\ell}^{(n)},\underline{a}^{(n)},m_n}(j_n)$ s.t.:

(a$_n$) $d(\vartheta_i[\widehat{f}^n(\widehat{x})], \vartheta_i[\widehat{x}_n]) + |u(\widehat{f}^i(\widehat{f}^n(\widehat{x})))^{-1} - u(\widehat{f}^i(\widehat{x}_n))^{-1}| < e^{-8(j_n+2)}$ for $|i| \leq 1$.
(b$_n$) $\frac{Q_\varepsilon(\widehat{f}^n(\widehat{x}))}{Q_\varepsilon(\widehat{x}_n)} = e^{\pm\frac{\varepsilon}{3}}$.

Define $p_n = \delta_\varepsilon \min\{e^{\varepsilon|k|} Q_\varepsilon(\widehat{x}_{n+k}) : k \leq 0\}$. We claim that $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$ is an ε–gpo in $\mathscr{A}^{\mathbb{Z}}$ that shadows $\widehat{x}$.

Claim 2: $\Psi_{\widehat{x}_n}^{p_n} \in \mathscr{A}$ for all $n \in \mathbb{Z}$.

(CG1) By definition, $\Gamma(\widehat{x}_n) \in Y_{\underline{k}^{(n)},\underline{\ell}^{(n)},\underline{a}^{(n)},m_n}(j_n)$.
(CG2) By (b$_n$), $\min\{e^{\varepsilon|k|} Q_\varepsilon(\widehat{x}_{n+k}) : k \leq 0\} = e^{\pm\frac{\varepsilon}{3}} \min\{e^{\varepsilon|k|} Q_\varepsilon(\widehat{f}^{n+k}(\widehat{x})) : k \leq 0\}$ hence p_n is well-defined and satisfies $p_n = e^{\pm\frac{\varepsilon}{3}} q_\varepsilon(\widehat{f}^n(\widehat{x}))$. Therefore $p_n \in I_\varepsilon$, $p_n \leq \delta_\varepsilon Q_\varepsilon(\widehat{x}_n)$, and $p_n \in [e^{-j_n-2}, e^{-j_n+2})$.

Claim 3: $\Psi_{\widehat{x}_n}^{p_n} \overset{\varepsilon}{\leftarrow} \Psi_{\widehat{x}_{n+1}}^{p_{n+1}}$ for all $n \in \mathbb{Z}$.

(E1) By (a$_{n+1}$) with $i = -1$ and (a$_n$) with $i = 0$,

$$\begin{aligned}
&d(\vartheta[\widehat{f}^{-1}(\widehat{x}_{n+1})], \vartheta[\widehat{x}_n]) + |u(\widehat{f}^{-1}(\widehat{x}_{n+1}))^{-1} - u(\widehat{x}_n)^{-1}|\\
&\leq d(\vartheta_{-1}[\widehat{x}_{n+1}], \vartheta_{-1}[\widehat{f}^{n+1}(\widehat{x})]) + |u(\widehat{f}^{-1}(\widehat{x}_{n+1}))^{-1} - u(\widehat{f}^n(\widehat{x}))^{-1}|\\
&\quad + d(\vartheta_0[\widehat{f}^n(\widehat{x})], \vartheta_0[\widehat{x}_n]) + |u(\widehat{f}^n(\widehat{x}))^{-1} - u(\widehat{x}_n)^{-1}|\\
&< e^{-8(j_{n+1}+2)} + e^{-8(j_n+2)}.
\end{aligned}$$

Note that

$$\begin{aligned}
&e^{-8(j_{n+1}+2)} + e^{-8(j_n+2)} \leq e^{-8}\left(q_\varepsilon(\widehat{f}^{n+1}(\widehat{x}))^8 + q_\varepsilon(\widehat{f}^n(\widehat{x}))^8\right)\\
&\overset{!}{\leq} e^{-8}(1 + e^{8\varepsilon}) q_\varepsilon(\widehat{f}^n(\widehat{x}))^8 \leq e^{-8+\frac{8\varepsilon}{3}}(1 + e^{8\varepsilon}) p_n^8 \overset{!!}{<} p_n^8,
\end{aligned}$$

where in $\overset{!}{\leq}$ we used Lemma 2.9(2) and in $\overset{!!}{<}$ we used that $e^{-8+\frac{8\varepsilon}{3}}(1 + e^{8\varepsilon}) < 1$ when $\varepsilon > 0$ is sufficiently small. Therefore $\Psi_{\widehat{f}^{-1}(\widehat{x}_{n+1})}^{p_n} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_n}^{p_n}$.

(E2) We will use the inequality $e^{-8(j_{n+1}+2)} + e^{-8(j_n+2)} < p_n^8$ proved above.

(E2.1) As remarked before, it follows directly from condition (E1).

(E2.2) By (a$_n$) with $i=1$ and (a$_{n+1}$) with $i=0$ we have $|u(\widehat{f}^{n+1}(\widehat{x}))^{-1}-u(\widehat{f}(\widehat{x}_n))^{-1}|<e^{-8(j_n+2)}<p_n^8$ and $|u(\widehat{f}^{n+1}(\widehat{x}))^{-1}-u(\widehat{x}_{n+1})^{-1}|<e^{-8(j_{n+1}+2)}<p_{n+1}^8$. Proceeding as in the proof of Proposition 2.5(1), this first inequality implies that $\frac{u(\widehat{f}^{n+1}(\widehat{x}))}{u(\widehat{f}(\widehat{x}_n))}=e^{\pm p_n^6}$ and the second implies that $\frac{u(\widehat{f}^{n+1}(\widehat{x}))}{u(\widehat{x}_{n+1})}=e^{\pm p_{n+1}^6}$. Since $p_n^6\ll\frac{p_n}{2}$ and $p_{n+1}^6\leq e^{6\varepsilon}p_n^6\ll\frac{p_n}{2}$, it follows that $\frac{u(\widehat{f}(\widehat{x}_n))}{u(\widehat{x}_{n+1})}=e^{\pm p_n}$.

(E2.3) The definition of p_n guarantees that $p_{n+1}=\min\{e^{\varepsilon}p_n,\delta_\varepsilon Q_\varepsilon(\widehat{x}_{n+1})\}$.

Claim 4: $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$ shadows $\widehat{x}$.

By (a$_n$) with $i=0$, we have $\Psi_{\widehat{f}^n(\widehat{x})}^{p_n}\overset{\varepsilon}{\approx}\Psi_{\widehat{x}_n}^{p_n}$, hence by Proposition 2.5(2) we have $\vartheta_n[\widehat{x}]=\vartheta[\widehat{f}^n(\widehat{x})]=\Psi_{\widehat{f}^n(\widehat{x})}(0)\in\Psi_{\widehat{x}_n}(R[p_n])$, therefore $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$ shadows $\widehat{x}$.

This concludes the proof of sufficiency. Note that if $\widehat{x}\in\mathrm{NUE}_\chi^\#$ then the ε–gpo constructed above belongs to $\Sigma^\#$. This observation will be used in the proof of Proposition 3.2 below.

Proof of relevance. The family $\mathscr{A}$ might not a priori satisfy the relevance condition, but we can easily reduce it to a sub-alphabet $\mathscr{A}'$ satisfying (1)–(3) as follows. Call $v\in\mathscr{A}$ relevant if there is $\underline{v}=\{v_n\}_{n\in\mathbb{Z}}\in\mathscr{A}^{\mathbb{Z}}$ with $v_0=v$ s.t. $\underline{v}$ shadows a point in NUE_χ^*. When this happens then every v_n is relevant (since NUE_χ^* is $\widehat{f}$–invariant). Therefore $\mathscr{A}'=\{v\in\mathscr{A}:v \text{ is relevant}\}$ is discrete because $\mathscr{A}'\subset\mathscr{A}$, and it is sufficient and relevant by definition. □

Let Σ be the TMS associated to the graph with vertex set $\mathscr{A}$ given by Theorem 3.1 and edges $v\overset{\varepsilon}{\leftarrow}w$. An element of Σ is an ε–gpo, hence we define $\pi:\Sigma\to\widehat{M}$ by

$$\{\pi[\{v_n\}_{n\in\mathbb{Z}}]\}:=V^s[\{v_n\}_{n\geq 0}]\cap V^u[\{v_n\}_{n\leq 0}].$$

Here are the main properties of the triple (Σ,σ,π).

Proposition 3.2. *The following holds for all $\varepsilon>0$ small enough.*

(1) $\pi:\Sigma\to\widehat{M}$ *is Hölder continuous.*
(2) $\pi\circ\sigma=\widehat{f}\circ\pi$.
(3) $\pi[\Sigma^\#]\supset\mathrm{NUE}_\chi^\#$.

Part (1) follows from Proposition 2.11(3), part (2) follows from Proposition 2.11(2), and part (3) is a direct consequence of the observation in the end of the proof of Theorem 3.1(2).

Remark 3.3. It is important noticing the difference between our (Σ,σ) and those constructed in [53,39,38]: while the later ones have finite ingoing and outgoing degrees (thus Σ is locally compact), our symbolic space does not necessarily satisfy this. The reason is that condition (E2.3) does not imply a lower bound on p_{n+1}. This non-finiteness property also holds in Hofbauer towers.

In general (Σ,σ,π) does *not* satisfy Theorem 1.1, since π might be infinite-to-one. We use π to induce a locally finite cover of $\mathrm{NUE}_\chi^\#$, which will then be refined to a partition of $\mathrm{NUE}_\chi^\#$ that will lead to the proof of Theorem 1.1.

4. The inverse problem

Our goal now is to analyze when π loses injectivity. More specifically, given that $\pi(\underline{v})=\pi(\underline{w})$ we want to compare v_n with w_n and show that one is defined by the other "up to bounded error". We do this under the additional assumption that $\underline{v},\underline{w}\in\Sigma^\#$. Remind that $\Sigma^\#$ is the *recurrent set* of Σ:

$$\Sigma^\#:=\left\{\underline{v}\in\Sigma:\exists v,w\in V \text{ s.t. } \begin{array}{l} v_n=v \text{ for infinitely many } n>0\\ v_n=w \text{ for infinitely many } n<0\end{array}\right\}.$$

The main result of this section is the following theorem.

Theorem 4.1 *(Inverse theorem). The following holds for $\varepsilon > 0$ small enough. If $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}, \{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}} \in \Sigma^\#$ satisfy $\pi[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}] = \pi[\{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}}]$ then for all $n \in \mathbb{Z}$:*

(1) $d(\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n]) \le 2\max\{p_n, q_n\}$.
(2) $\frac{u(\widehat{x}_n)}{u(\widehat{y}_n)} = e^{\pm 2\sqrt{\varepsilon}}$.
(3) $\frac{Q_\varepsilon(\widehat{x}_n)}{Q_\varepsilon(\widehat{y}_n)} = e^{\pm\sqrt[3]{\varepsilon}}$.
(4) $\frac{p_n}{q_n} = e^{\pm\sqrt[3]{\varepsilon}}$.
(5) $(\Psi_{\widehat{y}_n}^{-1} \circ \Psi_{\widehat{x}_n})(t) = t + \Delta_n(t) + \delta_n$ *for* $t \in R[10Q_\varepsilon(\widehat{x}_n)]$, *where* $\delta_n \in \mathbb{R}$ *with* $|\delta_n| < 3q_n$ *and* $\Delta_n : R[10Q_\varepsilon(\widehat{x}_n)] \to \mathbb{R}$ *with* $\Delta_n(0) = 0$ *and* $\|d\Delta_n\|_0 < 4\sqrt{\varepsilon}$.

The substantial differences of the above theorem from [53, Thm 5.2] rely on part (5): in our case we can only obtain estimates inside the smaller rectangle $R[10Q_\varepsilon(\widehat{x}_n)]$, and our estimate on δ_n is slightly weaker. This latter fact is the reason we introduced weak edges, as it will be clear in section 5 (see Remark 5.1). Part (1) is proved similarly to [53, Prop. 5.3], as follows. Write $\pi[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}] = \pi[\{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}}] = \widehat{x}$. For each $n \in \mathbb{Z}$, $\vartheta_n[\widehat{x}] = \Psi_{\widehat{x}_n}(t)$ for some $t \in R[p_n]$. Since $\Psi_{\widehat{x}_n}(0) = \vartheta[\widehat{x}_n]$ and $\Psi_{\widehat{x}_n}$ is 1–Lipschitz, $d(\vartheta_n[\widehat{x}], \vartheta[\widehat{x}_n]) \le p_n$. Similarly $d(\vartheta_n[\widehat{x}], \vartheta[\widehat{y}_n]) \le q_n$, and so $d(\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n]) \le p_n + q_n$, which is better than stated in part (1).

Before proceeding to the other parts, we discuss two consequences of part (1). The first one is that for $\varepsilon > 0$ small enough it holds:

$$\frac{1}{2} \le \frac{d(\vartheta_i[\widehat{x}_n], \mathscr{S})^a}{d(\vartheta_i[\widehat{y}_n], \mathscr{S})^a} \le 2, \ \forall |i| \le 1. \tag{4.1}$$

Take $i = 0$, and start noting that $d(\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n]) \le p_n + q_n < \varepsilon[d(\vartheta[\widehat{x}_n], \mathscr{S}) + d(\vartheta[\widehat{y}_n], \mathscr{S})]$ hence $d(\vartheta[\widehat{x}_n], \mathscr{S}) = d(\vartheta[\widehat{y}_n], \mathscr{S}) \pm d(\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n]) = d(\vartheta[\widehat{y}_n], \mathscr{S}) \pm \varepsilon[d(\vartheta[\widehat{x}_n], \mathscr{S}) + d(\vartheta[\widehat{y}_n], \mathscr{S})]$, thus $\frac{1-\varepsilon}{1+\varepsilon} \le \frac{d(\vartheta[\widehat{x}_n], \mathscr{S})}{d(\vartheta[\widehat{y}_n], \mathscr{S})} \le \frac{1+\varepsilon}{1-\varepsilon}$. If $\varepsilon > 0$ is small enough then $\frac{1+\varepsilon}{1-\varepsilon} < e^{3\varepsilon}$, and so

$$\frac{d(\vartheta[\widehat{x}_n], \mathscr{S})}{d(\vartheta[\widehat{y}_n], \mathscr{S})} = e^{\pm 3\varepsilon}. \tag{4.2}$$

Now take $i = -1$. By (E1), $d(\vartheta_{-1}[\widehat{x}_n], \vartheta[\widehat{x}_{n-1}]) \le p_{n-1}^8 < \varepsilon d(\vartheta[\widehat{x}_{n-1}], \mathscr{S})$ and so $d(\vartheta_{-1}[\widehat{x}_n], \mathscr{S}) = d(\vartheta[\widehat{x}_{n-1}], \mathscr{S}) \pm d(\vartheta_{-1}[\widehat{x}_n], \vartheta[\widehat{x}_{n-1}]) = (1 \pm \varepsilon) d(\vartheta[\widehat{x}_{n-1}], \mathscr{S})$. If $\varepsilon > 0$ is small enough then $e^{-2\varepsilon} < 1 - \varepsilon < 1 + \varepsilon < e^{2\varepsilon}$, therefore $\frac{d(\vartheta_{-1}[\widehat{x}_n], \mathscr{S})}{d(\vartheta[\widehat{x}_{n-1}], \mathscr{S})} = e^{\pm 2\varepsilon}$. Similarly $\frac{d(\vartheta_{-1}[\widehat{y}_n], \mathscr{S})}{d(\vartheta[\widehat{y}_{n-1}], \mathscr{S})} = e^{\pm 2\varepsilon}$, hence (4.2) above implies that

$$\frac{d(\vartheta_{-1}[\widehat{x}_n], \mathscr{S})}{d(\vartheta_{-1}[\widehat{y}_n], \mathscr{S})} = e^{\pm 7\varepsilon}. \tag{4.3}$$

It remains to take $i = 1$. By (E2.1), $d(\vartheta_1[\widehat{x}_n], \vartheta[\widehat{x}_{n+1}]) < p_n < \varepsilon d(\vartheta_1[\widehat{x}_n], \mathscr{S})$ and so $d(\vartheta[\widehat{x}_{n+1}], \mathscr{S}) = (1 \pm \varepsilon) d(\vartheta_1[\widehat{x}_n], \mathscr{S})$. Now proceed as in the case $i = -1$ to get

$$\frac{d(\vartheta_1[\widehat{x}_n], \mathscr{S})}{d(\vartheta_1[\widehat{y}_n], \mathscr{S})} = e^{\pm 7\varepsilon}. \tag{4.4}$$

It is clear that (4.1) follows from (4.2), (4.3), (4.4).

The second consequence of part (1) is that $\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n] \in D_{\vartheta[\widehat{x}_n]} \cap D_{\vartheta[\widehat{y}_n]}$. To see this, note that (4.2) and (4.4) imply $\frac{1}{2} \le \frac{\min\{d(\vartheta[\widehat{x}_n], \mathscr{S})^a, d(\vartheta_1[\widehat{x}_n], \mathscr{S})^a\}}{\min\{d(\vartheta[\widehat{y}_n], \mathscr{S})^a, d(\vartheta_1[\widehat{y}_n], \mathscr{S})^a\}} \le 2$ and so

$$\begin{aligned} d(\vartheta[\widehat{x}_n], \vartheta[\widehat{y}_n]) &< \varepsilon[\min\{d(\vartheta[\widehat{x}_n], \mathscr{S})^a, d(\vartheta_1[\widehat{x}_n], \mathscr{S})^a\} + \min\{d(\vartheta[\widehat{y}_n], \mathscr{S})^a, d(\vartheta_1[\widehat{y}_n], \mathscr{S})^a\}] \\ &< 3\varepsilon \min\{d(\vartheta[\widehat{y}_n], \mathscr{S})^a, d(\vartheta_1[\widehat{y}_n], \mathscr{S})^a\} < \mathfrak{r}(\vartheta[\widehat{y}_n]) \end{aligned}$$

and thus $\vartheta[\widehat{x}_n] \in D_{\vartheta[\widehat{y}_n]}$. Similarly $\vartheta[\widehat{y}_n] \in D_{\vartheta[\widehat{x}_n]}$. As a consequence, we can apply assumptions (A1)–(A3) with respect to either $\vartheta[\widehat{x}_n]$ or $\vartheta[\widehat{y}_n]$.

4.1. Control of $u(\widehat{x}_n)$

We now make use of the hyperbolicity of $\widehat{f}$ to show that u improves along an ε–gpo.

Proposition 4.2. *The following holds for all $\varepsilon > 0$ small enough. If $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}, \{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}} \in \Sigma^\#$ satisfy $\pi[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}] = \pi[\{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}}]$ then $\frac{u(\widehat{x}_n)}{u(\widehat{y}_n)} = e^{\pm 2\sqrt{\varepsilon}}$, $\forall n \in \mathbb{Z}$.*

Proof. When M is compact and f is a $C^{1+\beta}$ diffeomorphism, this follows from Lemma 7.2 and Proposition 7.3 of [53]. We employ similar methods. Let $\underline{v} = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$, $\underline{w} = \{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}}$, and $\pi[\underline{v}] = \pi[\underline{w}] = \widehat{x}$.

CLAIM 1 (IMPROVEMENT LEMMA): The following holds for all $\varepsilon > 0$ small enough. For $\xi \geq \sqrt{\varepsilon}$, if $\frac{u(\widehat{f}^n(\widehat{x}))}{u(\widehat{x}_n)} = e^{\pm\xi}$ then $\frac{u(\widehat{f}^{n+1}(\widehat{x}))}{u(\widehat{x}_{n+1})} = e^{\pm(\xi - Q_\varepsilon(\widehat{x}_n)^{\beta/4})}$.

Note that the ratio improves.

Proof of Claim 1. It is enough to prove the claim for $n = 0$, so assume $\frac{u(\widehat{x})}{u(\widehat{x}_0)} = e^{\pm\xi}$ with $\xi \geq \sqrt{\varepsilon}$. We have $\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{x}_1)} = \frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{f}(\widehat{x}_0))} \cdot \frac{u(\widehat{f}(\widehat{x}_0))}{u(\widehat{x}_1)}$. By (E2.2) we have $\frac{u(\widehat{f}(\widehat{x}_0))}{u(\widehat{x}_1)} = e^{\pm p_0}$, and since $p_0 \ll Q_\varepsilon(\widehat{x}_0)^{\beta/4}$ it follows that $\frac{u(\widehat{f}(\widehat{x}_0))}{u(\widehat{x}_1)} = e^{\pm Q_\varepsilon(\widehat{x}_0)^{\beta/4}}$. Thus it is enough to show that $\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{f}(\widehat{x}_0))} = e^{\pm(\xi - 2Q_\varepsilon(\widehat{x}_0)^{\beta/4})}$. We show that $\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{f}(\widehat{x}_0))} \leq e^{\xi - 2Q_\varepsilon(\widehat{x}_0)^{\beta/4}}$ (the other side is proved similarly). By the proof of Lemma 2.1,

$$\frac{u(\widehat{f}(\widehat{x}))^2}{u(\widehat{f}(\widehat{x}_0))^2} = \frac{1 + e^{2\chi}|df_{\vartheta[\widehat{x}]}|^{-2}u(\widehat{x})^2}{1 + e^{2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2} \leq \frac{1 + e^{2\xi+2\chi}|df_{\vartheta[\widehat{x}]}|^{-2}u(\widehat{x}_0)^2}{1 + e^{2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2}$$
$$\leq \underbrace{\left(\frac{1+e^{2\xi+2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2}{1+e^{2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2}\right)}_{=\,\mathrm{I}} \underbrace{\exp\left(2\left|\log|df_{\vartheta[\widehat{x}]}| - \log|df_{\vartheta[\widehat{x}_0]}|\right|\right)}_{=\,\mathrm{II}}.$$

Note that I can be written as:

$$\mathrm{I} = \frac{1+e^{2\xi+2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2}{1+e^{2\chi}|df_{\vartheta[\widehat{x}_0]}|^{-2}u(\widehat{x}_0)^2} = e^{2\xi} - \frac{e^{2\xi}-1}{u(\widehat{f}(\widehat{x}_0))^2} = e^{2\xi}\left[1 - \frac{1-e^{-2\xi}}{u(\widehat{f}(\widehat{x}_0))^2}\right].$$

Using that $\xi \geq \sqrt{\varepsilon}$ and that $u(\widehat{f}(\widehat{x}_0)) < e^{\chi+1}\rho(\widehat{x}_0)^{-a}u(\widehat{x}_0)$ (see the proof of discreteness of Theorem 3.1), it follows that for $\varepsilon > 0$ small enough it holds:

$$\frac{1-e^{-2\xi}}{u(\widehat{f}(\widehat{x}_0))^2} > \varepsilon^{\frac{1}{2}}e^{-2\chi-2}\rho(\widehat{x}_0)^{2a}u(\widehat{x}_0)^{-2} > \varepsilon^{\frac{1}{2}}e^{-2\chi-2}\left[\varepsilon^{-\frac{1}{12}}Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{36}}\right]\left[\varepsilon^{-\frac{1}{4}}Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{12}}\right]$$
$$= \varepsilon^{\frac{1}{6}}e^{-2\chi-2}Q_\varepsilon(\widehat{x}_0)^{-\frac{5\beta}{36}}Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{4}} > \varepsilon^{\frac{1}{6}}e^{-2\chi-2}\varepsilon^{-\frac{5}{12}}Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{4}} = \varepsilon^{-\frac{1}{4}}e^{-2\chi-2}Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{4}} > 5Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{4}}.$$

Using that $1 - t < e^{-t}$ for $t \in \mathbb{R}$, we obtain that $\mathrm{I} < e^{2\xi - 5Q_\varepsilon(\widehat{x}_0)^{\beta/4}}$.

We now estimate II. By (A2)–(A3), if $x, y \in M$ with $y \in D_x$ then

$$\left|\frac{|df_x|}{|df_y|} - 1\right| \leq |df_y|^{-1}|df_x - df_y| \leq \mathfrak{K}d(x, \mathscr{S})^{-a}|x - y|^\beta.$$

Since $\log t \leq t - 1$ for $t > 1$, we get $|\log|df_x| - \log|df_y|| \leq \mathfrak{K}d(x, \mathscr{S})^{-a}|x - y|^\beta$ whenever $y \in D_x$, hence for $\varepsilon > 0$ small enough it holds:

$$2\left|\log|df_{\vartheta[\widehat{x}]}| - \log|df_{\vartheta[\widehat{x}_0]}|\right| \leq 2\mathfrak{K}d(\vartheta[\widehat{x}_0], \mathscr{S})^{-a}d(\vartheta[\widehat{x}], \vartheta[\widehat{x}_0])^\beta$$
$$\leq 2\mathfrak{K}\rho(\widehat{x}_0)^{-a}Q_\varepsilon(\widehat{x}_0)^\beta < 2\mathfrak{K}\varepsilon^{\frac{1}{24}}Q_\varepsilon(\widehat{x}_0)^{\frac{71\beta}{72}} < Q_\varepsilon(\widehat{x}_0)^{\frac{\beta}{4}}.$$

The estimates of I and II above imply that $\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{f}(\widehat{x}_0))} \leq e^{\xi - 2Q_\varepsilon(\widehat{x}_0)^{\beta/4}}$, as claimed. □

Remark 4.3. The proof above also works to show that if $\widehat{f}(\widehat{x}) \in V^u[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\le 1}]$ and if $\frac{u(\widehat{x})}{u(\widehat{x}_0)} = e^{\pm\xi}$ for $\xi \ge \sqrt{\varepsilon}$, then $\frac{u(\widehat{f}(\widehat{x}))}{u(\widehat{x}_1)} = e^{\pm(\xi - Q_\varepsilon(\widehat{x}_0)^{\beta/4})}$. In particular, the ratio does not get worse.

At this point, we do not yet know that $u(\widehat{x}) < \infty$. The next claim gives this.

Claim 2: $\exists \xi \ge \sqrt{\varepsilon}$ and a sequence $n_k \to -\infty$ s.t. $\frac{u(\widehat{f}^{n_k}(\widehat{x}))}{u(\widehat{x}_{n_k})} = e^{\pm\xi}$ for all k.

In particular $u(\widehat{f}^{n_k}(\widehat{x})) < \infty$ and so, by the proof of Lemma 2.1, $u(\widehat{x}) < \infty$.

Proof of Claim 2. The statement and its proof are very similar to Claim 1 in the proof of [53, Prop. 7.3]. Since $\underline{v} \in \Sigma^\#$, there is a sequence $n_k \to -\infty$ s.t. $v_{n_k} = v = \Psi_{\widehat{y}}^q$ for all k. Thus it is enough to show that $u(\widehat{f}^{n_k}(\widehat{x}))$ is uniformly bounded. Since v is relevant, Theorem 3.1(3) implies that there is $\underline{w} = \{w_n\}_{n\in\mathbb{Z}} \in \Sigma$ with $w_0 = v$ and $\pi[\underline{w}] = \widehat{z} \in \mathrm{NUE}_\chi^*$. In particular, $u(\widehat{z}) < \infty$. Let $V^u := V^u[\{w_n\}_{n\le 0}]$, then by Proposition 2.11(3)(b) we have $u(\widehat{z}') \le \exp[\varepsilon^{3/4}]u(\widehat{z})$ for all $\widehat{z}' \in V^u$. Define $\xi_0 > 0$ by $e^{\xi_0} := \max\left\{\frac{u(\widehat{y})}{u(\widehat{z})}, \frac{u(\widehat{z})}{u(\widehat{y})}, e^{\sqrt{\varepsilon}}\right\}$ and $\xi := \xi_0 + \varepsilon^{3/4}$. Below we will show that $u(\widehat{f}^{n_k}(\widehat{x})) \le e^\xi u(\widehat{y})$ for all k.

Fix k. For $\ell < k$ define $\widehat{z}_\ell \in \widehat{M}$ by $\widehat{f}^{n_\ell}(\widehat{z}_\ell) := V^u \cap V^s[\{v_m\}_{m\ge n_\ell}]$. This definition makes sense because $v_{n_\ell} = v$ and V^u is an unstable set at v. It is clear that $\widehat{z}_\ell \to \widehat{x}$. Since $\widehat{z}, \widehat{f}^{n_\ell}(\widehat{z}_\ell) \in V^u$, we have that $\frac{u(\widehat{f}^{n_\ell}(\widehat{z}_\ell))}{u(\widehat{y})} = \frac{u(\widehat{f}^{n_\ell}(\widehat{z}_\ell))}{u(\widehat{z})} \cdot \frac{u(\widehat{z})}{u(\widehat{y})} = e^{\pm\xi}$, hence by Remark 4.3 above we have $\frac{u(\widehat{f}^{n_k}(\widehat{z}_\ell))}{u(\widehat{y})} = e^{\pm\xi}$. For each fixed $N > 0$ we have $\widehat{df}^{(-n)}_{\widehat{f}^{n_k}(\widehat{z}_\ell)} \xrightarrow[\ell\to\infty]{} \widehat{df}^{(-n)}_{\widehat{f}^{n_k}(\widehat{x})}$ for all $n = 0, \dots, N$, therefore

$$\sum_{n=0}^{N} e^{2n\chi} |\widehat{df}^{(-n)}_{\widehat{f}^{n_k}(\widehat{x})}|^2 \le \lim_{\ell\to\infty} \sum_{n=0}^{N} e^{2n\chi} |\widehat{df}^{(-n)}_{\widehat{f}^{n_k}(\widehat{z}_\ell)}|^2 \le \lim_{\ell\to\infty} u(\widehat{f}^{n_k}(\widehat{z}_\ell))^2 \le e^{2\xi} u(\widehat{y})^2$$

and so $u(\widehat{f}^{n_k}(\widehat{x})) \le e^\xi u(\widehat{y})$. This completes the proof of the claim. □

Now we complete the proof of the proposition. By Claim 2, we can apply Claim 1 along $\underline{v}$ and the orbit of $\widehat{x}$ for some $\xi \ge \sqrt{\varepsilon}$: if $v_m = v$ for infinitely many $m < n$, then the ratio improves a definite amount at each of these indices. The conclusion is that $\frac{u(\widehat{f}^n(\widehat{x}))}{u(\widehat{x}_n)} = e^{\pm\sqrt{\varepsilon}}$, and by the same reason $\frac{u(\widehat{f}^n(\widehat{x}))}{u(\widehat{y}_n)} = e^{\pm\sqrt{\varepsilon}}$. Therefore $\frac{u(\widehat{x}_n)}{u(\widehat{y}_n)} = e^{\pm 2\sqrt{\varepsilon}}$. □

4.2. Control of $Q_\varepsilon(\widehat{x}_n)$ and p_n

Recall that $Q_\varepsilon(\widehat{x}) := \max\{q \in I_\varepsilon : q \le \widetilde{Q}_\varepsilon(\widehat{x})\}$ where

$$\widetilde{Q}_\varepsilon(\widehat{x}) = \varepsilon^{3/\beta} \min\left\{u(\widehat{x})^{-24/\beta}, u(\widehat{f}^{-1}(\widehat{x}))^{-12/\beta} \rho(\widehat{x})^{72a/\beta}\right\},$$

so we first control $\widetilde{Q}_\varepsilon(\widehat{x}_n)$. By part (2), $\frac{u(\widehat{x}_n)}{u(\widehat{y}_n)} = e^{\pm 2\sqrt{\varepsilon}}$ for all $n \in \mathbb{Z}$. Since $\Psi_{\widehat{f}^{-1}(\widehat{x}_n)}^{p_{n-1}} \overset{\varepsilon}{\approx} \Psi_{\widehat{x}_{n-1}}^{p_{n-1}}$, Proposition 2.5(1) implies that $\frac{u(\widehat{f}^{-1}(\widehat{x}_n))}{u(\widehat{x}_{n-1})} = e^{\pm\sqrt{\varepsilon}}$, and similarly $\frac{u(\widehat{f}^{-1}(\widehat{y}_n))}{u(\widehat{y}_{n-1})} = e^{\pm\sqrt{\varepsilon}}$. Hence

$$\frac{u(\widehat{f}^{-1}(\widehat{x}_n))}{u(\widehat{f}^{-1}(\widehat{y}_n))} = \frac{u(\widehat{f}^{-1}(\widehat{x}_n))}{u(\widehat{x}_{n-1})} \cdot \frac{u(\widehat{x}_{n-1})}{u(\widehat{y}_{n-1})} \cdot \frac{u(\widehat{y}_{n-1})}{u(\widehat{f}^{-1}(\widehat{y}_n))} = e^{\pm 4\sqrt{\varepsilon}}.$$

Now we estimate the ratio $\frac{\rho(\widehat{x}_n)}{\rho(\widehat{y}_n)}$. By (4.2), (4.3), (4.4) we have $\frac{\rho(\widehat{x}_n)}{\rho(\widehat{y}_n)} = e^{\pm 7\varepsilon}$, hence $\frac{\rho(\widehat{x}_n)^{72a}}{\rho(\widehat{y}_n)^{72a}} = e^{\pm\sqrt{\varepsilon}}$. The conclusion is that $\frac{\widetilde{Q}_\varepsilon(\widehat{x}_n)}{\widetilde{Q}_\varepsilon(\widehat{y}_n)} = \exp[\pm(\frac{49}{\beta}\sqrt{\varepsilon})]$, thus $\frac{Q_\varepsilon(\widehat{x}_n)}{Q_\varepsilon(\widehat{y}_n)} = \exp[\pm(\frac{2}{3}\varepsilon + \frac{49}{\beta}\sqrt{\varepsilon})]$. For $\varepsilon > 0$ small enough we get $\frac{Q_\varepsilon(\widehat{x}_n)}{Q_\varepsilon(\widehat{y}_n)} = e^{\pm\sqrt[3]{\varepsilon}}$.

We now prove part (4). We use the lemma below, whose proof is the same as in [53, Prop. 8.3].

Lemma 4.4. *If* $\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}} \in \Sigma^\#$ *then* $p_n = \delta_\varepsilon Q_\varepsilon(\widehat{x}_n)$ *for infinitely many* $n < 0$.

By symmetry, it is enough to prove that $p_n \ge e^{-\sqrt[3]{\varepsilon}} q_n$ for all $n \in \mathbb{Z}$. We have:

- If $p_n = \delta_\varepsilon Q_\varepsilon(\widehat{x}_n)$ then part (3) gives $p_n = \delta_\varepsilon Q_\varepsilon(\widehat{x}_n) \geq e^{-\sqrt[3]{\varepsilon}} \delta_\varepsilon Q_\varepsilon(\widehat{y}_n) \geq e^{-\sqrt[3]{\varepsilon}} q_n$.
- If $p_n \geq e^{-\sqrt[3]{\varepsilon}} q_n$ then (E2.3) and part (3) imply that

$$p_{n+1} = \min\{e^\varepsilon p_n, \delta_\varepsilon Q_\varepsilon(\widehat{x}_{n+1})\} \geq e^{-\sqrt[3]{\varepsilon}} \min\{e^\varepsilon q_n, \delta_\varepsilon Q_\varepsilon(\widehat{y}_{n+1})\} = e^{-\sqrt[3]{\varepsilon}} q_{n+1}.$$

By Lemma 4.4, we conclude that $p_n \geq e^{-\sqrt[3]{\varepsilon}} q_n$ for all $n \in \mathbb{Z}$.

4.3. *Control of* $\Psi_{\widehat{y}_n}^{-1} \circ \Psi_{\widehat{x}_n}$

We can assume that $n = 0$ and write $\Psi_{\widehat{x}_0}^{p_0} = \Psi_{\widehat{x}}^p$, $\Psi_{\widehat{y}_0}^{q_0} = \Psi_{\widehat{y}}^q$. Firstly note that, since $1 - 2t < e^{-t} < e^t < 1 + 2t$ for small $t > 0$, part (2) implies $\left|\frac{u(\widehat{y})}{u(\widehat{x})} - 1\right| < 4\sqrt{\varepsilon}$. As in the proof of Proposition 2.5(3), we have $(\Psi_{\widehat{y}}^{-1} \circ \Psi_{\widehat{x}} - \mathrm{Id})(t) = \Delta(t) + \delta$, where $\Delta(t) := \left(\frac{u(\widehat{y})}{u(\widehat{x})} - 1\right) t$. Thus:

- $\Delta(0) = 0$ and $\|d\Delta\|_0 = \left|\frac{u(\widehat{y})}{u(\widehat{x})} - 1\right| < 4\sqrt{\varepsilon}$.
- There are $t \in R[p], t' \in R[q]$ s.t. $\Psi_{\widehat{x}}(t) = \Psi_{\widehat{y}}(t') = \vartheta[\widehat{x}]$, hence $t' = \Psi_{\widehat{y}}^{-1} \circ \Psi_{\widehat{x}}(t) = t + \Delta(t) + \delta$. If $\varepsilon > 0$ is small enough, part (4) implies:

$$|\delta| \leq |t| + \|d\Delta\|_0 |t| + |t'| \leq (1 + 4\sqrt{\varepsilon}) p + q \leq [(1 + 4\sqrt{\varepsilon}) e^{\sqrt[3]{\varepsilon}} + 1] q < 3q.$$

This completes the proof of part (5), and hence of Theorem 4.1.

5. Symbolic dynamics

5.1. *A countable Markov partition*

Remind that (Σ, σ) is the TMS constructed from Theorem 3.1, and $\pi : \Sigma \to \widehat{M}$ is the map defined in the end of section 3. We now employ Theorem 4.1 to build a cover of $\mathrm{NUE}_\chi^\#$ that is locally finite and satisfies a (symbolic) Markov property. We will use the constructions of [53,39,38] to build a Markov partition for $\widehat{f}$, paying attention to the following facts:

- Our stable and unstable sets are not curves, but they do have good descriptions in terms of the coordinates of $\widehat{x}$ (Lemma 2.10), and they do satisfy properties analogous to their smooth versions (Proposition 2.11).
- Most of the methods of [53,39,38] used to construct the Markov partition are abstract and rely on the properties of stable and unstable sets that will be stated in this section.

THE MARKOV COVER $\mathscr{Z}$: Let $\mathscr{Z} := \{Z(v) : v \in \mathscr{A}\}$, where

$$Z(v) := \{\pi(\underline{v}) : \underline{v} \in \Sigma^\# \text{ and } v_0 = v\}.$$

In other words, we take the natural partition of $\Sigma^\#$ into cylinders at the zeroth position and use π to induce a cover $\mathscr{Z}$. Stable/unstable sets allow us to define "invariant fibres" inside each $Z \in \mathscr{Z}$, as follows. Let $Z = Z(v)$.

s/u–FIBRES IN $\mathscr{Z}$: Given $\widehat{x} \in Z$, let $W^s(\widehat{x}, Z) := V^s[\{\Psi_{\widehat{x}_n}^{100p_n}\}_{n\geq 0}] \cap Z$ be the *s–fibre* of $\widehat{x}$ in Z for some (any) $\underline{v} = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}} \in \Sigma^\#$ s.t. $\pi(\underline{v}) = \widehat{x}$ and $\Psi_{\widehat{x}_0}^{p_0} = v$. Similarly, let $W^u(\widehat{x}, Z) := V^u[\{\Psi_{\widehat{x}_n}^{100p_n}\}_{n\leq 0}] \cap Z$ be the *u–fibre* of $\widehat{x}$ in Z.

We also define $V^s(\widehat{x}, Z) := V^s[\{\Psi_{\widehat{x}_n}^{100p_n}\}_{n\geq 0}]$ and $V^u(\widehat{x}, Z) := V^u[\{\Psi_{\widehat{x}_n}^{100p_n}\}_{n\leq 0}]$. These sets are well-defined because $\{\Psi_{\widehat{x}_n}^{100p_n}\}_{n\in\mathbb{Z}}$ is a weak ε–gpo. By Proposition 2.11(4), the definitions of $V^{s/u}(\widehat{x}, Z)$, $W^{s/u}(\widehat{x}, Z)$ do not depend on the choice of $\underline{v}$, and any two s–fibres (u–fibres) either coincide or are disjoint. It is important noticing the relation between $W^{s/u}(\widehat{x}, Z)$ and $V^{s/u}(\widehat{x}, Z)$:

- $W^s(\widehat{x}, Z) = V^s(\widehat{x}, Z) = \vartheta^{-1}[x]$, where $x \in M$ is uniquely defined by $f^n(x) \in \Psi_{\widehat{x}_n}(R[p_n])$ for all $n \geq 0$, see Lemma 2.10(1).
- $V^u(\widehat{x}, Z)$ is isomorphic to the interval $\Psi_{\widehat{x}_0}(R[100p_0])$, while $W^u(\widehat{x}, Z)$ is isomorphic to a (usually fractal) subset of $\Psi_{\widehat{x}_0}(R[p_0])$.

Remark 5.1. In [53,39,38], $Z(\Psi_{\widehat{x}}^p) \subset \Psi_{\widehat{x}}(R[10^{-2}p])$ and the estimate for δ in Theorem 4.1(5) is $|\delta| < 10^{-1}q$. In particular, defining $V^u(\widehat{x}, Z) := V^u[\{\Psi_{\widehat{x}_n}^{p_n}\}_{n\leq 0}]$ (without taking larger domains) is enough to guarantee that Smale brackets of points in Z, Z' are well-defined whenever $Z \cap Z' \neq \emptyset$. In our case, $Z(\Psi_{\widehat{x}}^p) \subset \Psi_{\widehat{x}}(R[p])$ and the estimate for δ on Theorem 4.1(5) is $|\delta| < 3q$. That is why we define $V^u(\widehat{x}, Z)$ in a larger domain. This change is merely technical, since the statements and proofs about $V^{s/u}(\widehat{x}, Z)$ in our case are essentially the same as in [53,39,38].

Below we collect the main properties of $\mathscr{Z}$.

Proposition 5.2. *The following holds for all $\varepsilon > 0$ small enough.*

(1) Covering property: *$\mathscr{Z}$ is a cover of* $\mathrm{NUE}_\chi^\#$.
(2) Local finiteness: *For every $Z \in \mathscr{Z}$, $\#\{Z' \in \mathscr{Z} : Z \cap Z' \neq \emptyset\} < \infty$.*
(3) Product structure: *For every $Z \in \mathscr{Z}$ and every $\widehat{x}, \widehat{y} \in Z$, the intersection $W^s(\widehat{x}, Z) \cap W^u(\widehat{y}, Z)$ consists of a single element of Z.*
(4) Symbolic Markov property: *If $\widehat{x} = \pi(\underline{v})$ with $\underline{v} = \{v_n\}_{n\in\mathbb{Z}} \in \Sigma^\#$ then*

$$\widehat{f}(W^s(\widehat{x}, Z(v_0))) \subset W^s(\widehat{f}(\widehat{x}), Z(v_1)) \text{ and } f^{-1}(W^u(\widehat{f}(\widehat{x}), Z(v_1))) \subset W^u(\widehat{x}, Z(v_0)).$$

Part (1) follows from Theorem 3.1(2), part (2) follows from Theorem 4.1(4) and Theorem 3.1(1), part (3) follows from Proposition 2.11(1) and the product structure of $\Sigma^\#$, and part (4) is proved as in [53, Prop. 10.9]. For $\widehat{x}, \widehat{y} \in Z$, let $[\widehat{x}, \widehat{y}]_Z$ denote the intersection element of $W^s(\widehat{x}, Z)$ and $W^u(\widehat{y}, Z)$.

Lemma 5.3. *The following holds for all $\varepsilon > 0$ small enough.*

(1) Compatibility: *If $\widehat{x}, \widehat{y} \in Z(v_0)$ and $\widehat{f}(\widehat{x}), \widehat{f}(\widehat{y}) \in Z(v_1)$ with $v_0 \overset{\varepsilon}{\leftarrow} v_1$ then $\widehat{f}([\widehat{x}, \widehat{y}]_{Z(v_0)}) = [\widehat{f}(\widehat{x}), \widehat{f}(\widehat{y})]_{Z(v_1)}$.*
(2) Overlapping charts properties: *If $Z = Z(\Psi_{\widehat{x}}^p)$, $Z' = Z(\Psi_{\widehat{y}}^q) \in \mathscr{Z}$ with $Z \cap Z' \neq \emptyset$ then:*
   (a) *$Z \subset \Psi_{\widehat{y}}(R[100q])$.*
   (b) *If $\widehat{x} \in Z \cap Z'$ then $W^{s/u}(\widehat{x}, Z) \subset V^{s/u}(\widehat{x}, Z')$.*
   (c) *If $\widehat{z} \in Z$, $\widehat{w} \in Z'$ then $V^s(\widehat{z}, Z)$ and $V^u(\widehat{w}, Z')$ intersect at a unique element.*

Proof. (1) Proceed as in [53, Lemma 10.7].

(2) By Theorem 4.1, we have $\frac{p}{q} = e^{\pm\sqrt[3]{\varepsilon}}$ and $(\Psi_{\widehat{y}}^{-1} \circ \Psi_{\widehat{x}})(t) = t + \Delta(t) + \delta$ for $t \in R[10Q_\varepsilon(\widehat{x})]$, where $\delta \in \mathbb{R}$ with $|\delta| < 3q$ and $\Delta : R[10Q_\varepsilon(\widehat{x})] \to \mathbb{R}$ with $\Delta(0) = 0$ and $\|d\Delta\|_0 < 4\sqrt{\varepsilon}$. Let us prove (a). Since $Z \subset \Psi_{\widehat{x}}(R[p])$, it is enough to show that $\Psi_{\widehat{x}}(R[p]) \subset \Psi_{\widehat{y}}(R[100q])$. Here is the proof: if $x = \Psi_{\widehat{x}}(t)$ with $|t| \leq p$ then $x = \Psi_{\widehat{y}}(t')$ with $t' = (\Psi_{\widehat{y}}^{-1} \circ \Psi_{\widehat{x}})(t)$, and $|t'| \leq |t| + |\Delta(t)| + |\delta| \leq (1 + 4\sqrt{\varepsilon})p + 3q \leq [(1 + 4\sqrt{\varepsilon})e^{\sqrt[3]{\varepsilon}} + 3]q < 100q$.

Now we prove (b). Let $\underline{v}, \underline{w} \in \Sigma^\#$ s.t. $\widehat{x} = \pi(\underline{v}) = \pi(\underline{w})$, where $\underline{v} = \{\Psi_{\widehat{x}_n}^{p_n}\}_{n\in\mathbb{Z}}$ and $\underline{w} = \{\Psi_{\widehat{y}_n}^{q_n}\}_{n\in\mathbb{Z}}$ with $\Psi_{\widehat{x}_0}^{p_0} = \Psi_{\widehat{x}}^p$ and $\Psi_{\widehat{y}_0}^{q_0} = \Psi_{\widehat{y}}^q$. Proceeding as in the last paragraph, Theorem 4.1 implies that $\Psi_{\widehat{x}_n}(R[p_n]) \subset \Psi_{\widehat{y}_n}(R[100q_n])$ and $\Psi_{\widehat{y}_n}(R[q_n]) \subset \Psi_{\widehat{x}_n}(R[100p_n])$ for all $n \in \mathbb{Z}$. By the discussion before Lemma 2.3, we get that $g_{\widehat{x}_n}$ and $g_{\widehat{y}_n}$ are the same inverse branch of f, for every $n \in \mathbb{Z}$. Since

$$W^u(\widehat{x}, Z) \subset \{(x_n)_{n\in\mathbb{Z}} \in \widehat{M} : x_0 \in \Psi_{\widehat{x}_0}(R[p_0]) \text{ and } x_{n-1} = g_{\widehat{x}_n}(x_n), \forall n \leq 0\},$$
$$V^u(\widehat{x}, Z') = \{(x_n)_{n\in\mathbb{Z}} \in \widehat{M} : x_0 \in \Psi_{\widehat{y}_0}(R[100q_0]) \text{ and } x_{n-1} = g_{\widehat{y}_n}(x_n), \forall n \leq 0\}$$

and $\Psi_{\widehat{x}_0}(R[p_0]) \subset \Psi_{\widehat{y}_0}(R[100q_0])$, it follows that $W^u(\widehat{x}, Z) \subset V^u(\widehat{x}, Z')$. The other inclusion is actually an equality: $W^s(\widehat{x}, Z) = V^s(\widehat{x}, Z) = \vartheta^{-1}[\vartheta[\widehat{x}]] = V^s(\widehat{x}, Z')$.

Now we prove (c). Write $V^s(\widehat{z}, Z) = \vartheta^{-1}[z]$ and $V^u(\widehat{w}, Z') = V^u[\{\Psi_{\widehat{y}_n}^{100q_n}\}_{n\le 0}]$, where $\Psi_{\widehat{y}_0}^{q_0} = \Psi_{\widehat{y}}^{q}$. Define $(x_n)_{n\in\mathbb{Z}}$ by $x_n = f^n(z)$ for $n \ge 0$ and $x_{n-1} = g_{\widehat{y}_n}(x_n)$ for $n \le 0$. We have $(x_n)_{n\in\mathbb{Z}} \in V^s(\widehat{z}, Z)$. Since $z \in \Psi_{\widehat{x}}(R[p]) \subset \Psi_{\widehat{y}}(R[100q])$, we also have that $(x_n)_{n\in\mathbb{Z}} \in V^u(\widehat{w}, Z')$, hence $V^s(\widehat{z}, Z) \cap V^u(\widehat{w}, Z')$ contains $(x_n)_{n\in\mathbb{Z}}$. Clearly, any element in this intersection must be equal to $(x_n)_{n\in\mathbb{Z}}$. □

The next step is to apply a refinement method to destroy non-trivial intersections in $\mathscr{Z}$. The result is a pairwise disjoint cover of $\text{NUE}_\chi^\#$ with a (geometrical) Markov property. This idea, originally developed by Sinaĭ and Bowen for finite covers [54,55,10], works equally well for locally finite countable covers [53]. Let $\mathscr{Z} = \{Z_1, Z_2, \ldots\}$.

THE MARKOV PARTITION $\mathscr{R}$: For every $Z_i, Z_j \in \mathscr{Z}$, define a partition of Z_i by:

$$\begin{aligned}
T_{ij}^{su} &= \{\widehat{x} \in Z_i : W^s(\widehat{x}, Z_i) \cap Z_j \neq \emptyset, W^u(\widehat{x}, Z_i) \cap Z_j \neq \emptyset\}\\
T_{ij}^{s\emptyset} &= \{\widehat{x} \in Z_i : W^s(\widehat{x}, Z_i) \cap Z_j \neq \emptyset, W^u(\widehat{x}, Z_i) \cap Z_j = \emptyset\}\\
T_{ij}^{\emptyset u} &= \{\widehat{x} \in Z_i : W^s(\widehat{x}, Z_i) \cap Z_j = \emptyset, W^u(\widehat{x}, Z_i) \cap Z_j \neq \emptyset\}\\
T_{ij}^{\emptyset\emptyset} &= \{\widehat{x} \in Z_i : W^s(\widehat{x}, Z_i) \cap Z_j = \emptyset, W^u(\widehat{x}, Z_i) \cap Z_j = \emptyset\}.
\end{aligned}$$

Let $\mathscr{T} := \{T_{ij}^{\alpha\beta} : i, j \ge 1, \alpha \in \{s, \emptyset\}, \beta \in \{u, \emptyset\}\}$, and let $\mathscr{R}$ be the partition generated by $\mathscr{T}$.

Since $T_{ii}^{su} = Z_i$, $\mathscr{R}$ is a pairwise disjoint cover of $\text{NUE}_\chi^\#$. Clearly, $\mathscr{R}$ is finer than $\mathscr{Z}$. The local finiteness of $\mathscr{Z}$ (Proposition 5.2(2)) implies two local finiteness properties for $\mathscr{R}$:

- For every $Z \in \mathscr{Z}$, $\#\{R \in \mathscr{R} : R \subset Z\} < \infty$.
- For every $R \in \mathscr{R}$, $\#\{Z \in \mathscr{Z} : Z \supset R\} < \infty$.

Now we show that $\mathscr{R}$ is a Markov partition in the sense of Sinaĭ [55]. For that, we define s/u–fibres in $\mathscr{R}$.

s/u–FIBRES IN $\mathscr{R}$: Given $\widehat{x} \in R \in \mathscr{R}$, we define the *s–fibre* and *u–fibre* of $\widehat{x}$ by:

$$W^s(\widehat{x}, R) := \bigcap_{\substack{T_{ij}^{\alpha\beta} \in \mathscr{T}\\ T_{ij}^{\alpha\beta} \supset R}} W^s(\widehat{x}, Z_i) \cap T_{ij}^{\alpha\beta} \quad \text{and} \quad W^u(\widehat{x}, R) := \bigcap_{\substack{T_{ij}^{\alpha\beta} \in \mathscr{T}\\ T_{ij}^{\alpha\beta} \supset R}} W^u(\widehat{x}, Z_i) \cap T_{ij}^{\alpha\beta}.$$

It is clear that any two s–fibres (u–fibres) either coincide or are disjoint.

Proposition 5.4. *The following holds for $\varepsilon > 0$ small enough.*

(1) PRODUCT STRUCTURE: *For every $R \in \mathscr{R}$ and every $\widehat{x}, \widehat{y} \in R$, the intersection $W^s(\widehat{x}, R) \cap W^u(\widehat{y}, R)$ consists of a single element of R. Denote it by $[\widehat{x}, \widehat{y}]$.*
(2) HYPERBOLICITY: *If $\widehat{z}, \widehat{w} \in W^s(\widehat{x}, R)$ then $d(\widehat{f}^n(\widehat{z}), \widehat{f}^n(\widehat{w})) \xrightarrow[n\to\infty]{} 0$, and if $\widehat{z}, \widehat{w} \in W^u(\widehat{x}, R)$ then $d(\widehat{f}^n(\widehat{z}), \widehat{f}^n(\widehat{w})) \xrightarrow[n\to-\infty]{} 0$. The rates are exponential.*
(3) GEOMETRICAL MARKOV PROPERTY: *Let $R_0, R_1 \in \mathscr{R}$. If $\widehat{x} \in R_0$ and $\widehat{f}(\widehat{x}) \in R_1$ then*

$$\widehat{f}(W^s(\widehat{x}, R_0)) \subset W^s(\widehat{f}(\widehat{x}), R_1) \text{ and } \widehat{f}^{-1}(W^u(\widehat{f}(\widehat{x}), R_1)) \subset W^u(\widehat{x}, R_0).$$

When M is compact and f is a diffeomorphism, this is [53, Prop. 11.5 and 11.7] and the same proof works in our case.

5.2. A finite-to-one Markov extension

We construct a new coding for $\widehat{f}$. Let $\widehat{\mathscr{G}} = (\widehat{V}, \widehat{E})$ be the oriented graph with vertex set $\widehat{V} = \mathscr{R}$ and edge set $\widehat{E} = \{R \to S : R, S \in \mathscr{R} \text{ s.t. } \widehat{f}(R) \cap S \neq \emptyset\}$, and let $(\widehat{\Sigma}, \widehat{\sigma})$ be the TMS induced by $\widehat{\mathscr{G}}$.

For $\ell \in \mathbb{Z}$ and a path $R_m \to \cdots \to R_n$ on $\widehat{\mathscr{G}}$ define ${}_\ell[R_m, \dots, R_n] := \widehat{f}^{-\ell}(R_m) \cap \cdots \cap \widehat{f}^{-\ell-(n-m)}(R_n)$, which is the set of points whose itinerary between the iterates $\ell, \dots, \ell + (n-m)$ visits the rectangles $R_m, \dots, R_n$. The crucial property of these sets is that ${}_\ell[R_m, \dots, R_n] \neq \emptyset$. This follows by induction, using the Markov property of $\mathscr{R}$ (Proposition 5.4(3)).

The map π defines similar sets: for $\ell \in \mathbb{Z}$ and a path $v_m \overset{\varepsilon}{\leftarrow} \cdots \overset{\varepsilon}{\leftarrow} v_n$ on Σ let $Z_\ell[v_m, \dots, v_n] := \{\pi(\underline{w}) : \underline{w} \in \Sigma^\# \text{ and } w_\ell = v_m, \dots, w_{\ell+(n-m)} = v_n\}$. There is a relation between Σ and $\widehat{\Sigma}$ in terms of these sets.

Lemma 5.5. *If* $\{R_n\}_{n\in\mathbb{Z}} \in \widehat{\Sigma}$ *then there exists* $\{v_n\}_{n\in\mathbb{Z}} \in \Sigma$ *s.t.* $R_n \subset Z(v_n)$ *and* ${}_{-n}[R_{-n}, \dots, R_n] \subset Z_{-n}[v_{-n}, \dots, v_n]$ *for all* $n \geq 0$.

Proof. When M is compact and f is a diffeomorphism, this is [53, Lemma 12.2], whose proof applies a diagonal argument using the fact that each vertex of Σ has finite ingoing and outgoing degree. Since our Σ does do not necessarily satisfy this finiteness property (see Remark 3.3), the same proof does *not* work in our case. Instead, we use the local finiteness of $\mathscr{R}$ to apply the diagonal argument, as follows.

For $n \geq 0$, take $\widehat{x}_n \in {}_{-n}[R_{-n}, \dots, R_n]$, and let $\underline{v}^{(n)} = \{v_k^{(n)}\}_{k\in\mathbb{Z}} \in \Sigma^\#$ s.t. $\pi(\underline{v}^{(n)}) = \widehat{x}_n$. As in [53, Lemma 12.2], ${}_{-n}[R_{-n}, \dots, R_n] \subset Z_{-n}[v_{-n}^{(n)}, \dots, v_n^{(n)}]$. Since R_k is included in finitely many elements of $\mathscr{Z}$, there are finitely many choices for $(v_{-n}^{(n)}, \dots, v_n^{(n)})$. By a diagonal argument, there is $\underline{v}$ s.t. for all $n \geq 0$, $(v_{-n}, \dots, v_n) = (v_{-n}^{(m)}, \dots, v_n^{(m)})$ for infinitely many m. Now continue as in [53, Lemma 12.2]. $\square$

By Proposition 5.4(2), $\bigcap_{n\geq 0} \overline{{}_{-n}[R_{-n}, \dots, R_n]}$ is the intersection of a descending chain of nonempty closed sets with diameters converging to zero.

THE MAP $\widehat{\pi} : \widehat{\Sigma} \to \widehat{M}$: Given $\underline{R} = \{R_n\}_{n\in\mathbb{Z}} \in \widehat{\Sigma}$, $\widehat{\pi}(\underline{R})$ is defined by the identity

$$\{\widehat{\pi}(\underline{R})\} := \bigcap_{n\geq 0} \overline{{}_{-n}[R_{-n}, \dots, R_n]}.$$

The triple $(\widehat{\Sigma}, \widehat{\sigma}, \widehat{\pi})$ satisfies Theorem 1.1.

Theorem 5.6. *The following holds for all* $\varepsilon > 0$ *small enough.*

(1) $\widehat{\pi} : \widehat{\Sigma} \to \widehat{M}$ *is Hölder continuous.*
(2) $\widehat{\pi} \circ \widehat{\sigma} = \widehat{f} \circ \widehat{\pi}$.
(3) $\widehat{\pi}[\widehat{\Sigma}^\#] \supset \mathrm{NUE}_\chi^\#$.
(4) *Every point of* $\widehat{\pi}[\widehat{\Sigma}^\#]$ *has finitely many pre-images in* $\widehat{\Sigma}^\#$.

In particular, if μ is an f–adapted χ–expanding measure and $\widehat{\mu}$ is its lift to the natural extension, then $\widehat{\mu}(\widehat{\pi}[\widehat{\Sigma}^\#]) = 1$. When M is compact and f is a diffeomorphism, parts (1)–(3) are [53, Thm. 12.5] and part (4) is [39, Thm. 5.6(4)]. The same proofs work in our case, and the bound on the number of pre-images is exactly the same: there is a function $N : \mathscr{R} \to \mathbb{N}$ s.t. if $\widehat{x} = \widehat{\pi}(\underline{R})$ with $R_n = R$ for infinitely many $n > 0$ and $R_n = S$ for infinitely many $n < 0$ then $\#\{\underline{S} \in \widehat{\Sigma}^\# : \widehat{\pi}(\underline{S}) = \widehat{x}\} \leq N(R)N(S)$.

Declaration of competing interest

None declared under financial, general, and institutional competing interests.

References

[1] Jon Aaronson, An Introduction to Infinite Ergodic Theory, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997.
[2] R.L. Adler, B. Weiss, Entropy, a complete metric invariant for automorphisms of the torus, Proc. Natl. Acad. Sci. USA 57 (1967) 1573–1576.
[3] Roy L. Adler, Benjamin Weiss, Similarity of Automorphisms of the Torus, Memoirs of the American Mathematical Society, vol. 98, American Mathematical Society, Providence, R.I., 1970.

[4] Michael Benedicks, Lennart Carleson, On iterations of $1 - ax^2$ on $(-1, 1)$, Ann. Math. (2) 122 (1) (1985) 1–25.
[5] Michael Benedicks, Lennart Carleson, The dynamics of the Hénon map, Ann. Math. (2) 133 (1) (1991) 73–169.
[6] Henk Bruin, Gerhard Keller, Equilibrium states for S-unimodal maps, Ergod. Theory Dyn. Syst. 18 (4) (1998) 765–789.
[7] Snir Ben Ovadia, Symbolic dynamics for non-uniformly hyperbolic diffeomorphisms of compact smooth manifolds, J. Mod. Dyn. 13 (2018) 43–113.
[8] Rufus Bowen, Markov partitions for Axiom A diffeomorphisms, Am. J. Math. 92 (1970) 725–747.
[9] Rufus Bowen, Symbolic dynamics for hyperbolic flows, Am. J. Math. 95 (1973) 429–460.
[10] Rufus Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Lecture Notes in Mathematics, vol. 470, Springer-Verlag, Berlin, 1975.
[11] Rufus Bowen, Bernoulli maps of the interval, Isr. J. Math. 28 (1–2) (1977) 161–168.
[12] H. Bruin, Induced maps, Markov extensions and invariant measures in one-dimensional dynamics, Commun. Math. Phys. 168 (3) (1995) 571–580.
[13] Jérôme Buzzi, Omri Sarig, Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps, Ergod. Theory Dyn. Syst. 23 (5) (2003) 1383–1400.
[14] L.A. Bunimovich, Ya.G. Sinaĭ , N.I. Chernov, Markov partitions for two-dimensional hyperbolic billiards, Usp. Mat. Nauk 45 (3(273)) (1990) 97–134, 221.
[15] Henk Bruin, Mike Todd, Equilibrium states for interval maps: potentials with $\sup \phi - \inf \phi < h_{\text{top}}(f)$, Commun. Math. Phys. 283 (3) (2008) 579–611.
[16] Henk Bruin, Mike Todd, Equilibrium states for interval maps: the potential $-t \log |Df|$, Ann. Sci. Éc. Norm. Supér. (4) 42 (4) (2009) 559–600.
[17] J. Buzzi, Bowen Factors, Their Degree, and Codings of Surface Diffeomorphisms, 2019, arXiv preprint https://arxiv.org/abs/1807.040170.
[18] Jérôme Buzzi, Intrinsic ergodicity of affine maps in $[0, 1]^d$, Monatshefte Math. 124 (2) (1997) 97–118.
[19] Jérôme Buzzi, Intrinsic ergodicity of smooth interval maps, Isr. J. Math. 100 (1997) 125–161.
[20] Jérôme Buzzi, Ergodicité intrinsèque de produits fibrés d'applications chaotiques unidimensionelles, Bull. Soc. Math. Fr. 126 (1) (1998) 51–77.
[21] Jérôme Buzzi, Markov extensions for multi-dimensional dynamical systems, Isr. J. Math. 112 (1999) 357–380.
[22] Manfred Denker, Gerhard Keller, Mariusz Urbański, On the uniqueness of equilibrium states for piecewise monotone mappings, Stud. Math. 97 (1) (1990) 27–36.
[23] Neil Dobbs, Pesin theory and equilibrium measures on the interval, Fundam. Math. 231 (1) (2015) 1–17.
[24] Katrin Gelfert, Repellers for non-uniformly expanding maps with singular or critical points, Bull. Braz. Math. Soc. 41 (2) (2010) 237–257.
[25] Jacek Graczyk, Grzegorz Świątek, Generic hyperbolicity in the logistic family, Ann. Math. (2) 146 (1) (1997) 1–52.
[26] B.M. Gurevič, Topological entropy of a countable Markov chain, Dokl. Akad. Nauk SSSR 187 (1969) 715–718.
[27] B.M. Gurevič, Shift entropy and Markov measures in the space of paths of a countable graph, Dokl. Akad. Nauk SSSR 192 (1970) 963–965.
[28] Franz Hofbauer, Gerhard Keller, Equilibrium states for piecewise monotonic transformations, Ergod. Theory Dyn. Syst. 2 (1) (1982) 23–43.
[29] Franz Hofbauer, β-shifts have unique maximal measure, Monatshefte Math. 85 (3) (1978) 189–198.
[30] Franz Hofbauer, On intrinsic ergodicity of piecewise monotonic transformations with positive entropy, Isr. J. Math. 34 (3) (1979) 213–237 (1980).
[31] Franz Hofbauer, On intrinsic ergodicity of piecewise monotonic transformations with positive entropy. II., Isr. J. Math. 38 (1–2) (1981) 107–115.
[32] Godofredo Iommi, Mike Todd, Natural equilibrium states for multimodal maps, Commun. Math. Phys. 300 (1) (2010) 65–94.
[33] M.V. Jakobson, Absolutely continuous invariant measures for one-parameter families of one-dimensional maps, Commun. Math. Phys. 81 (1) (1981) 39–88.
[34] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, Publ. Math. IHES (51) (1980) 137–173.
[35] Anatole Katok, Jean-Marie Strelcyn, F. Ledrappier, F. Przytycki, Invariant Manifolds, Entropy and Billiards; Smooth Maps with Singularities, Lecture Notes in Mathematics, vol. 1222, Springer-Verlag, Berlin, 1986.
[36] Svetlana Katok, Ilie Ugarcovici, Symbolic dynamics for the modular surface and beyond, Bull. Am. Math. Soc. 44 (1) (2007) 87–132.
[37] François Ledrappier, Some properties of absolutely continuous invariant measures on an interval, Ergod. Theory Dyn. Syst. 1 (1) (1981) 77–93.
[38] Yuri Lima, Carlos Matheus, Symbolic dynamics for non-uniformly hyperbolic surface maps with discontinuities, Ann. Sci. Éc. Norm. Supér. (4) 51 (1) (2018) 1–38.
[39] Yuri Lima, Omri M. Sarig, Symbolic dynamics for three-dimensional flows with positive topological entropy, J. Eur. Math. Soc. 21 (1) (2019) 199–256.
[40] A. Lasota, James A. Yorke, On the existence of invariant measures for piecewise monotonic transformations, Trans. Am. Math. Soc. 186 (1973) 481–488 (1974).
[41] Mikhail Lyubich, Dynamics of quadratic polynomials. I, II, Acta Math. 178 (2) (1997) 185–247, 247–297.
[42] Ricardo Mañé, Hyperbolicity, sinks and measure in one-dimensional dynamics, Commun. Math. Phys. 100 (4) (1985) 495–524.
[43] Vilton Pinheiro, Expanding measures, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 28 (6) (2011) 889–939.
[44] Feliks Przytycki, Lyapunov characteristic exponents are nonnegative, Proc. Am. Math. Soc. 119 (1) (1993) 309–317.
[45] Yakov Pesin, Samuel Senti, Thermodynamical formalism associated with inducing schemes for one-dimensional maps, Mosc. Math. J. 5 (3) (2005) 669–678, 743–744.
[46] Ya.B. Pesin, S. Senti, K. Zhang, Lifting measures to inducing schemes, Ergod. Theory Dyn. Syst. 28 (2) (2008) 553–574.
[47] M.E. Ratner, Markov decomposition for an U-flow on a three-dimensional manifold, Mat. Zametki 6 (1969) 693–704.
[48] M. Ratner, Markov partitions for Anosov flows on n-dimensional manifolds, Isr. J. Math. 15 (1973) 92–114.

[49] Juan Rivera-Letelier, Asymptotic Expansion of Smooth Interval Maps, 2012, 26 pages, arXiv preprint arXiv:1204.3071.
[50] V.A. Rohlin, Exact endomorphisms of a Lebesgue space, Izv. Akad. Nauk SSSR, Ser. Mat. 25 (1961) 499–530.
[51] Sylvie Ruette, Mixing C^r maps of the interval without maximal measure, Isr. J. Math. 127 (2002) 253–277.
[52] Omri M. Sarig, Bernoulli equilibrium states for surface diffeomorphisms, J. Mod. Dyn. 5 (3) (2011) 593–608.
[53] Omri M. Sarig, Symbolic dynamics for surface diffeomorphisms with positive entropy, J. Am. Math. Soc. 26 (2) (2013) 341–426.
[54] Ja.G. Sinaĭ, Construction of Markov partitionings, Funkc. Anal. Prilozh. 2 (3) (1968) 70–80 (Loose errata).
[55] Ja.G. Sinaĭ, Markov partitions and U-diffeomorphisms, Funkc. Anal. Prilozh. 2 (1) (1968) 64–89.
[56] M. Smorodinsky, β-automorphisms are Bernoulli shifts, Acta Math. Acad. Sci. Hung. 24 (1973) 273–278.
[57] Yōichirō Takahashi, Isomorphisms of β-automorphisms to Markov automorphisms, Osaka J. Math. 10 (1973) 175–184.
[58] J.-C. Yoccoz, A proof of Jakobson's theorem, 34 pages, available at https://www.college-de-france.fr/media/jean-christophe-yoccoz/UPL7416254474776698194_Jakobson_jcy.pdf.
[59] Lai-Sang Young, Statistical properties of dynamical systems with some hyperbolicity, Ann. Math. (2) 147 (3) (1998) 585–650.