Stretched exponential estimate on growth of the number of periodic points for prevalent diffeomorphisms

Vadim Yu. Kaloshin[1], Brian R. Hunt[2]

1 Mathematics 253-37, Caltech, Pasadena, CA 91125;
e-mail: kaloshin@its.caltech.edu

2 Department of Mathematics, Mathematics Building, University of Maryland College Park, MD 20742-4015;
e-mail: bhunt@math.umd.edu

Abstract

For diffeomorphisms of smooth compact finite-dimensional manifolds, we consider the problem of how fast the number of periodic points with period n grows as a function of n. In many familiar cases (*e.g.*, Anosov systems) the growth is exponential, but arbitrarily fast growth is possible; in fact, the first author has shown that arbitrarily fast growth is topologically (Baire) generic for C^2 or smoother diffeomorphisms. In the present work we show that, by contrast, for a measure-theoretic notion of genericity we call "prevalence", the growth is not much faster than exponential. Specifically, we show that for each $\rho, \delta > 0$, there is a prevalent set of $C^{1+\rho}$ (or smoother) diffeomorphisms for which the number of period n points is bounded above by $\exp(Cn^{1+\delta})$ for some C independent of n. We also obtain a related bound on the decay of hyperbolicity of the periodic points as a function of n, and obtain the same results for 1-dimensional endomorphisms. The contrast between topologically generic and measure-theoretically generic behavior for the growth of the number of periodic points and the decay of their hyperbolicity shows this to be a subtle and complex phenomenon, reminiscent of KAM theory. Here in Part I we state our results and describe the methods we use. We complete most of the proof in the 1-dimensional C^2-smooth case and outline the remaining steps, deferred to Part II, that are needed to establish the general case.

The novel feature of the approach we develop in this paper is an introduction of Newton Interpolation Polynomials as a tool of perturbing trajectories of iterated maps.

Table of Contents

Chapter 1

A Problem of the Growth of the Number of Periodic Points and Decay of Hyperbolicity for Generic Diffeomorphisms.

1.1 Introduction

Let $\mathrm{Diff}^r(M)$ be the space of C^r diffeomorphisms of a finite-dimensional smooth compact manifold M with the uniform C^r–topology, where $\dim M \geq 2$, and let $f \in \mathrm{Diff}^r(M)$. Consider the number of periodic points of period n

$$P_n(f) = \#\{x \in M : \quad x = f^n(x)\}. \tag{1.1}$$

The main question of this paper is:

Question 1.1.1. *How quickly can $P_n(f)$ grow with n for a "generic" C^r diffeomorphism f?*

We put the word "generic" in brackets because as the reader will see the answer depends on notion of genericity.

For technical reasons one sometimes counts only *isolated* points of period n; let

$$\begin{aligned} P_n^i(f) = \#\{x \in M : \quad & x = f^n(x) \text{ and } y \neq f^n(y) \\ & \text{for } y \neq x \text{ in some neighborhood of } x\}. \end{aligned} \tag{1.2}$$

We call a diffeomorphism $f \in \mathrm{Diff}^r(M)$ an *Artin-Mazur diffeomorphism* (or simply *A-M diffeomorphism*) if the number of isolated periodic orbits of f grows at most exponentially fast, *i.e.* for some number $C > 0$ we have

$$P_n^i(f) \leq \exp(Cn) \quad \text{for all } n \in \mathbb{Z}_+. \tag{1.3}$$

Artin & Mazur [AM] proved the following result.

Theorem 1.1.2. *For $0 \leq r \leq \infty$, A-M diffeomorphisms are dense in* $\mathrm{Diff}^r(M)$ *with the uniform C^r–topology.*

We say that a point $x \in M$ of period n for f is hyperbolic if $df^n(x)$, the linearization of f^n at x, has no eigenvalues with modulus 1. (Notice that a hyperbolic solution to $f^n(x) = x$ must also be isolated.) We call $f \in \mathrm{Diff}^r(M)$ a strongly Artin-Mazur diffeomorphism if for some number $C > 0$,

$$P_n(f) \leq \exp(Cn) \quad \text{for all } n \in \mathbb{Z}_+, \tag{1.4}$$

and all periodic points of f are hyperbolic (whence $P_n(f) = P_n^i(f)$). In [K1] an elementary proof of the following extension of the Artin-Mazur result is given.

Theorem 1.1.3. *For $0 \leq r < \infty$, strongly A-M diffeomorphisms are dense in* $\mathrm{Diff}^r(M)$ *with the uniform C^r–topology.*

According to the standard terminology a set in $\mathrm{Diff}^r(M)$ is called residual if it contains a countable intersection of open dense sets and a property is called *(Baire) generic* if diffeomorphisms with that property form a residual set. It turns out the A-M property is not generic, as is shown in [K2]. Moreover:

Theorem 1.1.4. *[K2] For any $2 \leq r < \infty$ there is an open set $\mathcal{N} \subset$* $\mathrm{Diff}^r(M)$ *such that for any given sequence $a = \{a_n\}_{n \in \mathbb{Z}_+}$ there is a Baire generic set $\mathcal{R}_a$ in $\mathcal{N}$ depending on the sequence a_n with the property if $f \in \mathcal{R}_a$, then for infinitely many $n_k \in \mathbb{Z}_+$ we have $P_{n_k}^i(f) > a_{n_k}$.*

Of course since $P_n(f) \geq P_n^i(f)$, the same statement can be made about $P_n(f)$. But in fact it is shown in [K2] that $P_n(f)$ is infinite for n sufficiently large, due to a continuum of periodic points, for at least a dense set of $f \in \mathcal{N}$.

The proof of this Theorem is based on a result of Gonchenko-Shilnikov-Turaev [GST1]. Two slightly different detailed proofs of their result are given in [K2] and [GST2]. The proof in [K2] relies on a strategy outlined in [GST1].

However, it seems unnatural that if you pick a diffeomorphism at random then it may have an arbitrarily fast growth of number of periodic points. Moreover, Baire generic sets in Euclidean spaces can have zero Lebesgue measure. Phenomena that are Baire generic, but have a small probability are well-known in dynamical systems, KAM theory, number theory, etc. (see [O], [HSY], [K3] for various examples). This partially motivates the problem posed by Arnold [A]:

Problem 1.1.5. *Prove that "with probability one" $f \in \mathrm{Diff}^r(M)$ is an A-M diffeomorphism.*

Arnold suggested the following interpretation of "with probability one": *for a (Baire) generic finite parameter family of diffeomorphisms $\{f_\varepsilon\}$, for Lebesgue almost every ε we have that f_ε is A-M* (compare with [K3]). As Theorem 1.3 shows, a result on the genericity of the set of A-M diffeomorphisms based on (Baire) topology is likely to be extremely subtle, if possible at all [1]. We use instead a notion of "probability one" based on prevalence [HSY, K3], which is independent of Baire genericity. We also are able to state the result in the form Arnold suggested for generic families using this measure-theoretic notion of genericity.

For a rough understanding of prevalence, consider a Borel measure μ on a Banach space V. We say that a property holds "μ–almost surely for perturbations" if it holds on a Borel set $P \subset V$ such that *for all $v \in V$ we have $v + w \in P$ for almost every w with respect to μ*. Notice that if $V = \mathbb{R}^k$ and μ is Lebesgue measure, then "almost surely with respect to perturbations by μ" is equivalent to "Lebesgue almost everywhere". Moreover, the Fubini/Tonelli Theorem implies that if μ is any Borel probability measure on $\mathbb{R}^k$, then a property that holds almost surely with respect to perturbations by μ must also be hold Lebesgue almost everywhere. Based on this observation, we call a property on a Banach space "prevalent" if it holds almost surely with respect to perturbations by μ for some Borel probability measure μ on V, which for technical reasons we require to have compact support. In order to apply this notion to the Banach manifold $\mathrm{Diff}^r(M)$, we must describe how we make perturbations in this space, which we will do in the next Section.

Our first main result is a partial solution to Arnold's problem. It says that *for a prevalent diffeomorphism $f \in \mathrm{Diff}^r(M)$, with $1 < r \leq \infty$, and all $\delta > 0$ there exists*

[1] For example, using technique from [GST2] and [K2] one can prove that for a (Baire) generic finite-parameter family $\{f_\varepsilon\}$ and a (Baire) generic parameter value ε the corresponding diffeomorphism f_ε is not A-M. Unfortunately, how to estimate the measure of non-A-M diffeomorphisms from below is a so far unreachable question

$C = C(\delta) > 0$ *such that for all* $n \in \mathbb{Z}_+$,

$$P_n(f) \leq \exp(Cn^{1+\delta}). \tag{1.5}$$

The results of this paper have been announced in [KH].

The Kupka-Smale Theorem (see e.g.[PM]) states that for a generic diffeomorphism all periodic points are hyperbolic and all associated stable and unstable manifolds intersect one another transversally. [K3] shows that the Kupka-Smale Theorem also holds on a prevalent set. So, the Kupka-Smale Theorem, in particular, says that a Baire generic (resp. prevalent) diffeomorphism has only hyperbolic periodic points, but *how hyperbolic are the periodic points, as function of their period, for a Baire generic (resp. prevalent) diffeomorphism* f*?* This is the second main problem we deal with in this paper.

Recall that a linear operator $L : \mathbb{R}^N \to \mathbb{R}^N$ is *hyperbolic* if it has no eigenvalues on the unit circle $\{|z| = 1\} \subset \mathbb{C}$. Denote by $|\cdot|$ the Euclidean norm in $\mathbb{C}^N$. Then we define the *hyperbolicity* of a linear operator L by

$$\gamma(L) = \inf_{\phi \in [0,1)} \inf_{|v|=1} |Lv - \exp(2\pi i \phi)v|. \tag{1.6}$$

We also say that L is γ–hyperbolic if $\gamma(L) \geq \gamma$. In particular, if L is γ–hyperbolic, then its eigenvalues $\{\lambda_j\}_{j=1}^N \subset \mathbb{C}$ are at least γ–distant from the unit circle, i.e. $\min_j ||\lambda_j| - 1| \geq \gamma$. The *hyperbolicity* of a periodic point $x = f^n(x)$ of period n, denoted by $\gamma_n(x, f)$, equals the hyperbolicity of the linearization $df^n(x)$ of f^n at points x, *i.e.* $\gamma_n(x, f) = \gamma(df^n(x))$. Similarly the number of periodic points $P_n(f)$ of period n is defined, one can define

$$\gamma_n(f) = \min_{\{x:\ x = f^n(x)\}} \gamma_n(x, f). \tag{1.7}$$

The idea of Gromov [G] and Yomdin [Y] of measuring hyperbolicity is that a γ–hyperbolic point of period n of a C^2 diffeomorphism f has an $M_2^{-2n}\gamma$–neighborhood (where $M_2 = \|f\|_{C^2}$) free from periodic points of the same period[2]. In Appendix A we prove the following result.

Proposition 1.1.6. *Let* M *be a compact manifold of dimension* N*, let* $f : M \to M$ *be a* $C^{1+\rho}$ *diffeomorphism, where* $0 < \rho \leq 1$*, that has only hyperbolic periodic points,*

[2]In [Y] hyperbolicity is introduced as the minimal distance of eigenvalues to the unit circle. This way of defining hyperbolicity does not guarantee the existence of a $M_2^{-2n}\gamma$–neighborhood free from periodic points of the same period; see Appendix A

and let $M_{1+\rho} = \max\{\|f\|_{C^{1+\rho}}, 2^{1/\rho}\}$*. Then there is a constant* $C = C(M) > 0$ *such that for each* $n \in \mathbb{Z}_+$ *we have*

$$P_n(f) \leq C\,(M_{1+\rho})^{nN(1+\rho)/\rho}\,\gamma_n(f)^{-N/\rho}. \tag{1.8}$$

Proposition 1.1.6 implies that a lower estimate on a decay of hyperbolicity $\gamma_n(f)$ gives an upper estimate on growth of number of periodic points $P_n(f)$. Therefore, a natural question is:

Question 1.1.7. *How quickly can* $\gamma_n(f)$ *decay with* n *for a "generic"* C^r *diffeomorphism* f*?*

For a Baire generic $f \in \mathrm{Diff}^r(M)$, the existence of lower bound on a rate of decay of $\gamma_n(f)$ would imply the existence of an upper bound on a rate of growth of the number of periodic points $P_n(f)$, whereas no such bound exists by Theorem 1.1.4. Thus again we consider genericity in the measure-theoretic sense of prevalence. Our second main result, which in view of Proposition 1.1.6 implies the first main result, is that *for a prevalent diffeomorphism* $f \in \mathrm{Diff}^r(M)$*, with* $1 < r \leq \infty$*, and for any* $\delta > 0$ *there exists* $C = C(\delta) > 0$ *such that*

$$\gamma_n(f) \geq \exp(-Cn^{1+\delta}). \tag{1.9}$$

Now we shall discuss in more detail our definition of prevalence ("probability one") in the space of diffeomorphisms $\mathrm{Diff}^r(M)$.

1.2 Prevalence in the space of diffeomorphisms Diff$^r(M)$

The space of C^r diffeomorphisms $\mathrm{Diff}^r(M)$ of a compact manifold M is a Banach manifold. Locally we can identify it with a Banach space, which gives it a local linear structure in the sense that we can perturb a diffeomorphism by "adding" small elements of the Banach space. As we described in the previous Section, the notion of prevalence requires us to make additive perturbations with respect to a probability measure that is independent of the place where we make the perturbation. Thus although there is not a unique way to put a linear structure on $\mathrm{Diff}^r(M)$, it is important to make a choice that is consistent throughout the Banach manifold.

The way we make perturbations on $\mathrm{Diff}^r(M)$ by small elements of a Banach space is as follows. First we embed M into the interior of the closed unit ball $B^N \subset \mathbb{R}^N$, which we can do for N sufficiently large by the Whitney Embedding Theorem [W]. We emphasize that our results hold for *every* possible choice of an embedding of M into $\mathbb{R}^N$. We then consider a closed neighborhood $U \subset B^N$ of M and Banach space $C^r(U, \mathbb{R}^N)$ of C^r functions from U to $\mathbb{R}^N$. Next we *extend* every element $f \in \mathrm{Diff}^r(M)$ to an element $F \in C^r(U, \mathbb{R}^N)$ that is strongly contracting in *all* the directions transverse to M [3]. Again the particular choice of how we make this extension is not important to our results; in Appendix C we describe how to extend a diffemorphism and what conditions we need to ensure that the results of Sacker [Sac] and Fenichel [F] apply as follows. Since F has M as an invariant manifold, if we add to F a small perturbation in $g \in C^r(U, \mathbb{R}^N)$, the perturbed map $F+g$ has an invariant manifold in U that is C^r-close to M. Then $F+g$ restricted to its invariant manifold corresponds in a natural way to an element of $\mathrm{Diff}^r(M)$, which we consider to be the perturbation of $f \in \mathrm{Diff}^r(M)$ by $g \in C^r(U, \mathbb{R}^N)$. The details of this construction are described in Appendix C.

In this way we reduce the problem to the study of maps in $\mathrm{Diff}^r(U)$, the open subset of $C^r(U, \mathbb{R}^N)$ consisting of those elements that are diffeomorphisms from U to some subset of its interior. The construction we described in the previous paragraph ensures that the number of periodic points $P_n(f)$ and their hyperbolicity $\gamma_n(f)$ for elements of $\mathrm{Diff}^r(M)$ are the same for the corresponding elements of $\mathrm{Diff}^r(U)$, so the bounds that we prove on these quantities for almost every perturbation of any element of $\mathrm{Diff}^r(U)$ hold as well for almost every perturbation of any element of $\mathrm{Diff}^r(M)$. Another justification for considering diffeomorphisms in Euclidean space is that the problem of exponential/superexponential growth of the number of periodic points $P_n(f)$ for a prevalent $f \in \mathrm{Diff}^r(M)$ is a *local problem* on M and is not affected by a global shape of M.

The results stated in the next Section apply to any compact domain $U \subset \mathbb{R}^N$, but for simplicity we state them for the closed unit ball B^N. In the previous Section, we said that a property is *prevalent* on a Banach space such as $C^r(B^N)$ if it holds on a Borel subset S for which there exists a Borel probability measure μ on $C^r(B^N)$ with compact support such that for all $f \in C^r(B^N)$ we have $f + g \in S$ for almost every g with respect to μ. The complement of a prevalent set is said to be *shy*. We then say that a property is prevalent on an open subset of $C^r(B^N)$ such as $\mathrm{Diff}^r(B^N)$ if the exceptions to the property in $\mathrm{Diff}^r(B^N)$ form a shy subset of $C^r(B^N)$.

In this paper the perturbation measure μ that we use is supported within the analytic functions in $C^r(B^N)$. In this sense we foliate $\mathrm{Diff}^r(B^N)$ by analytic leaves

[3] The existence of such extension is not obvious, as pointed out by C. Carminati.

that are compact and overlapping. The main result then says that *for* **every** *analytic leaf $L \subset \mathrm{Diff}^r(B^N)$ and every $\delta > 0$, for almost every diffeomorphism $f \in L$ in the leaf L both (1.5) and (1.9) are satisfied.* Now we define an analytic leaf as a "Hilbert Brick" in the space of analytic functions and a natural Lebesgue product probability measure μ on it.

1.3 Formulation of the main result in the multidimensional case

Fix a coordinate system $x = (x_1, \dots, x_N) \in \mathbb{R}^N \supset B^N$ and the scalar product $\langle x, y\rangle = \sum_i x_i y_i$. Let $\alpha = (\alpha_1, \dots, \alpha_N)$ be a multiindex from $\mathbb{Z}_+^N$, and let $|\alpha| = \sum_i \alpha_i$. For a point $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ we write $x^\alpha = \prod_{i=1}^N x_i^{\alpha_i}$. Associate to a real analytic function $\phi : B^N \to \mathbb{R}^N$ the set of coefficients of its expansion:

$$\phi_{\vec{\varepsilon}}(x) = \sum_{\alpha \in \mathbb{Z}_+^N} \vec{\varepsilon}_\alpha x^\alpha. \tag{1.10}$$

Denote by $W_{k,N}$ the space of N–component homogeneous vector-polynomials of degree k in N variables and by $\nu(k, N) = \dim W_{k,N}$ the dimension of $W_{k,N}$. According to the notation of the expansion (1.10), denote coordinates in $W_{k,N}$ by

$$\vec{\varepsilon}_k = \left(\{\vec{\varepsilon}_\alpha\}_{|\alpha|=k}\right) \in W_{k,N}. \tag{1.11}$$

In $W_{k,N}$ we use a scalar product that is invariant with respect to orthogonal transformation of $\mathbb{R}^N \supset B^N$ (see Appendix B), defined as follows:

$$\langle \vec{\varepsilon}_k, \vec{\nu}_k\rangle_k = \sum_{|\alpha|=k} \binom{k}{\alpha}^{-1} \langle \vec{\varepsilon}_\alpha, \vec{\nu}_\alpha\rangle, \quad \|\vec{\varepsilon}_k\|_k = \left(\langle \vec{\varepsilon}_k, \vec{\varepsilon}_k\rangle_k\right)^{1/2}. \tag{1.12}$$

Denote by

$$B_k^N(r) = \{\vec{\varepsilon}_k \in W_{k,N} : \ \|\vec{\varepsilon}_k\|_k \le r\} \tag{1.13}$$

the closed r–ball in $W_{k,N}$ centered at the origin. Let $Leb_{k,N}$ be Lebesgue measure on $W_{k,N}$ induced by the scalar product (1.12) and normalized by a constant so that the volume of the unit ball is one: $Leb_{k,N}(B_k^N(1)) = 1$.

Fix a nonincreasing sequence of positive numbers $\vec{\mathbf{r}} = (\{r_k\}_{k=0}^{\infty})$ such that $r_k \to 0$ as $k \to \infty$ and define a Hilbert Brick of size $\vec{\mathbf{r}}$

$$\begin{aligned} HB^N(\vec{\mathbf{r}}) =& \{\vec{\varepsilon} = \{\vec{\varepsilon}_\alpha\}_{\alpha \in \mathbb{Z}_+^N} : \text{ for all } k \in \mathbb{Z}_+, \|\vec{\varepsilon}_k\|_k \le r_k\} \\ =& B_0^N(r_0) \times B_1^N(r_1) \times \cdots \times B_k^N(r_k) \times \dots \\ \subset & \ W_{0,N} \quad \times W_{1,N} \quad \times \cdots \times W_{k,N} \quad \times \dots . \end{aligned} \tag{1.14}$$

Define *a Lebesgue product probability measure* $\mu_{\vec{\mathbf{r}}}^N$ *associated to the Hilbert Brick* $HB^N(\vec{\mathbf{r}})$ *of size* $\vec{\mathbf{r}}$ by normalizing for each $k \in \mathbb{Z}_+$ the corresponding Lebesgue measure $Leb_{k,N}$ on $W_{k,N}$ to the Lebesgue probability measure on the r_k–ball $B_k^N(r_k)$:

$$\mu_{k,r}^N = r^{-\nu(k,N)}\ Leb_{k,N} \quad \text{and} \quad \mu_{\vec{\mathbf{r}}}^N = \times_{k=0}^{\infty} \mu_{k,r_k}^N. \tag{1.15}$$

Definition 1.3.1. *Let $f \in \mathrm{Diff}^r(B^N)$ be a C^r diffeomorphism of B^N into its interior. We call $HB^N(\vec{\mathbf{r}})$ a Hilbert Brick of an admissible size $\vec{\mathbf{r}} = (\{r_k\}_{k=0}^{\infty})$ with respect to f if*
A) for each $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$, the corresponding function $\phi_{\vec{\varepsilon}}(x) = \sum_{\alpha \in \mathbb{Z}_+^N} \vec{\varepsilon}_\alpha x^\alpha$ is analytic on B^N;
B) for each $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$, the corresponding map $f_{\vec{\varepsilon}}(x) = f(x) + \phi_{\vec{\varepsilon}}(x)$ is a diffeomorphism from B^N into its interior, i.e. $\{f_{\vec{\varepsilon}}\}_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})} \subset \mathrm{Diff}^r(B^N)$;
C) for all $\delta > 0$ and all $C > 0$, the sequence $r_k \exp(Ck^{1+\delta}) \to \infty$ as $k \to \infty$.

Remark 1.3.2. *The first and second conditions ensure that the family $\{f_{\vec{\varepsilon}}\}_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})}$ lie inside an analytic leaf within the class of diffeomorphisms $\mathrm{Diff}^r(B^N)$. The third condition provides us enough freedom to perturb. It is important for our method to have infinitely many parameters to perturb. If r_k's were decaying too fast to zero it would make our family of perturbations essentially finite-dimensional.*

An example of an admissible sequence $\vec{\mathbf{r}} = (\{r_k\}_{k=0}^{\infty})$ is $r_k = \tau/k!$, where τ depends on f and is chosen sufficiently small to ensure that condition (B) holds. Notice that the diameter of $HB^N(\vec{\mathbf{r}})$ is then propotional to τ, so that τ can be chosen as some multiple of the distance from f to the boundary of $\mathrm{Diff}^r(B^N)$.

Main Theorem. *For any $0 < \rho \le \infty$ (or even $1 + \rho = \omega$) and any $C^{1+\rho}$ diffeomorphism $f \in \mathrm{Diff}^{1+\rho}(B^N)$, consider a Hilbert Brick $HB^N(\vec{\mathbf{r}})$ of an admissible size $\vec{\mathbf{r}}$ with respect to f and the family of analytic perturbations of f*

$$\{f_{\vec{\varepsilon}}(x) = f(x) + \phi_{\vec{\varepsilon}}(x)\}_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})} \tag{1.16}$$

with the Lebesgue product probability measure $\mu_{\vec{\mathbf{r}}}^N$ associated to $HB^N(\vec{\mathbf{r}})$. Then for every $\delta > 0$ and $\mu_{\vec{\mathbf{r}}}^N$–a.e. $\vec{\varepsilon}$ there is $C = C(\vec{\varepsilon}, \delta) > 0$ such that for all $n \in \mathbb{Z}_+$

$$\gamma_n(f_{\vec{\varepsilon}}) > \exp(-Cn^{1+\delta}), \quad P_n(f_{\vec{\varepsilon}}) < \exp(Cn^{1+\delta}). \tag{1.17}$$

Remark 1.3.3. *The fact that the measure $\mu^N_{\vec{\mathbf{r}}}$ depends on f does not conform to our definition of prevalence. However, we can decompose $\mathrm{Diff}^r(B^N)$ into a nested countable union of sets $\mathcal{S}_j$ that are each a positive distance from the boundary of $\mathrm{Diff}^r(B^N)$ and for each $j \in \mathbb{Z}^+$ choose an admissible sequence $\vec{\mathbf{r}}_j$ that is valid for all $f \in \mathcal{S}_j$. Since a countable intersection of prevalent subsets of a Banach space is prevalent [HSY], the Main Theorem implies the results stated in terms of prevalence in the introduction.*

Remark 1.3.4. *Recently the first author along with A. Gorodetski [GK] applied the technique developed here and obtained partial solution of Palis' conjecture about finiteness of the number of coexisting sinks for surface diffeomorphisms.*

In Appendix C we deduce from the Main Theorem the following result.

Theorem 1.3.5. *Let $\{f_\varepsilon\}_{\varepsilon \in B^m} \subset \mathrm{Diff}^{1+\rho}(M)$ be a generic m–parameter family of $C^{1+\rho}$ diffeomorphisms of a compact manifold M for some $\rho > 0$. Then for every $\delta > 0$ and a.e. $\varepsilon \in B^m$ there is a constant $C = C(\varepsilon, \delta)$ such that (1.17) is satisfied for every $n \in \mathbb{Z}_+$.*

In Appendix C we also give a precise meaning to the term *generic*. Now we formulate the most general result we shall prove.

Definition 1.3.6. *Let $\gamma \geq 0$ and $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$ diffeomorphism for some $\rho > 0$. A point $x \in B^N$ is called (n, γ)–periodic if $|f^n(x) - x| \leq \gamma$ and (n, γ)–hyperbolic if $\gamma_n(x, f) = \gamma(df^n(x)) \geq \gamma$.*

(Notice that a point can be (n, γ)–hyperbolic regardless of its periodicity, but this property is of interest primarily for (n, γ)–periodic points.) For positive C and δ let $\gamma_n(C, \delta) = \exp(-Cn^{1+\delta})$.

Theorem 1.3.7. *Given the hypotheses of the Main Theorem, for every $\delta > 0$ and for $\mu^N_{\vec{\mathbf{r}}}$–a.e. $\vec{\varepsilon}$ there is $C = C(\vec{\varepsilon}, \delta) > 0$ such that for all $n \in \mathbb{Z}_+$, every $(n, \gamma_n^{1/\rho}(C, \delta))$–periodic point $x \in B^N$ is $(n, \gamma_n(C, \delta))$–hyperbolic. (Here we assume $0 < \rho \leq 1$; in a space $\mathrm{Diff}^{1+\rho}(B^N)$ with $\rho > 1$, the statement holds with ρ replaced by 1.)*

This result together with Proposition 1.1.6 implies the Main Theorem, because any periodic point of period n is (n, γ)–periodic for any $\gamma > 0$.

Remark 1.3.8. *In the statement of the Main Theorem and Theorem 1.3.7 the unit ball B^N can be replaced by an bounded open set $U \subset \mathbb{R}^N$. After scaling U can be considered as a subset of the unit ball B^N.*

One can define a distance on a compact manifold M and almost periodic points of diffeomorphisms of M. Then one can cover $M = \cup_i U_i$ by coordinate charts and define hyperbolicity for almost periodic points using these charts $\{U_i\}_i$ (see [Y] for details). This gives a precise meaning to the following result.

Theorem 1.3.9. *Let $\{f_\varepsilon\}_{\varepsilon \in B^m} \subset \mathrm{Diff}^{1+\rho}(M)$ be a generic m–parameter family of diffeomorphisms of a compact manifold M for some $\rho > 0$. Then for every $\delta > 0$ and almost every $\varepsilon \in B^m$ there is a constant $C = C(\varepsilon, \delta)$ such that every $(n, \gamma_n^{1/\rho}(C, \delta))$–periodic point x in B^N is $(n, \gamma_n(C, \delta))$–hyperbolic. (Here again we assume $0 < \rho \leq 1$, replacing ρ with 1 in the conclusion if $\rho > 1$.)*

The meaning of the term generic is the same as in Theorem 1.3.5 and is discussed in Appendix C.

1.4 Formulation of the main result in the 1–dimensional case

The proof of the main multidimensional result (Theorem 1.3.7) is quite long and complicated. In order to describe the general approach we develop in this paper we apply our method to the 1–dimensional maps which represent a nontrivial simplified model for the multidimensional problem. The statement of the main result for the 1–dimensional maps has another important feature: it clarifies the statement of the main multidimensional result.

Fix the interval $I = [-1, 1]$. Associate to a real analytic function $\phi : I \to \mathbb{R}$ the set of coefficients of its expansion

$$\phi_\varepsilon(x) = \sum_{k=0}^{\infty} \varepsilon_k x^k. \tag{1.18}$$

For a nonincreasing sequence of positive numbers $\vec{\mathbf{r}} = (\{r_k\}_{k=0}^{\infty})$ such that $r_k \to 0$ as $k \to \infty$ following the multidimensional notations we define a Hilbert Brick of size $\vec{\mathbf{r}}$

$$HB^1(\vec{\mathbf{r}}) = \{\varepsilon = \{\varepsilon_k\}_{k=0}^{\infty} : \quad \text{for all} \quad k \in \mathbb{Z}_+, \quad |\varepsilon_k| \leq r_k\} \tag{1.19}$$

and the product probability measure $\mu^1_{\vec{\mathbf{r}}}$ associated to the Hilbert Brick $HB^1(\vec{\mathbf{r}})$ of size $\vec{\mathbf{r}}$ which considers each ε_k as a random variable independent from the other ε_k's and uniformly distributed on $[-r_k, r_k]$.

Main 1–dimensional Theorem. *For any $0 < \rho \le \infty$ (or even $1 + \rho = \omega$) and any $C^{1+\rho}$ map $f : I \to I$ of the interval $I = [-1, 1]$ consider a Hilbert Brick $HB^1(\vec{\mathbf{r}})$ of an admissible size $\vec{\mathbf{r}}$ with respect to f and the family of analytic perturbations of f*

$$\{f_\varepsilon(x) = f(x) + \phi_\varepsilon(x)\}_{\varepsilon \in HB^1(\vec{\mathbf{r}})} \tag{1.20}$$

with the Lebesgue product probability measure $\mu^1_{\vec{\mathbf{r}}}$ associated to $HB^1(\vec{\mathbf{r}})$. Then for every $\delta > 0$ and $\mu^1_{\vec{\mathbf{r}}}$–a.e. ε there is $C = C(\varepsilon, \delta) > 0$ such that for all $n \in \mathbb{Z}_+$

$$\gamma_n(f_\varepsilon) > \exp(-Cn^{1+\delta}), \quad P_n(f_\varepsilon) < \exp(Cn^{1+\delta}). \tag{1.21}$$

Moreover, for $\mu^1_{\vec{\mathbf{r}}}$–a.e. ε, every $(n, \exp(-Cn^{1+\delta}))$–periodic point is $(n, \exp(-Cn^{1+\delta}))$–hyperbolic.

In [MMS] Martens-de Melo-Van Strien prove a stronger statement for C^2 maps. They show that for any C^2 map f of an interval without flat critical points there are $\gamma > 0$ and $n_0 \in \mathbb{Z}_+$ such that for any $n > n_0$ we have $|\gamma_n(f)| > 1 + \gamma$. This also implies that the number of periodic points is bounded by an exponential function of the period. The notion of a flat critical point used in [MMS] is a nonstandard one from a point of view of singularity theory, in particular, if 0 is a critical point, then distance of $f(x)$ to $f(0)$ does not have to decay to 0 as $x \to 0$ faster than any degree of x.

In [KK] an example of a C^r–unimodal map with a critical point having tangency of order $2r + 2$ and an arbitrary fast rate of growth of the number of periodic points is presented.

Let us point out again that the main purpose of discussing the 1-dimensional case in details is to highlight ideas and explain the general method without overloading presentation by technical details. The general N-dimensional case is highly involved and excessive amount of technical details make understanding of general ideas of the method not easily accessible.

Acknowledgments: It is great pleasure to thank the thesis advisor of the first author, John Mather, who regularly spent hours listening oral expositions of various parts of the proof for nearly two years[4]. Without his patience and support this project would have never been completed. The authors are truly grateful to Anton Gorodetski, Giovanni Forni, and an antonymous referee who read the manuscript carefully and has made many useful remarks. The first author thanks the Institute for Physical Science and Technology, University of Maryland and, in particular, James Yorke for hospitality. The first author is grateful to Anatole Katok for providing an opportunity

[4]This paper is based on the first author's Ph.D. thesis.

to give a minicourse on the subject of this paper at Penn State University during the fall of 2000. The authors have profited from conversations with Carlo Carminati, Bill Cowieson, Dima Dolgopyat, Anatole Katok, Michael Lyubich, Michael Shub, Yakov Sinai, Marcelo Viana, Jean-Christophe Yoccoz, Lai-Sang Young, and many others. The first author acknowledges the support of Sloan dissertation fellowship during his final year at Princeton, when significant part of the work had been done. The second author is grateful to the Institute for Advanced Study at Princeton respectively for its hospitality. The first author is supported by the AIM and Sloan fellowships, and NSF-grant DMS-0300229, and the second by NSF grant DMS-0104087.

Chapter 2

Strategy of the Proof

Here we describe the strategy of the proof of the Main Result (Theorem 1.3.7). The general idea is to fix $C > 0$ and prove an upper bound on the measure of the set of "bad" parameter values $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ for which the conclusion of the Theorem does not hold. The upper bound we obtain will approach zero as $C \to \infty$, from which it follows immediately that the set of $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ that are "bad" for all $C > 0$ has measure zero. For a given $C > 0$, we bound the measure of "bad" parameter values inductively as follows.

Stage 1. We delete all parameter values $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ for which the corresponding diffeomorphism $f_{\vec{\varepsilon}}$ has an almost fixed point which is not sufficiently hyperbolic and bound the measure of the deleted set.

Stage 2. We consider only parameter values for which each almost fixed point is sufficiently hyperbolic. Then we delete all parameter values $\vec{\varepsilon}$ for which $f_{\vec{\varepsilon}}$ has an almost periodic point of period 2 which is not sufficiently hyperbolic and bound the measure of that set.

Stage n. We consider only parameter values for which each almost periodic point of period at most $n - 1$ is sufficiently hyperbolic (we shall call this *the Inductive Hypothesis*). Then we delete all parameter values $\vec{\varepsilon}$ for which $f_{\vec{\varepsilon}}$ has an almost periodic point of period n which is not sufficiently hyperbolic and bound the measure of that set.

The main difficulty in the proof is then to prove a bound on the measure of "bad" parameter values at stage n such that the bounds are summable over n and that the sum approaches zero as $C \to \infty$. Let us formalize the problem. Fix positive ρ, δ, and C, and recall that $\gamma_n(C, \delta) = \exp(-Cn^{1+\delta})$ for $n \in \mathbb{Z}_+$. Assume $\rho \leq 1$; if not, change its value to 1.

Definition 2.0.1. *A diffeomorphism $f \in \mathrm{Diff}^{1+\rho}(B^N)$ satisfies the Inductive Hypothesis of order n with constants (C, δ, ρ), denoted $f \in IH(n, C, \delta, \rho)$, if for all $k \le n$, every $(k, \gamma_k^{1/\rho}(C, \delta))$-periodic point is $(k, \gamma_k(C, \delta))$-hyperbolic.*

For $f \in \mathrm{Diff}^{1+\rho}(M)$, consider the sequence of sets in the parameter space $HB^N(\vec{\mathbf{r}})$

$$\begin{aligned} &B_n(C, \delta, \rho, \vec{\mathbf{r}}, f) = \{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}}) : \\ &f_{\vec{\varepsilon}} \in IH(n-1, C, \delta, \rho) \text{ but } f_{\vec{\varepsilon}} \notin IH(n, C, \delta, \rho)\} \end{aligned} \tag{2.1}$$

In other words, $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)$ is the set of "bad" parameter values $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ for which all almost periodic points of $f_{\vec{\varepsilon}}$ with period strictly less than n are sufficiently hyperbolic, but there is an almost periodic point of period n that is not sufficiently hyperbolic. Let

$$\begin{aligned} M_1 &= \sup_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})} \max\{\|f_{\vec{\varepsilon}}\|_{C^1}, \|f_{\vec{\varepsilon}}^{-1}\|_{C^1}\}; \\ M_{1+\rho} &= \sup_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})} \max\{\|f_{\vec{\varepsilon}}\|_{C^{1+\rho}}, M_1, 2^{1/\rho}\}. \end{aligned} \tag{2.2}$$

Our goal is to find an upper bound

$$\mu_{\vec{\mathbf{r}}}^N \{B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)\} \le \mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho}). \tag{2.3}$$

for the measure of the set of "bad" parameter values. Then the sum over n of (2.3) gives an upper bound

$$\mu_{\vec{\mathbf{r}}}^N \{\cup_{n=1}^{\infty} B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)\} \le \sum_{n=1}^{\infty} \mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho}) \tag{2.4}$$

on the measure of the set of all parameters $\vec{\varepsilon}$ for which $f_{\vec{\varepsilon}}$ has for at least one n an $(n, \gamma_n^{1/\rho}(C, \delta))$-periodic point that is not $(n, \gamma_n(C, \delta))$-hyperbolic. If this sum converges and

$$\sum_{n=1}^{\infty} \mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho}) = \mu(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho}) \to 0 \text{ as } C \to \infty \tag{2.5}$$

for every positive ρ, δ, and $M_{1+\rho}$, then Theorem 1.3.7 follows. In the remainder of this Chapter we describe the key construction we use to obtain a bound $\mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho})$ that meets condition (2.5).

2.1 Various perturbations of recurrent trajectories by Newton interpolation polynomials

The approach we take to estimate the measure of "bad" parameter values in the space of perturbations $HB^N(\vec{\mathbf{r}})$ is to choose a coordinate system for this space and for a finite subset of the coordinates to estimate the amount that we must change a particular coordinate to make a "bad" parameter value "good". Actually we will choose a coordinate system that depends on a particular point $x_0 \in B^N$, the idea being to use this coordinate system to estimate the measure of "bad" parameter values corresponding to initial conditions in some neighborhood of x_0, then cover B^N with a finite number of such neighborhoods and sum the corresponding estimates. For a particular set of initial conditions, a diffeomorphism will be "good" if every point in the set is either sufficiently nonperiodic or sufficiently hyperbolic.

In order to keep the notations and formulas simple as we formalize this approach, we consider the case of 1-dimensional maps, but the reader should always have in mind that our approach is designed for multidimensional diffeomorphisms. Let $f : I \to I$ be a C^1 map on the interval $I = [-1, 1]$. Recall that a trajectory $\{x_k\}_{k\in\mathbb{Z}}$ of f is called *recurrent* if it returns arbitrarily close to its initial position – that is, for all $\gamma > 0$ we have $|x_0 - x_n| < \gamma$ for some $n > 0$. A very basic question is how much one should perturb f to make x_0 periodic. Here is an elementary Closing Lemma that gives a simple partial answer to this question.

Closing Lemma. *Let $\{x_k = f^k(x_0)\}_{k=0}^n$ be a trajectory of length $n+1$ of a map $f : I \to I$. Let $u = (x_0 - x_n)/\prod_{k=0}^{n-2}(x_{n-1} - x_k)$. Then x_0 is a periodic point of period n of the map*

$$f_u(x) = f(x) + u \prod_{k=0}^{n-2}(x - x_k) \tag{2.6}$$

Of course f_u is close to f if and only if u is sufficiently small, meaning that $|x_0 - x_n|$ should be small compared to $\prod_{k=0}^{n-2}|x_{n-1} - x_k|$. However, this product is likely to contain small factors for recurrent trajectories. In general, it is difficult to control the effect of perturbations for recurrent trajectories. The simple reason why it is so us because *one can not perturb f at two nearby points independently*.

The Closing Lemma above also gives an idea of how much we must change the parameter u to make a point x_0 that is (n,γ)-periodic not be (n,γ)-periodic for a given $\gamma > 0$, which as we described above is one way to make a map that is "bad" for the initial condition x_0 become "good". To make use of the other part of our alternative we must determine how much we need to perturb a map f to make a given x_0 be (n,γ)-hyperbolic for some $\gamma > 0$.

Perturbation of hyperbolicity. *Let $\{x_k = f^k(x_0)\}_{k=0}^{n-1}$ be a trajectory of length n of a C^1 map $f : I \to I$. Then for the map*

$$f_v(x) = f(x) + v(x - x_{n-1}) \prod_{k=0}^{n-2} (x - x_k)^2 \tag{2.7}$$

such that $v \in \mathbb{R}$ and

$$||(f_v^n)'(x_0)| - 1| = \left| \left| \prod_{k=0}^{n-1} f'(x_k) + v \prod_{k=0}^{n-2} (x_{n-1} - x_k)^2 \prod_{k=0}^{n-2} f'(x_k) \right| - 1 \right| > \gamma \tag{2.8}$$

we have that x_0 is an (n,γ)-hyperbolic point of f_v.

One more time we can see the product of distances $\prod_{k=0}^{n-2} |x_{n-1} - x_k|$ along the trajectory is an important quantitative characteristic of how much freedom we have to perturb.

The perturbations (2.6) and (2.7) are reminiscent of Newton interpolation polynomials. Let us put these formulas into a general setting using singularity theory.

Given $n > 0$ and a C^1 function $f : I \to \mathbb{R}$ we define an associated function $j^{1,n} f : I^n \to I^n \times \mathbb{R}^{2n}$ by

$$\begin{aligned} j^{1,n} f(x_0, \ldots, x_{n-1}) = \Big(& x_0, \ldots, x_{n-1}, f(x_0), \ldots, f(x_{n-1}), \\ & f'(x_0), \ldots, f'(x_{n-1}) \Big). \end{aligned} \tag{2.9}$$

In singularity theory this function is called the *n-tuple 1-jet* of f. The ordinary 1-jet of f, usually denoted by $j^1 f(x) = (x, f(x), f'(x))$, maps I to the 1-*jet space* $\mathcal{J}^1(I, \mathbb{R}) \simeq I \times \mathbb{R}^2$. The product of n copies of $\mathcal{J}^1(I,\mathbb{R})$, called the *multijet space*, is denoted by

$$\mathcal{J}^{1,n}(I,\mathbb{R}) = \underbrace{\mathcal{J}^1(I,\mathbb{R}) \times \cdots \times \mathcal{J}^1(I,\mathbb{R})}_{n \text{ times}}, \tag{2.10}$$

and is equivalent to $I^n \times \mathbb{R}^{2n}$ after rearranging coordinates. The n-tuple 1-jet of f associates with each n-tuple of points in I^n all the information necessary to determine how close the n-tuple is to being a periodic orbit, and if so, how close it is to being nonhyperbolic.

The set

$$\Delta_n(I) = \Big\{\{x_0, \dots, x_{n-1}\} \times I^n \times \mathbb{R}^n \subset \mathcal{J}^{1,n}(I, \mathbb{R}) : \\ \exists\, i \neq j \text{ such that } x_i = x_j\Big\} \tag{2.11}$$

is called the *diagonal* (or sometimes the *generalized diagonal*) in the space of multijets. In singularity theory the space of multijets is defined outside of the diagonal $\Delta_n(I)$ and is usually denoted by $\mathcal{J}^1_n(I,\mathbb{R}) = \mathcal{J}^{1,n}(I,\mathbb{R}) \setminus \Delta_n(I)$ (see [GG]). It is easy to see that *a recurrent trajectory* $\{x_k\}_{k\in\mathbb{Z}_+}$ *is located in a neighborhood of the diagonal* $\Delta_n(I) \subset \mathcal{J}^{1,n}(I,\mathbb{R})$ *in the space of multijets for a sufficiently large* n. If $\{x_k\}_{k=0}^{n-1}$ is a part of a recurrent trajectory of length n, then the product of distances along the trajectory

$$\prod_{k=0}^{n-2} |x_{n-1} - x_k| \tag{2.12}$$

measures how close $\{x_k\}_{k=0}^{n-1}$ to the diagonal $\Delta_n(I)$, or how independently one can perturb points of a trajectory. One can also say that (2.12) is a quantitative characteristic of how recurrent a trajectory of length n is. Introduction of this *product of distances along a trajectory into analysis of recurrent trajectories* is a new point of our paper.

2.2 Newton interpolation and blow-up along the diagonal in multijet space

Now we present a construction due to Grigoriev and Yakovenko [GY] which puts the "Closing Lemma" and "Perturbation of Hyperbolicity" statements above into a general framework. It is an interpretation of Newton interpolation polynomials as an algebraic blow-up along the diagonal in the multijet space. In order to keep the notations and formulas simple we continue in this section to consider only the 1-dimensional case.

Consider the $2n$-parameter family of perturbations of a C^1 map $f : I \to I$ by polynomials of degree $2n - 1$

$$f_\varepsilon(x) = f(x) + \phi_\varepsilon(x), \qquad \phi_\varepsilon(x) = \sum_{k=0}^{2n-1} \varepsilon_k x^k, \tag{2.13}$$

where $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_{2n-1}) \in \mathbb{R}^{2n}$. The perturbation vector ε consists of coordinates from the Hilbert Brick $HB^1(\vec{\mathbf{r}})$ of analytic perturbations defined in Section 1.3. Our goal now is to describe how such perturbations affect the n-tuple 1-jet of f, and since the operator $j^{1,n}$ is linear in f, for the time being we consider only the perturbations ϕ_ε and their n-tuple 1-jets. For each n-tuple $\{x_k\}_{k=0}^{n-1}$ there is a natural transformation $\mathcal{J}^{1,n} : I^n \times \mathbb{R}^{2n} \to \mathcal{J}^{1,n}(I, \mathbb{R})$ from ε-coordinates to jet-coordinates, given by

$$\mathcal{J}^{1,n}(x_0, \ldots, x_{n-1}, \varepsilon) = j^{1,n}\phi_\varepsilon(x_0, \ldots, x_{n-1}). \tag{2.14}$$

Instead of working directly with the transformation $\mathcal{J}^{1,n}$, we introduce intermediate u-coordinates based on Newton interpolation polynomials. The relation between ε-coordinates and u-coordinates is given implicitly by

$$\phi_\varepsilon(x) = \sum_{k=0}^{2n-1} \varepsilon_k x^k = \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j(\text{mod } n)}). \tag{2.15}$$

Based on this identity, we will define functions $\mathcal{D}^{1,n} : I^n \times \mathbb{R}^{2n} \to I^n \times \mathbb{R}^{2n}$ and $\pi^{1,n} : I^n \times \mathbb{R}^{2n} \to \mathcal{J}^{1,n}(I, \mathbb{R})$ so that $\mathcal{J}^{1,n} = \pi^{1,n} \circ \mathcal{D}^{1,n}$, or in other words the diagram in Figure 1 commutes. We will show later that $\mathcal{D}^{1,n}$ is invertible, while $\pi^{1,n}$ is invertible away from the diagonal $\Delta_n(I)$ and defines a blow-up along it in the space of multijets $\mathcal{J}^{1,n}(I, \mathbb{R})$.

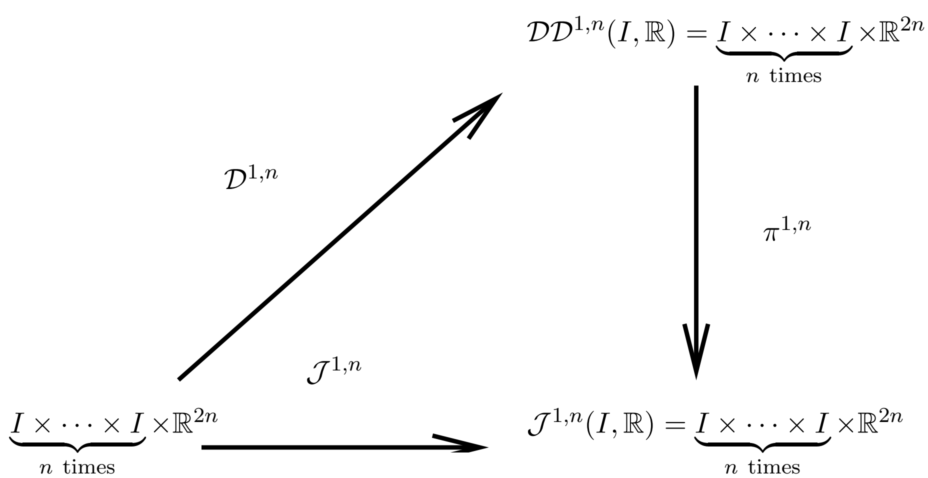

Figure 2.1: Algebraic Blow-up along the Diagonal $\Delta_n(I)$

The intermediate space, which we denote by $\mathcal{DD}^{1,n}(I, \mathbb{R})$, is called *the space of divided differences* and consists of n-tuples of points $\{x_k\}_{k=0}^{n-1}$ and $2n$ real coefficients $\{u_k\}_{k=0}^{2n-1}$. Here are explicit coordinate-by-coordinate formulas defining $\pi^{1,n}$:

$\mathcal{DD}^{1,n}(I,\mathbb{R}) \to \mathcal{J}^{1,n}(I,\mathbb{R})$. This mapping is given by

$$\begin{aligned}\pi^{1,n}(x_0, \ldots, x_{n-1}, &u_0, \ldots, u_{2n-1}) = \\ &\Big(x_0, \ldots, x_{n-1}, \phi_\varepsilon(x_0), \ldots, \phi_\varepsilon(x_{n-1}), \phi'_\varepsilon(x_0), \ldots, \phi'_\varepsilon(x_{n-1})\Big),\end{aligned} \tag{2.16}$$

where

$$\begin{aligned}
\phi_\varepsilon(x_0) =& u_0, \\
\phi_\varepsilon(x_1) =& u_0 + u_1(x_1 - x_0), \\
\phi_\varepsilon(x_2) =& u_0 + u_1(x_2 - x_0) + u_2(x_2 - x_0)(x_2 - x_1), \\
&\vdots \\
\phi_\varepsilon(x_{n-1}) =& u_0 + u_1(x_{n-1} - x_0) + \ldots \\
&+ u_{n-1}(x_{n-1} - x_0)\ldots(x_{n-1} - x_{n-2}), \\
\phi'_\varepsilon(x_0) =& \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k} (x - x_{j (\text{mod } n)})\right)\Big|_{x=x_0}, \\
&\vdots \\
\phi'_\varepsilon(x_{n-1}) =& \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k} (x - x_{j (\text{mod } n)})\right)\Big|_{x=x_{n-1}},
\end{aligned} \tag{2.17}$$

These formulas are very useful for dynamics. For a given base map f and initial point x_0, the image $f_\varepsilon(x_0) = f(x_0) + \phi_\varepsilon(x_0)$ of x_0 depends only on u_0. Furthermore the image can be set to any desired point by choosing u_0 appropriately – we say then that it depends only and nontrivially on u_0. If x_0, x_1, and u_0 are fixed, the image $f_\varepsilon(x_1)$ of x_1 depends only on u_1, and as long as $x_0 \neq x_1$ it depends nontrivially on u_1. More generally for $0 \leq k \leq n-1$, if distinct points $\{x_j\}_{j=0}^{k}$ and coefficients $\{u_j\}_{j=0}^{k-1}$ are fixed, then the image $f_\varepsilon(x_k)$ of x_k depends only and nontrivially on u_k.

Suppose now that an n-tuple of points $\{x_j\}_{j=0}^{n}$ not on the diagonal $\Delta_n(I)$ and Newton coefficients $\{u_j\}_{j=0}^{n-1}$ are fixed. Then derivative $f'_\varepsilon(x_0)$ at x_0 depends only and nontrivially on u_n. Likewise for $0 \leq k \leq n-1$, if distinct points $\{x_j\}_{j=0}^{n-1}$ and Newton coefficients $\{u_j\}_{j=0}^{n+k-1}$ are fixed, then the derivative $f'_\varepsilon(x_k)$ at x_k depends only and nontrivially on u_{n+k}.

As Figure 2 illustrates, these considerations show that for any map f and any desired trajectory of distinct points with any given derivatives along it, one can choose Newton coefficients $\{u_k\}_{k=0}^{2n-1}$ and explicitly construct a map $f_\varepsilon = f + \phi_\varepsilon$ with such a

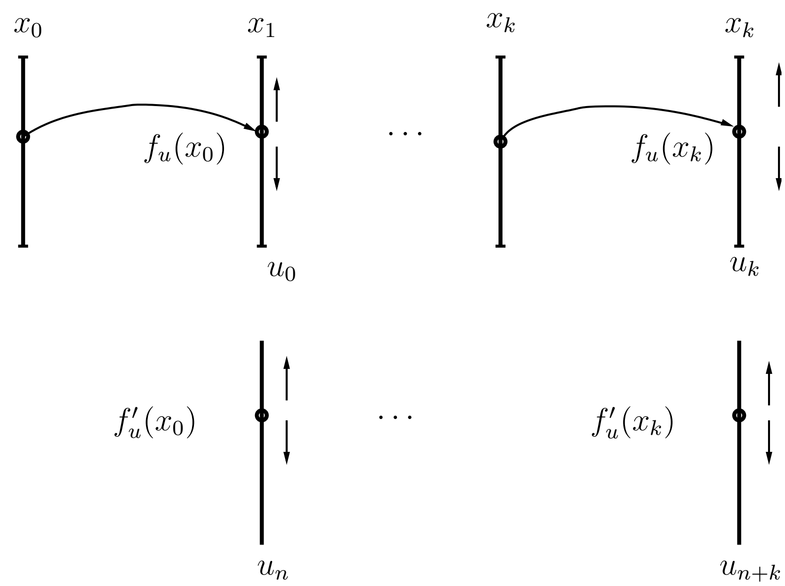

Figure 2.2: Newton Coefficients and their action

trajectory. Thus we have shown that $\pi^{1,n}$ is invertible away from the diagonal $\Delta_n(I)$ and defines a blow-up along it in the space of multijets $\mathcal{J}^{1,n}(I,\mathbb{R})$.

Next we define the function $\mathcal{D}^{1,n} : I^n \times \mathbb{R}^{2n} \to \mathcal{DD}^{1,n}(I,\mathbb{R})$ explicitly using so-called divided differences. Let $g : \mathbb{R} \to \mathbb{R}$ be a C^r function of one real variable.

Definition 2.2.1. *The first order divided difference of g is defined as*

$$\Delta g(x_0, x_1) = \frac{g(x_1) - g(x_0)}{x_1 - x_0} \tag{2.18}$$

for $x_1 \neq x_0$ and extended by its limit value as $g'(x_0)$ for $x_1 = x_0$. Iterating this construction we define divided differences of the m-th order for $2 \leq m \leq r$,

$$\Delta^m g(x_0, \ldots, x_m) = \frac{\Delta^{m-1} g(x_0, \ldots, x_{m-2}, x_m) - \Delta^{m-1} g(x_0, \ldots, x_{m-2}, x_{m-1})}{x_m - x_{m-1}} \tag{2.19}$$

for $x_{m-1} \neq x_m$ and extended by its limit value for $x_{m-1} = x_m$.

A function loses at most one derivative of smoothness with each application of Δ, so $\Delta^m g$ is at least C^{r-m} if g is C^r. Notice that Δ^m is linear as a function of g, and one can show that it is a symmetric function of $x_0, \dots, x_m$; in fact, by induction it follows that

$$\Delta^m g(x_0, \dots, x_m) = \sum_{i=0}^{m} \frac{g(x_i)}{\prod_{j \neq i}(x_i - x_j)} \tag{2.20}$$

Another identity that is proved by induction will be more important for us, namely

$$\Delta^m \; x^k(x_0, \dots, x_m) = p_{k,m}(x_0, \dots, x_m), \tag{2.21}$$

where $p_{k,m}(x_0, \dots, x_m)$ is 0 for $m > k$ and for $m \leq k$ is the sum of all degree $k - m$ monomials in $x_0, \dots, x_m$ with unit coefficients,

$$p_{k,m}(x_0, \dots, x_m) = \sum_{r_0 + \dots + r_m = k-m} \; \prod_{j=0}^{m} x_j^{r_j}. \tag{2.22}$$

The divided differences are the right coefficients for the Newton interpolation formula. For all C^∞ functions $g : \mathbb{R} \to \mathbb{R}$ we have

$$\begin{aligned} g(x) =& \Delta^0 g(x_0) + \Delta^1 g(x_0, x_1)(x - x_0) + \dots \\ &+ \Delta^{n-1} g(x_0, \dots, x_{n-1})(x - x_0) \dots (x - x_{n-2}) \\ &+ \Delta^n g(x_0, \dots, x_{n-1}, x)(x - x_0) \dots (x - x_{n-1}) \end{aligned} \tag{2.23}$$

identically for all values of $x, x_0, \dots, x_{n-1}$. All terms of this representation are polynomial in x except for the last one which we view as a remainder term. The sum of the polynomial terms is the degree $(n-1)$ *Newton interpolation polynomial* for g at $\{x_k\}_{k=0}^{n-1}$. To obtain a degree $2n-1$ interpolation polynomial for g and its derivative at $\{x_k\}_{k=0}^{n-1}$, we simply use (2.23) with n replaced by $2n$ and the $2n$-tuple of points $\{x_{k(\text{mod } n)}\}_{k=0}^{2n-1}$.

Recall that $\mathcal{D}^{1,n}$ was defined implicitly by (2.15). We have described how to use divided differences to construct a degree $2n-1$ interpolating polynomial of the form on the right-hand side of (2.15) for an arbitrary C^∞ function g. Our interest then is in the case $g = \phi_\varepsilon$, which as a degree $2n-1$ polynomial itself will have no remainder term and coincide exactly with the interpolating polynomial. Thus $\mathcal{D}^{1,n}$ is given coordinate-by-coordinate by

$$\begin{aligned} u_m =& \Delta^m \left(\sum_{k=0}^{2n-1} \varepsilon_k x^k \right) (x_0, \dots, x_{m \ (mod \ n)}) \\ =& \varepsilon_m + \sum_{k=m+1}^{2n-1} \varepsilon_k p_{k,m}(x_0, \dots, x_{m \ (mod \ n)}) \end{aligned} \tag{2.24}$$

for $m = 0, \dots, 2n-1$.

Equation (2.24) defines a transformation $(u_0, \dots, u_{2n-1}) = \mathcal{L}^1_{\mathbf{X}_n}(\varepsilon)$ on $\mathbb{R}^{2n}$, where $\mathbf{X}_n = (x_0, \dots, x_{n-1}) \in I^n$. We call $\mathcal{L}^1_{\mathbf{X}_n}$ the *Newton map*. This map is simply a restriction of $\mathcal{D}^{1,n}$ to its final $2n$ coordinates:

$$\mathcal{D}^{1,n}(\mathbf{X}_n, \varepsilon) = (\mathbf{X}_n, \mathcal{L}^1_{\mathbf{X}_n}(\varepsilon)). \tag{2.25}$$

Notice that for fixed $\mathbf{X}_n$, the Newton map is linear and given by an upper triangular matrix with units on the diagonal. Hence it is Lebesgue measure-preserving and invertible, whether or not $\mathbf{X}_n$ lies on the diagonal $\Delta_n(I)$.

Furthermore, the Newton map $\mathcal{L}^1_{\mathbf{X}_n}$ preserves the class of *scaled Lebesgue product measures* introduced in (1.15). In general, a measure μ on $\mathbb{R}^{2n}$ is a scaled Lebesgue product measure if it is the product $\mu = \mu_0 \times \cdots \times \mu_{2n-1}$, where each μ_j is Lebesgue measure on $\mathbb{R}$ scaled by a constant factor (which may depend on the coordinate j). Since the $\mathcal{L}^1_{\mathbf{X}_n}$ only shears in coordinate directions, we have the following lemma.

Lemma 2.2.2. *The Newton map $\mathcal{L}^1_{\mathbf{X}_n}$ given by (2.24) preserves all scaled Lebesgue product measures.*

This lemma will be used in Chapter 3. In the next section, we will introduce the particular scaled Lebesgue product measure that the lemma will be applied to.

We call the basis of monomials

$$\prod_{j=0}^{k}(x - x_{j (\text{mod } n)}) \quad \text{for} \quad k = 0, \dots, 2n-1 \tag{2.26}$$

in the space of polynomials of degree $2n-1$ the *Newton basis* defined by the n-tuple $\{x_k\}_{k=0}^{n-1}$. The Newton map and the Newton basis, and their analogues in dimension N, are useful tools for perturbing trajectories and estimating the measure $\mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho})$ of "bad" parameter values $\vec{\varepsilon} \in HB^N(\mathbf{r})$.

2.3 Estimates of the measure of "bad" parameters and Fubini reduction to finite-dimensional families

We return now to the the general case of $C^{1+\rho}$ diffeomorphisms on $\mathbb{R}^N$. In order to bound $\mu^N_{\vec{\mathbf{r}}}\{B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)\}$ we decompose the infinite-dimensional Hilbert Brick

$HB^N(\vec{\mathbf{r}})$ into the direct sum of a finite-dimensional brick of polynomials of degree $2n-1$ in N variables and its orthogonal complement.

Recall that $\vec{\mathbf{r}} = (\{r_m\}_{m=0}^{\infty})$ denotes the nonincreasing sequence $\{r_m\}_{m\in\mathbb{Z}_+}$ of sizes of the Hilbert Brick. With the notation (1.11) and (1.12), define

$$\begin{aligned} HB^N_{<k}(\vec{\mathbf{r}}) =& \big\{\{\vec{\varepsilon}_m\}_{m<k} : \text{ for every } 0 \le m < k, \ \|\vec{\varepsilon}_m\|_m \le r_m\big\} \\ =& B^N_0(r_0) \times \cdots \times B^N_{k-1}(r_{k-1}) \subset W_{0,N} \times W_{1,N} \times \cdots \times W_{k-1,N}; \\ HB^N_{\ge k}(\vec{\mathbf{r}}) =& \big\{\{\vec{\varepsilon}_m\}_{m\ge k} : \text{ for every } m \ge k, \ \|\vec{\varepsilon}_m\|_m \le r_m\big\} \\ =& B^N_k(r_k) \times B^N_{k+1}(r_{k+1}) \times \cdots \subset W_{k,N} \times W_{k+1,N} \times \dots; \\ HB^N(\vec{\mathbf{r}}) =& HB^N_{<k}(\vec{\mathbf{r}}) \oplus HB^N_{\ge k}(\vec{\mathbf{r}}). \end{aligned} \tag{2.27}$$

Each parameter $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ has a unique decomposition into

$$\begin{aligned} \vec{\varepsilon} =& (\vec{\varepsilon}_{<k}, \vec{\varepsilon}_{\ge k}) \in HB^N_{<k}(\vec{\mathbf{r}}) \oplus HB^N_{\ge k}(\vec{\mathbf{r}}), \\ \phi_{\vec{\varepsilon}}(x) =& \phi_{\vec{\varepsilon}_{<k}}(x) + \phi_{\vec{\varepsilon}_{\ge k}}(x) = \sum_{|\alpha|<k} \vec{\varepsilon}_\alpha x^\alpha + \sum_{|\alpha|\ge k} \vec{\varepsilon}_\alpha x^\alpha, \end{aligned} \tag{2.28}$$

where $\phi_{\vec{\varepsilon}_{<k}}(x)$ is a vector-polynomial of degree $k-1$ and $\phi_{\vec{\varepsilon}_{\ge k}}(x)$ is an analytic function with all Taylor coefficients of order less than k being equal to zero. Recall the notation (1.15), and decompose the measure $\mu^N_{\vec{\mathbf{r}}}$ on the brick $HB^N(\vec{\mathbf{r}})$ into the product

$$\mu^N_{<k,\vec{\mathbf{r}}} = \times_{m=0}^{k-1} \, \mu^N_{m,r_m}, \ \ \mu^N_{\ge k,\vec{\mathbf{r}}} = \times_{m=k}^{\infty} \, \mu^N_{m,r_m}, \ \ \mu^N_{\vec{\mathbf{r}}} = \mu^N_{<k,\vec{\mathbf{r}}} \times \mu^N_{\ge k,\vec{\mathbf{r}}}. \tag{2.29}$$

Thus, each component of the decomposition of the brick $HB^N_{<k}(\vec{\mathbf{r}})$ (resp. $HB^N_{\ge k}(\vec{\mathbf{r}})$) is supplied with the Lebesgue product probability measure $\mu^N_{<k,\vec{\mathbf{r}}}$ (resp. $\mu^N_{\ge k,\vec{\mathbf{r}}}$). Denote by

$$W_{<k,N} = \times_{m=0}^{k-1} \, W_{m,N}, \quad W_{\ge k,N} = \times_{m=k}^{\infty} \, W_{m,N} \tag{2.30}$$

the spaces to which the brick $HB^N_{<k}(\vec{\mathbf{r}})$ and the Hilbert Brick $HB^N_{\ge k}(\vec{\mathbf{r}})$ belong.

Consider the decomposition with $k = 2n-1$. Suppose we can get an estimate

$$\mu^N_{<2n,\vec{\mathbf{r}}}\{B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{\ge 2n})\} \le \mu_n(C,\delta,\rho,M_{1+\rho}) \tag{2.31}$$

of the measure of the "bad" set

$$\begin{aligned} B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{\ge 2n}) \ &= \ \{\vec{\varepsilon}_{<2n} \in HB^N_{<2n}(\vec{\mathbf{r}}) : \\ f_{\vec{\varepsilon}} \in IH(n-1,C,\delta,\rho) \text{ but } f_{\vec{\varepsilon}} &\notin IH(n,C,\delta,\rho)\}. \end{aligned} \tag{2.32}$$

in each slice $HB^N_{<2n}(\vec{\mathbf{r}}) \times \{\vec{\varepsilon}_{\ge 2n}\} \subset HB^N(\vec{\mathbf{r}})$, uniformly over $\vec{\varepsilon}_{\ge 2n} \in HB^N_{\ge 2n}(\vec{\mathbf{r}})$. Then by the Fubini/Tonelli Theorem and by the choice of the probability measure (2.29), estimate (2.31) implies (2.3). Thus we reduce the problem of estimating measure of the "bad" set (2.1) in the infinite-dimensional Hilbert Brick $HB^N(\vec{\mathbf{r}})$ to estimating measure of the "bad" set (2.32) in the finite-dimensional brick $HB^N_{<2n}(\vec{\mathbf{r}})$ of vector-polynomials of degree $2n-1$. Now our main goal is to get an estimate for the right-hand side of (2.31).

Fix a parameter value $\vec{\varepsilon}_{\geq 2n} \in HB^N_{\geq 2n}(\vec{\mathbf{r}})$ and the corresponding parameter slice $HB^N_{<2n}(\vec{\mathbf{r}}) \times \{\vec{\varepsilon}_{\geq 2n}\}$ in the Hilbert Brick $HB^N(\vec{\mathbf{r}})$. Let $\tilde{f} = f_{(0,\vec{\varepsilon}_{\geq 2n})}$ be the center of this slice. In this slice we have the finite-parameter family

$$\{\tilde{f}_{\vec{\varepsilon}_{<2n}}\}_{\vec{\varepsilon}_{<2n} \in HB^N_{<2n}(\vec{\mathbf{r}})} = \{f_{(\vec{\varepsilon}_{<2n},\vec{\varepsilon}_{\geq 2n})}\}_{\vec{\varepsilon}_{<2n} \in HB^N_{<2n}(\vec{\mathbf{r}})} \tag{2.33}$$

of perturbations by polynomials of degree $2n-1$. This is the family we consider at the n-th stage of the induction. We redenote the "bad" set of parameter values $B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{\geq 2n})$ by $B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f})$.

2.4 Simple trajectories and the Inductive Hypothesis

Based on the discussion in Section 2.1, we make the following definition.

Definition 2.4.1. *A trajectory $x_0,\dots,x_{n-1}$ of length n of a diffeomorphism $f \in \mathrm{Diff}^r(B^N)$, where $x_k = f^k(x_0)$, is called (n,γ)-simple if*

$$\prod_{k=0}^{n-2} |x_{n-1} - x_k| \geq \gamma^{1/4N}. \tag{2.34}$$

A point x_0 is called (n,γ)-simple if its trajectory $\{x_k = f^k(x_0)\}_{k=0}^{n-1}$ of length n is (n,γ)-simple. Otherwise a point (resp. a trajectory) is called non-(n,γ)-simple.

If a trajectory is simple, then perturbation of this trajectory by Newton Interpolation Polynomials is effective as the Closing Lemma and Perturbation of hyperbolicity examples of Section 2.1 show. To evaluate the product of distances it is important to choose *a "good" starting point* x_0 of an almost periodic trajectory $\{x_k\}_k$ in order to have the largest possible value of the product in (2.34); for some starting points the product of distances may be artificially small.

Consider the following example of a homoclinic intersection: Let $f : B^2 \hookrightarrow B^2$ be a diffeomorphism with a hyperbolic saddle point at the origin $f(0) = 0$. Suppose that the stable manifold $W^s(0)$ and the unstable manifold $W^u(0)$ intersect at some point $q \in W^s(0) \cap W^u(0)$. Then for a sufficiently large n there is a periodic point x of period n in a neighborhood of q going once nearby 0. It is clear that the trajectory $\{f^k(x)\}_{k=1}^n$ spends a lot of time in a neighborhood of the origin. Choose two starting points $x_0' = f^{k'}(x)$ and $x_0'' = f^{k''}(x)$ for the product (2.34). If x_0' is not in an $\exp(-\varepsilon n)$-neighborhood of the origin for some $\varepsilon > 0$, but x_0'' is, then it might happen that $\prod_{k=0}^{n-2} |f^{n-1}(x_0') - f^k(x_0')| \sim \exp(-\delta n)$ and $\prod_{k=0}^{n-2} |f^{n-1}(x_0'') - f^k(x_0'')| \sim \exp(-\delta' n^2)$ for some $\delta, \delta' > 0$. Indeed, if we pick out of n points of $\{f^k(x)\}_{k=1}^n$ only $n/2$ closest to the origin, then simple calculation shows that all of them are in $\exp(-\varepsilon n)$-neighborhood of the origin. ε is some positive number depending on the eigenvalues of $df(0)$. So the first product might be significantly larger than the second one. This is because the trajectory $\{f^k(x_0")\}_{k=0}^{n-1}$ has many points in a neighborhood of the origin and all of the corresponding terms in the product are small. This shows that sometimes the product of distances along a trajectory (2.34) might be small not because the trajectory is too recurrent, but because we choose a "bad" starting point. This motivates the following definition.

Definition 2.4.2. *A point x is called essentially (n, γ)-simple if for some nonnegative $j < n$, the point $f^j(x)$ is (n, γ)-simple. Otherwise a point is called essentially non-(n, γ)-simple.*

Let us return to the strategy of the proof of Theorem 1.3.7. At the n-th stage of the induction over the period we consider the family of polynomial perturbations $\{\tilde{f}_{\vec{\varepsilon}_{<2n}}\}_{\vec{\varepsilon}_{<2n} \in HB^N_{<2n}(\vec{\mathbf{r}})}$ of the form (2.33) of the diffeomorphism $\tilde{f} \in \mathrm{Diff}^{1+\rho}(B^N)$ by polynomials of degree $2n-1$. Consider among them only diffeomorphisms $\tilde{f}_{\vec{\varepsilon}_{<2n}}$ that satisfies Inductive Hypothesis of order $n-1$ with constants (C, δ, ρ), i.e. $\tilde{f}_{\vec{\varepsilon}_{<2n}} \in IH(n-1, C, \delta, \rho)$ as we proposed earlier. To simplify notations we redenote the set $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f, \vec{\varepsilon}_{\geq 2n})$ by $B_n(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f})$ with $\tilde{f} = f_{\vec{\varepsilon}_{\geq 2n}}$. Our main goal is to estimate measure of "bad" parameter values $B_n(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f})$, given by (2.32), for which the corresponding diffeomorphism has an $(n, \gamma_n^{1/\rho}(C, \delta))$-periodic, but not $(n, \gamma_n(C, \delta))$-hyperbolic, point $x \in B^N$.

We split the set of all possible almost periodic points of period n into two classes: *essentially* $(n, \gamma_n^{1/N}(C, \delta))$*-simple* and *essentially non-*$(n, \gamma_n^{1/N}(C, \delta))$*-simple*. Decompose the set of "bad" parameters $B_n(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f})$ into two sets of "bad" parameters with simple and nonsimple almost periodic points that are not sufficiently hyperbolic:

$$\begin{aligned} B_n^{\mathrm{sim}}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f}) = \{&\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}}) : \tilde{f}_{\vec{\varepsilon}_{<2n}} \in IH(n-1, C, \delta, \rho), \\ &\tilde{f}_{\vec{\varepsilon}_{<2n}} \text{ has an } (n, \gamma_n^{1/\rho}(C, \delta))\text{-periodic, essentially} \\ &(n, \gamma_n^{1/N}(C, \delta))\text{-simple, but not } (n, \gamma_n(C, \delta))\text{-hyperbolic point } x\} \end{aligned} \tag{2.35}$$

and

$$\begin{aligned}&B_n^{\rm non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f)=\{\vec\varepsilon\in HB^N(\vec{\mathbf{r}}):\tilde f_{\vec\varepsilon_{<2n}}\in IH(n-1,C,\delta,\rho),\\&\tilde f_{\vec\varepsilon_{<2n}}\text{ has an }(n,\gamma_n^{1/\rho}(C,\delta))\text{-periodic, essentially}\\&\text{non-}(n,\gamma_n^{1/N}(C,\delta))\text{-simple, but not }(n,\gamma_n(C,\delta))\text{-hyperbolic point }x\}\end{aligned} \tag{2.36}$$

It is clear that we have

$$B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde f)=B_n^{\rm sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f)\cup B_n^{\rm non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f). \tag{2.37}$$

We shall estimate the measures of the sets of simple orbits $B_n^{\rm sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f)$ and nonsimple orbits $B_n^{\rm non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f)$ separately, but using very similar methods.

Let $\tilde f_{\vec\varepsilon_{<2n}}\in IH(n-1,C,\delta,\rho)$ be a diffeomorphism that satisfies the Inductive Hypothesis of order $n-1$ with constants (C,δ,ρ). It turns out that if $\tilde f_{\vec\varepsilon_{<2n}}$ has an $(n,\gamma_n^{1/\rho}(C,\delta))$-periodic and essentially non-$(n,\gamma_n^{1/N}(C,\delta))$-simple point x_0, then the trajectory of x_0 has a close return $\tilde f^k_{\vec\varepsilon_{<2n}}(x_0)=x_k$ for $k<n$ such that distance $|x_0-x_k|$ is much smaller of all the previous $|x_0-x_j|,\ 1\le j<k$. Let us formulate more precisely what we mean here by “much smaller”.

Definition 2.4.3. *Let $g\in\mathrm{Diff}^{1+\rho}(B^N)$ be a diffeomorphism and let $D>1$ be some number. A point $x_0\in B^N$ (resp. a trajectory $x_0,\dots,x_{n-1}=g^{n-1}(x_0)\subset B^N$ of length n) has a weak (D,n)-gap at a point $x_k=g^k(x_0)$ if*

$$|x_k-x_0|\le D^{-n}\min_{0<j\le k-1}|x_0-x_j|. \tag{2.38}$$

and there is no $m<k$ such that x_0 has a weak (D,n)-gap at $x_m=g^m(x_0)$.

Remark 2.4.4. *The term “gap” arises by considering of the sequence $-\log|x_0-x_1|,\ -\log|x_0-x_2|,\dots,-\log|x_0-x_k|$. Definition 2.4.3 implies that the last term is significantly larger then all the previous terms.*

Let’s show that n should be divisible by k for an almost periodic point of period n with a weak gap at x_k. This feature of a weak gap allows us to treat almost periodic trajectories of length n with a weak gap at x_k as n/k almost identical parts of length k each.

Lemma 2.4.5. *Let $g\in\mathrm{Diff}^{1+\rho}(B^N)$ be a diffeomorphism, M_1 be an upper bound on C^1-norm of g and g^{-1}, $D>M_1^2$, and let x_0 have a weak (D,n)-gap at x_k and $|x_0-x_n|\le|x_0-x_k|$. Then n is divisible by k.*

Sketch of Proof: Denote by $\gcd(k,n)$ the greatest common divisor of k and n. Then using the bound on C^1-norm of g and g^{-1} for any $x, y \in B^N$ we have

$$M_1^{-1}\ |g^{-1}(x) - g^{-1}(y)| \le |x-y| \le M_1\ |g(x) - g(y)| \tag{2.39}$$

Using the Euclidean division algorithm developed in Part II of this paper, one can show that

$$|x_0 - x_{\gcd(k,n)}| \le M_1^{2n} D^{-n} \min_{0<j\le k-1} |x_0 - x_j|. \tag{2.40}$$

Since $D > M_1^2$, we cannot have $\gcd(k,n) < k$, so n must be divisible by k. Q.E.D.

In Part II of this paper we prove the following result.

Theorem 2.4.6. *Let $g \in \mathrm{Diff}^{1+\rho}(B^N)$ be a diffeomorphism for some $\rho > 0$ and satisfy the Inductive Hypothesis of order $n-1$ with constants (C,δ,ρ), i.e. $g \in IH(n-1, C, \delta, \rho)$ and let $M_{1+\rho} = \max\{\|g^{-1}\|_{C^1},\ \|g\|_{C^{1+\rho}},\ 2^{1/\rho}\}$, $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$, and $D = \max\{M_{1+\rho}^{30/\rho}, \exp(C/100)\}$. Suppose the diffeomorphism g has an $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic and essentially non-$(n, \gamma_n^{1/N}(C,\delta))$-simple point $x_0 \in B^N$. Then either x_0 is $(n, \gamma_n(C,\delta))$-hyperbolic or x_0 has a weak (D,n)-gap at $x_k = g^k(x_0)$ for some k dividing n and x_j is $(k, \gamma_n^{1/N}(C,\delta))$-simple for some $j < n$.*

Remark 2.4.7. *As a matter of fact we need a sharper result, but Theorem 2.4.6 is a nice starting point.*

Theorem 2.4.6 implies that the set of "bad" parameters with an essentially non-simple trajectory can be decomposed into the following finite union: Define the set of parameters with an almost periodic point of period n with a weak gap at the k-th point of its trajectory.

$$\begin{aligned} &B_n^{\mathrm{wgap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f; D) = \{\vec\varepsilon \in HB^N(\vec{\mathbf{r}}) : \tilde f_{\vec\varepsilon} \in IH(n-1,C,\delta,\rho),\\ &\tilde f_{\vec\varepsilon} \text{ has an } (n, \gamma_n^{1/\rho}(C,\delta))\text{-periodic, but not}\\ &(n,\gamma_n(C,\delta))\text{-hyperbolic point } x_0 \text{ with a weak } (D,n)\text{-gap at } x_k = \tilde f_{\vec\varepsilon}^k(x)\} \end{aligned} \tag{2.41}$$

Then for $D = \max\{M_{1+\rho}^{30/\rho}, \exp(C/100)\}$, Theorem 2.4.6 gives

$$B_n^{\mathrm{non}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f) \subseteq \left(\cup_{k|n} B_n^{\mathrm{wgap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f; D)\right). \tag{2.42}$$

Combining inclusions (2.35) and (2.42), we have

$$\begin{aligned} &B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde f) \subseteq\\ &B_n^{\mathrm{sim}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f) \cup \left(\cup_{k|n} B_n^{\mathrm{wgap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde f; D)\right). \end{aligned} \tag{2.43}$$

Thus we need to get estimates on the measures of bad parameters associated with essentially simple trajectories $B_n^{\mathrm{sim}}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f})$ and trajectories with a weak gap $B_n^{\mathrm{wgap}(k)}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f}; D)$, where k divides n. In Chapter 3, we describe the Discretization method for the 1-dimensional model problem. This method will allow us to estimate the measure of parameters $B_n^{\mathrm{sim}}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f})$ associated with simple almost periodic points. At the end of Chapter 3, we show how using the Discretization method one can estimate the measure of parameters $B_n^{\mathrm{wgap}(k)}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f}; D)$ associated with almost periodic trajectories with a weak gap. Loosely speaking, it is because those trajectories have simple parts of length k (see the end of Theorem 2.4.6), and hyperbolicity of simple part of length k enforces hyperbolicity of the trajectories of length n.

Chapter 3

A Model Problem: C^2-smooth Maps of the Interval $I = [-1, 1]$

In Section 2.4 we concluded that the key to the proof of Theorem 1.3.7 (which implies the Main Theorem) is to get an estimate of the measure of "bad" parameters. Recall that the set of "bad" parameters (2.32) consists of those parameters $\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})$ for which the corresponding diffeomorphism $f_{\vec{\varepsilon}} : B^N \hookrightarrow B^N$ has an almost periodic point x of period n that is not sufficiently hyperbolic. In this chapter we present a detailed discussion of C^2-smooth 1-dimensional noninvertible maps ($N = 1$ and $\rho = 1$) with a Hilbert Brick of a "nice" size. This 1-dimensional model gives a useful insight into the general approach of estimating the measure of "bad" parameters for the N-dimensional $C^{1+\rho}$-smooth diffeomorphisms and allows us to avoid several technical complications that will arise in Part II of this paper [K5]. These complications are outlined in the next chapter.

3.1 Setting up of the model

Let $C^2(I, I)$ be the space of C^2-smooth maps of the interval $I = [-1, 1]$ into its interior. Consider a C^2-smooth map of the interval $f \in C^2(I, I)$ and the family of perturbations of f by analytic functions represented as their power series

$$f_\varepsilon(x) = f(x) + \sum_{k=0}^{\infty} \varepsilon_k x^k. \tag{3.1}$$

Fix a positive $\tau > 0$. Define a range of parameters of this family in the form of a Hilbert Brick

$$HB^{\mathrm{st}}(\tau) = \Big\{ \{\varepsilon_m\}_{m=0}^{\infty} : \ \forall\ m \geq 0,\ |\varepsilon_m| < \frac{\tau}{m!} \Big\}. \tag{3.2}$$

We call $HB^{\mathrm{st}}(\tau)$ *a Hilbert Brick of standard thickness* with width τ. If we choose τ small enough, then the whole family $\{f_\varepsilon\}_{\varepsilon \in HB^{\mathrm{st}}(\tau)} \subset C^2(I,I)$ consists of C^2-smooth maps of the interval I.

Define the Lebesgue product probability measure, denoted by μ_τ^{st}, on the Hilbert Brick of parameters $HB^{\mathrm{st}}(\tau)$ by normalizing the 1-dimensional Lebesgue measure along each component to the 1-dimensional Lebesgue probability measure

$$\mu_{m,\tau}^{\mathrm{st}} = \left(\frac{m!}{2\tau}\right) Leb_1, \quad \mu_{<k,\tau}^{\mathrm{st}} = \times_{m=0}^{k-1}\ \mu_{m,\tau}^{\mathrm{st}} \quad \mu_{\tau}^{\mathrm{st}} = \times_{m=0}^{\infty}\ \mu_{m,\tau}^{\mathrm{st}}. \tag{3.3}$$

The main result of this Chapter is the following 1-dimensional analogue of Theorem 1.3.7.

Theorem 3.1.1. *Let $f \in C^2(I,I)$ be a C^2-smooth map of the interval I into its interior and let $\tau > 0$ be so small that the family of analytic perturbations $\{f_\varepsilon\}_{\varepsilon \in HB^{\mathrm{st}}(\tau)} \subset C^2(I,I)$ consists of C^2-smooth maps of the interval I. Then for any $\delta > 0$ and μ_τ^{st}-a.e. $\varepsilon \in HB^{\mathrm{st}}(\tau)$ there exists $C = C(\varepsilon,\delta) > 0$ such that the number of periodic points $P_n(f_\varepsilon)$ of f_ε of period n and their minimal hyperbolicity $\gamma_n(f_\varepsilon)$, defined in (1.7), for all $n \in \mathbb{Z}_+$ satisfy*

$$\gamma_n(f_\varepsilon) > \exp(-Cn^{1+\delta}), \quad P_n(f_\varepsilon) < \exp(Cn^{1+\delta}). \tag{3.4}$$

The strategy for the proof of this Theorem is the same as the strategy of the proof of Theorem 1.3.7 described in Chapter 2. Denote the supremum C^2 and C^1-norms of the family (3.1)

$$M_1 = \sup_{\varepsilon \in HB^{\mathrm{st}}(\tau)} \{\|f_\varepsilon\|_{C^1}\}, \quad M_2 = \sup_{\varepsilon \in HB^{\mathrm{st}}(\tau)} \{\|f_\varepsilon\|_{C^2}, M_1, 2\} \tag{3.5}$$

By analogy with the direct decomposition of the Hilbert Brick in the N-dimensional case (2.27), for each positive integer $k \in \mathbb{Z}_+$ define the direct decomposition of the Hilbert Brick of standard thickness $HB^{\mathrm{st}}(\tau)$

$$\begin{aligned} HB_{<k}^{\mathrm{st}}(\tau) =& \Big\{ \{\varepsilon_m\}_{m=0}^{k-1} : \ \forall\ 0 \leq m < k,\ |\varepsilon_m| < \frac{\tau}{m!} \Big\}. \\ HB_{\geq k}^{\mathrm{st}}(\tau) =& \Big\{ \{\varepsilon_m\}_{m=k}^{\infty} : \ \forall\ m \geq k,\ |\varepsilon_m| < \frac{\tau}{m!} \Big\}. \end{aligned} \tag{3.6}$$

We call $HB^{\rm st}_{<k}(\tau)$ *a (k-dimensional) Brick of standard thickness* with width τ. The product measure $\mu^{\rm st}_\tau$ on the whole Hilbert Brick $HB^{\rm st}(\tau)$ induces the measure of product of Lebesgue probability $\mu^{\rm st}_{<k,\tau}$ on the k-dimensional Brick $HB^{\rm st}_{<k}(\tau)$.

Fix $n \in \mathbb{Z}_+$ and consider the n-th stage of the induction over the period (see the beginning of Chapter 2). Let

$$\tilde f(x) = f(x) + \sum_{k=2n}^{\infty} \varepsilon_k x^k \tag{3.7}$$

for some $\{\varepsilon_k\}_{k=2n}^\infty \in HB^{\rm st}_{\ge 2n}(\tau)$, and consider the $2n$-parameter family of perturbations by polynomials of degree $2n-1$ with coefficients in the brick of standard thickness $HB^{\rm st}_{<2n}(\tau)$,

$$\tilde f_\varepsilon(x) = \tilde f(x) + \sum_{k=0}^{2n-1} \varepsilon_k x^k, \quad \varepsilon = (\varepsilon_0, \dots, \varepsilon_{2n-1}) \in HB^{\rm st}_{<2n}(\tau). \tag{3.8}$$

The bounds M_1 and M_2 from (3.5) apply to this subfamily of (3.1).

Using the Fubini reduction to finite-dimensional families from Section 2.3 right after (2.31), for the proof of Theorem 3.1.1 it is sufficient to estimate the measure of "bad" parameters in each such a family.

To fit the notations of our model we choose a sufficiently small positive γ_n and we introduce sets of all "bad" parameters (compare with (2.32))

$$\begin{aligned} &B_{n,\tau}(C,\delta,\tilde f,\gamma_n) = \{\varepsilon \in HB^{\rm st}_{<2n}(\tau) : \tilde f_\varepsilon \in IH(n-1,C,\delta,1), \\ &\tilde f_\varepsilon \text{ has an } (n,\gamma_n)\text{-periodic, but not } (n,\gamma_n)\text{-hyperbolic point } x_0\}, \end{aligned} \tag{3.9}$$

and define the sets $B^{\rm sim}_{n,\tau}(C,\delta,\tilde f,\gamma_n)$ and $B^{\rm non}_{n,\tau}(C,\delta,\tilde f,\gamma_n)$ of "bad" parameters with essentially simple (respectively nonsimple) trajectories as in (2.35) and (2.36):

$$\begin{aligned} &B^{\rm sim}_{n}(C,\delta,\tilde f,\gamma_n) = \{\vec\varepsilon \in HB^{\rm st}_{<2n}(\tau) : \tilde f_\varepsilon \in IH(n-1,C,\delta,1), \\ &\tilde f_\varepsilon \text{ has an } (n,\gamma_n)\text{-periodic, essentially} \\ &(n,\gamma_n)\text{-simple, but not } (n,\gamma_n)\text{-hyperbolic point } x_0\}, \end{aligned} \tag{3.10}$$

and

$$\begin{aligned} &B^{\rm non}_{n}(C,\delta,\tilde f,\gamma_n) = \{\vec\varepsilon \in HB^{\rm st}_{<2n}(\tau) : \tilde f_\varepsilon \in IH(n-1,C,\delta,1), \\ &\tilde f_\varepsilon \text{ has an } (n,\gamma_n)\text{-periodic, essentially} \\ &\text{non-}(n,\gamma_n)\text{-simple, but not } (n,\gamma_n)\text{-hyperbolic point } x_0\}. \end{aligned} \tag{3.11}$$

For sufficiently small γ_n, e.g., $\gamma_n \leq \gamma_n(C, \delta)$, similarly to (2.37) we have the following decomposition

$$B_{n,\tau}(C, \delta, \tilde{f}, \gamma_n) \;=\; B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n) \cup B^{\rm non}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n). \tag{3.12}$$

The main result of the next three sections is the following estimate.

Proposition 3.1.2. *Let $\{\tilde{f}_\varepsilon\}_{\varepsilon \in HB^{\rm st}_{<2n}(\tau)}$ be the family of polynomial perturbations (3.8) with bound M_2 on the C^2-norm. Then with the notation above, for any $C > 2$, $\delta > 0$, and $\tau > 0$ and a sufficiently small positive γ_n, e.g., $\gamma_n \leq \gamma_n(C, \delta)$, the following estimate on the measure of parameters associated with maps $\tilde{f}_\varepsilon$ with an (n, γ_n)-periodic, essentially (n, γ_n)-simple, but not (n, γ_n)-hyperbolic, point holds*

$$\mu^{\rm st}_{<2n,\tau}\big\{B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)\big\} \leq 6^{2n} M_2^{6n+1}\, \frac{(n-1)!}{\tau}\, \frac{(2n-1)!}{\tau}\, \gamma_n^{1/4}. \tag{3.13}$$

It is clear that for any $C > 0$ and $\delta > 0$, if $\gamma_n = \exp(-Cn^{1+\delta})$, then the right-hand side of (3.13) tends to 0 as $n \to \infty$ superexponentially fast in n. An estimate on the measure of essentially nonsimple trajectories $\mu^{\rm st}_{<2n,\tau}\big\{B^{\rm non}_{n}(C, \delta, \tilde{f}, \gamma_n)\big\}$ is obtained in Section 3.5, Proposition 3.5.1. Application of these two Propositions and arguments (2.3–2.4) will prove Theorem 3.1.1. The method of obtaining an estimate for $\mu^{\rm st}_{<2n,\tau}\big\{B^{\rm non}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)\big\}$ is similar to the one we shall develop now to prove (3.13).

3.2 Decomposition into pseudotrajectories

In this section, we decompose the set of "bad" parameters $B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)$ for which there exists a simple, almost periodic, but not sufficiently hyperbolic trajectory into a finite union of sets of "bad" parameters. Each set will be associated with a particular simple, almost periodic, but not sufficiently hyperbolic pseudotrajectory. In the next section we will estimate the measure of "bad" parameters associated with a *particular* trajectory, and in the subsequent section we will extend this estimate to the set of "bad" parameters associated with *all possible* simple trajectories, obtaining estimate (3.13).

Fix a sufficiently small $\gamma_n > 0$ and $\tilde{\gamma}_n = \gamma_n M_2^{-2n}$. Consider the $2\tilde{\gamma}_n$-grid in the interval I

$$I_{\tilde{\gamma}_n} = \{x \in I : \ \exists k \in \mathbb{Z} \text{ such that } x = (2k+1)\tilde{\gamma}_n\} \subset I \tag{3.14}$$

Definition 3.2.1. *We call a k-tuple* $\{x_j\}_{j=0}^{k-1} \in I_{\tilde{\gamma}_n}^k$ *a* $\tilde{\gamma}_n$*-pseudotrajectory associated to* ε *(or to the map* $\tilde{f}_\varepsilon$*) if for each* $j = 1, \dots, k-1$ *we have* $|\tilde{f}_\varepsilon(x_{j-1}) - x_j| \le \tilde{\gamma}_n$*, and we call it a* $\tilde{\gamma}_n$*-pseudotrajectory associated to* $HB^{\rm st}_{<2n}(\tau)$ *(or to the family* $\{\tilde{f}_\varepsilon\}_{\varepsilon \in HB^{\rm st}_{<2n}(\tau)}$*) if it is associated to some* $\varepsilon \in HB^{\rm st}_{<2n}(\tau)$*.*

A $\tilde{\gamma}_n$*-pseudotrajectory* $x_0, \dots, x_{n-1}$ *of length n associated to some parameter* $\varepsilon \in HB^{\rm st}_{<2n}(\tau)$ *for some* $\gamma > 0$ *is*

- (n, γ)*-periodic if* $|\tilde{f}_\varepsilon(x_{n-1}) - x_0| \le \gamma$*,*
- (n, γ)*-simple if* $\prod_{j=0}^{n-2} |x_{n-1} - x_j| \ge \gamma^{1/4}$*,*
- (n, γ)*-hyperbolic if* $\left| \prod_{j=0}^{n-1} \left| (\tilde{f}_\varepsilon)'(x_j) \right| - 1 \right| \ge \gamma$*.*

Remark 3.2.2. *For fixed* $\varepsilon \in HB^{\rm st}_{<2n}(\tau)$*, each initial point* $x_0 \in I_{\tilde{\gamma}_n}$ *generates a* $\tilde{\gamma}_n$*-pseudotrajectory* $\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_{n-1}$ *of length n as follows. For each successive* $k = 1, \dots, n-1$*, choose* $\tilde{x}_k \in I_{\tilde{\gamma}_n}$ *such that* $|\tilde{x}_k - f_\varepsilon(\tilde{x}_{k-1})| \le \tilde{\gamma}_n$*. Notice that this choice is unique unless* $\tilde{f}_\varepsilon(\tilde{x}_{k-1})$ *happens to lie halfway between two points of* $I_{\tilde{\gamma}_n}$*. It may be helpful in understanding the upcoming arguments to think of each initial point* $x_0 \in I_{\tilde{\gamma}_n}$ *as generating a unique* $\tilde{\gamma}_n$*-pseudotrajectory for a given* f_ε*, though for a measure zero set of* ε *there are exceptions to this rule. In fact, for our estimates it is important only that the number of* $\tilde{\gamma}_n$*-pseudotrajectories per initial point be bounded by an exponential function of n, which is true in this case even if there is a choice of two grid points at each iteration.*

We would like to contain the set of "bad" parameters $B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)$ in a finite collection of subsets each of "bad" parameters corresponding to a single $\tilde{\gamma}_n$-pseudotrajectory

$$\begin{aligned} &B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{n-1}) = \{\varepsilon \in HB^{\rm st}_{<2n}(\tau) : \\ &\{x_k\}_{k=0}^{n-1} \text{ is a } \tilde{\gamma}_n\text{-pseudotrajectory associated to } \varepsilon \text{ and is} \\ &\left(n, \frac{\gamma_n}{2}\right)\text{-simple and } (n, M_2^n \gamma_n)\text{-periodic, but not } (n, M_2^{3n}\gamma_n)\text{-hyperbolic}\} \end{aligned} \tag{3.15}$$

Introduce the union of all "bad" sets associated with $\tilde{\gamma}_n$-pseudotrajectories

$$B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2) = \cup_{\{x_0,\dots,x_{n-1}\} \subset I^n_{\tilde{\gamma}_n}} B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{n-1}). \tag{3.16}$$

Most of the sets in the right hand side are empty, and one of our goals is to estimate the number of nonempty ones.

In comparison to the definition (3.10) of $B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)$, we increase periodicity and hyperbolicity for pseudotrajectories and decrease simplicity. This will allow us to prove that

$$B^{\rm sim}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n) \subset B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2). \tag{3.17}$$

Intuitively this is true because each trajectory of length n can be approximated by a pseudotrajectory of length n which has almost the same periodicity, simplicity, and hyperbolicity as the original one. We will make this argument precise at the end of Section 3.4.

Remark 3.2.3. *Unlike $B^{\text{sim}}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)$, we do not assume in the definition (3.15) of $B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2)$ that $\tilde{f}_\varepsilon \in IH(n-1, C, \delta, 1)$. This is because we only need the Inductive Hypothesis to estimate the measure of "bad" parameters in the case of nonsimple trajectories.*

Our goal is then to estimate the measure $\mu^{\text{st}}_{<2n,\tau}\{B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2)\}$ in order to prove Proposition 3.1.2. Loosely speaking, this measure will be estimated in two steps:

Step 1. Estimate the number of different $\tilde{\gamma}_n$-pseudotrajectories $\#_n(\tilde{\gamma}_n, \tau)$ associated to $HB^{\text{st}}_{<2n}(\tau)$;

Step 2. Estimate the measure

$$\mu^{\text{st}}_{<2n,\tau}\{B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{n-1})\} \leq \mu_n(M_2, \gamma_n, \tilde{\gamma}_n, \tau) \tag{3.18}$$

uniformly for an (n, γ_n)-simple $\tilde{\gamma}_n$-pseudotrajectory $\{x_0, \dots, x_{n-1}\} \in I^n_{\tilde{\gamma}_n}$.

Then the product of two numbers $\#_n(\tilde{\gamma}_n, \tau)$ and $\mu_n(M_2, \gamma_n, \tilde{\gamma}_n, \tau)$ that are obtained in Steps 1 and 2 gives the required estimate (3.13).

Actually the procedure of estimating $\mu^{\text{st}}_{<2n,\tau}\{B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2)\}$ is a little more complicated. Based on the definition (3.15) of the set $B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{n-1})$ of parameters ε for which the diffeomorphism f_ε has a prescribed $\tilde{\gamma}_n$-pseudotrajectory $\{x_0, \dots, x_{n-1}\} \in I^n_{\tilde{\gamma}_n}$ that is almost periodic and not sufficiently hyperbolic, define a set of parameters ε for which only a part of the $\tilde{\gamma}_n$-pseudotrajectory $\{x_0, \dots, x_{m-1}\} \in I^m_{\tilde{\gamma}_n}$ is prescribed for f_ε:

$$\begin{aligned} &B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{m-1}) = \{\varepsilon \in HB^{\text{st}}_{<2n}(\tau) : \text{ there exist} \\ &x_m, \dots, x_{n-1} \in I_{\tilde{\gamma}_n} \text{ such that } \{x_j\}_{j=0}^{n-1} \text{ is a } \tilde{\gamma}_n\text{-pseudotrajectory} \\ &\text{associated to } \varepsilon, \text{ and } \varepsilon \in B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{n-1})\}. \end{aligned} \tag{3.19}$$

For each $m = 1, 2, \dots, n-1$ it is clear that

$$B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_{m-1}) = \cup_{x_m \in I_{\tilde{\gamma}_n}} B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_m). \tag{3.20}$$

Inductive application of this formula to the definition (3.16) yields

$$B_{n,\tau}^{\mathrm{sim},\tilde{\gamma}_n}(\tilde{f}, \gamma_n, M_2) = \cup_{\tilde{x}_0 \in I_{\tilde{\gamma}_n}} B_{n,\tau}^{\mathrm{sim},\tilde{\gamma}_n}(\tilde{f}, \gamma_n, M_2; \tilde{x}_0). \tag{3.21}$$

The estimate of Step 1 then breaks down as follows:

$$\#_n(\tilde{\gamma}_n, \tau) \approx \boxed{\begin{array}{c}\text{\# of initial}\\ \text{points of } I_{\tilde{\gamma}_n}\end{array}} \times \boxed{\begin{array}{c}\text{\# of } \tilde{\gamma}_n\text{-pseudotrajectories}\\ \text{per initial point}\end{array}} \tag{3.22}$$

And up to an exponential function of n, the estimate of Step 2 breaks down like:

$$\mu_n(M_2, \gamma_n, \tilde{\gamma}_n, \tau) \approx \frac{\boxed{\begin{array}{c}\text{Measure of}\\ \text{periodicity}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{hyperbolicity}\end{array}}}{\boxed{\begin{array}{c}\text{\# of } \tilde{\gamma}_n\text{-pseudotrajectories}\\ \text{per initial point}\end{array}}} \tag{3.23}$$

(Roughly speaking, the terms in the numerator represent respectively the measure of parameters for which a given initial point will be (n, γ_n)-periodic and the measure of parameters for which a given n-tuple is (n, γ_n)-hyperbolic; they correspond to estimates (3.30) and (3.33) in the next section.) Thus after cancellation, the estimate of the measure of "bad" set $B_{n,\tau}^{\mathrm{sim}}(C, \delta, \tilde{f}, \gamma_n)$ associated to simple, almost periodic, not sufficiently hyperbolic trajectories becomes:

$$\begin{array}{c}\boxed{\begin{array}{c}\text{Measure of bad}\\ \text{parameters}\end{array}} \le \\ \boxed{\begin{array}{c}\text{\# of initial}\\ \text{points of } I_{\tilde{\gamma}_n}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{periodicity}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{hyperbolicity}\end{array}}\end{array} \tag{3.24}$$

The first term on the right hand side of (3.24) is of order γ_n^{-1} (up to an exponential function in n). In Section 3.3, we will show that the second term is at most of order $n!\gamma_n^{3/4}$, and the third term is at most of order $(2n)!\gamma_n^{1/2}$, so that the product on the right-hand side of (3.24) is of order at most $n!(2n)!\gamma_n^{1/4}$ (up to an exponential function in n) and is superexponentially small in n. These bounds use the change of parameter coordinates by Newton interpolation polynomials that was introduced in Section 2.2, and they *do not depend on whether the parameters are associated with the brick* $HB_{<2n}^{\mathrm{st}}(\tau)$, except in that we use the bound M_2 on the C^2 norm of the maps involved.

In Section 3.4, we complete the proof of Proposition 3.1.2 by bounding the total measure of "bad" parameters for all pseudotrajectories associated to $HB^{\mathrm{st}}_{<2n}(\tau)$. Since we use the Fubini theorem in the Newton coordinates $u_0, \ldots, u_{2n-1}$, we need to know the maximum range of each of these parameters in the image of $HB^{\mathrm{st}}_{<2n}(\tau)$ under this coordinate change. In the "Distortion Lemma", we show that the image of $HB^{\mathrm{st}}_{<2n}(\tau)$ is contained in a brick 3 times as large in each direction. Then, in the "Collection Lemma", we show in effect that the cancellation in going from (3.22) and (3.23) to (3.24) is valid. In fact, the number of $\tilde{\gamma}_n$-pseudotrajectories for a given initial point may depend significantly on the initial point, and we do not bound it explicitly. Rather, we show that in the decomposition (3.21), the measure of each term $B^{\mathrm{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; \tilde{x}_0)$ is bounded (up to a factor exponential in n) by the product of the "measure of periodicity" and "measure of hyperbolicity" derived in Section 3.3, thus yielding a final estimate of the form (3.24).

3.3 Application of Newton interpolation polynomials to estimate the measure of "bad" parameters for a single trajectory

In this section we fix an n-tuple of points $\{x_j\}_{j=0}^{n-1} \in I^n$, denoted by $\mathbf{X}_n$, and estimate the measure of "bad" parameters $B^{\mathrm{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \ldots, x_{n-1})\big\}$ associated with *this particular* trajectory. Recall that $\tilde{\gamma}_n = M_2^{-2n}\gamma_n$.

Problem 3.3.1. *Estimate the measure of $\varepsilon \in HB^{\mathrm{st}}_{<2n}(\tau)$ for which the n-tuple $\{x_j\}_{j=0}^{n-1}$ is*

$$\begin{aligned} &A)\ \text{a } \tilde{\gamma}_n\text{-pseudotrajectory, i.e., } |\tilde{f}_\varepsilon(x_j) - x_{j+1}| \le \tilde{\gamma}_n \quad \text{for} \quad j = 0, \ldots, n-2; \\ &B)\ (n, \gamma_n)\text{-periodic, i.e., } |\tilde{f}_\varepsilon(x_{n-1}) - x_0| \le \gamma_n; \text{ and} \\ &C)\ \text{not } (n, \gamma_n)\text{-hyperbolic, i.e., } \left| \prod_{j=0}^{n-1} |(f_\varepsilon)'(x_j)| - 1 \right| \le \gamma_n. \end{aligned} \tag{3.25}$$

Recall the definitions and notation of Sections 2.2 and 2.3. In particular, $W_{<2n,1}$ is the space of polynomials of degree $2n-1$ with the standard basis $\{x^m\}_{m=0}^{2n-1}$. The measure $\mu^{\mathrm{st}}_{<2n,\tau}$ defined on the brick $HB^{\mathrm{st}}_{<2n}(\tau) \in W_{<2n,1}$ by (3.3) extends naturally to $W_{<2n,1}$ using the same formulas. Denote by $W^{u,\mathbf{X}_n}_{<2n,1}$ the same space of polynomials of degree $2n-1$, but with the Newton basis (2.26). Lemma 2.2.2 implies that the Newton Map $\mathcal{L}^1_{\mathbf{X}_n} : W_{<2n,1} \to W^{u,\mathbf{X}_n}_{<2n,1}$ defined by (2.24) preserves the measure $\mu^{\mathrm{st}}_{<2n,\tau}$.

In other words, the definition (3.3) produces the same measure whether the standard basis or Newton basis is used.

Now we will estimate the measure of "bad" parameters for a particular trajectory using the Newton basis, without regard (except in the final hyperbolicity estimate) to whether the parameters $u = (u_0, u_1, \dots, u_{2n-1})$ lie in the image $\mathcal{L}^1_{\mathbf{X}_n}(HB^{\mathrm{st}}_{<2n}(\tau))$ of the brick we are concerned with. For a fixed n-tuple of points $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1}$, consider the Newton family of polynomial perturbations

$$\tilde{f}_{u,\mathbf{X}_n}(x) = \tilde{f}(x) + \sum_{m=0}^{2n-1} u_m \prod_{j=0}^{m-1} (x - x_{j \pmod n}). \tag{3.26}$$

Notice that in (2.17) and Figure 2.2, the image $\tilde{f}_{u,\mathbf{X}_n}(x_0)$ of x_0 is independent of u_k for all $k > 0$. Therefore, the position of $\tilde{f}_{u,\mathbf{X}_n}(x_0)$ *depends only on* u_0. Recall that $\mu^{\mathrm{st}}_{m,\tau}$ is 1-dimensional Lebesgue measure scaled by $m!/(2\tau)$. This gives

$$\mu^{\mathrm{st}}_{0,\tau}\left\{u_0 : \left|\tilde{f}_{u,\mathbf{X}_n}(x_0) - x_1\right| = \left|\tilde{f}(x_0) + u_0 - x_1\right| \le \tilde{\gamma}_n\right\} \le \frac{0!}{2\tau} 2\tilde{\gamma}_n = \frac{0!\ \tilde{\gamma}_n}{\tau} \tag{3.27}$$

Fix u_0. Similarly, the position of $\tilde{f}_u(x_1)$ *depends only on* u_1 (see (2.17) and Fig. 2.2). Thus, we have

$$\begin{gathered}\mu^{\mathrm{st}}_{1,\tau}\{u_1 : \left|\tilde{f}_{u,\mathbf{X}_n}(x_1) - x_2\right| = \\ \left|\tilde{f}(x_1) + u_0 + u_1(x_1 - x_0) - x_2\right| \le \tilde{\gamma}_n\} \le \frac{1!\ \tilde{\gamma}_n}{\tau\ |x_1 - x_0|}\end{gathered} \tag{3.28}$$

Inductively for $k = 2, \dots, n-1$, fix $u_0, \dots, u_{k-1}$. Then the position of $\tilde{f}_{u,\mathbf{X}_n}(x_k)$ *depends only on* u_k. Moreover, for $k = 2, \dots, n-2$ we have

$$\begin{gathered}\mu^{\mathrm{st}}_{k,\tau}\Big\{u_k : \left|\tilde{f}_{u,\mathbf{X}_n}(x_k) - x_{k+1}\right| = \\ \left|\tilde{f}(x_k) + \sum_{m=0}^{k} u_m \prod_{j=0}^{m-1} (x_k - x_j) - x_0\right| \le \tilde{\gamma}_n\Big\} \le \frac{k!}{\tau}\ \frac{\tilde{\gamma}_n}{\prod_{j=0}^{k-1} |x_k - x_j|},\end{gathered} \tag{3.29}$$

and for $k = n-1$ we have

$$\mu^{\mathrm{st}}_{n-1,\tau}\{u_{n-1} : \left|\tilde{f}_{u,\mathbf{X}_n}(x_{n-1}) - x_0\right| \le \gamma_n\} \le \frac{(n-1)!}{\tau} \frac{\gamma_n}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|}. \tag{3.30}$$

In particular, the parameter u_{n-1} *is responsible* for (n, γ_n)-periodicity of the n-tuple $\mathbf{X}_n$. Formula (3.30) estimates the "measure of periodicity" discussed in the previous section.

Choose $u_0, \ldots, u_{n-1}$ so that the n-tuple $\mathbf{X}_n$ is a $\tilde{\gamma}_n$-pseudotrajectory and is (n, γ_n)-periodic. Notice that parameters $u_n, u_{n+1}, \ldots, u_{2n-1}$ do not change the $\tilde{\gamma}_n$-pseudotrajectory $\{x_k\}_{k=0}^{n-1}$. Fix now parameters $u_0, \ldots, u_{2n-2}$ and vary only u_{2n-1}. Then for any C^1-smooth map $g : I \to I$, consider the 1-parameter family

$$g_{u_{2n-1}}(x) = g(x) + (x - x_{n-1}) \prod_{j=0}^{n-2} (x - x_j)^2 \tag{3.31}$$

Since the corresponding monomial $(x - x_{n-1}) \prod_{j=0}^{n-2} (x - x_j)^2$ has zeroes of the second order at all points x_k, except the last one x_{n-1}, we have

$$\prod_{j=0}^{n-1} (g_{u_{2n-1}})'(x_j) = \left(g'(x_{n-1}) + u_{2n-1} \prod_{j=0}^{n-2} |x_{n-1} - x_j|^2 \right) \prod_{j=0}^{n-2} g'(x_j). \tag{3.32}$$

To get the final estimate, we use the fact that we are interested only in maps from the family $\{\tilde{f}_\varepsilon\}_{\varepsilon \in HB^{\text{st}}_{<2n}(\tau)}$. Therefore, we have $|\tilde{f}'_{u,\mathbf{X}_n}(x_{n-1})| \leq M_1 \leq M_2$. For condition (C) of (3.25) to hold, $\left| \prod_{j=0}^{n-1} \tilde{f}'_{u,\mathbf{X}_n}(x_j) \right|$ must lie in $[1 - \gamma_n, 1 + \gamma_n]$. If this occurs for any $u_{2n-1} = \varepsilon_{2n-1} \in HB^{\text{st}}_{2n}(\tau)$, then $\left| \prod_{j=0}^{n-2} \tilde{f}'_{u,\mathbf{X}_n}(x_j) \right| \geq (1 - \gamma_n)/M_2$ for all u_{2n-1}, because this product does not depend on u_{2n-1}. Using (3.32) and the fact that $1 - \gamma_n \geq 1/2$, we get

$$\begin{gathered} \mu^{\text{st}}_{2n-1,\tau} \left\{ u_{2n-1} : \left| \prod_{j=0}^{n-1} |(\tilde{f}_{u,\mathbf{X}_n})'(x_j)| - 1 \right| \leq \gamma_n \right\} \leq \\ 2M_2 \frac{(2n-1)!}{2\tau} \frac{4\gamma_n}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|^2} = 4M_2 \frac{(2n-1)!}{\tau} \frac{\gamma_n}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|^2}. \end{gathered} \tag{3.33}$$

This formula estimates the "measure of hyperbolicity" discussed in the previous section.

By Lemma 2.2.2, we can combine all these estimates and get

$$\begin{gathered} \mu^{\text{st}}_{<n,\tau} \times \mu^{\text{st}}_{2n-1,\tau} \{ (u_0, \ldots, u_{n-1}, u_{2n-1}) \in W^{u,\mathbf{X}_n}_{<n,1} \times W^{u,\mathbf{X}_n}_{2n-1,1} : \\ \tilde{f}_{u,\mathbf{X}_n} \text{ satisfies conditions (3.25) and } \|\tilde{f}_{u,\mathbf{X}_n}\|_{C^2} \leq M_2 \} \leq \\ 4M_2 \frac{(n-1)!\, \gamma_n}{\tau \prod_{j=0}^{n-2} |x_{n-1} - x_j|} \frac{(2n-1)!\, \gamma_n}{\tau \prod_{j=0}^{n-2} |x_{n-1} - x_j|^2} \prod_{m=0}^{n-2} \frac{m!\, \tilde{\gamma}_n}{\tau \prod_{j=0}^{m-1} |x_m - x_j|}, \end{gathered} \tag{3.34}$$

where the spaces $W^{u,\mathbf{X}_n}_{<n,1}$ and $W^{u,\mathbf{X}_n}_{2n-1,1}$ are discussed in the beginning of this section. This estimate corresponds loosely to (3.23) in the previous section. The final term

is an upper bound on measure of parameters for which $\mathbf{X}_n$ is a $\tilde{\gamma}_n$-pseudotrajectory for $f_{u,\mathbf{X}_n}$. Roughly speaking, since almost every initial point x_0 has exactly one $\tilde{\gamma}_n$-pseudotrajectory $\mathbf{X}_n \in I^n_{\tilde{\gamma}_n}$ for each set of parameters, and the total measure of parameters in $HB^{\rm st}_{<2n}(\tau)$ is 1, the sum over all $\tilde{\gamma}_n$-pseudotrajectories $\mathbf{X}_n$ associated to x_0 and $HB^{\rm st}_{<2n}(\tau)$ of the parameter measure associated with $\mathbf{X}_n$ should be 1. Thus the final term on the right-hand side of (3.34) also represents an upper bound on the inverse of the number of $\tilde{\gamma}_n$-pseudotrajectories per initial point, which appears in (3.23). However, we need the upper bound to be sharp in order to cancel this term with that in (3.22), and the heuristic explanation of this paragraph is complicated by the fact that the parameterization we are using *depends on the $\tilde{\gamma}_n$-pseudotrajectory* $\mathbf{X}_n$. These challenges will be resolved in the Collection Lemma of the next section.

3.4 The Distortion and Collection Lemmas

In this section we formulate the Distortion Lemma for the Newton map $\mathcal{L}^1_{\mathbf{X}_n}$, and complete the estimate of the measure of all "bad" parameters with a simple, almost periodic, but not sufficiently hyperbolic trajectory (3.13) by collecting all possible "bad" pseudotrajectories (see the Collection Lemma below).

Consider an ordered n-tuple of points $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \in I^n$ and the Newton map $\mathcal{L}^1_{\mathbf{X}_n} : W_{<2n,1} \to W^{u,\mathbf{X}_n}_{<2n,1}$, defined by (2.24). We now estimate the distortion of the Newton map $\mathcal{L}^1_{\mathbf{X}_n}$ as the map from the standard basis $\{\varepsilon_k\}_{k=0}^{2n-1}$ in $W_{<2n,1}$ to the Newton basis $\{u_k\}_{k=0}^{2n-1}$ in $W^{u,\mathbf{X}_n}_{<2n,1}$. It helps to have in mind the following picture characterizing the distortion of the Newton map.

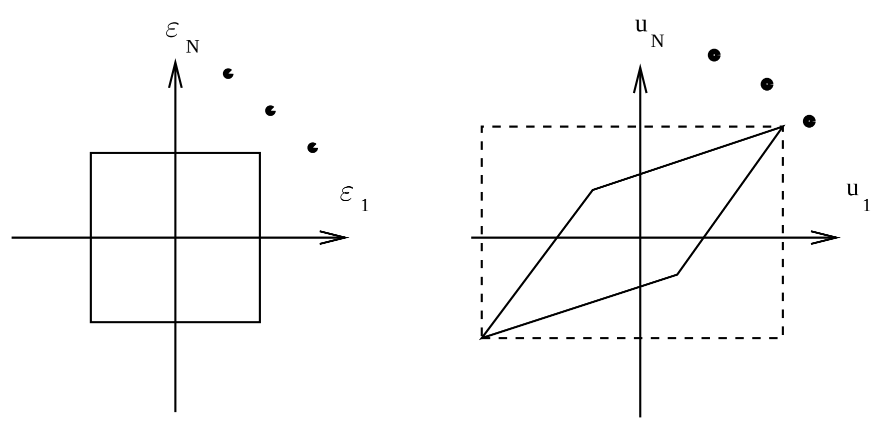

Figure 3.1: Distortion by the Newton map

The Distortion Lemma. *Let $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \in I^n$ be an ordered n-tuple of points in the interval $I = [-1, 1]$ and $\mathcal{L}^1_{\mathbf{X}_n} : W_{<2n,1} \to W^{u,\mathbf{X}_n}_{<2n,1}$ be the Newton map, defined by (2.24). Then the image of the Brick of standard thickness $HB^{\rm st}_{<2n}(\tau)$ with width $\tau > 0$ is contained in the Brick of standard thickness $HB^{\rm st}_{<2n}(3\tau)$ with width 3τ:*

$$\mathcal{L}^1_{\mathbf{X}_n}(HB^{\rm st}_{<2n}(\tau)) \subset HB^{\rm st}_{<2n}(3\tau) \subset W^{u,\mathbf{X}_n}_{<2n,1}. \tag{3.35}$$

In other words, independently of the choice of an n-tuple $\{x_j\}_{j=0}^{n-1} \in I^n$ for any $0 \leq m < 2n$, the coefficient u_m has at most the range of values $|u_m| \leq \frac{3\tau}{m!}$ in the image $\mathcal{L}^1_{\mathbf{X}_n}(HB^{\rm st}_{<2n}(\tau))$.

Remark 3.4.1. *For this lemma, the sides of the Brick $HB^{\rm st}_{<2n}(\tau)$ have to decay at least as fast as a factorial in the order of the side, i.e., $r_n \leq \frac{\tau}{n!}$ for some $\tau > 0$. If the sides of a Brick under investigation decay, say, as an exponential function, i.e., $r_n = \exp(-Kn)$ for some $K > 0$, then the Distortion Lemma fails and there is no uniform estimate on distortion. In terms of formula (2.24), if the range of values of ε_k does not decay fast enough with k, then u_m depends significantly on ε_k with k much larger than m.*

Proof: Recall that for $\{\varepsilon_m\}_{m=0}^{2n-1} \in HB^{\rm st}_{<2n}(\tau)$ we have for each m that $|\varepsilon_m| \leq \tau/m!$. By definition (2.24) of the Newton map $\mathcal{L}^1_{\mathbf{X}_n}$ we have

$$u_m = \varepsilon_m + \sum_{k=m+1}^{2n-1} \varepsilon_k p_{k,m}(x_0, \ldots, x_{m \pmod n}), \tag{3.36}$$

where $p_{k,m}$ is the homogeneous polynomial of degree $k - m$ defined by (2.22). Notice that every monomial of $p_{k,m}$ is uniformly bounded by 1, provided all points $\{x_j\}_{j=0}^{n-1}$ are bounded in absolute value by 1. Therefore, $|p_{k,m}|$ is uniformly bounded by the number of its monomials $\binom{k}{m}$. This implies that

$$\begin{aligned} |u_m| \leq |\varepsilon_m| + \sum_{k=m+1}^{2n-1} |\varepsilon_k| \binom{k}{m} \leq \\ \frac{\tau}{m!}\left(1 + \sum_{k=m+1}^{2n-1} \frac{1}{(k-m)!}\right) \leq \frac{3\tau}{m!}. \end{aligned} \tag{3.37}$$

This completes the proof of the Lemma. Q.E.D.

For a given n-tuple $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \in I^n$, the parallelepiped

$$\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau) = \mathcal{L}^1_{\mathbf{X}_n}(HB^{\rm st}_{<2n}(\tau)) \subset W^{u,\mathbf{X}_n}_{<2n,1} \tag{3.38}$$

is the set of parameters $(u_0, \dots, u_{2n-1})$ that correspond to parameters $(\varepsilon_0, \dots, \varepsilon_{2n-1}) \in HB^{\text{st}}_{<2n}(\tau)$. In other words, these are the Newton parameters *allowed by the family* (3.8) for the n-tuple $\mathbf{X}_n$. We already knew by Lemma 2.2.2 that $\mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)$ has the same volume as $HB^{\text{st}}_{<2n}(\tau)$, but the Distortion Lemma tells us in addition that the projection of $\mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)$ onto any coordinate axis is at most a factor of 3 longer than the projection of $HB^{\text{st}}_{<2n}(\tau)$.

Let $\mathbf{X}_m = \{x_j\}_{j=0}^{m-1}$ be the m-tuple of first m points of the n-tuple $\mathbf{X}_n$. We now consider which Newton parameters are allowed by the family (3.8) when $\mathbf{X}_m$ is fixed but $x_m, \dots, x_{n-1}$ are arbitrary. Since we will only be using the definitions below for discretized n-tuples $\mathbf{X}_n \in I^n_{\tilde{\gamma}_n}$, we consider only the (finite number of) possibilities $x_m, \dots, x_{n-1} \in I_{\tilde{\gamma}_n}$. Let

$$\pi^{u,\mathbf{X}_n}_{<2n,\leq m} : W^{u,\mathbf{X}_n}_{<2n,1} \to W^{u,\mathbf{X}_m}_{\leq m,1} \text{ and } \pi^{u,\mathbf{X}_n}_{<2n,m} : W^{u,\mathbf{X}_n}_{<2n,1} \to W^{u,\mathbf{X}_m}_{m,1}$$

be the natural projections onto the space $W^{u,\mathbf{X}_m}_{\leq m,1} \simeq \mathbb{R}^m$ of polynomials of degree m and the space $W^{u,\mathbf{X}_m}_{m,1} \simeq \mathbb{R}$ of homogeneous polynomials of degree m respectively. Denote the unions over all $x_m, \dots, x_{n-1} \in I_{\tilde{\gamma}_n}$ of the images of $\mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)$ under the respective projections $\pi^{u,\mathbf{X}_m}_{<2n,\leq m}$ and $\pi^{u,\mathbf{X}_n}_{<2n,m}$ by

$$\begin{aligned} \mathcal{P}^{\text{st}}_{<2n,\leq m,\mathbf{X}_m}(\tau) &= \cup_{x_m,\dots,x_{n-1}\in I_{\tilde{\gamma}_n}} \pi^{u,\mathbf{X}_n}_{<2n,\leq m}(\mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)) \subset W^{u,\mathbf{X}_m}_{\leq m,1} \\ \mathcal{P}^{\text{st}}_{<2n,m,\mathbf{X}_m}(\tau) &= \cup_{x_m,\dots,x_{n-1}\in I_{\tilde{\gamma}_n}} \pi^{u,\mathbf{X}_n}_{<2n,m}(\mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)) \subset W^{u,\mathbf{X}_m}_{m,1}. \end{aligned} \tag{3.39}$$

For each $m < n$, the set $\mathcal{P}^{\text{st}}_{<2n,\leq m,\mathbf{X}_m}(\tau)$ is a polyhedron and $\mathcal{P}^{\text{st}}_{<2n,m,\mathbf{X}_m}(\tau)$ is a segment of length at most $6\tau/m!$ by the Distortion Lemma. Both depend only on the m-tuple $\mathbf{X}_m$ and width τ. The set $\mathcal{P}^{\text{st}}_{<2n,\leq m,\mathbf{X}_m}(\tau)$ consists of all Newton parameters $\{u_j\}_{j=0}^m \in W^{u,\mathbf{X}_m}_{m,1}$ that are allowed by the family (3.8) for the m-tuple $\mathbf{X}_m$.

For each $m < n$, we introduce the family of diffeomorphisms

$$\tilde{f}_{u(m),\mathbf{X}_m}(x) = \tilde{f}(x) + \sum_{s=0}^{m} u_s \prod_{j=0}^{s-1}(x - x_j), \tag{3.40}$$

where $u(m) = (u_0, \dots, u_m) \in \mathcal{P}^{\text{st}}_{<2n,\leq m,\mathbf{X}_m}(\tau)$. For each possible continuation $\mathbf{X}_n$ of $\mathbf{X}_m$, the family $\tilde{f}_{u(m),\mathbf{X}_m}$ includes the subfamily of $\tilde{f}_{u,\mathbf{X}_n}$ (with $u \in \mathcal{P}^{\text{st}}_{<2n,\mathbf{X}_n}(\tau)$) corresponding to $u_{m+1} = u_{m+2} = \cdots = u_{2n-1} = 0$. However, the action of $\tilde{f}_{u,\mathbf{X}_n}$ on $x_0, \dots, x_m$ doesn't depend on $u_{m+1}, \dots, u_{2n-1}$, so for these points the family $\tilde{f}_{u(m),\mathbf{X}_m}$ is representative of the entire family $\tilde{f}_{u,\mathbf{X}_n}$. This motivates the definition

$$\begin{aligned} &T^{1,\tilde{\gamma}_n}_{<2n,\leq m,\tau}(\tilde{f}; x_0, \dots, x_{m-1}, x_m, x_{m+1}) = \\ &\Big\{u(m) \in \mathcal{P}^{\text{st}}_{<2n,\leq m,\mathbf{X}_m}(\tau) \subset W^{u,\mathbf{X}_m}_{\leq m,1} : \\ &|\tilde{f}_{u(m),\mathbf{X}_m}(x_{j-1}) - x_j| \leq \tilde{\gamma}_n \text{ for } j = 1, \dots, m+1\Big\}. \end{aligned} \tag{3.41}$$

The set $T^{1,\tilde{\gamma}_n}_{<2n,\leq m,\tau}(\tilde{f}; x_0, \dots, x_{m-1}, x_m, x_{m+1})$ represents the set of Newton parameters $u(m) = \big(u_0, \dots, u_m\big)$ allowed by the family (3.8) for which $x_0, \dots, x_{m+1}$ is a $\tilde{\gamma}_n$-pseudotrajectory of $\tilde{f}_{u(m),\mathbf{X}_m}$ (and hence of $\tilde{f}_{u,\mathbf{X}_n}$ for all valid extensions u and $\mathbf{X}_n$ of $u(m)$ and $\mathbf{X}_m$).

In the following lemma, we collect all possible $\tilde{\gamma}_n$-pseudotrajectories and estimates of "bad" measure corresponding to those $\tilde{\gamma}_n$-pseudotrajectories.

The Collection Lemma. *With the notations above, for all $x_0 \in I_{\tilde{\gamma}_n}$ the measure of the "bad" parameters satisfies*

$$\begin{aligned} &\mu^{\rm st}_{<2n,\tau}\big\{B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0)\big\} \leq \\ &6^{2n} M_2^{4n+1}\ \frac{(n-1)!}{\tau}\ \frac{(2n-1)!}{\tau}\ \gamma_n^{5/4} \quad . \end{aligned} \tag{3.42}$$

Proof of the Collection Lemma: We prove by backward induction on m that for $x_0, \dots, x_m \in I_{\tilde{\gamma}_n}$,

$$\begin{aligned} &\mu^{\rm st}_{<2n,\tau}\big\{B^{{\rm sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f}, \gamma_n, M_2; x_0, \dots, x_m)\big\} \\ &\leq 6^{2n-m} M_2^{4n+1}\ \frac{(n-1)!}{\tau}\ \frac{(2n-1)!}{\tau} \\ &\mu^{\rm st}_{<m,\tau}\big\{T^{1,\tilde{\gamma}_n}_{<2n,\leq m-1,\tau}(\tilde{f}; x_0, \dots, x_m)\big\}\ \gamma_n^{5/4}, \end{aligned} \tag{3.43}$$

resulting when $m = 0$ in (3.42).

Consider the case $m = n-1$. Fix an $(n, \gamma_n/2)$-simple n-tuple $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \in I^n_{\tilde{\gamma}_n}$. Using formulas (3.30) and (3.33), we get

$$\begin{aligned} &\mu^{\rm st}_{n-1,\tau}\{u_{n-1} : \left|\tilde{f}_{u,\mathbf{X}_n}(x_{n-1}) - x_0\right| \leq M_2^n \gamma_n\} \leq \\ &\frac{(n-1)!}{\tau} \frac{M_2^n \gamma_n}{\prod_{m=0}^{n-2} |x_{n-1} - x_m|} \leq\ \frac{2^{1/4} M_2^n (n-1)!}{\tau}\ \gamma_n^{3/4} \end{aligned} \tag{3.44}$$

and

$$\begin{aligned} &\mu^{\rm st}_{2n-1,\tau}\{u_{2n-1} : \left|\prod_{j=0}^{n-1} |(\tilde{f}_{u,\mathbf{X}_n})'(x_j)| - 1\right| \leq M_2^{3n}\gamma_n\} \leq \\ &4M_2 \frac{(2n-1)!}{\tau}\ \frac{M_2^{3n}\gamma_n}{\prod_{m=0}^{n-2} |x_{n-1} - x_m|^2} \leq \frac{2^{5/2} M_2^{3n+1} (2n-1)!}{\tau} \gamma_n^{1/2}. \end{aligned} \tag{3.45}$$

The Fubini Theorem, Lemma 2.2.2, and the definition of the product measure $\mu^{\rm st}_{<2n,\tau}$

imply that

$$\begin{aligned}
&\mu^{\rm st}_{<2n,\tau}\left\{B^{{\rm sim},\tilde\gamma_n}_{n,\tau}(\tilde f,\gamma_n,M_2;x_0,\dots,x_{n-1})\right\}\le \\
&\qquad \mu^{\rm st}_{<n-1,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le n-2,\tau}(\tilde f;x_0,\dots,x_{n-1})\right\}\times \\
&\qquad \mu^{\rm st}_{n-1,\tau}\left\{u_{n-1}:\left|\tilde f_{u,\mathbf{X}_n}(x_{n-1})-x_0\right|\le M_2^n\gamma_n\right\}\times\prod_{s=n}^{2n-2}\mu^{\rm st}_{s,\tau}\{\mathcal{P}^{\rm st}_{<2n,s,\mathbf{X}_n}(\tau)\}\times \qquad (3.46)\\
&\qquad \mu^{\rm st}_{2n-1,\tau}\left\{u_{2n-1}:\left|\prod_{j=0}^{n-1}|(f_{u,\mathbf{X}_n})'(x_j)|-1\right|\le M_2^{3n}\gamma_n\right\} \\
&\le 2^{11/4}3^{n-1}M_2^{4n+1}\ \frac{(n-1)!}{\tau}\frac{(2n-1)!}{\tau}\ \mu^{\rm st}_{<n-1,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le n-2,\tau}(\tilde f;x_0,\dots,x_{n-1})\right\}\ \gamma_n^{5/4}
\end{aligned}$$

The last inequality follows from the Distortion Lemma, which says that for each $s=n,n+1,\dots,2n-2$

$$\mu^{\rm st}_{s,\tau}\{\mathcal{P}^{\rm st}_{<2n,s,\mathbf{X}_n}(\tau)\}\le 3. \qquad (3.47)$$

Since $2^{11/4}3^{n-1}<6^{n+1}$, this yields the required estimate (3.43) for $m=n-1$.

Suppose now that for $m+1$, (3.43) is true and we would like to prove it for m. Denote by $G^{1,\tilde\gamma_n}_{<2n,m,\tau}(\tilde f,u(m-1);x_0,\dots,x_m)\subset I_{\tilde\gamma_n}$ the set of points x_{m+1} of the $2\tilde\gamma_n$-grid $I_{\tilde\gamma_n}$ such that the $(m+2)$-tuple $x_0,\dots,x_{m+1}$ is a $\tilde\gamma_n$-pseudotrajectory associated to some extension $u(m)\in\mathcal{P}^{\rm st}_{<2n,\le m,\mathbf{X}_m}(\tau)$ of $u(m-1)$. In other words, $G^{1,\tilde\gamma_n}_{<2n,m,\tau}(\tilde f,u(m-1);x_0,\dots,x_m)$ is the set of all possible continuations of the $\tilde\gamma_n$-pseudotrajectory $x_0,\dots,x_m$ using all possible Newton parameters u_m allowed by the family (3.8).

Now if $x_0,\dots,x_m$ is a $\tilde\gamma_n$-pseudotrajectory associated to $u(m)=(u_0,\dots,u_m)$, then at most 2 values of $x_{m+1}\in I_{\tilde\gamma_n}$ are within $\tilde\gamma_n$ of $\tilde f_{u(m),\mathbf{X}_m}(x_m)$. Thus for fixed $u(m-1)=(u_0,\dots,u_{m-1})\in\mathcal{P}^{\rm st}_{<2n,\le m-1,\mathbf{X}_n}(\tau)$, each value of $u_m\in\mathcal{P}^{\rm st}_{<2n,m,\mathbf{X}_n}(\tau)$ corresponds to at most 2 points in $G^{1,\tilde\gamma_n}_{<2n,m,\tau}(\tilde f,u(m-1);x_0,\dots,x_m)$. It follows that

$$\begin{aligned}
&\sum_{x_{m+1}\in G^{1,\tilde\gamma_n}_{<2n,m,\tau}(\tilde f,u(m-1);x_0,\dots,x_m)}\mu^{\rm st}_{\le m,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le m,\tau}(\tilde f;x_0,\dots,x_{m+1})\right\}\le \\
&\qquad 2\ \mu^{\rm st}_{m,\tau}\{\mathcal{P}^{\rm st}_{<2n,m,\mathbf{X}_n}(\tau)\}\ \mu^{\rm st}_{\le m-1,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le m-1,\tau}(\tilde f;x_0,\dots,x_m)\right\}.
\end{aligned} \qquad (3.48)$$

The Distortion Lemma then implies that

$$\begin{aligned}
&\sum_{x_{m+1}\in G^{1,\tilde\gamma_n}_{<2n,m,\tau}(\tilde f,u(m-1);x_0,\dots,x_m)}\mu^{\rm st}_{\le m,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le m,\tau}(\tilde f;x_0,\dots,x_{m+1})\right\}\le \\
&\qquad 6\ \mu^{\rm st}_{\le m-1,\tau}\left\{T^{1,\tilde\gamma_n}_{<2n,\le m-1,\tau}(\tilde f;x_0,\dots,x_m)\right\}.
\end{aligned} \qquad (3.49)$$

Inductive application of this formula completes the proof of the Collection Lemma. Q.E.D.

Proof of Proposition 3.1.2: The number of starting points $\tilde{x}_0 \subset I_{\tilde{\gamma}_n}$ for a $\tilde{\gamma}_n$-pseudotrajectory equals $1/\tilde{\gamma}_n$. Therefore, multiplying the estimate (3.42) by $1/\tilde{\gamma}_n = M_2^{2n}/\gamma_n$ and using (3.21) we get

$$\mu^{\text{st}}_{<2n,\tau}\big\{B^{\text{sim},\tilde{\gamma}_n}_{n,\tau}(\tilde{f},\gamma_n,M_2)\big\} \leq 6^{2n} M_2^{6n+1}\ \frac{(n-1)!}{\tau}\ \frac{(2n-1)!}{\tau}\ \gamma_n^{1/4}. \tag{3.50}$$

To prove the required inequality (3.13), it remains only to prove (3.17).

If a parameter ε belongs to $B^{\text{sim}}_{n,\tau}(C,\delta,\tilde{f},\gamma_n)$, then ε is associated with an essentially (n,γ_n)-simple, (n,γ_n)-periodic, and not (n,γ_n)-hyperbolic trajectory $\{x_k = f^k_\varepsilon(x_0)\}_{k=0}^{n-1}$. Our goal is to show that $\varepsilon \in B^{\text{sim}}_{n,\tau}(\tilde{f},\gamma_n,M_2)$; recall the definitions (3.15) and (3.16).

The bound M_2 on the C^2-norm of $\tilde{f}_\varepsilon$ implies that for all $x,y \in I$ we have $|\tilde{f}_\varepsilon(x)-\tilde{f}_\varepsilon(y)| < M_2|x-y|$. Essential simplicity of the trajectory $\{x_m\}_{m=0}^{n-1}$ implies that for some $j<n$, the shifted trajectory $\{x_{j+m} = \tilde{f}^{j+m}_\varepsilon(x_0)\}_{m=0}^{n-1}$ is (n,γ_n)-simple and $(n, M_2^j\gamma_n)$-periodic. Let's approximate the shifted trajectory $\{x_{j+m}\}_{m=0}^{n-1}$ by a $\tilde{\gamma}_n$-pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1} \in I^n_{\tilde{\gamma}_n}$ associated to the (fixed above) parameter ε. Consider the $\tilde{\gamma}_n$-pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1} \in I^n_{\tilde{\gamma}_n}$ starting at $\tilde{x}_j$. Choose $\tilde{x}_j$ such that $|x_j - \tilde{x}_j| \leq \tilde{\gamma}_n$ and choose $\tilde{x}_{j+1}$ such that $|\tilde{f}_\varepsilon(\tilde{x}_j) - \tilde{x}_{j+1}| \leq \tilde{\gamma}_n$. Then

$$\begin{aligned} |x_{j+1}-\tilde{x}_{j+1}| &\leq |\tilde{f}_\varepsilon(x_j)-\tilde{f}_\varepsilon(\tilde{x}_j)| + |\tilde{f}_\varepsilon(\tilde{x}_j)-\tilde{x}_{j+1}| \\ &\leq (M_2+1)\tilde{\gamma}_n = \frac{M_2^2-1}{M_2-1}\tilde{\gamma}_n \end{aligned} \tag{3.51}$$

By induction on m, choosing $\tilde{x}_{j+m}$ such that $|\tilde{f}_\varepsilon(\tilde{x}_{j+m-1}) - \tilde{x}_{j+m}| \leq \tilde{\gamma}_n$, we have

$$\begin{aligned} |x_{j+m}-\tilde{x}_{j+m}| &\leq |\tilde{f}_\varepsilon(x_{j+m-1})-\tilde{f}_\varepsilon(\tilde{x}_{j+m-1})| + |\tilde{f}_\varepsilon(\tilde{x}_{j+m-1})-\tilde{x}_{j+m}| \\ &\leq \frac{M_2^{m+1}-1}{M_2-1}\tilde{\gamma}_n \end{aligned} \tag{3.52}$$

Using this estimate with $m=n-1$, we have

$$\begin{aligned} |\tilde{f}_\varepsilon(\tilde{x}_{j+n-1})-\tilde{x}_j| &\leq |\tilde{f}_\varepsilon(\tilde{x}_{j+n-1})-\tilde{f}_\varepsilon(x_{j+n-1})| + \\ |\tilde{f}_\varepsilon(x_{j+n-1})-x_j| + |x_j-\tilde{x}_j| &\leq M_2^{n+1}\tilde{\gamma}_n + M_2^{n-1}\gamma_n \leq M_2^n\gamma_n. \end{aligned} \tag{3.53}$$

So, the $\tilde{\gamma}_n$-pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1}$ is $(n, M_2^n\gamma_n)$-periodic.

Next, since $\{x_{j+m}\}_{m=0}^{n-1}$ is (n,γ_n)-simple, this means $\prod_{m=0}^{n-2}|x_{j+n-1}-x_{j+m}| \geq \gamma_n^{1/4}$. Each term in the product must then be at least $2^{-(n-2)}\gamma_n^{1/4}$. For $m \leq n-1$, we have already shown that $|x_{j+m} - \tilde{x}_{j+m}| \leq M_2^n \tilde{\gamma}_n$. Then recalling that $\tilde{\gamma}_n = M_2^{-2n}\gamma_n$,

$$\begin{aligned} |\tilde{x}_{j+n-1} - \tilde{x}_{j+m}| &\geq |x_{j+n-1} - x_{j+m}| - 2M_2^n\tilde{\gamma}_n \\ &\geq \frac{2^{-(n-2)}\gamma_n^{1/4} - 2M_2^n\tilde{\gamma}_n}{2^{-(n-2)}\gamma_n^{1/4}}|x_{j+n-1} - x_{j+m}| \\ &= (1 - 2^{n-1}M_2^{-n}\gamma_n^{3/4})|x_{j+n-1} - x_{j+m}| \\ &\geq (1 - \gamma_n^{3/4})|x_{j+n-1} - x_{j+m}|. \end{aligned} \tag{3.54}$$

Since $\gamma_n \leq \gamma_n(C,\delta) = \exp(-Cn^{1+\delta})$ with $C > 2$, a simple calculation shows that $(1-\gamma_n^{3/4})^{n-1} \geq 2^{-1/4}$. Then taking the product of (3.54) over $m = 0, 1, \ldots, n-2$ proves that the approximating pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1}$ is $(n, \gamma_n/2)$-simple.

Now consider the difference of derivatives

$$\left| \prod_{m=0}^{n-j-1} \tilde{f}'_\varepsilon(\tilde{x}_{j+m}) \prod_{m=0}^{j-1} \tilde{f}'_\varepsilon(\tilde{x}_{n+m}) - \prod_{m=0}^{n-1} \tilde{f}'_\varepsilon(x_m) \right|. \tag{3.55}$$

Since $\|\tilde{f}\|_{C^2} \leq M_2$, for $0 \leq m \leq n-j-1$ we have

$$|\tilde{f}'_\varepsilon(\tilde{x}_{j+m}) - \tilde{f}'_\varepsilon(x_{j+m})| \leq M_2|\tilde{x}_{j+m} - x_{j+m}| \leq M_2^{n+1}\tilde{\gamma}_n \leq \gamma_n \tag{3.56}$$

and for $0 \leq m \leq j-1$ we have

$$\begin{gathered} |\tilde{f}'_\varepsilon(\tilde{x}_{n+m}) - \tilde{f}'_\varepsilon(x_m)| \leq M_2|\tilde{x}_{n+m} - x_m| \leq \\ M_2(|\tilde{x}_{n+m} - x_{n+m}| + |x_{n+m} - x_m|) \leq M_2^{2n}\tilde{\gamma}_n + M_2^n\gamma_n \leq 2M_2^n\gamma_n. \end{gathered} \tag{3.57}$$

Then since $|\tilde{f}'_\varepsilon(x)| \leq M_2$ for all $x \in I$, we get that the difference of derivatives (3.55) is bounded by $2nM_2^{2n-1}\gamma_n \leq M_2^{3n-1}\gamma_n$. Therefore, if the initial exact trajectory $\{x_k\}_{j=0}^{n-1}$ is not (n,γ_n)-hyperbolic, then the pseudotrajectory is not $(n, M_2^{3n}\gamma_n)$-hyperbolic, and the parameter ε belongs to the set $B_{n,\tau}^{\mathrm{sim},\tilde{\gamma}_n}(\tilde{f}, \gamma_n, M_2)$. Q.E.D.

3.5 Discretization Method for trajectories with a gap

Recall that we consider the C^2-smooth 1-dimensional model, defined in Section 3.1. Our goal is to estimate the measure of the set (3.9) of all "bad" parameters $B_{n,\tau}(C,\delta,\tilde{f},\gamma_n)$. This set belongs to the union (3.12) of "bad" parameters $B_{n,\tau}^{\mathrm{sim}}(C,\delta,\tilde{f},\gamma_n)$ and $B_{n,\tau}^{\mathrm{non}}(C,\delta,\tilde{f},\gamma_n)$

associated with essentially simple and essentially nonsimple almost periodic points. In the last three sections, we developed the Discretization Method and estimated the measure of the set $B^{\mathrm{sim}}_{n,\tau}(C,\delta,\tilde{f},\gamma_n)$ associated with essentially simple almost periodic trajectories (3.13). In this and the next section, we consider "bad" parameters $B^{\mathrm{non}}_{n,\tau}(C,\delta,\tilde{f},\gamma_n)$ associated with essentially nonsimple almost periodic trajectories (3.11) and get an estimate on the measure of this set.

Proposition 3.5.1. *Let $\{\tilde{f}_\varepsilon\}_{\varepsilon\in HB^{\mathrm{st}}_{<2n}(\tau)}$ be the family of polynomial perturbations (3.8) and M_2 be an upper bound on the C^2-norm of the family. Then with the notation above, for any $\delta>0$, $C>100\delta^{-1}\log M_2$, $\tau>0$ and a sufficiently small positive γ_n, e.g., $\gamma_n\le\gamma_n(C,\delta)$, we have the following estimate on the measure of parameters associated with maps $\tilde{f}_\varepsilon$ with an (n,γ_n)-periodic, essentially non-(n,γ_n)-simple, but not (n,γ_n)-hyperbolic point:*

$$\mu^{\mathrm{st}}_{<2n,\tau}\big\{B^{\mathrm{non}}_{n,\tau}(C,\delta,\tilde{f},\gamma_n)\big\}\le 16n^{1+\delta}\,6^{2n}\,M_2^{10n+1}\,\exp(-Cn/200). \tag{3.58}$$

Remark 3.5.2. *Though we have stated Proposition 3.5.1 only for the C^2 case in $\mathbb{R}$, we will make our upcoming definitions for the general $C^{1+\rho}$ case in $\mathbb{R}^N$ so that they may be used later. For the time being, the reader may keep in mind the case $\rho=N=1$. The scheme of the proof of this Proposition is in Section 3.5.2.*

According to the decomposition (2.42), the set $B^{\mathrm{non}}_{n,\tau}(C,\delta,\tilde{f},\gamma_n)$ of parameters associated with essentially nonsimple trajectories can be decomposed into a finite union of sets of parameters with a trajectory that has a weak gap. This decomposition follows from Theorem 2.4.6. We need to improve this theorem. Let us sharpen Definition 2.4.3 of a weak gap for almost periodic trajectories.

Definition 3.5.3. *Let $g\in\mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism of the ball B^N (respectively $g\in C^{1+\rho}(I,I)$ be a $C^{1+\rho}$-smooth map of the interval I) and let $\vec{\mathbf{r}}=\{r_k\}_{k=0}^\infty$ be a nonincreasing sequence of sizes of the Brick $HB^N(\vec{\mathbf{r}})$ that tend to zero, and $D>2$ be some number. A point $x_0\in B^N$ (respectively $x_0\in I$) or a trajectory $x_0,\dots,x_{n-1}=g^{n-1}(x_0)$ in B^N (respectively I) of length n has a (D,n,r_{2k})-gap at a point $x_k=g^k(x_0)$ if*

$$|x_k-x_0|\le \min\left\{D^{-n\log_2 n}\min_{0<j\le k-1}|x_0-x_j|,\ r_{2k}^{4(N+N^2)},\ \left(\prod_{j=0}^{k-2}|x_{k-1}-x_j|\right)^{4(N+2)}\right\}. \tag{3.59}$$

In the case of the model when $r_k=\frac{\tau}{k!}$, we denote a (D,n,r_{2k})-gap by a (D,n,τ)-gap.

This definition is designed to fit the induction over the period n outlined in Chapter 2. Recall that the first term in the minimum (3.59) corresponds to Definition 2.4.3 of a weak (D,n)-gap. For $D = M_{1+\rho}^{30/\rho}$, if a trajectory $\{x_j\}_{j=0}^{n-1}$ is $(n, |x_k - x_0|)$-periodic and has a weak (D,n)-gap at x_k, then by Lemma 2.4.5 the fraction $p = n/k$ is an integer, and one can split the trajectory of length n into p almost identical parts of length k each. Then by perturbing the linearization $dg_u^k(x_0)$ at x_{k-1} using the family $g_u(x) = g(x) + u(x - x_{k-1}) \prod_{j=0}^{k-2}(x - x_j)^2$, one can reach sufficient hyperbolicity for $dg_u^k(x_0)$. Moreover, the first and the third terms in the minimum (3.59) allow us to extract sufficient hyperbolicity of $dg_u^n(x_0)$ from sufficient hyperbolicity of $dg_u^k(x_0)$. Roughly, it is because we have $dg_u^n(x_0) \approx \left(dg_u^k(x_0)\right)^p$.

The second and the third terms guarantee that with respect to the normalized Lebesgue measure $\mu_{\vec{\mathbf{r}}}^N$, "bad" parameters occupy a small measure set; loosely speaking, they counteract terms like $(2n-1)!/\tau$ and $\prod_{j=0}^{n-2} |x_{n-1} - x_j|^{-1}$ in (3.30) and (3.33).

Recall that in our notation, $\gamma_n(C,\delta) = \exp(-Cn^{1+\delta})$ for each $n \in \mathbb{Z}_+$. It is important to have in mind that the scale is $\gamma_k(C,\delta)$ not $\gamma_n(C,\delta)$ in the definition below.

Definition 3.5.4. *Let $g \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism of the ball B^N (respectively $g \in C^{1+\rho}(I,I)$ be a $C^{1+\rho}$-smooth map of the interval I) for some $\rho > 0$. Let also $C > 0$, $\delta > 0$ and $k < n$ be positive integers. We say that a point x_0 has a (k,n,C,δ,ρ)-leading saddle if $|x_0 - x_k| \le n^{-1/\rho}\gamma_k^{4N/\rho}(C,\delta)$. Also if x_0 is $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic, we say that x_0 has no (n,C,δ,ρ)-leading saddles if for all $k < n$ we have that x_0 has no (k,n,C,δ,ρ)-leading saddles.*

Lemma 3.5.5. *Let a $C^{1+\rho}$-smooth diffeomorphism (respectively $C^{1+\rho}$-smooth map) g satisfy the Inductive Hypothesis of some order $n-1$ with some constants $\rho > 0$, $C > 30\rho^{-1}\log M_{1+\rho}$, and $\delta > 0$, i.e., $g \in IH(n-1, C, \delta, \rho)$. Assume that g has a point $x_0 \in B^N$ (respectively $x \in I$) that has a (k,n,C,δ,ρ)-leading saddle. Then there is a periodic point $x^* = g^k(x^*)$ of period k such that $|x^* - x_0| \le 2n^{-1/\rho}\gamma_k^{3N/\rho}(C,\delta)$. Moreover, by the Inductive Hypothesis, x^* is $(k, \gamma_k(C,\delta))$-hyperbolic.*

This lemma follows from a lemma in Part II of this paper [K5] that is used to prove Theorem 2.4.6 and the Shift Theorem below. It turns out that Theorem 2.4.6 can be improved to

The Shift Theorem. *Let $g \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism (respectively $g \in C^{1+\rho}(I,I)$ be a $C^{1+\rho}$-smooth map of the interval I) with some $\rho > 0$, $M_{1+\rho} = \max\{\|g^{-1}\|_{C^1}, \|g\|_{C^{1+\rho}}, 2^{1/\rho}\}$, and let $\vec{\mathbf{r}} = \{r_k\}_{k=0}^{\infty}$ be an admissible sequence* *(see Definition 1.3.1). Assume that g satisfies the Inductive Hypothesis of some order* $n-1$ *with some constants* $\delta > 0$*,* $\rho > 0$*, and* $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$*, i.e.,* $g \in IH(n-1, C, \delta, \rho)$*. If a point* $x_0 \in B^N$ *(respectively* $x_0 \in I$*) is* $(n, \gamma_n^{1/\rho}(C,\delta))$*-periodic, then either* x_0 *is* $(n, \gamma_n(C,\delta))$*-hyperbolic, or for some* $j < n\log_2 n$ *the point* $x_j = g^j(x_0)$ *has no* (n, C, δ, ρ)*-leading saddles.*

Moreover, if x_0 *has no* (n, C, δ, ρ)*-leading saddles, then either it is* $(n, \gamma_n(C,\delta))$*-simple or it is* $(k, \gamma_n(C,\delta))$*-simple and has a* (D, n, r_{2k})*-gap at* x_k *for some* k *dividing* n *and* $D = \max\{M_{1+\rho}^{30/\rho}, \exp(C/100)\}$*.*

Remark 3.5.6. *The Shift Theorem is the key for splitting all almost periodic trajectories of period n into groups and is* **the main reason why** *the estimate on the number of periodic points is* $\exp(Cn^{1+\delta})$*, not say* $\exp(Cn\log n)$*. In view of the importance of the theorem, we present an outline of its proof.*

Outline of the proof of the Shift Theorem: Let x_0 be $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic.

- If x_0 has a (k, n, C, δ, ρ)-leading saddle, then by Lemma 3.5.5 there is a periodic point $x^* = g^k(x^*)$ such that $|x_0 - x^*| \le 2n^{-1/\rho}\gamma_n^{3N/\rho}(C,\delta)$, and by the Inductive Hypothesis x^* is $(k, \gamma_k(C,\delta))$-hyperbolic. If x_{pk} remains in the $2n^{-1/\rho}\gamma_n^{3N/\rho}(C,\delta)$-neighborhood of x^* for $1 \le p < n/k$, then the approximation $df^n(x_0) \approx df^n(x^*) = (df^k(x^*))^{n/k}$ is good enough to extract from $(k, \gamma_k(C,\delta))$-hyperbolicity of x^* sufficient hyperbolicity of $df^n(x_0)$. (If n/k is not an integer, we use a similar argument with k replaced by $\gcd(n,k)$; the Euclidean Algorithm in Part II of this paper [K5] shows that x_0 is almost $\gcd(n,k)$-periodic.)

- If x_0 has a (k, n, C, δ, ρ)-leading saddle, but x_{pk} leaves the $2n^{-1/\rho}\gamma_n^{3N/\rho}(C,\delta)$-neighborhood of x^* for some $p < n/k$, then we show (assuming k to be the smallest almost-period of x_0 with this property) that x_{pk} has no (k', n, C, δ, ρ)-leading saddles for $k' < 2k$. We then proceed inductively, either proving $(k', \gamma_{k'}(C,\delta))$-hyperbolicity of x_{pk} using a k'-leading saddle at x_{pk} for some $k' \ge 2k$, or concluding that with a further shift we can eliminate all leading saddles of period up to $2k'$. Sufficient $(k', \gamma_{k'}(C,\delta))$-hyperbolicity of x_{pk} implies the necessary $(n, \gamma_n(C,\delta))$-hyperbolicity of x_0, because x_{pk} is still quite close to x_0. Either we prove the necessary hyperbolicity at some step of the induction, or after at most $\log_2 n$ shifts we eliminate all (n, C, δ, ρ)-leading saddles.

- In the case we can not prove $(n, \gamma_n(C,\delta))$-hyperbolicity of x_0 using the arguments above, for some $j < p\log_2 n < n\log_2 n$ we have that $x_j = \tilde{x}_0$ has no

(n, C, δ, ρ)-leading saddles. Notice that $\tilde{x}_0$ is $(n, M_{1+\rho}^{n \log_2 n} \gamma_n^{1/\rho}(C, \delta))$-periodic. Put $D = \max\{M_{1+\rho}^{30/\rho}, \exp(C/100)\}$. It turns out that $\tilde{x}_0$ can have at most N weak (D, n)-gaps at $\tilde{x}_{k_1}, \dots, \tilde{x}_{k_s}$ for $k_1 < \cdots < k_s < n$, $s < N$. In the 1-dimensional case it trivializes to have at most one weak (D, n)-gap. The reason is that each weak (D, n)-gap x_{k_j} after the first one at k_1 implies that the linearization $df^{k_1}(\tilde{x}_0)$ has to have an almost eigenvalue that is a k_j/k_1-root of unity with almost eigenvector $(\tilde{x}_{k_j} - \tilde{x}_0)$. Heuristically, this follows from replacing f^{k_1} in a neighborhood of $\tilde{x}_0$ by its linearization. Once we know that there are no weak (D, n)-gaps, the Euclidian Algorithm gives that

$$\min_{0<j\le n-1} |\tilde{x}_0 - \tilde{x}_j| \ge D^{-n \log_2 n}. \tag{3.60}$$

If there exist weak (D, n)-gap, then there are at most N of them. If we can prove that for any $k < n$ absence of $(D, n, r_{2k'})$-gaps for $k' < k$ implies x_0 is $(k, \gamma_n(C, \delta))$-simple, then occurrence of (D, n, r_{2k})-gap at x_k completes the proof of this part of the Theorem. Assuming now that each of weak (D, n)-gaps is not a (D, n, r_{2k})-gap for the corresponding $k < n$ we get a lower bound on $\min_{0<j\le n-1} |\tilde{x}_0 - \tilde{x}_j|$.

At this point the proof follows the following scheme. First, we discuss the case of no weak (D, n)-gaps and show that x_0 is $(n, \gamma_n(C, \delta))$-simple. It turns out that the proof of $(n, \gamma_n(C, \delta))$-simplicity is stable enough to handle the case when x_0 has no (C, n, δ, ρ)-leading saddles and $s(\le N)$ weak (D, n)-gaps which are not (D, n, r_{2k})-gap for the corresponding $k < n$. Thus, we extend calculations of the former one to the latter case.

- Consider concentric balls B_k centered at $\tilde{x}_0$ of radii $r_k = D^{-k \log_2 k}$. Pigeon-hole arguments show that the number of visits m_k of $\{\tilde{x}_j\}_{j=1}^{n-1}$ to B_k is not too large. Otherwise, there are consecutive visits $\tilde{x}_j$ and $\tilde{x}_{j+m}$ to B_k with m being small. Thus, $\tilde{x}_j$ has an (m, n, C, δ, ρ)-leading saddle and, therefore, a periodic point $\tilde{x}_j^* = g^k(\tilde{x}_j^*)$ close to $\tilde{x}_j$ and also to $\tilde{x}_0$. This implies that $\tilde{x}_0$ has an (m, n, C, δ, ρ)-leading saddle too. This is a contradiction.

  To make estimates more precise let

  $$A_k = \{1 \le j < n : \ |\tilde{x}_0 - \tilde{x}_j|\}$$

  be the collection of indices of points of the trajectory $\{\tilde{x}_j\}_{j=0}^n$ which visit the

r_k-ball B_k around $\tilde{x}_0$. In the part II of this paper [K5] we show that

$$m_k = \#\{A_k\} \le \frac{16NC}{\rho \log D} \; \frac{n}{(k \log_2 k)^{\frac{1}{1+\delta}}}. \tag{3.61}$$

- Knowing that there are no visits to B_n by (3.60) and a bound on the number of visits m_k to B_k for $1 \le k < n$, we can get a lower estimate on $\prod_{j=1}^{n-1} |\tilde{x}_0 - \tilde{x}_j|$ according the following scheme. By (3.60) we have $A_k = \emptyset$ and $m_k = 0$. Rewrite the product of distance as follows

$$\prod_{j=1}^{n-1} |\tilde{x}_j - \tilde{x}_0| = \prod_{j \notin A_1} |\tilde{x}_j - \tilde{x}_0| \prod_{j \in A_1 \setminus A_2} |\tilde{x}_j - \tilde{x}_0| \cdots \prod_{j \in A_{n-1}} |\tilde{x}_j - \tilde{x}_0|. \tag{3.62}$$

By definition of A_k for each $j \in A_{k-1} \setminus A_k$ we have $|\tilde{x}_0 - \tilde{x}_j| \ge r_k$. Put $a_k = k \log_2 k$. Then the product (5.87) admits the following lower bound

$$\begin{gathered} r_1^{k-1-m_1} r_2^{m_1-m_2} \dots r_k^{m_{k-1}-m_k} = \\ \exp\Big(-\log D \, [a_0(k-1-m_1)+ \\ a_1(m_1-m_2) + \cdots + a_k(m_{k-1}-m_k)]\Big) \end{gathered} \tag{3.63}$$

Using Abel's resummation we can rewrite the last expression in the form

$$\exp\left(-\log D \left[m_1(a_1 - a_0) + m_2(a_2 - a_1) + \cdots + m_k(a_k - a_{k-1})\right]\right). \tag{3.64}$$

By definition of a_k we have $a_{k+1} - a_k \le 2 \log_2 k$. Using inequality (3.61) above we get the following lower bound for the product (5.89)

$$\begin{aligned} \exp\left(-\frac{32NC}{\rho} n \log_2 n \sum_{k=1}^{n} k^{-\frac{1}{1+\delta}}\right) &\ge \\ \ge \frac{1}{n} \exp\left(-\frac{64NC}{\rho} n^{1+\frac{\delta}{1+\delta}} \log_2 n\right) &\ge \\ \ge \exp\left(-Cn^{1+\delta}\right), & \end{aligned} \tag{3.65}$$

where $16N \, n^{-\frac{\delta^2}{1+\delta}} \log_2 n < \rho$. This show that $\tilde{x}_0$ is $(n, \gamma_n(C, \delta))$-simple.

- Similar estimates to (3.61) hold in the case $\tilde{x}_0$ has a weak (D,n)-gap at $\tilde{x}_k$. It provides a lower bound on

$$\prod_{j=0}^{k-1} |\tilde{x}_0 - \tilde{x}_j| \geq \exp\left(-\frac{6C}{N\rho} k n^{\frac{\delta}{1+\delta}} \log_2 n\right). \tag{3.66}$$

It turns out that if $k^{1+\delta} < \frac{\rho \log D}{2C} n$ and $\tilde{x}_0$ has a weak (D,n)-gap at $\tilde{x}_k$, then $\tilde{x}_0$ is $(n, \gamma_k(C,\delta))$-hyperbolic. Indeed, in this case $|\tilde{x}_0 - \tilde{x}_k| < D^{-n} < \gamma_k^{1/\rho}(C, dt)$ and, therefore, by inductive hypothesis $\tilde{x}_0$ is $(k, \gamma_k(C,\delta))$-hyperbolic. Points $\tilde{x}_k$ and $\tilde{x}_0$ are so close that n is divisible by k we can approximate $df^n(\tilde{x}_0)$ with $(df^k(\tilde{x}_0))^{n/k}$ as we do in the first item of the proof.

Suppose we have weak (D,n)-gaps at $\tilde{x}_{k_1}$ and $\tilde{x}_{k_2}$ for some $k_1 < k_2 \leq n$ and there is no (D,n,r_{2k})-gap at $\tilde{x}_{k_1}$. Then $\tilde{x}_{k_1}$ is so close to $\tilde{x}_0$ that the Euclidean Algorithm from Part II of this paper [K5] implies that k_2 is divisible by k_1. Denote $p_1 = k_2/k_1$. Moreover, $\tilde{x}_{k_1}$ is so close to $\tilde{x}_0$ that the following approximation holds true:

$$\prod_{j=1}^{k_2} |\tilde{x}_0 - \tilde{x}_j| \geq \frac{1}{2} \prod_{s=1}^{p_1} |\tilde{x}_0 - \tilde{x}_{sk_1}| \left(\prod_{j=1}^{k_1} |\tilde{x}_0 - \tilde{x}_j|\right)^{p_1} \tag{3.67}$$

Absence of (D,n,r_{2k_1})-gap at x_{k_1} shows that

$$\begin{aligned} &|\tilde{x}_0 - \tilde{x}_{sk_1}| \geq \\ &\geq \min\left\{ D^{-n\log_2 n} \min_{0<j\leq k_1-1} |\tilde{x}_0 - \tilde{x}_j|,\ r_{2k_1}^{4(N+N^2)},\ \left(\prod_{j=0}^{k_1-2} |\tilde{x}_{k-1} - \tilde{x}_j|\right)^{4(N+2)} \right\} \end{aligned}$$

Since there is a (D,n)-weak gap at $\tilde{x}_{k_2}$ we have

$$\min_{1\leq s\leq p_1-1} |\tilde{x}_0 - \tilde{x}_{sk_1}| \geq D^{-n} |\tilde{x}_0 - \tilde{x}_{k_1}| \tag{3.68}$$

Combining (3.66), (3.67), and (3.68) we get a lower bound on $\prod_{j=1}^{k_2-1} |\tilde{x}_0 - \tilde{x}_j|$. To see that such a lower bound has the same form as (3.66) notice that both terms in the right-hand side of (3.67) are bounded from below by $\exp(-CKkn^{\frac{\delta}{1+\delta}}\log_2 n)$ for some constant $K = K(N,\rho,\delta)$ and $k_2^{1+\delta} > \frac{\rho \log D}{2C} n$. Therefore, $\prod_{j=1}^{k_2-1} |\tilde{x}_0 - \tilde{x}_j|$ has a lower bound of the same form $\exp(-CK'kn^{\frac{\delta}{1+\delta}}\log_2 n)$ for some constant $K' = K'(N,\rho,\delta)$. Iterating this arguments for weak (D,n)-gaps at $\tilde{x}_{k_1}, \tilde{x}_{k_2}, \dots, \tilde{x}_{k_s} = \tilde{x}_0$ which are not (D,n,r_{2k})-gaps we obtain lower bound on

$$\prod_{j=1}^{n-1} |\tilde{x}_0 - \tilde{x}_j| \geq \exp(-CK''kn^{\frac{\delta}{1+\delta}}\log_2 n) \tag{3.69}$$

for some constant $K'' = K''(N,\rho,\delta)$. This proves $(n,\gamma_n(C))$-simplicity of $\tilde{x}_0$.

- Unfortunately, using this scheme *only for* $\gamma_n(C,\delta) = \exp(-Cn^{1+\delta})$ *we can prove that above* $\tilde{x}_0 = x_j$ *is* $(n,\gamma_n(C,\delta))$*-simple.* For example, if $\gamma_n(C,\delta) = \exp(-Cn\log n)$, then $\tilde{x}_0$ *as above might be non-*$(n,\gamma_n(C,\delta))$*-simple.* In Appendix D, we construct examples of trajectories of a full shift on two symbols that have no (n,C,δ,ρ)-leading saddles and are not $(n,\gamma_n(C,\delta))$-simple for $\gamma_n(C,\delta) = \exp(-Cn\log n)$. This shows that additional ideas are required to improve $\gamma_n(C,\delta) = \exp(-Cn^{1+\delta})$ to $\gamma_n(C,\delta) = \exp(-Cn\log n)$ or better.

3.5.1 Decomposition of nonsimple parameters into groups

With the notations of the general problem (2.42) and (2.43), for any $D > 2$ we introduce the set of parameters associated with an almost periodic point of period n having a gap at the k-th point of its trajectory:

$$\begin{aligned}&B_n^{\mathrm{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D) = \{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}}) : \tilde{f}_{\vec{\varepsilon}} \in IH(n-1,C,\delta,\rho),\\ &\tilde{f}_{\vec{\varepsilon}} \text{ has an } (n,\gamma_n^{1/\rho}(C,\delta))\text{-periodic, but not } (n,\gamma_n(C,\delta))\text{-hyperbolic}\\ &\text{point } x_0 \text{ with a } (D,n,r_{2k})\text{-gap at } x_k = \tilde{f}_{\vec{\varepsilon}}^k(x)\}\end{aligned} \tag{3.70}$$

Choose $D = \max\{M_{1+\rho}^{30/\rho}, \exp(C/100)\}$. Theorem 2.4.6 implies existence of inclusions (2.42) and (2.43). Similarly, the Shift Theorem implies that the following inclusions hold:

$$\begin{aligned}&B_n^{\mathrm{non}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}) \subseteq \cup_{k|n} B_n^{\mathrm{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D);\\ &B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}) \subseteq B_n^{\mathrm{sim}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}) \cup \left(\cup_{k|n} B_n^{\mathrm{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D)\right).\end{aligned} \tag{3.71}$$

Return to our C^2-smooth 1-dimensional model, $N = \rho = 1$. To fit the notation of the model, for a sufficiently small γ_n, e.g. $\gamma_n \le \gamma_n(C,\delta)$, we introduce the set

$$\begin{aligned} & B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D) = \{\varepsilon \in HB^{\mathrm{st}}_{<2n}(\tau) : \tilde f_\varepsilon \in IH(n-1,C,\delta,1), \\ & \tilde f_\varepsilon \text{ has an } (n,\gamma_n)\text{-periodic, but not } (n,\gamma_n)\text{-hyperbolic} \\ & \text{point } x_0 \text{ with a } (D,n,\tau)\text{-gap at } x_k = \tilde f^k_\varepsilon(x)\} \end{aligned} \tag{3.72}$$

and rewrite inclusion (3.71) in the notation of the 1-dimensional model

$$\begin{aligned} & B^{\mathrm{non}}_{n,\tau}(C,\delta,\tilde f,\gamma_n) \subseteq \cup_{k|n} B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D); \\ & B_{n,\tau}(C,\delta,\tilde f,\gamma_n) \subseteq B^{\mathrm{sim}}_{n,\tau}(C,\delta,\tilde f,\gamma_n) \cup \left(\cup_{k|n} B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D)\right). \end{aligned} \tag{3.73}$$

This inclusion shows that to prove Proposition 3.5.1, it is sufficient to prove

Proposition 3.5.7. *With the conditions of Proposition 3.5.1, let k be some integer that divides n. Then for $D = \max\{M_2^{30}, \exp(C/100)\}$ and a sufficiently small positive γ_n, e.g., $\gamma_n \le \gamma_n(C,\delta)$, we have the following estimate on the measure of maps $\tilde f_\varepsilon$'s associated with an (n,γ_n)-periodic, but not (n,γ_n)-hyperbolic point that has a (D,n,τ)-gap at the k-th point of its trajectory.*

$$\begin{aligned} & \mu^{\mathrm{st}}_{<2n,\tau}\{B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D)\} \le \\ & \quad \le 128 n^\delta\, 6^{2n}\, M_2^{10n+1} \exp(-Cn\log_2 n/200). \end{aligned} \tag{3.74}$$

Let us give a name to the right-hand side of the inequality (3.59) for the Brick of parameters $HB^{\mathrm{st}}_{<2k}(\tau)$ of the standard thickness. Let $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ be a k-tuple and $D > 2$ be some number. We call the number

$$\begin{aligned} & \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\} = \\ & \min\left\{ D^{-n\log_2 n} \min_{0<j\le k-1} |x_0 - x_j|,\ \frac{\tau^{4(N+N^2)}}{((2k)!)^{4(N+N^2)}},\ \prod_{j=0}^{k-2} |x_{k-1} - x_j|^{4(N+2)} \right\} \end{aligned} \tag{3.75}$$

the (D,n,τ)-gap number associated to the k-tuple $\mathbf{X}_k$ and the Brick $HB^{\mathrm{st}}_{<2n}(\tau)$ of standard thickness. In the case under current consideration, $N = 1$. Similarly, one can define the (D,n,r_{2k})-gap number for a nonincreasing sequence $\vec{\mathbf{r}} = \{r_m\}_{m=0}^\infty$ replacing $\tau/(2k)!$ by r_{2k}.

Let $\mathbf{X}_k(x_0,g) = \{g^j(x_0)\}_{j=0}^{k-1}$. By the definition of the (D,n,τ)-gap number, if x_0 has a (D,n,τ)-gap at x_k, then x_0 is $(k, \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k(x_0,g), D\})$-periodic. Introduce the set of "bad" parameters

$$\begin{aligned}
&B^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D)=\{\varepsilon\in HB^{\mathrm{st}}_{<2n}(\tau):\ \tilde f_\varepsilon\in IH(n-1,C,\delta,1),\\
&\tilde f_\varepsilon \text{ has a point } x_0 \text{ that is } (n,\gamma_n(C,\delta))\text{-periodic},\\
&\ \text{has a } (D,n,\tau)\text{-gap at } x_k=\tilde f^k_\varepsilon(x),\ \text{is } (k,\gamma_n(C,\delta))\text{-simple},\\
&\text{and is not } (k,M_2^{3n}\Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k(x_0,\tilde f_\varepsilon),D\})\text{-hyperbolic}\}.
\end{aligned}\tag{3.76}$$

In Section 3.6 using the Shift Theorem above we shall prove the following lemma.

Lemma 3.5.8. *Let $C>0$, $\delta>0$, and $k,n\in\mathbb{Z}_+$ be some positive integer, and let k divide n. Then with the notation above for any $\gamma_n\le\gamma_n(C,\delta)$ and $D=\max\{M_2^{30},\exp(C/100)\}$ we have*

$$B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D)\subset B^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D).\tag{3.77}$$

The proof is postponed until Section 3.6.

Remark 3.5.9. $B^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D)$ *depends only on properties of trajectories of length k, whereas $B^{\mathrm{gap}(k)}_{n,\tau}(C,\delta,\tilde f,\gamma_n;D)$ depends on properties of those of length n. This reduces consideration of sufficiently recurrent, nonsimple, and not hyperbolic, trajectories of length n to studying only their initial "simple" parts of length k, where x_k is a point of the first sufficiently close return to x_0 with n divisible by k. In this case one can effectively perturb such a trajectory at the k-th return using the Newton family $f(x)+u_0+u_1(x-x_0)+\cdots+u_{2k-1}(x-x_{k-1})\prod_{j=0}^{k-2}(x-x_j)^2$.*

3.5.2 Decomposition into i-th recurrent pseudotrajectories

We now split the set $B^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D)$ into a finite union. Let $\tilde\gamma_{n,i}(C,\delta,\rho)=M_2^{4ni}\gamma_n^{1/4\rho}(C,\delta)$.

Definition 3.5.10. *Let $g\in\mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism (respectively $g\in C^{1+\rho}(I,I)$ be a $C^{1+\rho}$-smooth map), $C>0$, $\delta>0$, $\tau>0$, and $D>2$ be some constants. We call a point $x_0\in B^N$ (respectively $x_0\in I$) i-th recurrent with constants (C,δ,ρ,τ,D) if for $\mathbf{X}_k(x_0,g)=\{g^j(x_0)\}_{j=0}^{n-1}$ we have*

$$\tilde\gamma_{n,i}(C,\delta,\rho)\ \le\ \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k(x_0,g),D\}\ <\ \tilde\gamma_{n,i+1}(C,\delta,\rho).\tag{3.78}$$

Remark 3.5.11. *Recall that the (D,n,r_{2k})-gap number is defined by replacing in the definition of the (D,n,τ)-gap number $\tau/(2k)!$ by r_{2k}. Similarly, we define an i-th recurrent point with constants (C,δ,ρ,r_{2k},D) by replacing in Definition 3.5.10 above the (D,n,τ)-gap number $\Delta^{st}_{k,n,\tau}\{\mathbf{X}_k,D\}$ by the (D,n,r_{2k})-gap number.*

For the purpose of this section, $\rho = 1$, and for brevity redenote $\tilde{\gamma}_{n,i} = \tilde{\gamma}_{n,i}(C, \delta, 1)$. Define the set of parameters from $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D, i)$ associated to i-th order recurrent with respect to $(C, \delta, 1, D)$ trajectories that satisfy conditions (3.76):

$$B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D, i) = \{\varepsilon \in HB^{\text{st}}_{<2n}(\tau) : \tilde{f}_\varepsilon \text{ has a point } x_0 \text{ as in (3.76) and } i\text{-th recurrent with constants } (C, \delta, 1, \tau, D)\}. \tag{3.79}$$

Lemma 3.5.12. *With the notations of Lemma 3.5.8, for* $L = Cn^\delta/(4 \log M_2)$ *we have*

$$B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D) \subseteq \bigcup_{i=0}^{L-1} B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}; M_2, D, i). \tag{3.80}$$

Lemma 3.5.13. *Let* $C > 0$, $\delta > 0$, *and* $n \in \mathbb{Z}_+$ *be some numbers, and let* $k \in \mathbb{Z}_+$ *divide* n. *Then with the notation as above, for* $D = \max\{M_2^{30}, \exp(C/100)\}$, *any* $\gamma_n \le \gamma_n(C, \delta)$, *and any* $i \in \mathbb{Z}_+$ *such that* $0 \le i < L$ *we have*

$$\mu^{\text{st}}_{<2n,\tau}\{B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D, i)\} \le 128 \cdot 6^{2k}\ M_2^{10n+1}\ D^{-n \log_2 n/2}. \tag{3.81}$$

We postpone the proof of this Lemma until the next Section.

Proofs of Propositions 3.5.1 and 3.5.7: Since there are at most n numbers $k < n$ dividing n, Proposition 3.5.1 is an easy corollary of Proposition 3.5.7. To prove Proposition 3.5.7, we combine the Shift Theorem, Lemmas 3.5.8, 3.5.12, and 3.5.13 as follows. Let γ_n be sufficiently small, e.g., $\gamma_n \le \gamma_n(C, \delta)$, and let $D = \max\{M_2^{30}, \exp(C/100)\}$. Then

- By the Shift Theorem, the set of all parameter values $B^{\text{non}}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n)$ associated to maps $\tilde{f}_\varepsilon$ with an (n, γ_n)-periodic, essentially non-(n, γ_n)-simple, but not (n, γ_n)-hyperbolic point, is contained in the union (3.73) over all k dividing n of $B^{\text{gap}(k)}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n; D)$ associated to maps $\tilde{f}_\varepsilon$ with an (n, γ_n)-periodic, but non-(n, γ_n)-hyperbolic point with a (D, n, τ)-gap at x_k.
- By Lemma 3.5.8 we have that $B^{\text{gap}(k)}_{n,\tau}(C, \delta, \tilde{f}, \gamma_n; D)$ is contained in the set of parameters $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D)$ associated to maps $\tilde{f}_\varepsilon$ which have a non-$(k, \Delta^{\text{st}}_{k,n,\tau}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\})$-hyperbolic point x_0 with a (D, n, τ)-gap at x_k.
- By Lemma 3.5.12, in turn, $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D)$ is contained in the union of $\{B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D, i)\}_{i=0}^{L-1}$ such that the i-th set $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C, \delta, \tilde{f}, M_2; D, i)$ is associated to maps $\tilde{f}_\varepsilon$ that have a non-$(k, \Delta^{\text{st}}_{k,n,\tau}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\})$-hyperbolic point x_0 with a (D, n, τ)-gap at x_k and such that the k-tuple $\{x_j = \tilde{f}^j_\varepsilon(x_0)\}_{j=0}^{k-1}$ is i-th recurrent.

• Application of Lemma 3.5.13 provides well enough estimates the measures of $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D,i)$ to prove Proposition 3.5.7 and hence Proposition 3.5.1. Q.E.D.

In the next section, we shall prove Lemma 3.5.13 using the Discretization Method and then we shall prove Lemmas 3.5.8 and 3.5.12 using simple approximation arguments similar to the one given in the proof of Proposition 3.1.2.

3.6 The measure of maps $\tilde f_\varepsilon$ having i-th recurrent, non sufficiently hyperbolic trajectories with a gap and proofs of auxiliary lemmas

We shall prove Lemma 3.5.13 in three steps.

Step 1. Reduction to polynomial perturbations of degree $2k-1$.

The measure $\mu^{\text{st}}_{<2n,\tau}$ is Lebesgue product probability measure and each of its component $\mu^{\text{st}}_{m,\tau}$ is Lebesgue probability measure (see (3.3))

$$\mu^{\text{st}}_{<2n,\tau} = \mu^{\text{st}}_{<2k,\tau} \times \left(\times_{m=2k}^{2n-1}\mu^{\text{st}}_{m,\tau}\right). \tag{3.82}$$

Therefore, by Fubini/Tonelli Theorem it is sufficient to prove that

$$\mu^{\text{st}}_{<2k,\tau}\left\{B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D,i)\cap\{(\varepsilon_{2k},\dots,\varepsilon_{2n-1})\}\right\} \\ \le 128\cdot 6^{2k}\ M_2^{10n+1}\ D^{-n\log_2 n/2}.$$

uniformly over $\{(\varepsilon_{2k},\dots,\varepsilon_{2n-1})\} \in HB^{\text{st}}_{2k}(\tau)\times HB^{\text{st}}_{2k+1}(\tau)\times\cdots\times HB^{\text{st}}_{2n-1}(\tau)$. To simplify notations we omit $(\varepsilon_{2k},\dots,\varepsilon_{2n-1})$ and write as if the set of parameters $B^{\text{gap}(k),\Delta^{\text{st}}}_{n,\tau}(C,\delta,\tilde f,M_2;D,i)$ is a subset of the Brick $HB^{\text{st}}_{<2k}(\tau)$.

From now on we consider the $2k$-parameter family of polynomial perturbations

$$\left\{\tilde f_{\varepsilon(2k-1)} = \tilde f(x) + \sum_{m=0}^{2k-1}\varepsilon_m x^m\right\}, \tag{3.83}$$

where $\varepsilon(2k-1)=(\varepsilon_0,\dots,\varepsilon_{2k-1})\in HB^{\text{st}}_{<2k}(\tau)$. Recall that $HB^{\text{st}}_{<2k}(\tau)$ is supplied with the Lebesgue product probability measure $\mu^{\text{st}}_{<2k,\tau}$.

Step 2. An estimate of the measure of parameters associated with a trajectory $\{x_j\}_{j=0}^{n-1}$ *with a gap at* x_k *that is* i*-th recurrent and not sufficiently hyperbolic.*

With the notations of Section 3.2 we give the following

Definition 3.6.1. *Let $D > 2$. We say that a $\tilde{\gamma}_n$-pseudotrajectory $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1} \subset I^k_{\tilde{\gamma}_n}$ associated to some map $g : I \hookrightarrow I$ has a (D, n, τ)-gap at a point $x_k = g(x_{k-1})$ if (3.59) is satisfied and there is no $m < k$ such that x_0 has a (D, n, τ)-gap at x_m.*

By analogy with (3.15), for each i satisfying $0 \le i < L$ of Lemma 3.5.12 we define

$$\begin{aligned}
&B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(\tilde{f}, \tilde{\gamma}_{n,i}, M_2; D, i; x_0, \dots, x_{k-1}) = \{\varepsilon \in HB^{\mathrm{st}}_{<2k}(\tau) : \\
&\mathbf{X}_k = \{x_m\}_{m=0}^{k-1} \subset I^k_{\tilde{\gamma}_{n,i}} \text{ is a } \tilde{\gamma}_{n,i}\text{-pseudotrajectory associated to } \varepsilon \\
&\text{that is } i\text{-th recurrent with constants } (C, \delta, 1, \tau, D), \\
&\left(k, 2M_2^n \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\}\right)\text{-periodic and not } \left(k, 2M_2^{3n} \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\}\right)\text{-hyperbolic}\}
\end{aligned} \tag{3.84}$$

To show an analogy with the case simple trajectory case consider the following

Problem 3.6.2. *Estimate the measure of $\varepsilon \in HB^{\mathrm{st}}_{<2k}(\tau)$ for which the k-tuple $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ is*

$$\begin{aligned}
&A)\ \text{a } \tilde{\gamma}_{n,i}\text{-pseudotrajectory, i.e., } |\tilde{f}_\varepsilon(x_j) - x_{j+1}| \le \tilde{\gamma}_{n,i} \ \text{ for } \ j = 0, \dots, k-2; \\
&B)\ i\text{-th recurrent with constants } (C, \delta, 1, \tau, D), \text{i.e.,} \\
&\qquad \tilde{\gamma}_{n,i} \le \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\} < \tilde{\gamma}_{n,i+1}; \\
&C)\ (k, 2M^n \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\})\text{-periodic, i.e.,} \\
&\qquad |\tilde{f}_\varepsilon(x_{k-1}) - x_0| < 2M^n \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\}; \\
&D)\ \text{not } (n, 2M_2^{3n} \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\})\text{-hyperbolic, i.e.,} \\
&\qquad \left| \left| \prod_{j=0}^{k-1} (\tilde{f}_\varepsilon)'(x_j) \right| - 1 \right| \le 2M_2^{3n} \Delta^{\mathrm{st}}_{k,n,\tau}\{\mathbf{X}_k, D\};
\end{aligned} \tag{3.85}$$

For a fixed k-tuple of points $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$, consider the Newton family of polynomial perturbations

$$\tilde{f}_{u,\mathbf{X}_k}(x) = \tilde{f}(x) + \sum_{m=0}^{2k-1} u_m \prod_{j=0}^{m-1} (x - x_{j \pmod{k}}). \tag{3.86}$$

Recalling (2.17) and Fig. 2.2, we notice that for any $m < k$ and $s > 0$ the image $\tilde{f}_{u,\mathbf{X}_k}(x_m)$ (respectively derivative $\tilde{f}'_{u,\mathbf{X}_k}(x_m)$) of (respectively at) the point x_m is independent of the Newton coefficients u_{m+s} (respectively u_{m+k+s}) with $s > 0$. This implies that the Newton coefficients $u_0, \dots, u_{k-2}$ determine if property (A) holds.

Fix $u_0, \ldots, u_{k-2}$. Now the Newton coefficient u_{k-1} determines if property (C) of almost periodicity holds. Once u_{k-1} is fixed, the Newton coefficient u_{2k-1} determines if property (D) of almost nonhyperbolicity holds.

Following formulas (3.27), (3.28), and formulas (3.30), (3.33) with n replaced by k and γ_n replaced by $2M^n \Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}$ for periodicity and $2M^{3n}\Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}$ for hyperbolicity and, we have that for fixed $\mathbf{X}_k$, the measure of $(u_0, \ldots, u_{k-1}, u_{2k-1})$ with conditions (3.85) is bounded as in (3.34). Then we apply the Distortion Lemma with n replaced by k. This gives an additional factor 6^k. Thus we get a bound

$$\begin{gathered}\mu^{\rm st}_{<2k,\tau}\Big\{u(2k-1) \in W^{u,\mathbf{X}_n}_{<2k,1} :\ \tilde{f}_{u,\mathbf{X}_k} \text{ satisfies conditions (3.85)}\Big\} \le \\ 4\ M_2\ 6^k\ \prod_{m=0}^{k-2} \frac{m!\ \tilde{\gamma}_{n,i}}{\tau\ \prod_{j=0}^{m-1}|x_m - x_j|} \\ \frac{(k-1)!}{\tau}\ \frac{2M_2^n \Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}}{\prod_{j=0}^{k-2}|x_{k-1}-x_j|}\ \frac{(2k-1)!}{\tau}\ \frac{2M_2^{3n}\ \Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}}{\prod_{j=0}^{k-2}|x_{k-1}-x_j|^2}.\end{gathered} \tag{3.87}$$

Consider the last two terms in the right-hand side product. Definition of the (D, n, τ)-gap number and the inequality $\min(a, b, c) \le a^{1/2}\ b^{1/4}\ c^{1/4}$ show that these two terms are bounded by

$$M_2^{4n} \frac{k!}{\tau} \frac{(2k)!}{\tau} \frac{\left(\Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}\right)^2}{(\prod_{j=0}^{k-2}|x_{k-1}-x_j|)^3} \le 2M_2^{4n} D^{-n\log_2 n/2} \Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\} \tag{3.88}$$

Therefore,

$$\begin{gathered}\mu^{\rm st}_{<2k,\tau}\Big\{u(2k-1) \in W^{u,\mathbf{X}_n}_{<2k,1} :\ \tilde{f}_{u,\mathbf{X}_k} \text{ satisfies conditions (3.85)}\Big\} \\ \le 32\ M_2^{4n+1}\ 6^k\ \prod_{m=0}^{k-2} \frac{m!\ \tilde{\gamma}_n}{\tau\ \prod_{j=0}^{m-1}|x_m - x_j|}\ M_2^{4n}\ D^{-n\log_2 n/2}\ \Delta^{\rm st}_{k,n,\tau}\{\mathbf{X}_k, D\}.\end{gathered} \tag{3.89}$$

The next step of the proof of Lemma 3.5.13 as in the case of simple trajectories (Section 3.2) is to collect all possible "bad" pseudotrajectories and make sure that those pseudotrajectories indeed approximate sufficiently well all "bad" true trajectories.

Step 3. Collecting of "bad" i-th recurrent, not sufficiently hyperbolic trajectories $\{x_j\}_{j=0}^{k-1}$ *with a gap at* x_k *using grids* $\{I_{\tilde{\gamma}_{n,i}}\}_i$ *of variable size in* i.

In the case of simple trajectories in Section 3.2, we considered only one $\tilde{\gamma}_n$-grid of a fixed size and collected all simple "bad" trajectories in the Collection Lemma (Section 3.4). In the case of i-th recurrent trajectories with a gap at x_k we define grids $\{I_{\tilde{\gamma}_{n,i}}\}_i$ of i dependent size $\tilde{\gamma}_{n,i}$. Then we prove that $\tilde{\gamma}_{n,i}$-pseudotrajectories approximate real i-th recurrent trajectories with a gap at x_k sufficiently well. Finally, we collect all possible i-th recurrent $\tilde{\gamma}_{n,i}$-pseudotrajectories with a gap at x_k and sum the estimates of the measures of "bad" sets. Let's realize this program. Recall that $\tilde{\gamma}_{n,i} = M_2^{4ni}\gamma_n(C,\delta)$ and call the i-th grid

$$I_{\tilde{\gamma}_{n,i}} = \{x \in I : \exists k \in \mathbb{Z} \text{ such that } x = (2k+1)\tilde{\gamma}_{n,i}\} \subset I. \tag{3.90}$$

Definition 3.6.3. *Let $\{x_j\}_{j=0}^{k-1} \in I_{\tilde{\gamma}_{n,i}}^k$ be a k-tuple for some $i \in \mathbb{Z}_+$. Then the k-tuple $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ is called i-th recurrent with constants $D > 2$ and $M_2 > 0$ if*

$$\frac{1}{2}\ \tilde{\gamma}_{n,i} \le \Delta_{k,n,\tau}^{\mathrm{st}}\{\mathbf{X}_k, D\} \le 2\ \tilde{\gamma}_{n,i+1}. \tag{3.91}$$

For simplicity if the k-tuple $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ is a $\tilde{\gamma}_{n,i}$-pseudotrajectory associated to some parameter $\varepsilon \in HB_{<2k}^{\mathrm{st}}(\tau)$ which is i-th recurrent with some constants $D > 2$ and the C^2-norm M_2 of the family (3.7), then we say that $\mathbf{X}_k$ is an i-th recurrent $\tilde{\gamma}_{n,i}$-pseudotrajectory.

Remark 3.6.4. *Definition of an i-th recurrent pseudotrajectory is reasonable only for a grid of a sufficiently small size. Indeed, for any i-th recurrent trajectory we need to find a $\tilde{\gamma}_{n,i}$-pseudotrajectory whose periodicity, hyperbolicity, and product of distances along itself approximate well enough those of the trajectory.*

Define a discretized version of the set $B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f},M_2;D,i)$, given by (3.79) associated with all i-th recurrent trajectories of length k which has a weak (D,n,τ)-gap at x_k and is not $M_2^{3n}\Delta_{k,n,\tau}^{\mathrm{st}}\{\mathbf{X}_k(x_0,\tilde{f}_\varepsilon),D\}$-hyperbolic.

$$\begin{aligned}&B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f},M_2,\tilde{\gamma}_{n,i};D,i;x_0) =\\ &\cup_{\mathbf{X}_k=\{x_j\}_{j=1}^{k-1}\in I_{\tilde{\gamma}_{n,i}}^{k-1}} B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f},M_2,\tilde{\gamma}_{n,i};D,i;x_0,\dots,x_{k-1}).\end{aligned} \tag{3.92}$$

Similarly to the case of simple trajectories, we prove that after discretization all real "bad" trajectories can be sufficiently well approximated by pseudotrajectories of a certain grid so that quantities of periodicity, hyperbolicity, existence of a gap, and product of distances along the trajectory are almost the same. Namely,

Lemma 3.6.5. *In the notations as above we have*

$$B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f};D,i) \subset \cup_{x_0\in I_{\tilde{\gamma}_{n,i}}} B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f},M_2,\tilde{\gamma}_{n,i};D,i;x_0). \tag{3.93}$$

The proof of this Lemma is very similar to the proof of Lemma 3.5.8 and is omitted.

To complete the proof of Lemma 3.5.13 we apply the Collection Lemma to (3.89) to estimate the measure of $B_{n,\tau}^{\text{gap}(k),\Delta^{\text{st}}}(C, \delta, \tilde{f}, M_2, \tilde{\gamma}_{n,i}; D, i; x_0)$. We get

$$\mu_{<2k,\tau}^{st}\{B_{n,\tau}^{\text{gap}(k),\Delta^{\text{st}}}(C, \delta, \tilde{f}, M_2, \tilde{\gamma}_{n,i}; D, i; x_0)\} \leq$$

$$\leq 32 \cdot 6^{2k} M_2^{4n+1} D^{-n\log_2 n/2} \Delta_{k,n,\tau}^{st}\{\mathbf{X}_k, D\}.$$

The number of grid points is $2\gamma_{n,i}^{-1}$. By definition of i-th recurrent pseudotrajectory we have $2\gamma_{n,i}^{-1}\Delta_{k,n,\tau}^{st}\{\mathbf{X} - k, D\} \leq 4M_2^{6n}$. This implies (3.81), which in turn implies Lemma 3.5.13. Remarks at the end of the last section also show that this proves Proposition 3.5.7 too. This Proposition implies Proposition 3.5.1.

Proof of Lemma 3.5.8: Fix a parameter $\varepsilon \in B_{n,\tau}^{\text{gap}(k)}(C, \delta, \tilde{f}, \gamma_n; D)$; we wish to show that $\varepsilon \in B_{n,\tau}^{\text{gap}(k),\Delta^{\text{st}}}(C, \delta, \tilde{f}, M_2; D)$. By definition (3.72) there is an (n, γ_n)-periodic, non-(n, γ_n)-hyperbolic point x_0 with a (D, n, τ)-gap at $x_k = \tilde{f}_\varepsilon^k(x_0)$. By definition (3.76), we want to show that x_0 is not $(k, M_2^{3n}\Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\})$-hyperbolic.

Since n is divisible by k, we can split the trajectory $\{x_j = \tilde{f}_\varepsilon^j(x_0)\}_{j=0}^{n-1}$ of length n into $p = n/k$ parts of length k each. Consider the linearization

$$(\tilde{f}_\varepsilon^n)'(x_0) = (\tilde{f}_\varepsilon^k)'(x_{(p-1)k}) \cdot \dots \cdot (\tilde{f}_\varepsilon^k)'(x_k) \cdot (\tilde{f}_\varepsilon^k)'(x_0). \tag{3.94}$$

Definition 3.5.3 of a (D, n, τ)-gap at x_k says that $|x_0 - x_k| \leq \Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}$, and hence for every $1 \leq j \leq p-1$ and $0 \leq s < k$ we have

$$|x_s - x_{jk+s}| \leq M_1^n \ \Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}. \tag{3.95}$$

This implies that the trajectory $\{x_j\}_{j=0}^{n-1}$ of length n consists of p almost identical parts of length k each. Thus, for each $0 \leq m < n$ we have

$$\left|(\tilde{f}_\varepsilon)'(x_m) - (\tilde{f}_\varepsilon)'(x_{m \pmod{k}})\right| \leq M_2^{n+1}\Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}. \tag{3.96}$$

Considering the product over $m = 0, 1, \dots, n-1$ of each term on the left hand side above, we get that

$$\left|(\tilde{f}_\varepsilon^n)'(x_0) - \left((\tilde{f}_\varepsilon^k)'(x_0)\right)^p\right| \leq nM_2^{2n}\Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}. \tag{3.97}$$

Therefore, since $\left|\left|(\tilde{f}_\varepsilon^n)'(x_0)\right| - 1\right| \leq \gamma_n \leq \gamma_n(C, \delta)$, and by the proof of Lemma 3.5.12 below, $\gamma_n(C, \delta) < \Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}$, it follows that

$$\left||(\tilde{f}_\varepsilon^k)'(x_0)| - 1\right| \leq \left||\left((\tilde{f}_\varepsilon^k)'(x_0)\right)^p| - 1\right| \leq 2nM_2^{2n}\Delta_{k,n,\tau}^{\text{st}}\{\mathbf{X}_k(x_0, \tilde{f}_\varepsilon), D\}, \tag{3.98}$$

and ε belongs to $B_{n,\tau}^{\mathrm{gap}(k),\Delta^{\mathrm{st}}}(C,\delta,\tilde{f},M_2;D)$. Q.E.D.

Proof of Lemma 3.5.12: It follows from the Shift Theorem (last sentence) with $N=1$ that if a point x_0 has no $(n,C,\delta,1)$-leading saddles and a (D,n,τ)-gap at x_k, then x_0 is $(k,\gamma_n(C,\delta))$-simple. Therefore, for the (D,n,τ)-gap number (3.75) we have the following bounds

$$\gamma_n(C,\delta)=\exp\left(-Cn^{1+\delta}\right)\le \Delta_{k,n,\tau}^{\mathrm{st}}\{\mathbf{X}_k(x_0,\tilde{f}_\varepsilon),D\}\le 2D^{-n}. \tag{3.99}$$

Thus since $D=M_2^{30}\ge 2$, each such x_0 with a (D,n,τ)-gap at x_k is i-th recurrent with constants (C,δ,ρ,τ,D) for some $i\ge 0$. If $i\ge L=Cn^\delta/(4\log M_2)$, then $\Delta_{k,n,\tau}^{\mathrm{st}}\{\mathbf{X}_k(x_0,\tilde{f}_\varepsilon),D\}\ge M_2^{4ni}\gamma_n(C,\delta)\ge 1>2D^{-n}$. This contradicts the above inequality. Q.E.D.

Proof of Lemma 3.6.5: The proof is by approximation arguments (3.94)-(3.98) similar to the proof of Lemma 3.5.8 above. Q.E.D.

Chapter 4

Comparison of the Discretization Method in 1–dimensional and N–dimensional Cases

In this section we discuss changes that one has to make and difficulties that arise while generalizing the Discretization Method for C^2–smooth 1–dimensional maps to a Discretization Method for the $C^{1+\rho}$–smooth N–dimensional diffeomorphisms. Many of the arguments, such as the decomposition into pseudotrajectories of Section 3.2, remain essentially the same. We focus primarily on the obstacles to generalizing the estimates of Sections 3.3 and 3.4 for simple trajectories.

4.1 Dependence of the main estimates on N and ρ

Recall that the Inductive Hypothesis (Definition 2.0.1) asserts that points that are $(n, \gamma_n^{1/\rho}(C, \delta))$-periodic are also $(n, \gamma_n(C, \delta))$-hyperbolic. The reason for the exponent $1/\rho$ is as follows. In the Discretization Method, we approximate trajectories of length n with $\tilde{\gamma}_n$-pseudotrajectories on a grid. In order that γ_n-hyperbolicity of the pseudotrajectory imply hyperbolicity of the true trajectory for a $C^{1+\rho}$ map, it requires that $\tilde{\gamma}_n \leq \gamma_n^{1/\rho}$ (up to a factor exponential in n). In our heuristic estimate (3.24) on the measure of "bad" parameters, the number of initial points in the $\tilde{\gamma}_n$-grid is (again up to an exponential factor) $\tilde{\gamma}_n^{-N} \geq \gamma_n^{-N/\rho}$. The best possible bounds one can get (for nonrecurrent trajectories) on the "measure of periodicity" and the "measure of hyperbolicity" in (3.24) are respectively the Nth power of the periodicity and the hyperbolicity γ_n. Thus, in order for the right side of (3.24) to be small, we need that the periodicity be at most roughly $\gamma_n^{1/\rho}$.

Next, we explain the exponent $1/4N$ in Definition 2.34 of simple trajectories. The actual estimate we obtain on the "measure of periodicity" discussed in the previous paragraph is not the Nth power of the periodicity $\gamma_n^{N/\rho}$, but instead $\gamma_n^{N/\rho}\ r_{n-1}^{-N}\ (\prod_{j=0}^{n-2}|x_{n-1}-x_j|)^{-N}$, where $\{x_j\}_{j=0}^{n-1}$ is a trajectory and r_{n-1} is the width of an admissible Hilbert Brick in the direction of degree $n-1$ polynomials (see Definition 1.3.1). This estimate reduces to (3.30) in the case $N=\rho=1$ and $r_{n-1}=\tau/(n-1)!$, and is obtained in a similar fashion, treating each of the N coordinates independently. In Appendix A, we will show (Proposition A.5) that the analogue of the bound (3.33) on the "measure of hyperbolicity" is, up to a factor exponential in n, $\gamma_n\ r_{2n-1}^{-N^2}\ (\prod_{j=0}^{n-2}|x_{n-1}-x_j|)^{-2}$. Again ignoring exponential factors, with $\tilde{\gamma}_n=\gamma_n^{1/\rho}$ the number of initial points in (3.24) is $\gamma_n^{-N/\rho}$, making the right side of (3.24)

$$\gamma_n\ r_{n-1}^{-N}\ r_{2n-1}^{-N^2}\ \left(\prod_{j=0}^{n-2}|x_{n-1}-x_j|\right)^{-N-2} \tag{4.1}$$

By Definition 1.3.1, for admissible Hilbert Bricks, γ_n decays faster than any power of r_{2n}. Thus to make the bound (3.24) on the measure of "bad" parameters small, we basically need γ_n to dominate $(\prod_{j=0}^{n-2}|x_{n-1}-x_j|)^{-N-2}$. This is certainly true if $\prod_{j=0}^{n-2}|x_{n-1}-x_j|\geq\gamma_n^{1/4N}$. Though our choice of the particular exponent $1/4N$ is somewhat arbitrary, the factor of N in the exponent is necessary.

4.2 The multidimensional space of divided differences and dynamically essential parameters

In the 1-dimensional case, with a fixed n-tuple of points $\{x_j\}_{j=0}^{n-1}\subset I$, the space of Newton Interpolation Polynomials (the Divided Differences Space) is $2n$–dimensional. In Section 2.2, in formulas (2.17), we noticed that there is a simple relation between parameters of Newton Interpolation Polynomials and dynamical properties of the trajectory. In terms of the family (3.86), we have that u_0 determines the position of $f_{u(2n-1);\mathbf{x}_n}(x_0)$, that u_1 determines the position of $f_{u(2n-1);\mathbf{x}_n}(x_1)$, provided that u_0 (and hence $f_{u(2n-1);\mathbf{x}_n}(x_0)$) is fixed, and so on — for $k=2,\dots,n-1$, we have that u_k determines the position of $f_{u(2n-1);\mathbf{x}_n}(x_k)$, provided that $\{u_j\}_{j=0}^{k-1}$ (hence

$\{f_{u(2n-1);\mathbf{X}_n}(x_j)\}_{j=0}^{k-1}$) are fixed. Similarly, for $k = 0, \ldots, n-1$, we have that u_{n+k} determines the derivative $(f_{u(2n-1);\mathbf{X}_n})'(x_k)$, provided that $\{u_j\}_{j=0}^{n+k-1}$ (hence the positions $\{f_{u(2n-1);\mathbf{X}_n}(x_j)\}_{j=0}^{n-1}$ and derivatives $\{(f_{u(2n-1);\mathbf{X}_n})'(x_j)\}_{j=0}^{k-1}$) are fixed.

This correspondence makes transparent estimates of the measure of parameters associated with a particular (pseudo-)trajectory having a given property of periodicity and hyperbolicity (see (3.27-3.33)). In the multidimensional case ($N > 1$), such a correspondence between dynamical properties of trajectories and coefficients of Newton Interpolation Polynomials becomes much less transparent.

In Section 2.2, we define the space of Divided Differences $\mathcal{DD}^{1,n}(I, \mathbb{R}) = I^n \times \mathbb{R}^{2n}$ in the 1–dimensional case, where $I = [-1, 1]$. In this case, we estimated the measure of "bad" parameters in Sections 3.3–3.4.

Recall now the notation of Section 1.3. In N dimensions, we define the space of Divided Differences

$$\begin{aligned}\mathcal{DD}^{N,n}(B^N, \mathbb{R}^N) =& \Big\{ \big(x_0, \ldots, x_{n-1}; \{\vec{u}_\alpha\}_{|\alpha|=0}, \ldots, \{\vec{u}_\alpha\}_{|\alpha|=2n-1}\big) \in \\ & \underbrace{B^N \times \cdots \times B^N}_{n \text{ times}} \times \mathbb{R}^{\nu(0,N)} \times \cdots \times \mathbb{R}^{\nu(2n-1,N)} \Big\} = \\ & \underbrace{B^N \times \cdots \times B^N}_{n \text{ times}} \times W_{0,N}^{u,\mathbf{X}_0} \times W_{1,N}^{u,\mathbf{X}_1} \cdots \times W_{n-1,N}^{u,\mathbf{X}_n} \times W_{n,N}^{u,\mathbf{X}_n} \times \cdots \times W_{2n-1,N}^{u,\mathbf{X}_n},\end{aligned} \tag{4.2}$$

where B^N is the N-dimensional unit ball, $\nu(k, N)$ is N times the number of N-dimensional multiindices α with $|\alpha| = k$, $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$, and $W_{k,N}^{u,\mathbf{X}_{k (\mathrm{mod}\ n)}}$ is the space of homogeneous polynomials of degree k from $\mathbb{R}^N$ to $\mathbb{R}^N$ with the Newton basis defined below. There are two issues we face that were not a concern for the 1-dimensional Newton basis (2.26).

Nonuniqueness. It turns out that the choice of a basis in the space of Divided Differences $\mathcal{DD}^{N,n}(B^N, \mathbb{R}^N)$ and the definition of the Newton map

$$\mathcal{L}_{\mathbf{X}_n}^N : \big(\{\vec{\varepsilon}_\alpha\}_{|\alpha|\le 2n-1}\big) \to \big(\{\vec{u}_\alpha\}_{|\alpha|\le 2n-1}\big) \tag{4.3}$$

(defined by (2.24) for $N = 1$) for a multiindex $\alpha \in \mathbb{Z}_+^N$ is far from unique. In the 1-dimensional case, the standard basis is $\{x^k\}_{k=0}^{2n-1}$ and the Newton basis is $\Big\{\prod_{j=0}^{k-1} |x - x_j|\Big\}_{k=0}^{2n-1}$. In the N–dimensional case, $(x - x_j) \in \mathbb{R}^N$ is an N–dimensional vector. For a fixed basis in $\mathbb{R}^N$, let $(x - x_j)_s$ denote the s–th coordinate of the vector $(x - x_j)$. The number of different monomials of the form

$$\left\{ \prod_{j=0}^{k-1} (x_k - x_j)_{i(j)} \right\}_{\{i(0),\dots,i(k-1)\}\in\{1,\dots,N\}^k} \tag{4.4}$$

is N^k, while the number of homogeneous monomials in N variables of degree k, i.e. $\{x^\alpha\}_{|\alpha|=k}$, is bounded above by k^N, which is much smaller than N^k for $k \gg N$.

Therefore, among the monomials (4.4) we need to choose an appropriate basis and define an appropriate Newton map $\mathcal{L}^N_{\mathbf{X}_n}$. The *standard* way to choose a Newton basis [GY] is as follows. For $\alpha \in \mathbb{Z}^N_+$, let the Newton basis monomial for the multiindex α be

$$\begin{aligned}\big(x; x_0, \dots, x_{(|\alpha|-1)\ (mod\ n)}\big)^\alpha = \prod_{i_1=0}^{\alpha_1-1} (x - x_{i_1})_1 \times \\ \prod_{i_2=0}^{\alpha_2-1} (x - x_{\alpha^1+i_2})_2 \times \dots \times \prod_{i_N=0}^{\alpha_N-1} (x - x_{\alpha^{N-1}+i_N})_N,\end{aligned} \tag{4.5}$$

where $\alpha^j = \sum_{i=0}^{j} \alpha_i$ for $j = 1, \dots, N-1$. The Newton basis for $W^{u,\mathbf{X}_n}_{k,N}$, the space of homogeneous vector-polynomials of degree k, consists of N such monomials (one for each basis vector of $\mathbb{R}^N$) for each α with $|\alpha| = k$.

Dynamically essential coordinates. After a Newton basis is chosen, one needs to make sure that it is effective for dynamical purposes. In Section 3.3, we noticed in (3.30) and (3.33) that in order to perturb by Newton Interpolation Polynomials in an effective way, we need to make sure that the product of distances $\prod_{j=0}^{n-2} |x_{n-1} - x_j|$ is not too small. Similarly, in the multidimensional case we need at least one Newton monomial $(x; x_0, \dots, x_{n-2})^\alpha$ with $|\alpha| = n-1$ not to be too small. The most natural way to choose a "good" monomial is by taking the maximal coordinates of corresponding vectors. Let $v \in \mathbb{R}^N$ be a nonzero vector and $\|v\| = (\sum_{i=1}^N v_i^2)^{1/2}$. Denote by

$$m(v) = \min\{i :\ 1 \le i \le N,\ |v_i| = \max_{j=1}^{N} |v_j|\}$$

the minimal index of one of the largest components v_i of v. Then

$$\left| \prod_{j=0}^{n-2} (x_{n-1} - x_j)_{m(x_{n-1}-x_j)} \right| \ge N^{\frac{n-1}{2}} \prod_{j=0}^{n-2} |x_{n-1} - x_j|. \tag{4.6}$$

This is a satisfactory estimate, because $\gamma_n(C,\delta)$ is a stretched exponential in n, so we can neglect factors that are exponential in n. For given $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1}$, we call the monomials

$$\prod_{j=0}^{k-1}(x - x_j)_{m(x_k - x_j)}, \qquad k = 0, \ldots, n-1 \tag{4.7}$$

dynamically essential. Using these monomials, we shall imitate estimates (3.27)–(3.34) in the multidimensional Discretization Method.

The dynamically essential monomials do not necessarily belong to the standard Newtion basis (4.5), so we will need to define the basis differently depending on which monomials are dynamically essential for the given pseudo-trajectory $\mathbf{X}_n$. This is not a major obstacle, since we already use a different basis for each $\mathbf{X}_n$ in the Discretization Method, but it does further complicate the argument.

The necessity of altering the basis is illustrated by the following example for $N = 2$: $x_0 = (1,0)$, $x_1 = (0,1)$, $x_2 = (1,1)$. Then for all α with $|\alpha| = 2$, we have $(x_2; x_0, x_1)^\alpha = 0$. Thus, the monomial $(x; x_0, x_1)^\alpha$ is useless to perturb the image of x_2.

4.3 The multidimensional Distortion Lemma

One of the key ingredients of the Discretization Method is the Distortion Lemma from Section 3.4. Fix an n-tuple of points $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \subset B^N$. The Distortion Lemma in dimension 1 shows that the Newton map $\mathcal{L}^1_{\mathbf{X}_n} : W_{<2n,1} \to W^{u,\mathbf{X}_n}_{<2n,1}$ expands a brick $HB^{st}_{<2n}(\tau)$ of standard thickness along each direction by at most a factor of 3 in each direction. Since, the space of 1-dimensional divided differences $W_{<2n,1}$ is $2n$-dimensional, this gives a total volume distortion of at most 3^{2n}. This factor is part of the estimates (3.13) and (3.42), but is ultimately unimportant since it is exponential in n.

In the multidimensional case, this naive approach does not work. One can, indeed, show that the Newton map expands a brick of standard thickness along each direction by at most a factor 3^N. However, the space of divided differences $W_{<2n,N}$ has dimension of the order of $(2n)^N$. So, the naive estimate of distortion is $3^{N(2n)^N}$, which is highly unaffordable. In the multidimensional case, we need a more precise estimate on distortion.

First, we define a Cubic Brick of at most standard thickness, which differs from the brick $HB^{st}_{<2n}(\vec{\mathbf{r}})$ defined by (2.27) in two ways. First, it is a parallelipiped for all

N, whereas $HB^{st}_{<2n}(\vec{\mathbf{r}})$ is a product of balls whose dimension is greater than 1 when $N > 1$. Second, we require that the side lengths of a Cubic Brick decay very rapidly. We will estimate distortion of such Cubic Bricks by the Newton map, then in the next section we will cover a brick with Cubic Bricks,

Definition 4.3.1. *Let $\vec{\lambda}_{<k} = (\lambda_0, \dots, \lambda_{k-1}) \in \mathbb{R}^k_+$ be a vector with strictly positive components. If for every $0 \le m < k$ we have $\lambda_m \ge \lambda_{m+1}(m+1)^{4N}$, we call the rectangular parallelepiped*

$$\begin{aligned} CHB^{N,st}_{<k}(\vec{\lambda}_{<k}, \vec{\delta}) = \big\{ \{\vec{\varepsilon}_\alpha\}_{|\alpha|<k} :\ \forall\ 0 \le m < k, \\ \forall\ \alpha \in \mathbb{Z}^N_+,\ |\alpha| = m,\ \forall\ j = 1, \dots, N,\ |\varepsilon^j_\alpha - \delta^j_\alpha| \le \lambda_m \big\} \end{aligned} \tag{4.8}$$

a Cubic Brick of at most standard thickness centered at $\vec{\delta}$ (here the superscript j denotes the jth coordinate). Similarly, for $\lambda_k \in \mathbb{R}_+$ let

$$\begin{aligned} CHB^{N,st}_{k}(\vec{\lambda}_{k}, \vec{\delta}) = \big\{ \{\vec{\varepsilon}_\alpha\}_{|\alpha|=k} :\ \forall\ \alpha \in \mathbb{Z}^N_+, \\ \forall\ |\alpha| = k,\ \forall\ j = 1, \dots, N,\ |\varepsilon^j_\alpha - \delta^j_\alpha| \le \lambda_k \big\} \end{aligned} \tag{4.9}$$

An example of a suitable thickness vector $\vec{\lambda}_{<k}$ is $\{\lambda_m = \frac{\tau}{(m!)^{4N}}\}_{m=0}^{k-1}$.

To formulate a multidimensional version of the Distortion Lemma, we extend the definition of the parameters (3.38) and (3.39) allowed by the family (3.8) in Section 3.4. Consider a $C^{1+\rho}$-smooth diffeomorphism $f \in \mathrm{Diff}^{1+\rho}(B^N)$ of the unit ball B^N into its interior and some positive integer n. Let $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})$ be a Cubic Brick of at most standard thickness, which defines the family of diffeomorphisms of the unit ball B^N into its interior

$$\left\{ f_{\vec{\varepsilon}}(x) = f(x) + \sum_{|\alpha| \le 2n-1} \vec{\varepsilon}_\alpha x^\alpha \right\}_{\{\vec{\varepsilon}_\alpha\}_{|\alpha|\le 2n-1} \in CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})} \tag{4.10}$$

Fix an n-tuple of points $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \subset B^N$, and fix the standard Newton basis $\big\{(x; x_0, \dots, x_{(|\alpha|-1)\ (mod\ n)})^\alpha\big\}_{|\alpha| \le 2n-1}$ in the multidimensional space of divided differences $\mathcal{DD}^{N,n}(B^N, \mathbb{R}^N)$, defined by (4.5). Denote by

$$\mathcal{L}^N_{\mathbf{X}_n} : \big(\{\vec{\varepsilon}_\alpha\}_{|\alpha| \le 2n-1}\big) \to \big(\{\vec{u}_\alpha\}_{|\alpha| \le 2n-1}\big) \tag{4.11}$$

the Newton map that corresponds to rewriting an N-vector polynomial of degree $2n-1$ given in the basis $\{x^\alpha\}_{|\alpha|\le 2n-1}$ in the Newton basis (4.5).

Consider the reparametrization of family (4.10) by the Newton parameters

$$f_{\vec{u}(2n-1),\mathbf{X}_n}(x) = f(x) + \sum_{|\alpha|\le 2n-1} \vec{u}_\alpha(x; x_0, \dots, x_{(|\alpha|-1)\ (mod\ n)})^\alpha. \tag{4.12}$$

Denote by $\vec{u}(m) = (\{\vec{u}_\alpha\}_{|\alpha|\le m})$ (respectively $\vec{u}_m = (\{\vec{u}_\alpha\}_{|\alpha|=m})$) the set of all Newton parameters of degree at most m (respectively of degree exactly m).

In order to state the Discretization Lemma in the N-dimensional case, we have to define the set of values of Newton parameters $\vec{u}(2n-1)$ that are allowed by the family (4.10) and the Cubic Brick (4.8).

Let $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ for $0 \le k \le n$, and for convenience below let $\mathbf{X}_k = \mathbf{X}_n$ for $n < k \le 2n-1$. Let

$$\pi^{N,u,\mathbf{X}_k}_{<2n,\le k} : W^u_{<2n,N} \to W^{u,\mathbf{X}_k}_{\le k,N} \quad \text{and} \quad \pi^{N,u,\mathbf{X}_k}_{<2n,k} : W^u_{<2n,N} \to W^{u,\mathbf{X}_k}_{k,N}$$

be the natural projection onto the space $W^{u,\mathbf{X}_k}_{\le k,N}$ of N-vector polynomials in N variables of degree k and the space $W^{u,\mathbf{X}_k}_{k,N}$ of homogeneous N–vector polynomials in N variables of degree k. As for $N=1$, this projection depends only on $\mathbf{X}_k$, not on $\mathbf{X}_n \setminus \mathbf{X}_k$.

Notice that the image of the Cubic Brick $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})$ under the Newton map $\mathcal{L}^N_{\mathbf{X}_n}$ is a parallelepiped

$$\mathcal{P}^{N,st}_{<2n,\mathbf{X}_n}(\vec{\lambda}_{<2n}, \vec{\delta}) = \mathcal{L}^N_{\mathbf{X}_n}(CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})), \tag{4.13}$$

because the map $\mathcal{L}^N_{\mathbf{X}_n}$ is linear.

We call the parallelepiped $\mathcal{P}^{N,st}_{<2n,\mathbf{X}_n}(\vec{\lambda}_{<2n}, \vec{\delta})$ the set of parameters *allowed by the family* (4.10). Notice that the values of $\vec{u}(2n-1) = (\vec{u}_0, \dots, \vec{u}_{2n-1}) \in W^{u,\mathbf{X}_n}_{<2n,N}$ that do not belong to $\mathcal{P}^{N,st}_{<2n,\mathbf{X}_n}(\vec{\lambda}_{<2n}, \vec{\delta})$ are of *no interest* for us, because they lie outside of the range of the family (4.10) under consideration. Now we define the range of allowed parameters $\vec{u}(k) \in W^{u,\mathbf{X}_k}_{\le k,N}$ for each $0 \le k \le 2n-1$.

Denote the images of the Cubic Brick $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})$ under composition of the Newton map $\mathcal{L}^N_{\mathbf{X}_n}$ and the projections $\pi^{N,u,\mathbf{X}_k}_{<2n,\le k}$ and $\pi^{N,u,\mathbf{X}_k}_{<2n,k}$ by

$$\begin{aligned} \mathcal{P}^{N,st}_{<2n,\le k,\mathbf{X}_k}(\vec{\lambda}_{<2n}, \vec{\delta}) &= \pi^{N,u,\mathbf{X}_k}_{<2n,\le k} \circ \mathcal{L}^N_{\mathbf{X}_n}(CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})) \subset W^u_{\le k,N} \\ \mathcal{P}^{N,st}_{<2n,k,\mathbf{X}_k}(\vec{\lambda}_{<2n}, \vec{\delta}) &= \pi^{N,u,\mathbf{X}_k}_{<2n,k} \circ \mathcal{L}^N_{\mathbf{X}_n}(CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}, \vec{\delta})) \ \subset W^u_{k,N}. \end{aligned} \tag{4.14}$$

It follows from the definition of the Newton map $\mathcal{L}^N_{\mathbf{X}_n}$ that for each $0 \le k \le 2n-1$, the sets $\mathcal{P}^{N,st}_{<2n,\le k,\mathbf{X}_k}(\vec{\lambda}_{<2n},\vec{\delta})$ and $\mathcal{P}^{N,st}_{<2n,k,\mathbf{X}_k}(\vec{\lambda}_{<2n},\vec{\delta})$ are polyhedrons depending on $\mathbf{X}_k$. We call $\mathcal{P}^{N,st}_{<2n,\le k,\mathbf{X}_k}(\vec{\lambda}_{<2n},\vec{\delta})$ the set of parameters *of order k allowed by the family* (4.10).

Similarly to the 1-dimensional Distortion Lemma from Section 3.4, its N-dimensional generalization gives an estimate on the ratio of volumes of the polyhedra $\mathcal{P}^{N,st}_{<2n,\le k,\mathbf{X}_k}(\vec{\lambda}_{<2n},\vec{\delta})$ of allowed Newton parameters $\vec{u}(k)$ and the corresponding Cubic Bricks $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n},\vec{\delta})$.

The Preliminary N-dimensional Distortion Lemma. *Let $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \subset B^N$ be an n–tuple of points in the unit ball B^N, and let $\mathcal{L}^N_{\mathbf{X}_n} : W_{<2n,N} \to W^{u,\mathbf{X}_n}_{<2n,N}$ be the Newton map (4.11) corresponding to the basis (4.5) for $W^{u,\mathbf{X}_n}_{<2n,N}$. Let $\vec{\lambda}_{<2n} = (\lambda_0,\ldots,\lambda_{2n-1}) \in \mathbb{R}^{2n}_+$ be a vector with strictly positive components that defines a Cubic Brick of $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n},\vec{\delta})$ of at most standard thickness and $Leb_{\le k,N}$ (resp. $Leb_{k,N}$) be the Lebesgue measure on the space $W_{\le k,N}$ (respectively $W_{k,N}$) of (respectively of homogeneous) N-vector polynomials in N variable of degree k. Then for all $0 \le k \le 2n-1$, we have the volume ratio estimate*

$$\frac{Leb_{k,N}\left(\mathcal{P}^{N,st}_{<2n,k,\mathbf{X}_k}(\vec{\lambda}_{<2n},\vec{\delta})\right)}{Leb_{k,N}\left(CHB^{N,st}_{k}(\vec{\lambda}_{k},\vec{\delta})\right)} \le 3^{N^2 n}. \tag{4.15}$$

Remark 4.3.2. *As in the* 1*-dimensional case, the Newton map $\mathcal{L}^N_{\mathbf{X}_n}$ is volume-preserving, and hence the parallelepipeds $\mathcal{P}^{N,st}_{<2n,\mathbf{X}_n}(\vec{\lambda}_{<2n},\vec{\delta})$ and $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n},\vec{\delta})$ have the same volume. However, the estimate (4.15) concerns the volumes of projections of these parallelepipeds. Since $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n},\vec{\delta})$ is a rectangular parallelepiped (aligned with the coordinate axes) and $\mathcal{P}^{N,st}_{<2n,\mathbf{X}_n}(\vec{\lambda}_{<2n},\vec{\delta})$ is not, the projections of the latter set have larger volume according to the amount of shear in the Newton map. The Distortion Lemma bounds the effect of the shear.*

The lemma above will be proved in Part II of this paper in a more general form required for the proof of the Main Result (Theorem 1.3.7). The other main ingredient of Section 3.4, the Collection Lemma, proceeds in much the same way for general N as it does for $N = 1$.

4.4 From a brick of at most standard thickness to an admissible brick

In order to apply the method which we are developing in this paper to other problems about generic properties of dynamical systems, we need to have a sufficiently rich space of parameters (enough freedom to perturb). Various dynamical phenomena have a "size" that is exponential in the period or number of iterations. Since we perturb trajectories of length n with polynomials whose degree is proportional to n, it seems essential to have a Hilbert Brick of parameters with sides decaying at most exponentially in the period of the polynomials the respective parameters multiply. That is, with the notations of Definition 1.3.1, we would like to have $r_n \geq \exp(-Cn)$ for some $C > 0$. However, even for the 1-dimensional model, if we consider a brick of parameters with exponentially decaying sides, then we cannot control the distortion properties of the Newton map (see Remark 3.4.1).

In order to circumvent this problem, we do the following. Consider a Hilbert Brick $HB^N(\vec{\mathbf{r}})$ of an admissible size $\vec{\mathbf{r}} = \{r_k\}_{k=0}^{\infty}$ (see Definition 1.3.1). Using Fubini's Reduction from Section 2.3 at the nth stage of induction over the period, we reduce consideration of an infinite-dimensional Hilbert Brick $HB^N(\vec{\mathbf{r}})$ to a finite-dimensional brick $HB^N_{<2n}(\vec{\mathbf{r}})$.

Recall that $HB^N_{<2n}(\vec{\mathbf{r}})$ belongs to the space of N-vector polynomials of degree $2n-1$, denoted by $W_{<2n,N}$.

Definition 4.4.1. *We call the Cubic Brick in $W_{<2n,N}$ of at most standard thickness centered at a point $\vec{\delta}_{<2n} \in W_{<2n,N}$ and associated to the brick $HB^N_{<2n}(\vec{\mathbf{r}})$ the one denoted by $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}), \vec{\delta}_{<2n})$ and defined by (4.8) where $\lambda_m(\vec{\mathbf{r}}) = r_m/(m!)^{4N}$ and $\vec{\lambda}_{<2n}(\vec{\mathbf{r}}) = (\lambda_0(\vec{\mathbf{r}}), \ldots, \lambda_{2n-1}(\vec{\mathbf{r}}))$.*

Remark 4.4.2. *It follows from the nonincreasing property of the sequence r_m (see Definition 1.3.1) that*

$$(m+1)^{4N}\lambda_{m+1}(\vec{\mathbf{r}}) = (m+1)^{4N}\frac{r_{m+1}}{((m+1)!)^{4N}} \leq \frac{r_m}{(m!)^{4N}} = \lambda_m(\vec{\mathbf{r}}), \tag{4.16}$$

so that the Cubic Brick $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}), \vec{\delta}_{<2n})$ is of at most standard thickness.

We wish to cover the brick $HB^N_{<2n}(\vec{\mathbf{r}})$ in $W_{<2n,N}$ by a collection of Cubic Bricks $\{CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}), \vec{\delta}_{<2n})\}_{\vec{\delta}_{<2n} \in G(\vec{\mathbf{r}})}$ of at most standard thickness associated with it. First we define a grid in $W_{<2n,N}$ on which the centers $\vec{\delta}_{<2n}$ should lie in order that the Cubic Bricks fit together without overlap.

Recall that $W_{<2n,N} = \times_{m=0}^{2n-1} W_{m,N}$, $\dim W_{m,N} = \nu(m,N)$, and denote $\eta(k,N) = \sum_{m=0}^{k} \nu(m,N)$. Let

$$\begin{aligned} \mathbb{Z}_{2\vec{\lambda}_{<2n}(\mathbf{\vec{r}})}^{\eta(2n-1,N)} &= \mathbb{Z}_{2\lambda_0(\mathbf{\vec{r}})}^{\nu(0,N)} \times \mathbb{Z}_{2\lambda_1(\mathbf{\vec{r}})}^{\nu(1,N)} \times \cdots \times \mathbb{Z}_{2\lambda_{2n-1}(\mathbf{\vec{r}})}^{\nu(2n-1,N)} \\ &\subset W_{0,N} \times W_{1,N} \times \cdots \times W_{2n-1,N} = W_{<2n,N}, \end{aligned} \tag{4.17}$$

where $\mathbb{Z}_{2\lambda_m(\mathbf{\vec{r}})}^{\nu(m,N)}$ is the grid in $\mathbb{R}^{\nu(m,N)}$ with spacing $2\lambda_m(\mathbf{\vec{r}})$ in each coordinate. Let

$$G(\mathbf{\vec{r}}) = \left\{ \vec{\delta}_{<2n} \in \mathbb{Z}_{2\vec{\lambda}_{<2n}(\mathbf{\vec{r}})}^{\eta(2n-1,N)} : HB_{<2n}^{N}(\mathbf{\vec{r}}) \cap CHB_{<2n}^{N,st}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}), \vec{\delta}_{<2n}) \neq \emptyset \right\}. \tag{4.18}$$

The Cubic Bricks with centers in $G(\mathbf{\vec{r}})$ are the ones needed for our covering:

$$HB_{<2n}^{N}(\mathbf{\vec{r}}) \subset \cup_{\vec{\delta} \in G(\mathbf{\vec{r}})} CHB_{<2n}^{N,st}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}), \vec{\delta}_{<2n}). \tag{4.19}$$

The Covering Lemma. *The ratio of the volume of the covering (4.19) by Cubic Bricks and the volume of the brick $HB_{<2n}^{N}(\mathbf{\vec{r}})$ is bounded by e^{6N}.*

Proof: We claim that the covering (4.19) is contained in the slightly larger brick $HB_{<2n}^{N}(\{(1 + \frac{2\sqrt{N}}{(m!)^{3N}}) r_m\}_{m=0}^{2n-1})$. The amount that the covering extends beyond $HB_{<2n}^{N}(\mathbf{\vec{r}})$ in the direction of a given multiindex α with $|\alpha| = m$ is at most the diameter in the α direction of one of the Cubic Bricks $CHB_{<2n}^{N,st}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}), \vec{\delta}_{<2n}^{\,j})$, which by Definition 4.4.1 is $2\sqrt{N}\lambda_m(\mathbf{\vec{r}}) = 2\sqrt{N} r_m/(m!)^{4N}$. Recall the definitions (2.27) of $HB_{<2n}^{N}(\mathbf{\vec{r}})$ and (1.12) of the norm used therein. In this norm, the diameter of $CHB_{<2n}^{N,st}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}), \vec{\delta}_{<2n}^{\,j})$ in the directions with $|\alpha| = m$ is $2\sqrt{N} r_m/(m!)^{4N}$ times the square root of $\sum_{|\alpha|=m} \binom{m}{\alpha}^{-1} < (m!)^N$, from which our claim follows. The ratio of volumes that we wish to bound is then at most $\prod_{m=0}^{2n-1} \exp(2\sqrt{N}/(m!)^{3N}) \leq e^{2e\sqrt{N}} < e^{6N}$. Q.E.D.

Recall now that our main goal in the proof of Theorem 1.3.7 is to get an estimate on the measure of the "bad" set of parameters (2.1) inside of the Hilbert Brick $HB^N(\mathbf{\vec{r}})$. Using the Fubini reduction from Section 2.3, we know that it is sufficient to get an estimate on the measure of the "bad" set in a finite-dimensional slice of the form $HB_{<2n}^{N}(\mathbf{\vec{r}}) \times \{\vec{\varepsilon}_{\geq 2n}\}$ that is uniform over $\vec{\varepsilon}_{\geq 2n} \in HB_{\geq 2n}^{N}(\mathbf{\vec{r}})$ see (2.31-2.32). Notice now that if we can prove that for each Cubic Brick slice $CHB_{<2n}^{N,st}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}), \vec{\delta}_{<2n}) \times \{\vec{\varepsilon}_{\geq 2n}\}$ from the covering collection $\vec{\delta}_{<2n} \in G(\mathbf{\vec{r}})$ that the fraction of parameters in the slice that are "bad" is uniformly small over $\vec{\delta}_{<2n} \in G(\vec{\mathbf{r}})$, then the same fraction is small in the whole slice $HB^N_{<2n}(\vec{\mathbf{r}}) \times \{\vec{\varepsilon}_{\geq 2n}\}$. (By the Covering Lemma, we must increase the bound on the fraction of "bad" parameters only by the factor e^{6N}, which is independent of n.) By the Fubini reduction, this shows that the measure of the "bad" set in $HB^N(\vec{\mathbf{r}})$ is small too. Thus it is sufficient to prove the following estimate.

$$\frac{\mu^N_{<2n,\vec{\mathbf{r}}}\left(B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{\geq 2n}) \cap CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}),\vec{\delta}_{<2n})\right)}{\mu^N_{<2n,\vec{\mathbf{r}}}\left(CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}),\vec{\delta}_{<2n})\right)} \leq \mu'_n(C,\delta,\rho,M_{1+\rho}). \tag{4.20}$$

uniformly over all $\delta_{<2n} \in J(\vec{\mathbf{r}})$ and so that $\sum_{n=1}^{\infty} \mu'_n(C,\delta,\rho,M_{1+\rho})$ converges for all positive C, δ, ρ, and $M_{1+\rho}$, and tends to zero as C tends to infinity.

4.5 The main estimate on the measure of "bad" parameters

In this subsection we formulate the main theorem of the rest of the paper which implies (the Main) Theorem 1.3.7. It will be proved in Part II of this paper.

Theorem 4.5.1. *For any $\rho > 0$ and any diffeomorphism $f \in \mathrm{Diff}^{1+\rho}(B^N)$, consider a Hilbert Brick $HB^N(\vec{\mathbf{r}})$ of an admissible size $\vec{\mathbf{r}}$ with respect to f and the family of analytic perturbations of f*

$$\{f_{\vec{\varepsilon}}(x) = f(x) + \phi_{\vec{\varepsilon}}(x)\}_{\vec{\varepsilon} \in HB^N(\vec{\mathbf{r}})} \tag{4.21}$$

(see (1.10)) with the Lebesgue product probability measure $\mu^N_{\vec{\mathbf{r}}}$ (see (1.15)) associated to $HB^N(\vec{\mathbf{r}})$.

Then for any positive integer n and any $\vec{\varepsilon}_{\geq 2n} \in HB^N_{\geq 2n}(\vec{\mathbf{r}})$, consider a slice (2.27) of the Hilbert Brick $HB^N(\vec{\mathbf{r}})$ of the form

$$HB^N_{<2n}(\vec{\mathbf{r}}) \times \{\vec{\varepsilon}_{\geq 2n}\} \subset HB^N_{<2n}(\vec{\mathbf{r}}) \times HB^N_{\geq 2n}(\vec{\mathbf{r}}) = HB^N(\vec{\mathbf{r}}). \tag{4.22}$$

Inside of this slice fix a grid point $\vec{\delta}_{<2n} \in G(\vec{\mathbf{r}})$, and consider the Cubic Brick

$$CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\vec{\mathbf{r}}),\vec{\delta}_{<2n}) \times \{\vec{\varepsilon}_{\geq 2n}\} \subset HB^N_{<2n}(\vec{\mathbf{r}}) \times \{\vec{\varepsilon}_{\geq 2n}\}. \tag{4.23}$$

from the covering (4.19) associated to the brick $HB^N_{<2n}(\vec{\mathbf{r}})$.

Let $\tilde{f} = f_{(\vec{\delta}_{<2n},\vec{\varepsilon}_{\geq 2n})}$ be a diffeomorphism corresponding to the center of the Cubic Brick $CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}),\vec{\delta}_{<2n}) \times \{\vec{\varepsilon}_{\geq 2n}\}$. Consider the family of diffeomorphisms

$$\begin{aligned} &\{\tilde{f}_{\vec{\varepsilon}_{<2n}}\}_{\vec{\varepsilon}_{<2n}\in CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}),0)} \supset \\ &\{f_{(\vec{\varepsilon}_{<2n}+\vec{\delta}_{<2n},\vec{\varepsilon}_{\geq 2n})}\}_{(\vec{\varepsilon}_{<2n}+\vec{\delta}_{<2n}\in HB^{N}_{<2n}(\mathbf{\vec{r}})\cap CHB^{N,st}_{<2n}(\vec{\lambda}_{<2n}(\mathbf{\vec{r}}),\vec{\delta}_{<2n})}. \end{aligned} \tag{4.24}$$

Then for $C \geq 30/\rho \log M_{1+\rho}$, the fraction of the measure of "bad" parameters $B_n(C,\delta,\rho,\mathbf{\vec{r}},f,\vec{\varepsilon}_{\geq 2n})$, defined in (2.32), inside $CHB^{N,st}_{<2n}(\vec{\lambda}_{2n-1}(\mathbf{\vec{r}}),\vec{\delta}_{<2n}) \times \{\vec{\varepsilon}_{\geq 2n}\}$ satisfies the bound

$$\begin{aligned} &\frac{\mu^{N}_{<2n,\mathbf{\vec{r}}}\Big(B_n(C,\delta,\rho,\mathbf{\vec{r}},f,\vec{\varepsilon}_{\geq 2n})\cap CHB^{N,st}_{<2n}(\vec{\lambda}_{2n-1}(\mathbf{\vec{r}}),\vec{\delta}_{<2n})\Big)}{\mu^{N}_{<2n,\mathbf{\vec{r}}}\Big(CHB^{N,st}_{<2n}(\vec{\lambda}_{2n-1}(\mathbf{\vec{r}}),\vec{\delta}_{<2n})\Big)} \leq \\ &\qquad 24Ce^{N}n^{1+\delta}3^{2nN^2}M_{1+\rho}\max\{M_{1+\rho},\exp C/100\}^{-n/2} \end{aligned} \tag{4.25}$$

The discussion from the end of the previous section shows that this theorem implies (the Main) Theorem 1.3.7.

Chapter 5

Analysis of Simple and Nonsimple Trajectories

In this Chapter we investigate properties of trajectories $\{x_j = f^j(x_0)\}_j$ of $C^{1+\rho}$–smooth diffeomomorphisms the unit N–dimensional ball B^N into itself $f : B^N \hookrightarrow B^N$ and behavior of the product of distances along such trajectories $\prod_{j=1}^{n-1} |x_j - x_0|$ associated to them.

In Chapter 2 of Part I we described the strategy of the proof of (the main) Theorem 1.3.7. The key point of the proof is to estimate the measure (2.3) of the bad set $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f) \subset HB(\vec{\mathbf{r}})$, defined by (2.1). To estimate the measure of $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)$ we split it into the union of two sets $B_n^{sim}(C, \delta, \rho, \vec{\mathbf{r}}, f)$ and $B_n^{non}(C, \delta, \rho, \vec{\mathbf{r}}, f)$ corresponding to essentially simple and essentially nonsimple almost periodic trajectories respectively. These sets are defined by (2.35) and (2.36) respectively.

A method of estimating the measure of essentially simple almost periodic trajectories $B_n^{sim}(C, \delta, \rho, \vec{\mathbf{r}}, f)$ in the model C^2–smooth 1–dimensional case is described in Sections 3.3-3.4. The general $C^{1+\rho}$–smooth N–dimensional case goes along the same lines and will be discussed in Chapter 9. However, in order to get an estimate on the measure of essentially nonsimple almost periodic trajectories we need to understand properties of those. Loosely speaking, the main result of this Chapter is, as we called it in Section 3.5 *The Shift Theorem*, says that *if a diffeomorphism* $g \in \mathrm{Diff}^{1+\rho}(B^N)$ *satisfies Inductive Hypothesis of some order* $n-1$ *and* $x_0 \in B^N$ *is almost periodic of period* n*, then the following trichotomy is satisfied:*

· *either sufficient hyperbolicity of* x_0 *can be extracted from the Inductive Hypothesis,*

· *or after a possible shift of the initial point* $x_j = g^j(x_0)$ *with* $j < n \log_2 n$ *we have*

that x_j is nonrecurrent (simple),
· or x_j has a very close return x_{j+k} to x_j (has a gap at x_{j+k}) and x_j's trajectory of length k is nonrecurrent (simple). This trichotomy is encoded in Fig. 5.

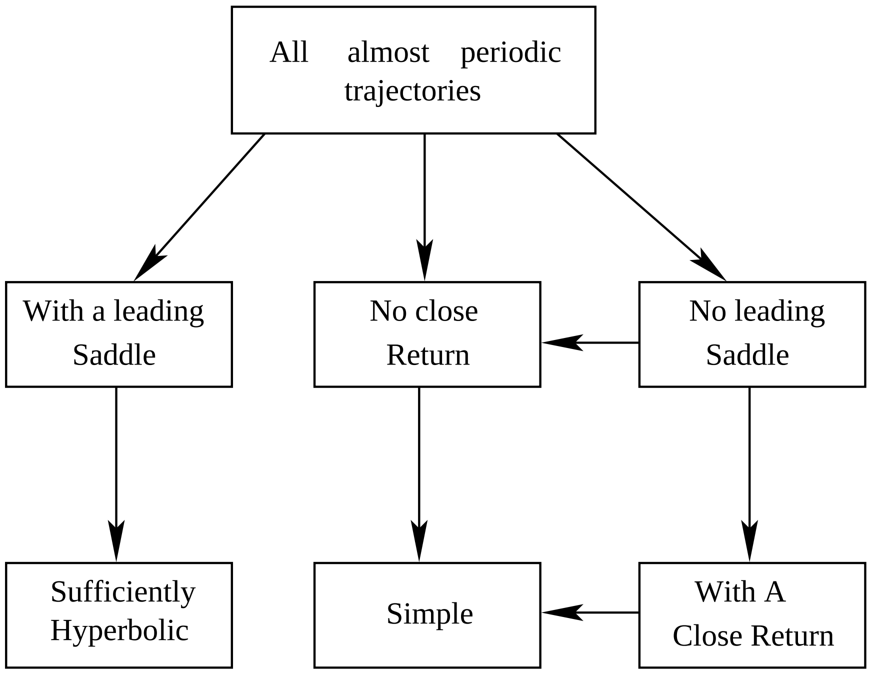

Figure 5.1: Various types of almost periodic periodic trajectories

A precise definition of a very close return, which we call a weak (D, n)–gap (resp. (D, n, r_{2k})–gap), is given by definition 2.4.3 (resp. by definition 3.5.3).

The main result of this Chapter implies the inclusions (2.42) and (3.71), i.e.

$$\begin{aligned} B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}) &\subseteq \cup_{k|n} B_n^{wgap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D), \\ B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f},\gamma_n) &\subseteq \cup_{k|n} B_n^{gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f},\gamma_n;D), \end{aligned} \tag{5.1}$$

where $D \geq M_{1+\rho}^{30/\rho}$. If these inclusions are proven, then we can proceed with estimating the measure of parameters associated with essentially nonsimple almost periodic trajectories $B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f})$ along the strategy described in Sections 3.5-3.6 for the C^2–smooth 1–dimensional case.

In Section 3.5 for the C^2–smooth 1–dimensional case we exhibited one important ingredient of the proof of the Shift Theorem: notion of (k, n, C, δ, ρ)*–leading saddle* given by definition 3.5.4. As we pointed out in Remark 3.5.5 that if a $C^{1+\rho}$–smooth diffeomorphism g satisfies Inductive Hypothesis of some order n with some constants $C > 0$, $\delta > 0$, and $0 < \rho \le 1$, i.e. $g \in IH(n, C, \delta, \rho)$, and has a point $x_0 \in B^N$ (resp. $x \in I$) which has a (k, n, C, δ, ρ)–leading saddle, then there is a periodic point $x_0^* = g^k(x_0^*)$ of period k nearby x_0 such that $|x_0^* - x_0| \le \gamma_k^{3/\rho}(C, \delta)$. By Inductive Hypothesis x_0^* is $(k, \gamma_k(C, \delta))$–hyperbolic and, therefore, there are three possibilities

• if x_0^* is a sink (resp. source)[1] periodic point $\{x_j^* = g^j(x_0^*)\}_{j=0}^{k-1}$ attracts (resp. repells) the trajectory $\{x_j\}_{j=0}^{n-1}$ is nearby and is absorbed (resp. repelled) by $\{x_j^*\}_{j=0}^{k-1}$ and inherits hyperbolicity from $\{x_j^*\}_{j=0}^{k-1}$ (resp. is either not almost periodic or also inherits hyperbolicity from the trajectory $\{x_j^*\}_{j=0}^{k-1}$) (see Lemma **??**). The only interesting case is

• x_0^* is a saddle periodic point and the trajectory $\{x_j\}_{j=0}^{n-1}$ can be recurrent (a homoclinic intersection of a stable manifold $W^s(x_0^*)$ and an unstable manifold $W^u(x_0^*)$ of the periodic point x_0^*). This is a motivation for the name "leading saddle", because it leads the trajectory $\{x_j\}_{j=0}^{n-1}$ for some time[2].

In order to clarify presentation we first prove Theorem 2.4.6, which is a simplified version of the Shift Theorem and then prove the Shift Theorem itself.

5.1 Consequtive close returns and the spectrum of the linearization

Definition 5.1.1. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$–smooth diffeomorphism for some $0 < \rho \le 1$, $T > 1$, and $\{x_j = f^j(x_0)\}_{j=0}^{\infty}$ be a trajectory of f. Then the trajectory $\{x_j = f^j(x_0)\}_{j=0}^{\infty}$ has the first T–ladder moment at moment l_1 if*

$$l_1 = \min\{i \ge 1 : \ |x_i - x_0| \le T^{-1}\}. \tag{5.2}$$

Assume that the $(t-1)$–st T–ladder moment of $\{x_j = f^j(x_0)\}_{j=0}^{\infty}$ is at $x_{l_{t-1}}$, then the t–th T–ladder moment at l_t if

$$l_t = \min\{i \ge l_{t-1} : \ |x_{l_t} - x_0| \le T^{-1}|x_{l_{t-1}} - x_0|\}. \tag{5.3}$$

Remark 5.1.2. *This definition is useful for analysis of almost periodic trajectories $\{x_j = f^j(x_0)\}_{j=0}^{n}$ of length n with T being greater or equal to some multiple of the*

---

[1] a periodic point is a sink (resp. a source) if all eigenvalues of the linearization $df^k(x_0^*)$ are inside (resp. outside) of the unit disk $\{|z| < 1\} \subset \mathbb{C}$

[2] As a matter of fact we are unable to prove this idealistic picture and use more involved arguments (see Sect. ??)

n–th power of $C^{1+\rho}$–norm of f and f^{-1}. The name comes from an analogous notion from probability theory: ladder moments for 1–dimensional random walks.

It turns out that *any* trajectory $\{x_j = f^j(x_0)\}_{j=0}^n$ of length n of any $C^{1+\rho}$ diffemorphism can have at most a *finite* number of $M_{1+\rho}^{2n}$–ladder moments. This is interesting and important fact about behavior of recurrent trajectories. It is also useful to introduce the following

Definition 5.1.3. *A collection of increasing positive integer $\{l_t\}_{t=1}^s$ is called multiplicatively ordered if for each $t = 1, \dots, s-1$ ratio $l_t/l_{t-1} = p_{t-1}$ is integer, i.e. l_t is divisible by l_{t-1}.*

Theorem 5.1.4. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$–smooth diffeomorphism for some $0 < \rho \le 1$, $M_1 = \max\{\|f\|_{C^1}, \|f^{-1}\|_{C^1}\}$, $M_{1+\rho} = \max\{\|f\|_{C^{1+\rho}}, M_1, 2^{1/\rho}\}$, and $T = M_{1+\rho}^{8N/\rho}$. Suppose that x_0 is (n, T^{-Nn})–periodic and $x_{l_1}, x_{l_2}, \dots, x_{l_s}$ are the first, the second, …, and the s–th T^n–ladder moments respectively before n, i.e. $l_s \le n$. Then the number of T^n–ladder moments s is always bounded by dimension N of the ball B^N.*

Corollary 5.1.5. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$–smooth diffeomorphism for some $0 < \rho \le 1$ with $M_{1+\rho}$ bound on the $C^{1+\rho}$-norm of f, defined as in Theorem 5.1.4, $T = M_{1+\rho}^{8N/\rho}$, and $D \ge T^3$. Then for any point $x_0 \in B^N$ if its trajectory $\{x_j\}_{j=0}^{k-1}$ of length k has no weak (D,n)–gaps, then the closest return satisfies the following lower bound*

$$\min_{1 \le j < k} |x_j - x_0| \ge T^{-Nn} D^{-Nn}. \tag{5.4}$$

Proof of the Corollary: Suppose that the trajectory of x_0 before x_k has T^n–ladder moments at $x_{l_1}, \dots, x_{l_s}$. By Theorem 5.1.4 we have $s \le N$. To estimate distance of the first ladder point x_{l_1} to x_0 notice that by definition of T^n–ladder moments we have

$$\min_{1 \le j < l_1} |x_j - x_0| \ge T^{-n} \tag{5.5}$$

and absence of weak (D,n)–gaps implies that

$$|x_{l_1} - x_0| \ge T^{-n} D^{-n}. \tag{5.6}$$

Repeat these agruments to find a lower estimate for the second ladder point x_2. By definition of T^n–ladder moments we have

$$\min_{1 \le j < l_2} |x_j - x_0| \ge T^{-n} |x_{l_1} - x_0| \ge T^{-2n} D^{-n}. \tag{5.7}$$

and absence of weak (D,n)–gaps implies that

$$|x_{l_2} - x_0| \geq T^{-2n} D^{-2n}. \tag{5.8}$$

Application of these arguments s times gives the required lower bound (5.4) for distance to x_0 of the closest return. This completes the proof of the Corollary. Q.E.D.

Before we start with a proof of this Theorem let's state three lemmas: one is a Euclidean Algorithm for trajectories with close returns and the other two are about properties of linear operators in R^N. Let $n_1, n_2, \dots, n_s$ be a set of positive integers. Denote by $gcd(n_1, n_2, \dots, n_s)$ their greatest common divisor.

Lemma 5.1.6. *(Euclidean Algorithm) Let $f \in \mathrm{Diff}^1(B^N)$ be a C^1-smooth diffeomorphism of the unit ball B^N (or let $f \in C^1(I,I)$ be a C^1-smooth selfmap of the interval $I = [-1,1]$), $M_1 = \max\{\|f\|_{C^1}, 2\}$ and $\lambda > 0$. Suppose that $\{x_j = f^j(x_0)\}_{j=0}^n \subset B^N$ is a trajectory of f such that for some $1 < n_1 < n_2 < \dots < n_s \leq n$ we have $|x_{n_j} - x_0| \leq \lambda$ for each $1 \leq j \leq s$. Then for $n^* = \gcd(n_1, n_2, \dots, n_s)$ we have $|x_{n^*} - x_0| \leq \lambda M_1^{2n}$.*

Lemma 5.1.7. *Let $A : \mathbb{R}^N \to \mathbb{R}^N$ be a linear operator. Suppose there are a multiplicatively ordered set of positive integers $\{l_t\}_{t=1}^s \subset \mathbb{Z}_+$ bounded by another integer $n \in \mathbb{Z}_+$, the set of ratios $\{p_t = l_{t+1}/l_t\}_{t=1}^{s-1} \subset \mathbb{Z}_+$, $T \geq M^{2N}$, and a collection of unit vectors $\{v_t \subset \mathbb{R}^N,\ \|v_t\| = 1\}_{t=0}^s$ such that for linear operators $\{B_t = B_{t-1}^{p_t}\}_{t=2}^s,\ B_1 = A,\ B_0 = Id,\ p_0 = 1\}$ we have*

$$\|(Id + B_t + \dots + B_t^{p_t - 1})v_{t+1}\| \leq T^{-n} \|(Id + B_{t-1} + \dots + B_{t-1}^{p_{t-1}-1})v_t\| \tag{5.9}$$

for $t = 2, \dots, s-1$. Then $s \leq N$.

This Lemma is proved with help of the following Lemma, which loosely speaking says that if for a unit vector $v \in \mathbb{R}^N$, a complex number $\lambda \in \mathbb{C}$, and a linear operator $A : \mathbb{R}^N \to \mathbb{R}^N$ we have that $(A - \lambda \cdot Id)v$ is small vector, then A has an eigenvector close to λ.

Lemma 5.1.8. *Let $B : \mathbb{R}^N \to \mathbb{R}^N$ be a linear operator, $M \geq \max\{2, \|B\|\}$. Suppose for some unit vector $v \in \mathbb{R}^N,\ \|v\| = 1,\ T > 0$, and a positive integer $p \in \mathbb{Z}_+$ we have*

$$\|(Id + B + \dots + B^{p-1})v\| \leq T^{-1}. \tag{5.10}$$

Then there is an eigenvalue $\lambda \in \mathbb{C}$ of B with the property for some $1 < k < p$ we have $|\lambda - \exp(2\pi i k/p)| \leq 2M^p/(pT^{1/N})$.

We shall prove both Lemmas at the end of Section 5.2.

Proof of Theorem 5.1.4: To explain an idea of the proof assume that the collection of T^{-n}–ladder moments is multiplicatively ordered and $\{p_t\}_{t=0}^{t-1}$ is the set of integer ratios, i.e. $l_t = p_{t-1} l_{t-1}$. Then consider the diffeomorphism f^{l_1} restricted to a neighbourhood U_{x_0} of x_0. To simplify explanation replace f^{l_1} in U_{x_0} by its linear part

$$f^{l_1}(x) = \ x_{l_1} + \ df^{l_1}(x_0)(x - x_0). \tag{5.11}$$

Then at the second ladder moment iterative application of this formula gives

$$\begin{gathered} f^{l_2}(x_0) = x_0 + (x_{l_2} - x_0) = x_0 + \\ \Big(Id + df^{l_1}(x_0) + \cdots + \big(df^{l_1}(x_0)\big)^{p_1 - 1} \Big)(x_{l_1} - x_0). \end{gathered} \tag{5.12}$$

Since $|x_{l_2} - x_0| < T^{-n}|x_{l_1} - x_0|$ we have that $(x_{l_1} - x_0)$, normalized to the unit length satisfies Lemma 5.1.7 with $B_1 = A = df^{l_1}(x_0)$ or indicates that $df^{l_1}(x_0)$ has an eigenvalue close to $\exp(2\pi i k_1/p_1)$ for some integer $1 < k_1 < p_1$. Similarly,

$$\begin{gathered} f^{l_3}(x_0) = x_0 + (x_{l_3} - x_0) = x_0 + \\ \Big(Id + df^{l_2}(x_0) + \cdots + \big(df^{l_2}(x_0)\big)^{p_2 - 1} \Big)(x_{l_2} - x_0). \end{gathered} \tag{5.13}$$

Since $|x_{l_3} - x_0| < T^{-n}|x_{l_2} - x_0|$ we have that $(x_{l_2} - x_0)$ normalized to the unit length satisfies Lemma 5.1.7 with $B_1 = df^{l_1}(x_0),\ B_2 = B_1^{p_1}$ or indicates that $df^{l_1}(x_0)$ has an eigenvalue close to $\exp(2\pi i \frac{k_2}{p_1 p_2})$ for some positive integer k_2.

Repeating these arguments $(s-1)$–times we shall prove that there are vectors $(x_{l_1} - x_0), (x_{l_2} - x_0), \ldots, (x_{l_s} - x_0)$ which after being normalized to the unit length satisfy Lemma 5.1.7. Application of this Lemma shows that $s \leq N$. In other words, each of those vectors $v_t = (x_{l_t} - x_0)/|x_{l_t} - x_0|$'s forces (by the Lemma) to have an eigenvalue close to a root of unity $\exp(2\pi i \frac{k_t}{p_1 \ldots p_t})$ and those roots are pairwise distinct. Now we prove that approximation by the linearization is good enough and our heuristic arguments work.

Step 1. Choice of multiplicatively ordered close returns.

In general T^{-n}-ladder moments $l_1, l_2, \ldots, l_s$ are not necessarily multiplicatively ordered. We shall construct now $r_1, r_2, \ldots, r_s$ that are multiplicatively ordered and fit to the scheme of almost eigenvalues (5.11-5.13) above.

Consider the trajectory $\{x_j\}_{j=0}^{n-1}$, and take the set of indices of those points that are T^{-n}-close to x_0: $L_1 = \{1 \leq j < n : |x_j - x_0| \leq T^{-n}\}$. By definition l_1 is the minimal element in L_1. By the Euclidean Algorithm (Lemma 5.1.6) for $r_1 = \gcd(L_1)$

we have that $|x_{r_1}-x_0|\le M_1^{2n}T^{-n}$. Since l_1 is divisible by r_1, $x_{l_1}=(f^{r_1})^{l_1/r_1}(x_0)$ and $|x_{l_1}-x_0|\le M_1^n|x_{r_1}-x_0|$.

Consider indices of points of $\{x_j\}_{j=0}^{n-1}$ that are $T^{-n}|x_{l_1}-x_0|$-close to x_0: $L_2=\{1\le j<n : |x_j-x_0|\le T^{-n}|x_{l_1}-x_0|\}$. By the Euclidean Algorithm (Lemma 5.1.6) we have

$$|x_{r_2}-x_0|\le M_1^{2n}T^{-n}|x_{l_1}-x_0|\le M_1^{3n}T^{-n}|x_{r_1}-x_0|,\ \text{where } r_2=\gcd(L_2). \tag{5.14}$$

Since l_2 is divisible by r_2, $x_{l_2}=(f^{r_2})^{l_2/r_2}(x_0)$ and $|x_{l_2}-x_0|\le M_1^n|x_{r_2}-x_0|$.

Similarly for each $m=1,2,\dots,s$, to define r_m consider indices of points $\{x_j\}_{j=0}^{n-1}$ that are $T^{-n}|x_{l_{m-1}}-x_0|$-close to x_0: $L_m=\{1\le j<n : |x_j-x_0|\le T^{-n}|x_{l_{t-1}}-x_0|\}$. Put $r_m=\gcd(L_m)$. By the Euclidean Algorithm (Lemma 5.1.6), we have that

$$|x_{r_m}-x_0|\le M_1^{2n}T^{-n}|x_{l_{m-1}}-x_0|\le M_1^{3n}T^{-n}|x_{r_{m-1}}-x_0|. \tag{5.15}$$

Since $L_1\supset L_2\supset\cdots\supset L_s$ it follows that $r_1,r_2,\dots r_s$ are multiplicatively ordered. Denote by p_{m-1} the ratio r_m/r_{m-1} for each $m=2,\dots,s$.

Step 2. Almost eigevalues along multiplicatively ordered iterates. In this step we reproduce calculations (5.11–5.13) in the nonlinear case. Denote for an integer m by $O_m(|y|^{1+\rho})$ a vector whose norm is bounded by $M_{1+\rho}^{m(1+\rho)}|y|^{1+\rho}$. Consider the Taylor expansion of $f^{r_1}(x_0)$ in a neighbourhood of x_0

$$f^{r_1}(x)=x_0+df^{r_1}(x_0)\,(x-x_0)+O_{r_1}\left(|x-x_0|^{1+\rho}\right). \tag{5.16}$$

Inductive application of this fomula with straightforward calculations shows that

$$\begin{aligned} f^{r_2}(x_0)&=x_0+(x_{r_2}-x_0)=x_0+(x_{r_1}-x_0)+\cdots+(x_{r_2}-x_{(p_1-1)r_1})=x_0\\ &+\left(Id+df^{r_1}(x_0)+\cdots+(df^{r_1}(x_0))^{p_1-1}\right)(x_{r_1}-x_0)+O_{r_2}\left(|x_{r_1}-x_0|^{1+\rho}\right). \end{aligned} \tag{5.17}$$

Since $T=M_{1+\rho}^{8N/\rho}$, $0<\rho\le 1$, x_{l_1} is the first T^n-ladder moment of the trajectory $\{x_j\}_{j=0}^{\infty}$. This, along with the estimate $|x_{r_1}-x_0|\le M_1^{2n}T^{-n}$, leads to an estimate of the remainder term

$$\left|O_{r_2}\left(|x_{r_1}-x_0|^{1+\rho}\right)\right|\le M_{1+\rho}^{2n\rho+(1+\rho)r_2}T^{-n\rho}|x_{r_1}-x_0|\le M_{1+\rho}^{4n\rho}T^{-n\rho}|x_{r_1}-x_0|. \tag{5.18}$$

By our construction (5.14), we have that $|x_{r_2}-x_0|\le M_{1+\rho}^{3n}T^{-n}|x_{r_1}-x_0|$. Therefore,

$$\begin{aligned} &\left|\left(Id+df^{r_1}(x_0)+(df^{r_1}(x_0))^{p_1-1}\right)(x_{r_1}-x_0)\right|\le\\ &(M_{1+\rho}^{3n}\,T^{-n}+M_{1+\rho}^{4\rho n}\,T^{-n\rho})\,|x_{r_1}-x_0|\le M_{1+\rho}^{-2Nn}\,|x_{r_1}-x_0|. \end{aligned} \tag{5.19}$$

and the unit vector $\frac{x_{r_1}-x_0}{|x_{r_1}-x_0|}$ satisfies Lemma 5.1.7 with $A = df^{r_1}(x_0),\ B_1 = A,\ p_1 = p_1$, and $T = M_{1+\rho}^{-2Nn}$. Thus, $df^{r_1}(x_0)$ has an eigenvalue close to $\exp(2\pi i k_1/p_1)$ for some $1 < k_1 < p_1$.

Consider now the Taylor expansion of $f^{r_2}(x_0)$ in a neighbourhood of x_0 similar to the one (5.16) for $f^{r_1}(x_0)$ and using similar calculations to those which lead to (5.17) we get

$$\begin{gathered} f^{r_3}(x) = x_0 + (x_{r_3} - x_0) = \ x_0 + \\ \big(Id + df^{r_2}(x_0) + \cdots + (df^{r_2}(x_0))^{p_2-1}\big)(x_{r_2} - x_0) + O_{r_3}\left(|x_{r_2} - x_0|^{1+\rho}\right). \end{gathered} \tag{5.20}$$

Using estimate (5.15) for $t = 3$ instead estimate (5.14) we can get the following estimate for the remainder term

$$\left|O_{r_3}\left(|x_{r_2} - x_0|^{1+\rho}\right)\right| \le M_{1+\rho}^{2\rho n+(1+\rho)r_3} T^{-n\rho}|x_{r_2} - x_0| \le M_{1+\rho}^{4n\rho} T^{-n\rho}|x_{r_2} - x_0|, \tag{5.21}$$

which is almost identical to (5.18). Therefore, we can get

$$\begin{gathered} \left|\left(Id + df^{r_2}(x_0) + (df^{r_2}(x_0))^{p_2-1}\right)(x_{r_2} - x_0)\right| \le \\ (M_{1+\rho}^{3n}\ T^{-n} + M_{1+\rho}^{4\rho n}\ T^{-n\rho})\ |x_{r_2} - x_0| \le M_{1+\rho}^{-2Nn}\ |x_{r_2} - x_0| \end{gathered} \tag{5.22}$$

and the unit vector $v_2 = \frac{x_{r_2}-x_0}{|x_{r_2}-x_0|}$ and $v_1,\ p_1,\ B_1$, and T as above satisfy Lemma 5.1.7. Thus, we almost proved that $df^{r_1}(x_0)$ has an eigenvalue close to $\exp(2\pi i \frac{k_2}{p_1 p_2})$ for some $1 < k_2$. However, in our considerations above we need to replace $(df^{r_2}(x_0))^{p_2}$ by $(df^{r_1}(x_0))^{p_1 p_2}$ with a remainder, because our goal is to show all vectors of the form $\{v_t = (x_{r_t} - x_0)/|x_{r_t} - x_0|\}_{t=1}^{s}$ satisfy Lemma 5.1.7 with operators $\{B_t\}_{t=1}^{s}$ being powers of one operator $A = df^{r_1}(x_0)$.

By our construction of multiplicatively ordered close returns (5.14-5.15) we have that $|x_{r_1} - x_0| \le M_1^{2n} T^{-n}$ and all the points $\{x_{l_t}, x_{r_t}\}_{t=1}^{s}$ have indices $\{l_t, r_t\}_{t=1}^{s}$ divisible by r_1. Therefore, the estimate M_1 on C^1-norm of f gives that

$$|f^{r_1}(x) - f^{r_1}(y)| \le M_1^{r_1}|x - y| \tag{5.23}$$

for each pair $x, y \in B^N$ and for any $d \le \frac{n}{r_1}$

$$|x_{dr_1} - x_0| \le M_1^{3n} T^{-n} \le M_{1+\rho}^{3n} T^{-n}. \tag{5.24}$$

By definition the linearization $df^{r_2}(x_0)$ can be written in the form

$$\begin{gathered} df^{r_2}(x_0) = df^{r_1}(x_{(p_1-1)r_1}) \circ \cdots \circ df^{r_1}(x_{r_1}) \circ df^{r_1}(x_0) = \\ \big(df^{r_1}(x_0) + \Delta_{r,f}(x_{(p_1-1)r_1}, x_0)\big) \circ \cdots \circ (df^{r_1}(x_0) + \Delta_{r,f}(x_{r_1}, x_0)) \circ df^{r_1}(x_0), \end{gathered} \tag{5.25}$$

where $\Delta_{r,f}(x,x_0)$ is a linear operator bounded by $M_{1+\rho}^{r_1}|x-x_0|^\rho$. By (5.24) we have that the norm of each $\Delta_{r,f}(x_{dr_1},x_0)$ is bounded by $M_{1+\rho}^{3n\rho+r_1}T^{-n\rho} \le M_{1+\rho}^{-5Nn+r_1}$. Therefore,

$$\|df^{r_2}(x_0)-(df^{r_1}(x_0))^{p_1}\| \le 2p_2 M_{1+\rho}^{-5Nn+r_1} \tag{5.26}$$

and applying this inequality to (5.22) we get

$$\left\|\left(Id+(df^{r_1}(x_0))^{p_1}+\cdots+((df^{r_1}(x_0))^{p_1})^{p_2-1}\right)v_2\right\| \le 4p_2^2 M_{1+\rho}^{-5Nn+r_2}. \tag{5.27}$$

The result of this calculations is that the unit vector $\frac{x_{r_2}-x_0}{|x_{r_2}-x_0|}$ satisfies Lemma 5.1.7 with parameters as those given after (5.19,5.22).

Inductive application of these calculations show that for each $t=1,\dots,s$ the unit vector $\frac{x_{r_t}-x_0}{|x_{r_t}-x_0|}$ satisfies Lemma 5.1.7 with $B_t=(df^{r_1}(x_0))^{p_1\dots p_t}$.

$$\left|\left(Id+(df^{r_1}(x_0))^{p_1\dots p_{t-1}}+\cdots+((df^{r_1}(x_0))^{p_1\dots p_{t-1}})^{p_2-1}\right)\right| \tag{5.28}$$

$$\le 2^{2t}\left(\frac{r_t}{r_1}\right)^2 M_{1+\rho}^{-5Nn+r_t}. \tag{5.29}$$

Application of Lemma refspectrum completes the proof of the Theorem. Q.E.D.

Proof of Lemma 5.1.6 (Euclidean Algorithm): As the reader can check, the direct application of the Euclidean Algorithm to the pair of points with largest indices m_{s-1} and m_s, then to the new pair with largest indices, and so on one arrives at the greatest common divisor of indices gives distortion $M_1^{sn} \gg M_1^{2n}$ in the conclusion of the Lemma. Also such an approach requires taking preimages. Below we shall only forward images (so it is applicable to noninvertible maps) and obtain required M_1^{2n} distortion.

Consider a set of positive integers $1<m_1<m_2<\cdots<m_p$, where p is some integer. Define a transformation of one increasing set of positive integers into another one:

$$\phi:(m_1,\dots,m_p)\to(m_1',\dots,m_{p'}'), \tag{5.30}$$

where ϕ is defined as follows. $r_p=\left[\frac{m_p}{m_{p-1}}\right]$, $\tilde m_p=m_p-\left[\frac{m_p}{m_{p-1}}\right]m_{p-1}$ and

· if $\tilde m_p$ equals m_j for some $1\le j<s$, then $\phi(m_1,\dots,m_p)=(m_1,\dots,m_{p-1})$

· if $\tilde m_p$ is different from all of m_j's for $1\le j<s$, then permute $m_1,\dots,m_{p-1},\tilde m_p$ so that they form an increasing sequence and denote this sequence by $(m_1',\dots,m_{p'}')$ with $p'=p$. In this case $\phi(m_1,\dots,m_p)=(m_1',\dots,m_{p'}')$.

Properties of ϕ: $\max(m'_1, \ldots, m'_{p'}) \leq m_{p-1} < m_p = \max(m_1, \ldots, m_p)$ and $\gcd(m_1, \ldots, m_p) = \gcd(m'_1, \ldots, m'_{p'})$. These properties follow directly from the definition. The first property implies that for some positive integer $k < n$ we have $\phi^k(m_1, \ldots, m_p) = \gcd(m_1, \ldots, m_p) = m^*$. Now we describe the corresponding procedure for points of the trajectory $\{x_j\}_{j=0}^{n-1}$ of a C^1-smooth diffeomorphism $f \in \text{Diff}^{1+\rho}(B^N)$ or a C^1-smooth map $f \in C^1(I, I)$ under consideration.

Suppose that for some indices m_p and m_{p-1}, points x_{m_p} and $x_{m_{p-1}}$ of the trajectory $\{x_j\}_{j=0}^{n-1}$ satisfy

$$|x_{m_p} - x_0| \leq \lambda' \quad \text{and} \quad |x_{m_{p-1}} - x_0| \leq \lambda'. \tag{5.31}$$

We need to show that for $\tilde{m}_p$, defined in (5.30), we have the following estimates

$$|x_{\tilde{m}_p} - x_0| \leq 3M_1^{m_p - m_{p-1}}\lambda'. \tag{5.32}$$

By definition $\tilde{m}_p = m_p - r_p m_{p-1}$. Using the estimate on distortion of distances under forward iterations, we get

$$\begin{aligned} |x_{r_p m_{p-1}} - x_0| \leq \sum_{j=0}^{r_p-1} |x_{(j+1)m_{p-1}} - x_{jm_{p-1}}| \leq \\ \frac{M_1^{r_p m_{p-1}} - 1}{M_1^{m_{p-1}} - 1}\lambda' \leq 2M_1^{(r_p-1)m_{p-1}}\lambda'. \end{aligned} \tag{5.33}$$

Now consider x_0 and $x_{r_p m_{p-1}}$ and iterate them $\tilde{m}_p = m_p - r_p m_{p-1}$ times forward. This gives

$$|x_{\tilde{m}_p} - x_0| \leq |x_{m_p} - x_0| + |x_{m_p} - x_{\tilde{m}_p}| \leq (1 + 2M_1^{m_p - m_{p-1}})\lambda'. \tag{5.34}$$

Let $m_p = n_s$ and $m_{p-1} = n_{s-1}$. Then for $\phi(n_1, \ldots, n_s) = (n'_1, \ldots, n'_{s'})$ we have $n'_{s'} = n_{s-1}$ and $\max_{1 \leq j \leq s'} |x_{n'_j} - x_0| \leq (3M_1^{n_s - n'_{s'}} + 1)\lambda$.

Let $m_p = n'_{s'}$ and $m_{p-1} = n'_{s'-1}$. Then for $\phi(n'_1, \ldots, n'_{s'}) = (\tilde{n}_1, \ldots, \tilde{n}_{\tilde{s}})$ we have $\tilde{n}_{\tilde{s}} \leq n'_{s'}$ and $\max_{1 \leq j \leq \tilde{s}} |x_{\tilde{n}_j} - x_0| \leq 4^2 M_1^{n'_{s'} - \tilde{n}_{\tilde{s}}}\lambda$.

Inductive application of these arguments show that for some $k < n$ we have $n^* = \phi^k(n_1, \ldots, n_s) = \gcd(n_1, \ldots, n_s)$ and

$$|x_{n^*} - x_0| \leq 3^{\log n} M_1^{n_s}\lambda \leq M_1^{2n}\lambda. \tag{5.35}$$

This completes the proof of Lemma 5.1.6. Q.E.D.

Remark 5.1.9. *We point out that we use only forward iterates so the proof works not only for diffeomorphisms, but also for noninvertible maps.*

To prove Lemma 5.1.7 we first prove Lemma 5.1.8.

Proof of Lemma 5.1.8: Let $v_1 = v, v_2, \dots, v_N \in \mathbb{R}^N$ be an orthonormal basis. Recall that for a linear operator $L : \mathbb{R}^N \to \mathbb{R}^N$ determinant $\det L$ equals oriented volume of the parallelogram span by $Lv_1, \dots, Lv_N$. Such a volume is bounded in absolute value by product of lengths $\prod_{j=1}^N |Lv_j|$. Put $L = Id + B + \cdots + B^{l-1}$. Since $\|B\| \le M$, we get $\|L\| \le (M^p - 1)/(M-1)$ and $|\det L| \le T^{-1}\left((M^p-1)/(M-1)\right)^{N-1}$.

Another way to express determinant is $\det L = \prod_{j=1}^N (\lambda_j^p - 1)/(\lambda_j - 1)$, where $\lambda_1, \dots, \lambda_N \in \mathbb{C}$ are eigenvalues of B. Since $\|B\| \le M$, each eigenvalue $|\lambda_j| \le M$. Denote $J_- = \{1 \le j \le N : \ |\lambda_j^p - 1| \le |\lambda_j - 1|\}$ and $J_+ = \{1, \dots, N\} \setminus J_-$. Then

$$|\det L| = \prod_{j \in J_-} \frac{\lambda_j^p - 1}{\lambda_j - 1} \prod_{j \in J_+} \frac{\lambda_j^p - 1}{\lambda_j - 1} \le T^{-1} \left(\frac{M^p - 1}{M - 1} \right)^{N-1}. \tag{5.36}$$

This implies that for some j^* we get

$$\left| \frac{\lambda_{j^*}^p - 1}{\lambda_j - 1} \right| \le \left(\frac{M^p - 1}{M - 1} \right) T^{-1/N}. \tag{5.37}$$

This in turn implies statement of Lemma 5.1.8. Q.E.D.

Proof of Lemma 5.1.7: Apply Lemma 5.1.8 to $B = B_t,\ v = v_t,\ p = p_t$ for each $t = 1, \dots, s$. It implies that each $B_t = A^{l_t}$ has an eigenvalue λ_t and $|\lambda_t - \exp(2\pi i k_t / p_t)| \le 2\|A\|^{-n}/p_t$ with $1 < k_t \le p_t < n$. Therefore, A has an eigenvalue close to $\exp(2\pi i k_t / l_{t+1})$. Notice that exponential in n smallness of the right-hand side above implies that $\{\lambda_t\}_{t=1}^s$ are pairwise distinct. To see that for some $t < t'$ suppose $\lambda_t^{1/l_t} = \lambda_{t'}^{1/l_{t'}}$. Then $\lambda_t^{p_t} = \lambda_{t'}^{l_{t+1}/l_{t'}}$. Since λ_t is close to $\exp(2\pi i k_t/p_t)$, then $\lambda_t^{p_t}$ is close to 1. Hovewer, $\lambda_{t'}^{l_{t+1}/l_{t'}}$ is close to $\exp(2\pi\ i k_{t'} l_{t+1}/l_{t'+1})$ with $k_{t'} > 1$ which is a contradiction. Since eigenvalues $\{\lambda_t\}_{t=1}^s$ of A are pairwise distinct, $s \le N$. This completes the proof of Lemma 5.1.7. Q.E.D.

5.2 One can get rid of leading saddles by shifting the starting point

In this section we prove a part of the Shift Theorem of Section 3.5 covering the left branch of Fig. 5.1, namely, *if for some $k < n$ a point x_0 and each its shifts x_j of length $0 \le j < n \log_2 n$ has a (k, n, C, δ, ρ)-leading saddle, then x_0 is $(n, \gamma_n(C, \delta))$-hyperbolic.* This section provides also an essential step toward completing a proof of the Shift Theorem (the right branch of the same Figure).

Lemma 5.2.1. *(The Shift Lemma) Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $0 < \rho \le 1$ with $M_{1+\rho}$ bound on the $C^{1+\rho}$-norm of f, defined as in Theorem 5.1.4. Let $\delta > 0$ and $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$ be positive constants (of Induction), let n be a positive integer (period), and let f satisfy Inductive Hypothesis of order $(n-1)$ with constants C, δ, and ρ, i.e. $f \in IH(C, \delta, \rho, n-1)$. Suppose f has a trajectory $\{x_s = f^s(x_0)\}_{s=0}^n$ such that x_0 is $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic and for some $k < n$ has a (k, n, C, δ, ρ)-leading saddle, i.e. $|x_k - x_0| \le n^{-1/\rho}\gamma_k^{4N/\rho}(C,\delta)$. Then either x_0 is $(n, \gamma_n(C,\delta))$-hyperbolic or for some $0 \le j < n\log_2 n$ there is a shift $x_j = f^j(x_0)$ such that the point x_j has no (n, C, δ, ρ)-leading saddles.*

First we formulate some auxiliary Lemmas and then prove the Shift Lemma.

Lemma 5.2.2. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $0 < \rho \le 1$ with $M_{1+\rho}$ bound on the $C^{1+\rho}$-norm of f, defined as in Theorem 5.1.4. For all $x, y \in B^N$ and $k \in \mathbb{Z}^+$,*

$$|df^k(x) - df^k(y)| \le (M_{1+\rho}^{k(1+\rho)} - 1)|x - y|^\rho. \tag{5.38}$$

Proof: Let $x_j = f^j(x)$ and $y_j = f^j(y)$. Then $|x_j - y_j| \le M_{1+\rho}^j|x - y|$, and

$$\begin{aligned}
&|df^k(x) - df^k(y)| \\
&\quad = |df(x_{k-1})df(x_{k-2})\cdots df(x_0) - df(y_{k-1})df(y_{k-2})\cdots df(y_0)| \\
&\quad = |(df(x_{k-1}) - df(y_{k-1}))df(x_{k-2})\cdots df(x_0) \\
&\qquad + df(y_{k-1})(df(x_{k-2}) - df(y_{k-2}))df(x_{k-3})\cdots df(x_0) \\
&\qquad + \cdots + df(y_{k-1})\cdots df(y_1)(df(x_0) - df(y_0))| \\
&\quad \le M_{1+\rho}^k(|x_{k-1} - y_{k-1}|^\rho + |x_{k-2} - y_{k-2}|^\rho + \cdots + |x_0 - y_0|^\rho) \\
&\quad \le M_{1+\rho}^k|x - y|^\rho(M_{1+\rho}^{(k-1)\rho} + M_{1+\rho}^{(k-2)\rho} + \cdots + 1) \\
&\quad \le M_{1+\rho}^k|x - y|^\rho(M_{1+\rho}^{k\rho} - 1) \\
&\quad \le (M_{1+\rho}^{k(1+\rho)} - 1)|x - y|^\rho
\end{aligned} \tag{5.39}$$

since $M_{1+\rho}^\rho \ge 2$. Q.E.D.

Lemma 5.2.3. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $0 < \rho \le 1$ with $M_{1+\rho}$ bound on the $C^{1+\rho}$-norm of f, defined as in Theorem 5.1.4. Then for any $k \in \mathbb{Z}_+$, $\alpha > 1 + 1/\rho$, and $\gamma \le (2M_{1+\rho}^k)^{-(1+\rho)/(\alpha\rho-\rho-1)}$, if a point $x_0 \in B^N$ is (k, γ^α)-periodic and (k,γ)-hyperbolic, then there is a periodic point x_0^* of period k, i.e., $f^k(x_0^*) = x_0^*$, such that $|x_0^* - x_0| \le 2\gamma^{\alpha-1}$.*

Proof: Consider the map $h(x) = f^k(x) - x$, $h : B^N \to \mathbb{R}^N$. Since x_0 is a (k, γ)-hyperbolic point of f, the norm of the inverse of the linearization matrix of h at x_0 satisfies the inequality $\|dh^{-1}(x_0)\| \le \gamma^{-1}$. For x near x_0, denote $\Delta(x) = dh(x) - dh(x_0)$, and consider the chain of inequalities

$$\begin{aligned} \|dh^{-1}(x)\| = \|(dh(x_0) + \Delta(x))^{-1}\| &\le \|dh^{-1}(x_0)(Id + dh^{-1}(x_0)\Delta(x))^{-1}\| \\ &\le \gamma^{-1}\|(Id + dh^{-1}(x_0)\Delta(x))^{-1}\| \end{aligned} \tag{5.40}$$

The right side is bounded by $2\gamma^{-1}$ if $\|dh^{-1}(x_0)\Delta(x))\| \le \frac{1}{2}$. This inequality in turn holds if $\gamma^{-1}\|\Delta(x)\| \le \frac{1}{2}$. Now by Lemma 5.2.2, if $|x - x_0| \le 2\gamma^{\alpha-1}$ then

$$\gamma^{-1}\|\Delta(x)\| \le \gamma^{-1} M_{1+\rho}^{k(1+\rho)} |x - x_0|^\rho \le M_{1+\rho}^{k(1+\rho)} 2^\rho \gamma^{(\alpha-1)\rho-1}. \tag{5.41}$$

The right side is at most $\frac{1}{2}$ by the hypothesis on γ. Thus, we have $\|dh^{-1}(x)\| \le 2\gamma^{-1}$ for $|x - x_0| \le 2\gamma^{\alpha-1}$. Therefore, the image of the $2\gamma^{\alpha-1}$-ball centered at x_0 under the map h contains the γ^α-ball centered at $h(x_0)$. Since $|h(x_0)| \le \gamma^\alpha$, there exists a point x_0^* in the $2\gamma^{\alpha-1}$-ball centered at x_0 such that $h(x_0^*) = 0$. Thus, $f^k(x_0^*) = x_0^*$. This completes the proof of the Lemma. Q.E.D.

Lemma 5.2.4. *Let x^* be a periodic point with period k that is $(k, \gamma_k(C, \delta))$-hyperbolic. If $|x - x^*| \le (M_{1+\rho}^{-k(1+\rho)} \gamma_k(C, \delta)/2)^{1/\rho}$, then*

$$|f^k(x) - x| \ge (\gamma_k(C, \delta)/2)|x - x^*| \tag{5.42}$$

Proof: We have

$$\begin{aligned} |f^k(x) - x| &= |f^k(x) - x - (f^k(x^*) - x^*)| \\ &\ge |df^k(x^*)(x - x^*) - (x - x^*)| - M_{1+\rho}^{k(1+\rho)} |x - x^*|^{1+\rho}. \end{aligned} \tag{5.43}$$

By hyperbolicity the first term on the right is bounded below by $\gamma_k(C, \delta)|x - x^*|$, and by hypothesis the second term is bounded above by half this quantity. Q.E.D.

Proof of Lemma 5.2.1 (Shift Lemma): Recall that $\gamma_k(C, \delta) = \exp(-Ck^{1+\delta})$. Let

$$\varepsilon_k(n, C, \delta, \rho) = n^{-1/\rho} \gamma_k^{4N/\rho}(C, \delta), \tag{5.44}$$

and

$$\beta_k(n, C, \delta, \rho) = 4\varepsilon_k(n, C, \delta, \rho)/\gamma_k(C, \delta). \tag{5.45}$$

Notice that x_0 has a (k, n, C, δ, ρ)-leading saddle if and only if $|x_k - x_0| \le \varepsilon_k(n, C, \delta, \rho)$. Also if $|x_k - x_0| \le 2\varepsilon_k(n, C, \delta, \rho)$, then application of Lemma 5.2.3 gives that $\beta_k(n, C, \delta, \rho)$-close to x_0 there is a periodic point $x_0^* = f^k(x_0^*)$. It will be important below that

$$\varepsilon_k(n, C, \delta, \rho) \ge M_{1+\rho}^{6k} \beta_{2k}(n, C, \delta, \rho) \tag{5.46}$$

and that

$$M^{3k}_{1+\rho}\beta_k(n, C, \delta, \rho) > 2M^{3(k+1)}_{1+\rho}\beta_{k+1}(n, C, \delta, \rho); \tag{5.47}$$

both of these inequalities follow from $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$.

Let k_0 be the smallest positive integer for which $|x_{k_0} - x_0| \leq 2\varepsilon_{k_0}(n, C, \delta, \rho)$; by hypothesis k_0 exists and is less than n. Let $m_0 = 0$. We will inductively define integer sequences $k_0, k_1, k_2, \ldots$ and $m_0, m_1, m_2, \ldots$ such that for all $i \geq 0$:

(a_i) $2^i \leq k_i < n$;

(b_i) $0 \leq m_i \leq in$;

(c_i) $|x_{m_i} - x_{m_i+k_i}| \leq 2\varepsilon_{k_i}(n, C, \delta, \rho)$;

(d_i) $|x_{m_i} - x_{m_i+k}| > 2\varepsilon_k(n, C, \delta, \rho)$ for all $k_i/2 \leq k < k_i$; and

(e_i) $|x_{m_i} - x_{m_i+k}| > \varepsilon_k(n, C, \delta, \rho) + 2M^{3k_i}_{1+\rho}\beta_{k_i}(n, C, \delta, \rho)$ for all $0 < k < k_i/2$.

The first four conditions are easily checked for $i = 0$, while the fifth follows from the fact that $2M^{3k_0}_{1+\rho}\beta_{k_0}(n, C, \delta, \rho) < M^{6k}_{1+\rho}\beta_{2k}(n, C, \delta, \rho) \leq \varepsilon_k(n, C, \delta, \rho)$ for $k < k_0/2$ (see (5.46)).

We proceed by induction on i, showing that if conditions (a_i)–(e_i) are satisfied, then either we can complete the proof of the Lemma or we can define k_{i+1} and m_{i+1} that satisfy conditions (a_{i+1})–(e_{i+1}). Notice that condition (a_i) implies that i is less than $\log_2 n$, so the Lemma is proved in fewer than $\log_2 n$ inductive steps.

Assume then that for some $i \geq 0$, conditions (a_i)–(e_i) are satisfied. Because $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$, it follows that the hypothesis of Lemma 5.2.3 are satisfied with $k = k_i$, $\gamma = \gamma_{k_i}(C, \delta)$, $\alpha = \log(2\varepsilon_{k_i}(n, C, \delta, \rho))/\log(\gamma_{k_i}(C, \delta))$, and x_0 replaced by x_{m_i}. Thus by Lemma 5.2.3 and condition (c_i), there is a periodic point x^*_i of period k_i such that $|x_{m_i} - x^*_i| \leq \beta_{k_i}(n, C, \delta, \rho)$. Let q_i be the smallest integer for which $q_i k_i \geq n$.

The remainder of the proof goes by induction in i according to the following scheme: Pick a point x_{m_i}. It satisfies (c_i). We consider iterates of x_{m_i} by f^{k_i}, i.e. $x_{m_i}, x_{m_i+k_i}, \ldots$. There are two cases:

A) for some $0 \leq p < q_i$ the corresponding $x_{m_i+pk_i}$ leaves the $\beta_{k_i}(n, C, \delta, \rho)$-neighborhood of x^*_i. In this case, we will conclude that $x_{m_i+pk_i}$ does not have a (k, n, C, δ, ρ)-leading saddle for any $k < 2k_i$. We will then let $m_{i+1} = m_i + pk_i$ and either be able to define k_{i+1} or conclude that $x_{m_{i+1}}$ has no (n, C, δ, ρ)-leading saddles.

B) every $x_{m_i+pk_i}$ for $0 \leq p < q_i$ is in the $\beta_{k_i}(n, C, \delta, \rho)$-neighborhood of the periodic point x_i^*. From this, we will conclude using the Inductive Hypothesis of order k_i and Lemma A.4 in Appendix A, Part I that x_0 is $(n, \gamma_n(C, \delta))$-hyperbolic.

Case A): For some $0 < p < q_i$ we have $|x_{m_i+pk_i} - x_i^*| > \beta_{k_i}(n, C, \delta, \rho)$. Let p be the smallest such value. Then $|x_{m_i+(p-1)k_i} - x_i^*| \leq \beta_{k_i}(n, C, \delta, \rho)$; if $p = 1$ this follows from the definition of x_i^* above. It follows that

$$\beta_{k_i}(n, C, \delta, \rho) < |x_{m_i+pk_i} - x_i^*| \leq M_{1+\rho}^{k_i} \beta_{k_i}(n, C, \delta, \rho). \tag{5.48}$$

By the inductive hypothesis, x_i^* is $(k, \gamma_k(C, \delta))$-hyperbolic.

Since $C > 100\rho^{-1}\delta^{-1} \log M_{1+\rho}$, we have that

$$M_{1+\rho}^{k_i} \beta_{k_i}(n, C, \delta, \rho) < (M_{1+\rho}^{-k_i(1+\rho)} \gamma_{k_i}(C, \delta)/2)^{1/\rho}. \tag{5.49}$$

Thus we can apply Lemma 5.2.4 with $x^* = x_i^*$ and $x = x_{m_i+pk_i}$, which yields

$$|x_{m_i+(p+1)k_i} - x_{m_i+pk_i}| \geq (\gamma_{k_i}(C, \delta)/2)|x_{m_i+pk_i} - x_i^*| > 2\varepsilon_{k_i}(n, C, \delta, \rho). \tag{5.50}$$

In particular, $x_{m_i+pk_i}$ does not have a $(k_i, n, C, \delta, \rho)$-leading saddle.

Let $m_{i+1} = m_i + pk_i$; condition (b$_{i+1}$) then follows from condition (b$_i$). We have just shown that $|x_{m_{i+1}} - x_{m_{i+1}+k_i}| > 2\varepsilon_{k_i}(n, C, \delta, \rho)$; let k_{i+1} be the smallest integer such that $k_i < k_{i+1} < n$ and $|x_{m_{i+1}} - x_{m_{i+1}+k_{i+1}}| \leq 2\varepsilon_{k_{i+1}}(n, C, \delta, \rho)$; then condition (c$_{i+1}$) is satisfied. If no such k_{i+1} exists, then the proof is complete with $j = m_{i+1} \leq (i+1)n < n \log_2 n$; the last inequality follows from condition (a$_i$).

We claim that $k_{i+1} \geq 3k_i/2$. If not, let $k = k_{i+1} - k_i < k_i/2$. Then

$$\begin{aligned}
&|x_{m_{i+1}} - x_{m_{i+1}+k}| \\
&\qquad \leq |x_{m_{i+1}} - x_{m_{i+1}+k_{i+1}}| + |x_{m_{i+1}+k_{i+1}} - x_{m_{i+1}+k}| \\
&\qquad \leq 2\varepsilon_{k_{i+1}}(n, C, \delta, \rho) + M_{1+\rho}^{k} |x_{m_{i+1}+k_i} - x_{m_{i+1}}| \\
&\qquad \leq 2\varepsilon_{k_{i+1}}(n, C, \delta, \rho) + M_{1+\rho}^{k} (|x_{m_{i+1}+k_i} - x^*| + |x_{m_{i+1}} - x^*|) \\
&\qquad \leq 2\varepsilon_{k_{i+1}}(n, C, \delta, \rho) + 2M_{1+\rho}^{k+2k_i} \beta_{k_i}(n, C, \delta, \rho) \\
&\qquad < \varepsilon_k(n, C, \delta, \rho) + 2M_{1+\rho}^{3k_i} \beta_{k_i}(n, C, \delta, \rho).
\end{aligned} \tag{5.51}$$

But this contradicts condition (e$_i$); thus $k_{i+1} \geq 3k_i/2$ as claimed.

Next, we claim that $k_{i+1} \geq 2k_i$. If not, let $k = 2k_i - k_{i+1} \leq k_i/2$. Then

$$\begin{aligned}
&|x_{m_{i+1}} - x_{m_{i+1}+k}| \\
&\qquad \leq |x_{m_{i+1}} - x_{m_{i+1}+2k_i}| + |x_{m_{i+1}+2k_i} - x_{m_{i+1}+k}| \\
&\qquad \leq |x_{m_{i+1}} - x^*| + |x_{m_{i+1}+2k_i} - x^*| + M_{1+\rho}^{k} |x_{m_{i+1}+k_{i+1}} - x_{m_{i+1}}| \\
&\qquad \leq 2M_{1+\rho}^{3k_i} \beta_{k_i}(n, C, \delta, \rho) + M_{1+\rho}^{k} 2\varepsilon_{k_{i+1}}(n, C, \delta, \rho) \\
&\qquad < 2M_{1+\rho}^{3k_i} \beta_{k_i}(n, C, \delta, \rho) + \varepsilon_k(n, C, \delta, \rho);
\end{aligned} \tag{5.52}$$

the last inequality follows from $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$. Again, this contradicts condition (e_i); thus $k_{i+1} \geq 2k_i$ as claimed.

Condition (a_{i+1}) then follows from condition (a_i), and condition (d_{i+1}) follows from the definition of k_{i+1} and the fact that $k_{i+1}/2 \geq k_i$. It remains to prove condition (e_{i+1}). For $k_i \leq k < k_{i+1}/2$, we verify condition (e_{i+1}) as follows:

$$\begin{aligned} |x_{m_{i+1}} - x_{m_{i+1}+k}| &\geq 2\varepsilon_k(n, C, \delta, \rho) \\ &\geq \varepsilon_k(n, C, \delta, \rho) + M_{1+\rho}^{6k}\beta_{2k}(n, C, \delta, \rho) \\ &> \varepsilon_k(n, C, \delta, \rho) + 2M_{1+\rho}^{3k_{i+1}}\beta_{k_{i+1}}(n, C, \delta, \rho). \end{aligned} \tag{5.53}$$

For $0 < k < k_i$, we use the estimates

$$|x_{m_{i+1}} - x_{m_i}| \leq |x_{m_{i+1}} - x_i^*| + |x_{m_i} - x_i^*| \leq 2M_{1+\rho}^{k_i}\beta_{k_i}(n, C, \delta, \rho), \tag{5.54}$$

and consequently

$$\begin{aligned} |x_{m_{i+1}} - x_{m_{i+1}+k}| &\geq |x_{m_i} - x_{m_i+k}| - |x_{m_{i+1}} - x_{m_i}| - |x_{m_i+k} - x_{m_{i+1}+k}| \\ &\geq |x_{m_i} - x_{m_i+k}| - 4M_{1+\rho}^{k+k_i}\beta_{k_i}(n, C, \delta, \rho) \\ &\geq |x_{m_i} - x_{m_i+k}| - M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho); \end{aligned} \tag{5.55}$$

the last inequality follows from $M_{1+\rho} \geq 2$. If $0 < k < k_i/2$, then by condition (e_i) we have

$$\begin{aligned} &|x_{m_{i+1}} - x_{m_{i+1}+k}| \\ &\quad \geq \varepsilon_k(n, C, \delta, \rho) + 2M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho) - M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho) \\ &\quad > \varepsilon_k(n, C, \delta, \rho) + 2M_{1+\rho}^{3k_{i+1}}\beta_{k_{i+1}}(n, C, \delta, \rho), \end{aligned} \tag{5.56}$$

as desired. If on the other hand $k_i/2 \leq k < k_i$, then by condition (d_i) we have

$$\begin{aligned} &|x_{m_{i+1}} - x_{m_{i+1}+k}| \\ &\quad \geq 2\varepsilon_k(n, C, \delta, \rho) - M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho) \\ &\quad \geq \varepsilon_k(n, C, \delta, \rho) + M_{1+\rho}^{6k}\beta_{2k}(n, C, \delta, \rho) - M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho) \\ &\quad > \varepsilon_k(n, C, \delta, \rho) + M_{1+\rho}^{3k_i}\beta_{k_i}(n, C, \delta, \rho) \\ &\quad > \varepsilon_k(n, C, \delta, \rho) + 2M_{1+\rho}^{3k_{i+1}}\beta_{k_{i+1}}(n, C, \delta, \rho), \end{aligned} \tag{5.57}$$

as desired. This completes ths inductive step in Case A).

Case B): For all $0 \leq p < q_i$ we have $|x_{m_i+pk_i} - x_i^*| \leq \beta_{k_i}(n, C, \delta, \rho)$. In this case we will complete the proof without continuing the induction, so to simpify notation let $k = k_i$, $m = m_i$, and $q = q_i$. Then by condition (a$_i$) we have $0 < k < n$, and by condition (b$_i$) we have $0 \le m < n \log_2 n$.

We claim then that x_m is $(n, \gamma_{n/2}(C, \delta)/2)$-hyperbolic. To prove this, we divide into two subcases of k being relatively small compare to n (see case B′)) and the other opposite (case B”) resp.). Then we will complete the proof by showing that x_0 is $(n, \gamma_n(C, \delta))$-hyperbolic.

Case B′): $2\beta_k(n, C, \delta, \rho) \le M_{1+\rho}^{n(\log_2 n)} \gamma_n^{1/\rho}(C, \delta)$. Let $k' = \gcd(k, n)$. Then by the Euclidean Algorithm (Lemma 5.1.6), since

$$|x_{m+k} - x_m| \le 2\beta_k(n, C, \delta, \rho) \le M_{1+\rho}^{n(\log_2 n)} \gamma_n^{1/\rho}(C, \delta) \tag{5.58}$$

and

$$|x_{m+n} - x_m| \le M_{1+\rho}^{n(\log_2 n)} |x_n - x_0| \le M_{1+\rho}^{n(\log_2 n)} \gamma_n^{1/\rho}(C, \delta), \tag{5.59}$$

we have $|x_{m+k'} - x_m| \le M_{1+\rho}^{n(\log_2 n+2)} \gamma_n^{1/\rho}(C, \delta)$. Also, $k'|n$, and in particular $k' \le n/2$.

Since $C > 100\rho^{-1}\delta^{-1} \log M_{1+\rho}$,

$$|x_{m+k'} - x_m| \le M_{1+\rho}^{n(\log_2 n+2)} \gamma_n^{1/\rho}(C, \delta) \le \gamma_{n/2}^{1/\rho}(C, \delta) \le \gamma_{k'}^{1/\rho}(C, \delta). \tag{5.60}$$

Thus x_m is $(k', \gamma_{k'}^{1/\rho}(C, \delta))$-periodic, and then by the Inductive Hypothesis, x_m is $(k', \gamma_{k'}(C, \delta))$-hyperbolic.

Let $q' = n/k'$. Using Lemma A.4 (Appendix A, Part I) with $A = df^{k'}(x^m)$ and $A_p = df^{k'}(x_{m+(p-1)k'})$,

$$\begin{aligned}
\gamma(df^n(x_m)) &= \gamma(A_{q'} A_{q'-1} \cdots A_1) \\
&\ge \gamma(A) - \sum_{p=1}^{q'} \|A - A_p\| \\
&\ge \gamma_{k'}(C, \delta) - \sum_{p=0}^{q'-1} M_{1+\rho}^{k'(1+\rho)} |x_{m+pk'} - x_m|^\rho \\
&\ge \gamma_{n/2}(C, \delta) - \sum_{p=0}^{q'-1} M_{1+\rho}^{k'(1+\rho)+p} |x_{m+k'} - x_m|^\rho \\
&\ge \gamma_{n/2}(C, \delta) - M_{1+\rho}^{k'(1+\rho)+q'} M_{1+\rho}^{\rho n(\log_2 n+2)} \gamma_n(C, \delta).
\end{aligned} \tag{5.61}$$

Again since $C > 100\rho^{-1}\delta^{-1} \log M_{1+\rho}$, we have $\gamma(df^n(x_m)) \ge \gamma_{n/2}(C, \delta)/2$, and hence x_m is $(n, \gamma_{n/2}(C, \delta)/2)$-hyperbolic as claimed.

Case B"): $2\beta_k(n, C, \delta, \rho) > M_{1+\rho}^{n(\log_2 n)} \gamma_n^{1/\rho}(C, \delta)$. In this case the term $M_{1+\rho}^{q'}$ in (5.61) is troublesome, because we can only bound $|x_{m+k'} - x_m|$ in terms of $\beta_k(n, C, \delta, \rho)$, and q' could be much larger than k. Thus we cannot ensure that the negative term in (5.61) is small. This is not only why Case B") is more complicated than Case B′), but it is why we needed to set up the whole induction on i; Case B′) could have started simply with the hypothesis that we have a (k, n, C, δ, ρ)-leading saddle, but works only when k is sufficiently large.

First we claim that k must be a factor of n. If not, then we can write $n = qk - r$ or $n = (q-1)k + r$ with $1 \leq r \leq k/2$. In the former case, $qk = n + r$, and

$$
\begin{aligned}
|x_m - x_{m+r}| &\leq |x_m - x_{m+qk}| + |x_{m+n+r} - x_{m+r}| \\
&\leq |x_m - x^*| + |x_{m+qk} - x^*| + M_{1+\rho}^{m+r} |x_n - x_0| \\
&\leq M_{1+\rho}^{k} (2\beta_k(n, C, \delta, \rho) + M_{1+\rho}^{n \log_2 n} \gamma_n^{1/\rho}(C, \delta)).
\end{aligned}
\tag{5.62}
$$

In the latter case,

$$
\begin{aligned}
|x_m - x_{m+r}| &\leq |x_m - x_{m+n}| + |x_{m+n} - x_{m+r}| \\
&\leq M_{1+\rho}^{m} \gamma_n^{1/\rho}(C, \delta) + M_{1+\rho}^{r} |x_{(q-1)k} - x_0| \\
&\leq M_{1+\rho}^{k} (M_{1+\rho}^{n \log_2 n} \gamma_n^{1/\rho}(C, \delta) + 2\beta_k(n, C, \delta, \rho)).
\end{aligned}
\tag{5.63}
$$

Both bounds are the same, and are bounded above by $4M_{1+\rho}^{k} \beta_k(n, C, \delta, \rho)$ by the hypothesis of Case B"). Therefore since $M_{1+\rho} \geq 2$,

$$
\begin{aligned}
|x_m - x_{m+r}| &\leq 4M_{1+\rho}^{k} \beta_k(n, C, \delta, \rho) \\
&\leq M_{1+\rho}^{3k} \beta_k(n, C, \delta, \rho) \\
&\leq M_{1+\rho}^{6r} \beta_{2r}(n, C, \delta, \rho) \\
&\leq \varepsilon_r(n, C, \delta, \rho)
\end{aligned}
\tag{5.64}
$$

by the inequality at the beginning of the proof. Recalling that $k = k_i$ and $m = m_i$, this contradicts condition (e_i), or condition (d_i) if $r = k/2$. Thus k is a factor of n as claimed.

Assume that x_m is not $(n, \gamma_{n/2}(C, \delta)/2)$-hyperbolic. Now by the Inductive Hypothesis, x^* is $(k, \gamma_k(C, \delta))$-hyperbolic. Then by Lemma A.4 (Appendix A, Part I) with $A = df^k(x^*)$ and $A_p = df^k(x_{m+(p-1)k})$,

$$
\begin{aligned}
\sum_{p=1}^{q} \|A - A_p\| &\geq \gamma(A) - \gamma(A_q A_{q-1} \cdots A_1) \\
&\geq \gamma_k(C, \delta) - \gamma_{n/2}(C, \delta)/2 \qquad \geq \gamma_k(C, \delta)/2.
\end{aligned}
\tag{5.65}
$$

It follows that

$$\sum_{p=0}^{q-1} M_{1+\rho}^{k(1+\rho)} |x_{m+pk} - x^*|^\rho \geq \gamma_k(C,\delta)/2. \tag{5.66}$$

The contribution to this sum from values of p with $2^{-t/\rho}\beta_k(n,C,\delta,\rho) < |x_{m+pk} - x^*| \leq 2^{(1-t)/\rho}\beta_k(n,C,\delta,\rho)$ must then be at least $2^{-t}\gamma_k(C,\delta)$ for some positive integer t; fix this value of t. Since each such point contributes at most $M_{1+\rho}^{k(1+\rho)} 2^{1-t}\beta_k^\rho(n,C,\delta\rho)$ to the sum, the number of such points must be at least

$$2^{-1} M_{1+\rho}^{-k(1+\rho)} \beta_k^{-\rho}(n,C,\delta\rho)\gamma_k(C,\delta) = 2^{-2\rho-1} M_{1+\rho}^{-k(1+\rho)} n\gamma_k^{1+\rho-4N}(C,\delta). \tag{5.67}$$

Lemma 5.2.5. *Given $P > 2^N$ points in a ball of radius R in N-dimensional space, some pair of points must be within $4RP^{-1/N}$ of each other.*

Proof: Consider the balls of radius $2RP^{-1/N}$ centered at each of the P points. Their total volume is equal to the volume of a ball of radius $2R$, and they all are contained in a ball of radius $R + 2RP^{-1/N} < 2R$. Therefore some pair of these balls must intersect, and their centers are withing $4RP^{-1/N}$ of each other as claimed. Q.E.D.

Let $P = 2^{-2\rho-1} M_{1+\rho}^{-k(1+\rho)} n\gamma_k^{1+\rho-4N}(C,\delta)$ and $R = 2^{(1-t)/\rho}\beta_k(n,C,\delta,\rho)$. Then by Lemma 5.2.5, there are two points x_{m+pk} and $x_{m+(p+r)k}$ with $0 \leq p < p+r < q \leq n$ such that $|x_{m+pk} - x^*|, |x_{m+(p+r)k} - x^*| \in (R/2, R]$ and $|x_{m+pk} - x_{m+(p+r)k}| \leq 4RP^{-1/N}$. Our goal is to show that this contradicts the $(k, \gamma_k(C,\delta))$-hyperbolicity of x^*. The argument is similar the proof of Lemma A.4 in Appendix A, Part I.

Let $v_s = x_{m+(p+s)k} - x^*$ and $\omega_s = \exp(-2\pi i s/r)$ for $s = 0, 1, \ldots, r$. Let

$$u_s = v_0 + \omega_s v_1 + \omega_s^2 v_2 + \cdots + \omega_s^{r-1} v_{r-1}. \tag{5.68}$$

Choose ℓ such that $|v_\ell| = \max(|v_0|, |v_1|, \ldots, |v_{r-1}|)$. Since

$$\sum_{s=0}^{r-1} \omega_s^{-\ell} u_s = r v_\ell, \tag{5.69}$$

there is an s for which $|u_s| \geq |v_\ell| > R/2$. Then

$$\begin{aligned}
\gamma_k(C,\delta) &\leq \gamma(df^k(x^*)) \\
&\leq |df^k(x^*)u_s - \omega_s^{r-1}u_s|/|u_s| \\
&= \frac{|df^k(x^*)v_0 - v_1 + df^x(x^*)\omega_s v_1 - \omega_s v_2 + \cdots + df^k(x^*)\omega_s^{r-1}v_{r-1} - \omega_s^{r-1}v_0|}{|u_s|} \\
&\leq \frac{M_{1+\rho}^{k(1+\rho)}(|v_0|^{1+\rho} + |v_1|^{1+\rho} + \cdots + |v_{r-1}|^{1+\rho}) + |v_r - v_0|}{|u_s|} \\
&\leq \frac{M_{1+\rho}^{k(1+\rho)}(|v_0|^{1+\rho} + |v_1|^{1+\rho} + \cdots + |v_{r-1}|^{1+\rho}) + 4RP^{-1/N}}{|u_s|} \\
&\leq M_{1+\rho}^{k(1+\rho)}(t-s)\beta_k^\rho(n,C,\delta,\rho) + 8P^{-1/N} \\
&\leq M_{1+\rho}^{k(1+\rho)} n(4n^{-1/\rho}\gamma_k^{4N/\rho-1}(n,C,\delta,\rho))^\rho \\
&\quad + 8(2^{-2\rho-1}M_{1+\rho}^{-k(1+\rho)} n\gamma_k(C,\delta)^{1+\rho-4N})^{-1/N} \\
&\leq M_{1+\rho}^{k(1+\rho)} 2^{2\rho}\gamma_k^{4N-\rho}(C,\delta) + 64M_{1+\rho}^{k(1+\rho)/N}\gamma_k^2(C,\delta).
\end{aligned} \tag{5.70}$$

Since $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$, this is a contradiction, and therefore x_m is indeed $(n, \gamma_{n/2}(C,\delta)/2)$-hyperbolic.

Finally, to show that x_0 is $(n, \gamma_n(C,\delta))$-hyperbolic, let $m = sn + r$ where $0 \leq s < \log_2 n$ and $0 \leq r < n$. Since $|x_0 - x_n| \leq \gamma_n^{1/\rho}(C,\delta)$ and $M_{1+\rho} \geq 2$, we have

$$\begin{aligned}
|x_0 - x_s n| &\leq |x_0 - x_n| + |x_n - x_{2n}| + \cdots + |x_{(s-1)n} - x_{sn}| \\
&\leq (1 + M_{1+\rho}^n + \cdots + M_{1+\rho}^{(s-1)n})\gamma_n^{1/\rho}(C,\delta) \\
&\leq M_{1+\rho}^{sn}\gamma_n^{1/\rho}(C,\delta) \\
&\leq M_{1+\rho}^{n\log_2 n}\gamma_n^{1/\rho}(C,\delta).
\end{aligned} \tag{5.71}$$

Thus $|x_m - x_r| \leq M_{1+\rho}^{r+n\log_2 n}\gamma_n^{1/\rho}(C,\delta) \leq M_{1+\rho}^{n(\log_2 n+1)}$. Then by Lemma 5.2.2,

$$\begin{aligned}
|df^{n-r}(x_m) - df^{n-r}(x_r)| &\leq M_{1+\rho}^{(n-r)(1+\rho)+\rho n(\log_2 n+1)}\gamma_n(C,\delta) \\
&\leq M_{1+\rho}^{n(\log_2 n+3)}\gamma_n(C,\delta).
\end{aligned} \tag{5.72}$$

Similarly, $|x_{m+n-r} - x_0| \leq M_{1+\rho}^{n(\log_2 n+1)}$, and by Lemma 5.2.2,

$$\begin{aligned}
|df^r(x_{m+n-r}) - df^r(x_0)| &\leq M_{1+\rho}^{r(1+\rho)+\rho n(\log_2 n+1)}\gamma_n(C,\delta) \\
&\leq M_{1+\rho}^{n(\log_2 n+3)}\gamma_n(C,\delta).
\end{aligned} \tag{5.73}$$

Let $A = df^{n-r}(x_m)$ and $B = df^r(x_{m+n-r})$, and notice that $df^n(x_m) = BA$. For all $v \in \mathbb{C}^n$ and all $\phi \in \mathbb{R}$,

$$\begin{aligned}
&|df^n(x_0)v - e^{i\phi}v| \\
&\qquad = |df^{n-r}(x_r)df^r(x_0)v - e^{i\phi}v| \\
&\qquad \geq |Adf^r(x_0)v - e^{i\phi}v| - |df^{n-r}(x_r) - A||df^r(x_0)||v| \\
&\qquad \geq |ABv - e^{i\phi}v| - |A||df^r(x_0) - B||v| - M_{1+\rho}^{n(\log_2 n+4)}\gamma_n(C,\delta)|v| \\
&\qquad \geq |B|^{-1}|BABv - e^{i\phi}Bv| - 2M_{1+\rho}^{n(\log_2 n+4)}\gamma_n(C,\delta)|v| \\
&\qquad \geq M_{1+\rho}^{-n}\gamma_{n/2}(C,\delta)|Bv|/2 - 2M_{1+\rho}^{n(\log_2 n+4)}\gamma_n(C,\delta)|v|.
\end{aligned} \tag{5.74}$$

If $|Bv| \geq M_{1+\rho}^{-n}|v|/2$, then

$$|df^n(x_0)v - e^{i\phi}v| \geq (M_{1+\rho}^{-2n}\gamma_{n/2}(C,\delta)/4 - 2M_{1+\rho}^{n(\log_2 n+4)}\gamma_n(C,\delta))|v| > \gamma_n(C,\delta)|v| \tag{5.75}$$

because $C > 100\rho^{-1}\delta^{-1}\log M_{1+\rho}$. If, on the other hand, $|Bv| < M_{1+\rho}^{-n}|v|/2$, then

$$\begin{aligned}
|df^n(x_0)v| &= |df^{n-r}(x_r)df^r(x_0)v| \\
&\leq |df^{n-r}(x_r)|(|Bv| + |df^r(x_0) - B||v|) \\
&\leq M_{1+\rho}^n(M_{1+\rho}^{-n}|v|/2 + M_{1+\rho}^{n(\log_2 n+3)}\gamma_n(C,\delta)|v|) \\
&\leq (1/2 + M_{1+\rho}^{n(\log_2 n+4)}\gamma_n(C,\delta))|v| \\
&\leq (3/4)|v|,
\end{aligned} \tag{5.76}$$

and hence $|df^n(x_0)v - e^{i\phi}v| \geq (1/4)|v| > \gamma_n(C,\delta)|v|$.

Since v is arbitrary, x_0 is $(n, \gamma_n(C,\delta))$-hyperbolic, and the proof of Lemma 5.2.1 is complete. Q.E.D.

5.3 A Proof of Theorem 2.4.6 about a weak gap

In this section we start investigation of the right branch of Fig. 5.1 from the beginning of this Chapter. Namely, we consider almost an periodic trajectories of length n of diffeomorphisms satisfying the Inductive Hypothesis of order $n-1$ with no leading saddles. The goal of this Section is to prove Theorem 2.4.6 about a weak gap. This Theorem says that if almost an periodic trajectory of described above type is not simple, then it has a very close return (a weak gap). In the next Section we need to sharpen this result to show that a nonsimple trajectory must have not only a weak gap, but also a (D, n, r_{2k})-gap, defined in Chapter 3.5.3.

Theorem 5.3.1. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism for some $0 < \rho \le 1$ with $M_{1+\rho}$ bound on the $C^{1+\rho}$-norm of f, defined as in Theorem 2.4.6 Let $\delta > 0$ and $C > 30\rho^{-1} \log M_{1+\rho}$ be positive constants (of Induction), let n be a positive integer (period), and let f satisfy the Inductive Hypothesis of order $(n-1)$ with constants $C, \delta,$ and ρ, i.e. $f \in IH(n-1, C, \delta, \rho)$. Suppose f has a trajectory $\{x_s = f^s(x_0)\}_{s=0}^n$ such that x_0 is $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic, has no (n, C, δ, ρ)-leading saddles, i.e. $|x_j - x_0| \ge n^{-1/\rho}\gamma_k^{4N/\rho}(C,\delta)$ for any $1 \le k < n$. Then for $D \ge M_{1+\rho}^{30N/\rho}$ if x_0 has no weak (D,n)-gap, then x_0 is $(n, \gamma_n(C,\delta))$-simple.*

The proof consists of several Lemmas.

Lemma 5.3.2. *With conditions and notations of Theorem 5.3.1 if k satisfies $k^{1+\delta} \ge \frac{\rho \log D}{2C} n$. Then for $\lambda = \frac{\log n}{\log k} - 1, \beta = \frac{1+\lambda}{1+\delta}$, and each $\left(\frac{4 \log n}{\rho \log D}\right) \le s < k$ we have*

$$\#\{1 \le j < k : |x_j - x_0| \le D^{-Ns^{1+\lambda}}\} \le \left(\frac{16NC}{\rho \log D}\right)^{1/(1+\delta)} \frac{k}{s^\beta}. \tag{5.77}$$

If $k^{1+\delta} < \frac{\rho \log D}{2C} n$, then x_0 is $(n, \gamma_k(C,\delta)/2)$-hyperbolic.

Proof of the Lemma: We start with a proof of the second statement, namely, let $k^{1+\delta} < \frac{\rho \log D}{2C} n$. It is easy to check that $D^{-n} \le \gamma_k^{2/\rho}(C,\delta) = \exp(-\frac{2Ck^{1+\delta}}{\rho})$. If x_0 has a weak (D,n)-gap, then $|x_0 - x_k| \le D^{-n} \le \gamma_k^{1/\rho}(C,\delta)$. By Lemmas 2.4.5 of Part I and 5.1.6 we have that n is divisible by k. Put $p = n/k \in \mathbb{Z}_+$. The Inductive Hypothesis of order $(n-1)$ $(> k)$ with constants $C, \delta,$ and ρ imply that x_0 is $(k, \gamma_k(C,\delta))$-hyperbolic. To show that x_0 is actually $(n, \gamma_k(C,\delta)/2)$-hyperbolic we approximate $df^n(x_0)$ with $(df^k(x_0))^p$. Arguments as in the proof of Lemma 5.2.2 show that

$$\max_{1 \le s \le p} |x_{sk} - x_0| \le M_1^n D^{-n} \le M_1^n \gamma_k^{2/\rho}(C,\delta). \tag{5.78}$$

Application of Lemma 5.2.2 gives that $|df^k(x_0) - df^k(x_{sk})| \le M_{1+\rho}^{k(1+\rho)+n\rho}\gamma_k^2(C,\delta)$ for any $1 \le s \le k$. Using $D \ge M_{1+\rho}^{30/\rho}$ and $M_{1+\rho} \ge 2^{1/\rho}$ we get that $M_{1+\rho}^{k(1+\rho)+n\rho}\gamma_k^2(C,\delta) \le M_{1+\rho}^{-2n}\gamma_k(C,\delta) \le 2^{-n}M_{1+\rho}^{-n}\gamma_k(C,\delta)$. Again following arguments of the proof of Lemma 5.2.2 we get $|df^n(x_0) - (df^k(x_0))^p| \le \gamma_k(C,\delta)/2$ (cf. arguments for (5.61)). Finally, application of Lemmas A.2 and A.3 in Appendix A, Part I gives that $df^n(x_0)$ is $gm_k(C,\delta)/2$-hyperbolic, which is the requires statement.

Recall that with our notations $\lambda \ge 0$ and such that $k^{1+\lambda} = n$. To prove the first part of the Lemma we introduce the sequence of positive numbers $\{r_s = D^{-s^{1+\lambda}}\}_{s=1}^k$

and consider the sequence of concentric balls $\{B_s(x_0)\}_{s=0}^k$ around x_0 of radii $\{r_s\}_{s=0}^k$ respectively. Let

$$A_s = \{1 \le j < k : |x_j - x_0| < r_s\} \tag{5.79}$$

be the collection of indices of points of the trajectory $\{x_j\}_{j=1}^{k-1}$ which visit the r_s-ball $B_s(x_0)$ around x_0. We prove estimate (5.77) by contradiction. Suppose

$$\#\{A_s\} > \left(\frac{16CN}{\rho \log D}\right)^{1/(1+\delta)} \frac{k}{s^\beta}. \tag{5.80}$$

Then by the pigeonhole principle there are two returns to the r_s-ball $B_s(x_0)$, say x_i and x_j, such that $p = j - i < \left(\frac{16CN}{\rho \log D}\right)^{1/(1+\delta)} s^\beta$. The last inequality can be rewritten as $s^{1+\lambda} \log D/2 \ge 8CNp^{1+\delta}/\rho$. Recall that we consider only $s \ge (4 \log n/\log D)^{1/(1+\lambda)}$, therefore, $s^{1+\lambda} \log D/2 \ge 2 \log n/\rho$. This gives $s^{1+\lambda} \log D \ge (8CNp^{1+\delta} + 2\log n)/\rho$. This inequality is equivalent to $r_s \le n^{-2/\rho} \gamma_p^{8N/\rho}(C, \delta)$. Thus, combining these inequalities with the pigeonhole principle we get

$$|x_i - x_j| < 2r_s \le 2n^{-2/\rho} \gamma_p^{8N/\rho}(C, \delta). \tag{5.81}$$

Application of Lemma 5.2.3 with $\alpha = 1 + 3/\rho$, $\gamma^\alpha = 2n^{-2/\rho}\gamma_p^{8N/\rho}(C, \delta)$, $k = p$, and x_0 replaced by x_i show that there is a periodic point $x_0^* = f^p(x_0^*)$ of period p such that

$$|x_i - x_0^*| < 4n^{-6/(3\rho+\rho^2)} \gamma_p^{24N/(3\rho+\rho^2)}(C, \delta) \le 4n^{-1/\rho}\gamma_p^{6N/\rho}(C, \delta). \tag{5.82}$$

The last inequality uses that $0 < \rho \le 1$. By triangle inequality his implies that

$$|x_0^* - x_0| \le |x_0 - x_i| + |x_i - x_0^*| \le 2r_s + 4n^{-1/\rho}\gamma_p^{6N/\rho}(C, \delta) \le n^{-1/\rho}\gamma_p^{5N/\rho}(C, \delta).$$

Notice that $|x_p - x_0^*| = |f^p(x_0) - f^p(x_0^*)| \le M_1^p |x_0 - x_0^*|$. By triangle inequality his implies that

$$|x_p - x_0| \le (M_1^p + 1) n^{-1/\rho}\gamma_p^{5N/\rho}(C, \delta) \le n^{-1/\rho}\gamma_p^{4N/\rho}(C, \delta) \tag{5.83}$$

and contradicts to the fact that x_0 has no (n, C, δ, ρ)-leading saddles. This completes the proof of Lemma 5.3.2. Q.E.D.

Lemma 5.3.3. *With conditions and notations of Theorem 5.3.1 suppose that a point $x_0 \in B^N$ is $(n, \gamma_n^{1/\rho})(C, \delta)$-periodic, with no (C, δ, ρ)-leading saddles, no weak (D, n)-gaps before k and a weak (D, n)-gap at x_k, where D and k satisfy conditions of Lemma*

5.3.2. Then for the product of distances along the trajectory $\{x_j\}_{j=0}^{k-1}$ we have

$$\begin{aligned}\prod_{j=1}^{k-1} |x_j - x_0| \ge \exp\left(-Ckn^{\frac{\delta}{1+\delta}} \frac{12\lambda(1+\delta)}{N\rho\delta(1+\lambda)}\right) \text{ for } k < n \\ \text{or} \quad \prod_{j=1}^{n-1} |x_j - x_0| \ge \exp(-Cn^{1+\frac{\delta}{1+\delta}} \frac{6}{N\rho}).\end{aligned} \tag{5.84}$$

Corollary 5.3.4. *With conditions and notations of Lemma 5.3.3 the second estimate of (5.84) implies that for a sufficiently large n the point x_0 is $(n, \gamma_n^{1/N}(C,\delta))$-simple, i.e. $\prod_{j=1}^{n-1} |x_j - x_0| \ge \exp(-Cn^{1+\frac{\delta}{1+\delta}}/(4N))$.*

Proof: By Corollary 5.1.5 we have $\min_{1\le j<n} |x_j - x_0| \ge D^{-Nn} M_{1+\rho}^{-30Nn/\rho}$. Since x_0 is $(n, \gamma_n^{1/\rho}(C,\delta))$-periodic, we have $|f(x_{n-1}) - x_0| \le \gamma_n^{1/\rho}(C,\delta))$. By the triangle inequality

$$||x_{n-1} - x_{j-1}| - |x_{n-1} - x_{j-1}|| \le |f^{-1}(x_0) - x_{n-1}|. \tag{5.85}$$

By distortion estimate we have $|f^{-1}(x_0) - x_{j-1}| < M_1\gamma_n^{1/\rho}(C,\delta)$ and $|x_{n-1} - x_{j-1}| \ge M_1^{-1} D^{-Nn} M_{1+\rho}^{-30Nn/\rho}$. So $|x_{n-1} - x_{j-1}| \gg |f^{-1}(x_0) - x_{j-1}|$. Therefore,

$$\begin{aligned}\prod_{j=1}^{n-2} |x_{n-1} - x_j| &\ge (2M_{1+\rho})^{-n} \exp\left(-Cn^{1+\frac{\delta}{1+\delta}} \frac{6}{N\rho}\right) \\ &\ge \exp(-\frac{Cn^{1+\delta}}{4N}).\end{aligned} \tag{5.86}$$

This completes the proof of the Corollary. Q.E.D.

Proof of Lemma 5.3.3: Recall that $\lambda \ge 0$ and $k^{1+\lambda} = n$, $\{r_s = D^{-2Ns^{1+\lambda}}\}_{s=1}^{k}$ is the sequence of radii of balls $\{B_s(x_0)\}_{s=1}^{k}$ around x_0, $A_s = \{1 \le j < k : |x_j - x_0| < 2r_s\}$ is a collection of returns to $B_s(x_0)$ of the trajectory $\{x_j\}_{j=0}^{k-1}$, and $m_s = \#\{A_s\}$ is the number of those returns.

By Corollary 5.1.5 $A_k = \emptyset$ and $m_k = 0$. Rewrite the product of distance as follows

$$\prod_{j=1}^{k-1} |x_j - x_0| = \prod_{j\notin A_1} |x_j - x_0| \prod_{j\in A_1\setminus A_2} |x_j - x_0| \cdots \prod_{j\in A_{k-1}} |x_j - x_0|. \tag{5.87}$$

By definition of A_s for each $j \in A_{s-1} \setminus A_s$ we have $|x_0 - x_j| \geq r_s$. Put $a_s = 2Ns^{1+\lambda}$. Then the product (5.87) admits the following lower bound

$$\begin{aligned} & r_1^{k-1-m_1} r_2^{m_1-m_2} \dots r_k^{m_{k-1}-m_k} = \\ & \exp\Big(-\log D\,[a_0(k-1-m_1)+ \\ & a_1(m_1-m_2)+\cdots+a_k(m_{k-1}-m_k)]\Big) \end{aligned} \tag{5.88}$$

Using Abel's resummation we can rewrite the last expression in the form

$$\exp\left(-\log D\,[m_1(a_1-a_0)+m_2(a_2-a_1)+\cdots+m_k(a_k-a_{k-1})]\right). \tag{5.89}$$

Since for $\lambda > 0$, we have $a_{s+1} - a_s \leq 2N\lambda s^{\lambda}$. Recall that $k^{1+\lambda} = n$. Using inequality (5.77) from Lemma 5.3.2 we get the following lower bound for the product (5.89)

$$\exp\left(\frac{6C}{N\rho} 2N\lambda k \sum_{s=1}^{k} s^{\lambda-\beta}\right) \leq \exp\left(\frac{12C\lambda}{\rho} k n^{\frac{\delta}{1+\delta}}\right). \tag{5.90}$$

This is a required estimate for $\lambda > 0$. In the case $k = n$ or $\lambda = 0$ we have $a_{s+1} - a_s = 2N$ and can get the required in (5.84) estimate. This completes the proof of Lemma 5.3.3. Q.E.D.

This completes the proof of Theorem 5.3.1. Q.E.D.

5.4 The last piece to the proof of the Shift Theorem

In this Section we complete a proof of the Shift Theorem by completing investigation of the right branch of diagram 3. In the previous Section we prove a weak form of the Shift Theorem which says that if a diffeomorphism f satisfying the Inductive Hypothesis of order $n-1$ and has almost an periodic point x_0 of period n with no leading saddles (x_0 is not a satellite of a periodic point of a lower period), then either x_0 is simple (nonrecurrent) or it has a weak gap (a close return) at x_k and $k < n$. However, for the proof of the main result (Theorem 1.3.7) we need to find for a nonsimple (recurrent) point an even closer return (a (D, n, r_{2k})-gap), defined in (3.5.3). This is the goal of this Section. Let's give a precise statement.

Theorem 5.4.1. *With conditions and notations of Theorem 5.3.1 consider a sequence of numbers* $\mathbf{r} = \{r_k\}_{k=0}^{\infty}$*, which defines a Hiblert Brick of an admissible size (see definition 1.3.1, then if an* $(n, \gamma_n^{1/\rho}(C,\delta))$*-periodic point* x_0 *has no* (D, n, r_{2k})*-gaps, then* x_0 *is* $(n, \gamma_n(C,\delta))$*-simple, i.e.* $\prod_{j=1}^{n-1} |x_j - x_0| \geq \exp(-Cn^{1+\delta}/(4N))$.

The proof, given below, consists of several Lemmas.

Lemma 5.4.2. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $\rho \in (0,1]$, $M_{1+\rho} = \max\{\|f\|_{C^{1+\rho}}, \|f^{-1}\|_{C^1}, 2^{1/\rho}\}$, and $D \geq M_{1+\rho}^{30/\rho}$. Suppose that for some positive integer n there is a point $x_0 \in B^N$ whose trajectory has weak (D,n)-gaps at $x_{n_1} = f^{n_1}(x_0), x_{n_2} = f^{n_2}(x_0), \dots, x_{n_s} = f^{n_s}(x_0)$ for numbers $n_1 < n_2 < \dots < n_s$. Then numbers $n_1, n_2, \dots, n_s$ are multiplicatively ordered, i.e. for each $i = 1, \dots, s-1$ ratio $n_{j+1}/n_j = p_j$ is integer.*

Proof By definition of a weak (D,n)-gap we know that

$$\begin{aligned} |x_{n_i} - x_0| \geq D^n |x_{n_{i+1}} - x_0| \quad \text{for each} \quad i = 1, \dots, s-1 \\ \min_{1 \leq l < n_1} |x_l - x_0| \geq D^n |x_{n_1} - x_0| = \lambda'. \end{aligned} \tag{5.91}$$

Consider the s-tuple of points $x_{n_1}, x_{n_2}, \dots, x_{n_s}$. Apply the Euclidean Algorithm (Lemma 5.1.6) to them with $K = D$ and $\lambda = \lambda'$. It gives that for $n^* = \gcd(n_1, \dots, n_s)$ $\leq n_1$ by the choice of λ' we have

$$|x_{n^*} - x_0| \leq M_{1+\rho}^{2n} D^{-n} \lambda' = M_{1+\rho}^{2n} |x_{n_1} - x_0|. \tag{5.92}$$

If $n^* \neq n_1$, then the second part of (5.91) contradicts (5.92). So $n^* = n_1$ and n_1 divides each n_i with $2 \leq i \leq s$.

Similar arguments show that n_2 divides each n_i with $2 \leq i \leq s$ and so on. This completes the proof of Lemma 5.4.2. Q.E.D.

Remark 5.4.3. *Notice that each weak (D,n)-gap point has to be a T^{-n}-ladder moment for $D \geq T$. More precisely, let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $\rho \in (0,1]$, let $M_{1+\rho} = \max\{\|f\|_{C^{1+\rho}}, 2^{1/\rho}\}$, and $D \geq T \geq M_{1+\rho}^{30/\rho}$ be positive numbers and, let n be some positive integer. Suppose that for some point $x_0 \in B^N$ its trajectory has a weak (D,n)-gap at x_k with $k < n$, i.e.*

$$|x_k - x_0| \leq D^{-n} \min_{1 \leq j < k} |x_j - x_0|. \tag{5.93}$$

Then x_k is a T^{-n}-ladder moment independently of behaviour of the trajectory $\{x_j = f^j(x_0)\}_{j=1}^{k-1}$ before k.

This follows from the definitions.

Corollary 5.4.4. *With conditions and notations of Remark 5.4.3 a trajectory $\{x_j\}_{j=0}^{n-1}$ of length n of any point $x_0 \in B^N$ has at most N weak (D,n)-gaps.*

This follows from Theorem 5.1.4 about a number of possible T^{-n}-ladder moments of a trajectory of length n.

Lemma 5.4.5. *Let $f \in \mathrm{Diff}^{1+\rho}(B^N)$ be a $C^{1+\rho}$-smooth diffeomorphism with $\rho \in (0,1]$, $M_{1+\rho} = \max\{\|f\|_{C^{1+\rho}}, \|f^{-1}\|_{C^1}, 2^{1/\rho}\}$, and $D \geq M_{1+\rho}^{30/\rho}$. Suppose that for some positive integer n there is a (n, D^{-2Nn})-periodic point $x_0 \in B^N$. If x_0 has a weak (D,n)-gap at x_k, then n is divisible by k, i.e. $n/k = p \in \mathbb{Z}_+$ and*

$$\prod_{j=1}^{n-1} |x_j - x_0| \geq \frac{1}{2} \left(\prod_{j=1}^{k-1} |x_j - x_0| \right)^p \prod_{i=1}^{p-1} |x_{ki} - x_0|. \tag{5.94}$$

Proof Since x_0 has a weak (D,n)-gap at x_k we have

$$|x_k - x_0| \leq D^{-n} \min_{1 \leq j < k} |x_j - x_0| \tag{5.95}$$

Decompose the product of distances in the form

$$\prod_{j=1}^{n-1} |x_j - x_0| = \prod_{i=1}^{p} \left(|x_{ki} - x_0| \prod_{j=1}^{k-1} |x_{ki+j} - x_0| \right). \tag{5.96}$$

Notice that for each $i = 1, \dots, p-1$ we have $|x_{ki} - x_{k(i-1)}| \leq M_1^{k(i-1)} |x_k - x_0|$ which implies

$$|x_{ki} - x_0| \leq M_1^n |x_k - x_0| \leq M_1^n D^{-n} \min_{1 \leq j < k} |x_j - x_0|. \tag{5.97}$$

By triangle inequality and the last inequality

$$||x_{ki+j} - x_0| - |x_j - x_0|| \leq |x_{ki+j} - x_j| \leq M_1^n D^{-n} |x_j - x_0|. \tag{5.98}$$

Therefore, ratio $|x_{ki+j} - x_0| / |x_j - x_0| \in [1 - M_1^n D^{-n}, 1 + M_1^n D^{-n}]$ differs from 1 by an exponentially small in n amount. This implies the required inequality (5.94) and completes the proof of Lemma 5.4.5. Q.E.D.

This Lemma along with Lemma 5.3.3 shows that investigating the product of distances (5.94) ande showing simplicity it is sufficient to take care of exponentially in n close returns (multiples of k). Indeed,

$$\left(\prod_{j=1}^{k-1} |x_j - x_0| \right)^p \geq \exp \left(- \frac{C n^{1+\delta}}{N} \frac{12 \tau(k,n,\delta) n^{-\frac{\delta^2}{1+\delta}}}{\rho} \right), \tag{5.99}$$

where $\tau(k,n,\delta) = \frac{(\log n - \log k)(1+\delta)}{\delta \log n} < 1$ by Remark 2.4.1 and for a sufficiently large n the second fraction in the right-hand side can be made arbitrary small.

Let's start investigation of the product of distances for exponentially in n close returns from the worst case.

Lemma 5.4.6. *With conditions and notations of Theorem 5.4.1 and* $D \geq T \geq M_{1+\rho}^{30/\rho}$ *suppose that a point* $x_0 \in B^N$ *is* $(n, \gamma_n^{1/\rho}(C,\delta))$*-periodic and its trajectory* $\{x_j = f^j(x_0)\}_{j=0}^{n-1}$ *of length n has s different* T^{-n}*-ladder moments at* $x_{n_1}, \ldots, x_{n_s} = x_n$ *with* $n_1 < n_2 < \cdots < n_s = n$*. Moreover, these* T^{-n}*-ladder moments are also weak* (D,n)*-gaps, but not* (D, n, r_{2n_t})*-gaps for* $t = 1, \ldots, s$ *respectively. Then*

$$\prod_{j=1}^{n-1} |x_j - x_0| \geq \exp\left(-\frac{Cn^{1+\delta}}{30N}\right) \tag{5.100}$$

and by Corollary 5.3.4 x_0 *is* $(n, \gamma_n(C,\delta))$*-simple.*

Proof By Theorem 5.1.4 the number of T^{-n}-ladder moments $s \leq N$.

By Lemma 5.4.2 numbers $n_1, \ldots, n_s$ are multiplicatively ordered, i.e. $n_{j+1}/n_j = p_j \in \mathbb{Z}_+$. Denote also n/n_1 by p and for any $1 \leq i < j \leq s$ denote $p_i \ldots p_j$ by $P_{i,j}$.

By Lemma 5.4.5 we have

$$\prod_{j=1}^{n-1} |x_j - x_0| \geq \frac{1}{2}\left(\prod_{j=1}^{k-1} |x_j - x_0|\right)^p \prod_{i=1}^{p-1} |x_{in_1} - x_0|. \tag{5.101}$$

Let's show by contradiction that for any i such that $n_1 < in_i < n$ and in_1 is not divisible by n_2 we have

$$|x_{in_1} - x_0| \geq D^{-n}|x_{n_1} - x_0|. \tag{5.102}$$

Suppose that $|x_{in_1} - x_0| < D^{-n}|x_{n_1} - x_0|$. Since x_{n_2} is a weak (D,n)-gap we have $|x_{n_2} - x_0| < D^{-n}|x_{n_1} - x_0|$. Apply Euclidean Algorithm (Lemma 5.1.6) to x_{in_1} and x_{n_2}. It gives that for $n_1' = \gcd(in_1, n_2)$

$$|x_{n_1'} - x_0| \geq M_1^{2n} D^{-n}|x_{n_1} - x_0| \leq T^{-n}|x_{n_1} - x_0|. \tag{5.103}$$

Nondivisibility of in_1 by n_2 implies that $n_1' < n_2$. Therefore, $x_{n_1'}$ is either a T^{-n}-ladder moment after x_{n_1} and before x_{n_2} along the trajectory $\{x_j\}_{j=0}^{n-1}$ or there is another T^{-n}-ladder moment before $x_{n_1'}$ and after x_{n_1}. This is a contradiction.

Similar arguments show that for any $t = 1, \ldots, s-1$ and any i such that $n_t < in_t < n$ and in_t is not divisible by n_{t+1} we have

$$|x_{in_t} - x_0| \geq D^{-n}|x_{n_t} - x_0|. \tag{5.104}$$

Since for any $t = 1, \ldots, s$ the corresponding point x_{n_t} is a point of a weak (D, n)-gap, but not a (D, n, r_{2n_t})-gap we have

$$
\begin{aligned}
|x_{n_t} - x_0| &\geq \min \left\{ r_{2n_t}^{4(N+N^2)}, \left(\prod_{j=1}^{n_t-1} |x_j - x_0| \right)^{4(N+2)} \right\} \\
&\geq D^{-n} r_{2n_t}^{4(N+N^2)} \left(\prod_{j=1}^{n_t-1} |x_j - x_0| \right)^{4(N+2)} .
\end{aligned} \tag{5.105}
$$

The last inequality since both multipliers $D^{-n} \left(\prod_{j=1}^{n_t-1} |x_j - x_0| \right)^{\sigma}$ and $r_{2n_t}^{4(N+N^2)}$ clearly less then 1. Recall that by definition of an admissible sequence $\vec{\mathbf{r}} = \{r_k\}_{k=0}^{\infty}$ from Section 1.3 there is a integer $n_0 = n_0(\delta, N)$ such that $r_{2n}^{4(N+N^2)} \geq \gamma_n(C, \delta/2)$. Therefore, for a sufficiently large n (5.105) implies that

$$
|x_{n_t} - x_0| \geq D^{-n} \left(\prod_{j=1}^{n_t-1} |x_j - x_0| \right)^{4(N+2)} \gamma_n(C, \delta/2) \tag{5.106}
$$

Denote $\prod_{j=1}^{k-1} |x_j - x_0|$ by $\Delta_{k,f}(x_0)$. Now we shall find a lower bound on $\Delta_{n_t,f}(x_0)$ by induction in t.

Basis of induction $t = 1$. By Lemma 5.3.3

$$
\Delta_{n_1,f}(x_0) \geq \exp \left(-\frac{C n_1 n^{\frac{\delta}{1+\delta}}}{N} \quad \frac{12(\log n - \log n_1)(1+\delta)}{\delta \log n} \right). \tag{5.107}
$$

Denote the right-hand side by $\Delta_{n_1}(C, \delta, N)$.

Inductive step. Suppose we have a lower bound for $\Delta_{n_t,f}(x_0)$. Since $x_{n_{t+1}}$ is a weak (D, n)-gap, by Lemma 5.4.5 we have

$$
\Delta_{n_{t+1},f}(x_0) \geq \frac{1}{2} \Delta_{n_t,f}^{p_t}(x_0) \prod_{i=1}^{p_t-1} |x_{in_t} - x_0|. \tag{5.108}
$$

By inequality (5.104) this implies

$$
\prod_{i=1}^{p_t-1} |x_{in_t} - x_0| \geq |x_{n_t} - x_0| D^{-n(p_t-1)} \geq \left(|x_{n_t} - x_0| D^{-n} \right)^{p_t}. \tag{5.109}
$$

Now we plug in the lower estimate (5.106) of $|x_{n_t} - x_0|$ via $\Delta_{n_t,f}(x_0)$ and $\gamma_{n_t}(C, \delta)$. The result is the following

$$
\Delta_{n_t,f}(x_0) \geq \frac{1}{2} \Delta_{n_t,f}^{(1+4(N+2))p_t}(x_0) \gamma_{n_t}^{p_t}(C, \delta/2) D^{-2np_t}. \tag{5.110}
$$

Inductive application of this estimate gives

$$\Delta_{n_t,f}(x_0) \geq \exp\left(-\log 2(1+(1+4(N+2))p_t + \cdots + (1+4(N+2))^{t-1}P_{2,t})\right)$$
$$\Delta_{n_1,f}^{(1+4(N+2))^t P_{1,t}}(x_0) \prod_{i=1}^{t} \gamma_{n_i}^{(1+4(N+2))^{t-i}P_{i,t}}(C,\delta/2) \tag{5.111}$$
$$\exp\Big(-2\log Dn(1+(1+4(N+2))p_{t-1}(1+\cdots+(1+4(N+2))p_1)\ldots\Big).$$

Recall that by Theorem 5.1.4 we know that $s \leq N$. So in the case $n_{t+1} = n_s = n$ we can provide the following estimate of the right-hand side

$$\exp\Big(-2\log 2(1+4(N+2))^N n\Big)\exp\left(\frac{-Cn^{1+\frac{dt}{1+\delta}}}{N}\frac{12(\log n - \log n_1)(1+\delta)}{\delta \log n}\right)$$
$$\exp\left(-C(2+4(N+2))^N n\sum_{i=1}^{s} n_i^{\delta/2}\right)\exp\Big(-\log D(1+4(N+2))^N n\rho\Big). \tag{5.112}$$

For $D = \exp\left(\frac{2C}{N\rho}\right)$ and n_1 (denoted k there) satysfying conditions of Lemma 5.3.2 we can get the following lower bound

$$\exp\left(-n\left(2\log 2(1+4(N+2))^N - \frac{2Cn^{\frac{\delta}{1+\delta}}(2+4(N+2))^N}{N}\right.\right.$$
$$\left.\left.\frac{12Cn^{\frac{\delta}{1+\delta}}}{N}\;2C\;(1+sg)^N\;n^{\delta/2}\right)\right) \tag{5.113}$$

In this form it is clear that for a sufficently large n this product is certainly bounded from below by $\exp(-\frac{Cn^{1+\delta}}{30N})$. This completes the proof of Lemma 5.4.6. Q.E.D.

The proof of in the general case, when we don't know that T^{-n}-ladder moments are weak (D,n)-gaps still goes along the same line as the one given in the last Lemma. Let's just indicate how it can be done.

Lemma 5.4.7. *Lemma 5.4.6 holds true without an assumption that T^{-n}-ladder moments are weak (D,n)-gaps.*

Proof Before reading the proof of this Lemma it is useful to study the proof of Lemma 5.4.6 first.

Recall that $x_{n_1}, x_{n_2}, \ldots, x_{n_s}$ are T^{-n}-ladder moments with $n_1 < n_2 < \cdots < n_s = n$, where x_{n_1} is also a weak (D,n)-gap. Denote by $x_{n_1} = x_{m_1}, x_{m_2}, \ldots, x_{m_r} = x_n$ those T^{-n}-ladder moments which also are weak (D,n)-gaps so $r \leq s$. By Lemma

5.4.2 numbers $m_1 < m_2 < \cdots < m_r$ are multiplicatively ordered so $l_t = m_{t+1}/m_t$ is an integer for each $t = 1, \ldots, r-1$.

Notice that if for some $1 \leq t < r$ and some $1 \leq i < n/m_t$ we have

$$|x_{im_t} - x_0| \leq D^{-2Nn}|x_{m_t} - x_0|, \tag{5.114}$$

then x_{im_t} has to be divisible by m_{t+1}. Consider two cases $im_t > m_{t+1}$ and $im_t < m_{t+1}$. In the first case if im_t is not divisible by m_{t+1}, then apply Euclidean Algorithm (Lemma 5.1.6) to x_{im_t} and $x_{m_{t+1}}$. The result is that $|x_{\gcd(im_t, m_{t+1})} - x_0| < M_1^{2n}|x_{m_{t+1}} - x_0|$ and, since $\gcd(im_t, m_{t+1}) < m_{t+1}$, this contradicts existence of a weak (D, n)-gap at $x_{m_{t+1}}$.

In the second case we know that $m_t = n_{t'}$ for some $1 \leq t' \leq s$. By Theorem 5.1.4 $s \leq N$ and between $n_{t'}$ and $in_{t'}$ there is at most N T^{-n}-ladder moments $x_{n_{t'+1}}, \ldots x_{n_{t"}}$, i.e. $n_{t"+1} > in_{t'}$. Since there are not weak (D, n)-gaps at $x_{n_{t'+1}}, \ldots x_{n_{t"}}$ we have

$$\begin{gathered} |x_{in_{t'}} - x_0| \geq T^{-n}|x_{n_{t"}} - x_0| \geq \\ T^{-n}D^{-(t"-t')n}|x_{n_{t'}} - x_0| \geq D^{-Nn}|x_{m_t} - x_0|. \end{gathered} \tag{5.115}$$

This is a contradiction.

This gives all necessary ingredients to go along the proof of Lemma 5.4.6. The only difference is that (5.104) should be replaced by

$$|x_{im_t} - x_0| \geq D^{-2Nn}|x_{m_t} - x_0| \tag{5.116}$$

for each im_t nondivisible by m_{t+1}. Straightforward calculations show that additional factor of $2N$ in the exponent of D does not affect the conclusion. This completes the proof of Lemma 5.4.7. Q.E.D.

This also completes the proof of Theorem 5.4.1 and the Shift Theorem. Q.E.D.

In this Section we completed analysis of recurrent trajectories and the upshot is this Chapter is that after a certain procedure for a nonsimple (recurrent), almost periodic, nonsufficiently hyperbolic trajectory of a diffeomorphism satisfying the Inductive Hypothesis we can find a very close return (a (D, n, r_{2k})-gap). In the next Chapter we shall investigate an effect of perturbation on such kind of trajectories $\{x_j = f^j(x_0)\}_{j=1}^{n-1}$ with a very close return at x_k by the family

$$f_{u_{k-1},U_{2k-1}}(x) = f(x) + u_{k-1}\prod_{j=0}^{k-2}|x - x_j| + U_{2k-1}(x - x_{k-1})\prod_{j=0}^{k-2}|x - x_j|^2, \tag{5.117}$$

where U_{k-1} is an $N \times N$ matrix whose parameters vary hyperbolicity of $df^n_{u_{k-1},U_{2k-1}}(x_0)$, $(x - x_{k-1}$ is a vector from $\mathbb{R}^N$ and u_{k-1} is a vector whose parameters give existence of a very close return (a (D, n, r_{2k})-gap) at x_k. Certainly, this is not a family of polynomial perturbation as in sections 2.2-9 we shall find an appropriate place for this family in the space of families of polynomial perturbations by polynomails of degree at most $2k-1$.

Chapter 6

Estimating the Measure of “bad” Parameters Corresponding to a Trajectory with a Gap

6.1 Statement and discussion of an estimate of the “bad” measure

To explain the goal of this Chapter we briefly outline the global strategy of the proof of Theorem 1.3.7 (extension of the Main Theorem). We have a family of diffeomorphisms of the unit ball $\{f_{\vec{\varepsilon}}\}_{\vec{\varepsilon}\in HB^N(\vec{\mathbf{r}})}$, defined in Section (1.3). In Chapter 2 we define the set of “bad” parameters $B_n(C,\delta,\rho,\vec{\mathbf{r}},f)$ given in (2.1). The key to the proof is to estimate the measure of $B_n(C,\delta,\rho,\vec{\mathbf{r}},f) \subset HB^N(\vec{\mathbf{r}})$. In order to get such an estimate we split this set into two (may be intersecting) parts: simple $B_n^{sim}(C,\delta,\rho,\vec{\mathbf{r}},f)$ and nonsimple $B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},f)$ (see (2.35-2.37) for definitions).

How to estimate the measure of the simple part $B_n^{sim}(C,\delta,\rho,\vec{\mathbf{r}},f)$ in the 1–dimensional case was described in Sections 3.3-3.4 and will be described in Chapter ?? in the general case. In order to estimate the measure of the nonsimple part $B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},f)$ in the 1–dimensional case we use additional arguments (Sections 3.5-3.6) and reduce this problem to the problem for the simple part. In other words an estimate of the measure of the nonsimple part $B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},f)$ is obtained using an estimate for the measure of the simple part $B_n^{sim}(C,\delta,\rho,\vec{\mathbf{r}},f)$.

This Chapter provides an important one trajectory type estimate in the N–dimensional case for trajectories with a close return. This estimate is required for an estimate for the measure of the nonsimple part $B_n^{non}(C,\delta,\rho,\vec{\mathbf{r}},f)$ [1]. More precisely, let $f : B^N \hookrightarrow B^N$ be a $C^{1+\rho}$–smooth diffeomorphism of the unit ball B^N into its interior with an almost periodic trajectory $x_0, x_1 = f(x_0), \dots, x_n = f^n(x_0)$ of length n, i.e. x_n and x_0 are close, which has a close return or in our terminology a "gap" at x_k for some $k < n$. We consider the family of perturbations

$$f_U(x) = f(x) + \vec{u}_1 \prod_{j=0}^{k-2} |x - x_j| + U_2(x - x_{k-1})^t \prod_{j=0}^{k-2} |x - x_j|^2, \tag{6.1}$$

where $(x - x_{k-1})^t$ is a column vector, $\prod_{j=0}^{k-2} |x - x_j|$ is a product of distances, $\vec{u}_1 \in \mathbb{R}^N$ is a vector, and U_2 is an $N \times N$ matrix form $M_N(\mathbb{R})$. It is easy to see that within the family first $(k-1)$ points of the trajectory $\{x_j\}_{j=0}^n$ stay fixed, i.e. $f_U(x_j) = f(x_j)$ for $j = 0, \dots, k-2$. For this family we estimate the measure of those U's for which x_0 has a "gap" at x_k and is non-(n,γ)–hyperbolic for a sufficiently small positive γ.

Since we alway perturb by polynomial families in this paper, we need a slightly different family from (6.1) which is polynomial. So we consider a coordinate system $x = (x_1, \dots, x_N)$ in $\mathbb{R}^N$ fixed at the beginning in Section 1.3. With the k–tuple of points $\mathbf{X}_k = \{x_j\}_{j=0}^{k-1}$ fixed denote by $\sigma_j(\mathbf{X}_k)$ the index of the minimal coordinate such that absolute value of $|(x_{k-1} - x_j)_{\sigma_j(\mathbf{X}_k)}|$ is maximal among all coodinates. This gives that

$$|(x_{k-1} - x_j)_{\sigma_j(\mathbf{X}_k)}| \geq \frac{1}{\sqrt{N}} |x_{k-1} - x_j| \tag{6.2}$$

Replace now a nonpolunomial family (6.1) by the following polynomial family

$$f_U(x) = f(x) + \vec{u}_1 \prod_{j=0}^{k-2} (x - x_j)_{\sigma_j(\mathbf{X}_k)} + U_2(x - x_{k-1})^t \prod_{j=0}^{k-2} (x - x_j)^2_{\sigma_j(\mathbf{X}_k)}, \tag{6.3}$$

where $(x - x_{k-1})^t$ is a column vector.

Now we introduce some notations to fomulate the main result of this section. Recall that $M_N(\mathbb{R})$ denotes the space of $N \times N$ matrices $U = \{u_{i,j}\}_{i,j=1}^N$ with real coefficients and $C(r, N^2)$ is the cube in $M_N(\mathbb{R})$ centered in the origin with edge $2r$, i.e.

$$C(r, N^2) = \{U = \{u_{i,j}\}_{i,j=1}^N \in M_N(\mathbb{R}) : \ \forall i,j \ \ |u_{i,j}| < r\}. \tag{6.4}$$

[1] One trajectory type estimates from Section 3.3 is a building block for the Discretization Method

For a diffeomorphism $f : B^N \hookrightarrow B^N$ and a point $x_0 \in B^N$ denote $x_j = f^j(x),\ j \in \mathbb{Z}_+$ its trajectory and

$$\Delta_{f,k}(x_0) = \prod_{j=0}^{k-2} (x_{k-1} - x_j)_{\sigma_j(\mathbf{x}_k)} \tag{6.5}$$

is the product of distances for the k–st point x_{k-1} to all the previous points $\{x_j\}_{j=0}^{k-2}$.

For positive r_1, r_2 consider two cubes $C(r_1, N) \subset \mathbb{R}^N$ and $C(r_2, N^2) \subset M_N(\mathbb{R})$. Suppose that r_1 and r_2 are small enough that the whole family

$$\{f_U = f_{\vec{u}_1, U_2}\}_{\vec{u}_1 \in C(r_1,N),\ U_2 \in C(r_2,N^2)}, \tag{6.6}$$

defined by (6.3) consists of diffeomorphisms. Denote the product of cubes

$$C(r_1, N; r_2, N^2) = C(r_1, N) \times C(r_2, N^2). \tag{6.7}$$

Introduce C^1 and $C^{1+\rho}$ norms of this family by

$$\begin{aligned} M_1 &= \sup_{\vec{u}_1 \in C(r_1,N),\ U_2 \in C(r_2,N^2)} \max\left(\|f_U\|_{C^1},\ \|f_U^{-1}\|_{C^1},\ 2^{1/\rho}\right) \\ M_{1+\rho} &= \max\left(M_1, \sup_{\vec{u}_1 \in C(r_1,N),\ U_2 \in C(r_2,N^2)} \|f_U\|_{C^{1+\rho}}\right). \end{aligned} \tag{6.8}$$

Theorem 6.1.1. *Let $f : B^N \hookrightarrow B^N$ be a $C^{1+\rho}$–smooth diffeomorphism of the unit ball B^N into its interior for some $0 < \rho \le 1$, $r_1,\ r_2$ be positive numbers satisfying conditions (6.6) and defining along with (6.3) the family $\{f_U\}_{U \in C^{N,N^2}(r_1,r_2)}$ of diffeomorphism, and let $M_1,\ M_{1+\rho}$ be C^1 and $C^{1+\rho}$–norms of this family respectively as in (6.8). Let $\delta > 0,\ \sigma > 1,\ C > \frac{8}{\rho} \ln M_{1+\rho}$, and $D = M_{1+\rho}^{8/\rho}$. Suppose x_0 in B^N is an $(n, \gamma_n(C, \delta))$ -periodic point, which has a (D, n, r_2, σ)–gap at $x_k = f^k(x_0)$, i.e.*

$$|x_k - x_0| \le g_{f,k} = \min\left\{ D^{-n\sigma} \min_{1 \le j \le k-1} |x_0 - x_j|, r_2^{4(N+N^2)}, (\Delta_{f,k}(x_0))^{\sigma} \right\}. \tag{6.9}$$

Moreover, for any $s = 1, \dots, k-1$ we have $\Delta_{f,k}(x_0) \ge \Delta_{f,k}(x_s)$. Consider the $(N + N^2)$ -parameter family of diffeomorphisms $\{f_U\}_{U \in C(r_1,N;r_2,N^2)}$ of the form (6.1). Let $\mu_{r_1,N;r_2,N^2}$ be the product of Lebesgue probability measures on the brick $C(r_1, N; r_2, N^2)$ (6.7). Then

$$\begin{aligned} \mu_{r_1,N;r_2,N^2}\{U \in C(r_1, N; r_2, N^2) : x_0 \text{ has a } (D, n, r_2, \sigma)\text{-gap at } \ x_k(U) \\ \text{and is not } \ (n, M_1^{8n} g_{f,k}^{\rho})\text{-hyperbolic for } f_U\} \le \\ M_1^{4(2+N+N^2)n}\ (8N)^{N^2}\ g_{f,k}^{(N(\sigma-1)+(\rho\sigma-2))/\sigma}\ r_1^{-N}\ r_2^{-N^2-1}. \end{aligned} \tag{6.10}$$

In the picture below we describe the situation which we consider now in this Chapter. A trajectory $\{x_j\}_{j=0}^{n-1}$ of length n returns close to itself after k iterations and starts nearly repeating itself, i.e. if we consider sufficiently small balls around points $\{x_j\}_{j=0}^{k-1}$, then the rest of the trajectory $\{x_j\}_{j=k}^{n-1}$ belongs to the union of those balls. Based on this simple observation we will do appropriate computations.

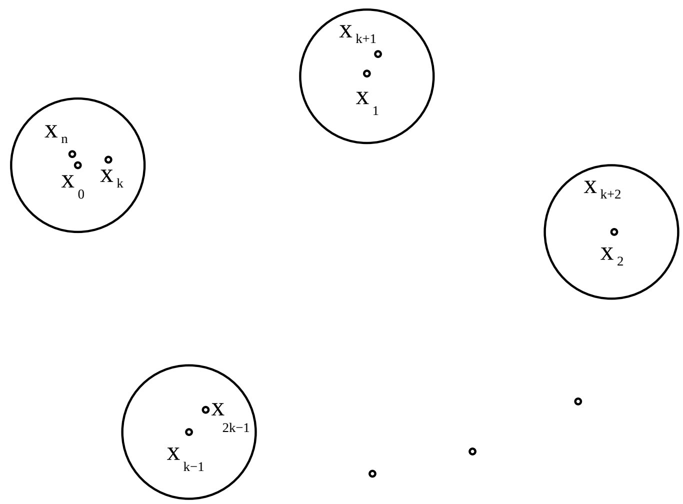

Figure 6.1: Localization of trajectory with a gap along itself

Remark 6.1.2. *1. By Euclidean Algorithm (Lemma 2.4.5) we have that* n *is divisible by* k*, i.e. there is* $p \in \mathbb{Z}_+$ *such that* $n = kp$*.*
2. The idea of considering the family by $\{f_U\}_{U \in C(r_1,N;r_2,N^2)}$ *is the following: the parameter* $\vec{u}_1$ *governs the position of the* k*–th iteration of* x_0 *under* f_U *and* $x_k(U) = x_k + \vec{u}_1 \Delta_{f,k}(x_0)$ *does not depend on* U_2*. So varying only* $\vec{u}_1$ *we estimate the measure of* $\vec{u}_1$ *for which* x_0 *has a* (D, n, r_2, σ)*–gap at* x_k*. Fix* $\vec{u}_1$*, then the parameter* U_2 *governs the linearization of* $df_U^k(x_0)$ *at* x_0 *which would allow us control hyperbolicity of* $df_U^k(x_0)$ *and hyperbolicity of* $df_U^n(x_0)$ *via an approximation formula* $df_U^n(x_0) = \left(df_U^k(x_0)\right)^p +$ *error.*

6.2 Sketch of the proof of the "bad" measure estimate (6.10)

In this section we start the proof of Theorem 3.6. The idea of the proof is as described in the second part of the above remark. First, we need to estimate the measure of $\vec{u}_1$'s for which x_0 has a (D, n, r_2, σ) -gap at $x_k(\vec{u}_1) = x_k + \vec{u}_1 \Delta_{f,k}(x_0)^2$. Second, for every $\vec{u}_1$ with the above property, i.e. there is a (D, n, r_2, σ) -gap at $x_k(\vec{u}_1)$ for x_0, we need to estimate measure of "bad" U_2's for which x_0 is not (n, g_k^ρ) -hyperbolic for $f_{(\vec{u}_1, U_2)}$ uniformly over $\vec{u}_1 \in C(r_1, N)$. Then the measure of $U = (\vec{u}_1, U_2)$ satisfying both conditions of (6.10) is the product of this two estimates on $\vec{u}_1$ and U_2 respectively. To estimate the measure of $\vec{u}_1$'s with a (D, n, r_2, σ)—-gap for x_0 notice that

$$f^k_{(\vec{u}_1, U_2)}(x_0) = f_k(x_0) + \vec{u}_1 \prod_{j=0}^{k-2} (x_{k-1} - x_j)_{\sigma_j(\mathbf{x}_k)} = f_k(x_0) + \vec{u}_1 \Delta_{f,k}(x_0) \tag{6.11}$$

Therefore, it is easy to see that

$$\begin{aligned} &\mu\{\vec{u}_1 \in C(r_1, N) : \ |f^k_{(\vec{u}_1, U_2)}(x_0) - x_0| \le g_{f,k}\} \le \\ &\qquad \left(r_1^{-1} g_{f,k} \Delta_{f,k}(x_0)^{-1}\right)^N \le \left(r_1^{-1} \ g_{f,k}^{(1-1/\sigma)}\right)^N, \end{aligned} \tag{6.12}$$

where μ is the Lebesgue probability measure. The last inequality follows form the definition (6.9) of $g_{f,k}$.

For simplicity of notations we denote $f_{(\vec{u}_1, U_2)}$ by $\tilde{f}_{U_2}$ and $f_{(\vec{u}_1, 0)}$ by $\tilde{f}$. Now we fix $\vec{u}_1$ satisfying the condition $|\tilde{f}^k(x_0) - x_0| \le g_{f,k}$ and consider the N^2–parameter family

$$\tilde{f}^k_{U_2}(x) = (f(x) + \vec{u}_1 \Delta_{f,k}(x)) + \ U_2(x - x_{k-1})^t \left(\Delta_{f,k}(x)\right)^2, \tag{6.13}$$

where U_2 varies in $C(r_2, N^2)$ and $\vec{u}_1$ is fixed. We need to estimate the measure of U_2's for which the linearization matrix $d\tilde{f}^n_{U_2}(x_0)$ is not sufficiently hyperbolic. The idea of finding such an estimate is the following:

• To compute the linearization matrix $d\tilde{f}^{kp}_{U_2}(x_0)$ and estimate the measure of U_2's for which it is not $\left(n, g^\sigma_{f,k}\right)$–hyperbolic, i.e. $\gamma(d\tilde{f}^{kp}_{U_2}(x_0)) \le M_1^{8n} g^\rho_{f,k}$ we approximate

$$\begin{aligned} d\tilde{f}^{kp}_{U_2}(x_0) = & d\tilde{f}^k_{U_2}(x_{(p-1)k}) \circ \cdots \circ d\tilde{f}^k_{U_2}(x_k) \circ d\tilde{f}^k_{U_2}(x_0) \\ & = \left(d\tilde{f}^k_{U_2}(x_0)\right)^p + \ error. \end{aligned} \tag{6.14}$$

• To compute the linearization matrix $d\tilde{f}^k_{U_2}(x_{jk})$ for every $j = 1, \ldots, p-1$ we approximate it by $d\tilde{f}^k_{U_2}(x_0)$

$$\begin{aligned} & d\tilde{f}^k_{U_2}(x_{jk}) = d\tilde{f}_{U_2}(x_{jk+k-1}) \circ \cdots \circ d\tilde{f}_{U_2}(x_{jk+1}) \circ d\tilde{f}_{U_2}(x_{jk}) = \\ & \left[d\tilde{f}(x_{k-1}) + U_2\left(\Delta_{f,k}(x)\right)^2 \right] \circ \cdots \circ d\tilde{f}(x_1) \circ d\tilde{f}(x_0) + \ error. \end{aligned} \tag{6.15}$$

the error should be small, because x_k is close to x_0 and, therefore, x_{jk} is close to x_0.

• To compute the linearization matrix $d\tilde{f}^k_{U_2}(x_{jk+s})$ for every $j = 1, \ldots, p-1$ and every $s = 0, \ldots, k-1$ we approximate $d\tilde{f}_{U_2}(x_{jk+s})$ by $d\tilde{f}(x_s)$.

Now we realize the above program of estimates going from the last bullet back to the first one.

6.3 Approximation $d\tilde{f}_{U_2}(x_{sk+j})$ by $d\tilde{f}(x_j)$.

Sublemma. *Let $\tilde{f}$ be a C^1 diffeomorphism, $M_1 \geq 2$ be an upper bound on C^1–norms of $\tilde{f}$ and $\tilde{f}^{-1}$. Then for any positive integer k and $p > 1$, $n = kp$, and any $s = 0, \ldots, p-1$, $j = 0, \ldots, k-1$ we have*

$$|x_0 - x_{sk}| \leq M_1^n |x_0 - x_k| \quad \text{and} \quad |x_0 - x_{sk+j}| \leq M_1^n |x_0 - x_j|, \tag{6.16}$$

provided that $M^n |x_0 - x_{sk}| \leq |x_0 - x_j|$.

Proof: Since M_1 is an upper bound on the C^1–norms of $\tilde{f}$ and $\tilde{f}^{-1}$, for any $x \neq y$ in B^N we have

$$M_1^{-1}|x - y| \leq |\tilde{f}(x) - \tilde{f}(y)| \leq M_1 |x - y|. \tag{6.17}$$

Thus, for any positive integer r we have $|x_{rk} - x_{(r+1)k}| \leq M_1^{rk} |x_0 - x_k|$. Summing over $r = 0, \ldots, j-1$ and applying the triangle inequality we get

$$|x_0 - x_{sk}| \leq \frac{M_1^{sk} - 1}{M_1^k - 1} |x_0 - x_k| \leq M^{sk} |x_0 - x_k| \tag{6.18}$$

which proves the first inequality. To prove the second one notice that

$$|x_0 - x_{sk+j}| \leq |x_{sk} - x_{sk+j}| + |x_{sk} - x_0| \leq M_1^{n-k}(|x_0 - x_k| + |x_0 - x_j|). \tag{6.19}$$

This completes the proof. Q.E.D.

The Sublemma is a reason to consider a point $\tilde{x}_j$ such that $|\tilde{x}_j - x_j| \leq M_1^n |x_0 - x_k|$ for $j = 0, \ldots, k-1$, because x_{sk+j} satisfy this condition for each $s = 0, \ldots p-1$.

Recall that $|x_0 - x_k| < g_{f,k}$ which is at least exponentially small in n and

$$d\tilde{f}_{U_2}(\tilde{x}_j) = d\tilde{f}(\tilde{x}_j) + d\left(U_2(\tilde{x}_j - x_{k-1})^t \prod_{s=0}^{k-2} (\tilde{x}_j - x_s)^2_{\sigma_s(\mathbf{X}_k)}\right), \tag{6.20}$$

where $U_2 = (\{U_{im}\}_{i,m=1}^N)$ is an $N \times N$ matrix. Denote by U^i the i–th row of U_2 for $i = 1, \dots, N$. Fixed some integer $\sigma_{k-1}(\mathbf{X}_k)$ from 1 to N and denote

$$\tilde{\Delta}_{f,k}(\tilde{x}_j, x_0) = \prod_{s=0, s\neq j}^{k-1} (\tilde{x}_j - x_s)_{\sigma_s(\mathbf{X}_k)}, \quad \min \mathbf{X}_k = \min_{1\leq j\leq k-1} |x_0 - x_j|, \tag{6.21}$$

by $\tilde{f}_{U_2}(x) = (\tilde{f}_{U_2,1}(x), \dots, \tilde{f}_{U_2,N}(x))$ coordinate functions, and by $\tilde{e}_j = \dfrac{(\tilde{x}_j - x_{k-1})}{|\tilde{x}_j - x_{k-1}|}$ unit vectors for $j = 0, \dots, k-1$. Introduce additional functionals of a trajectory $\{x_j = \tilde{f}^j\}_{j=0}^k$:

$$\begin{aligned} g_{f,k} &= \min\left\{D^{-n\sigma} \min_{1\leq j\leq k-1} |x_0 - x_j|,\ r_2^{4(N^2+N)},\ \Big(\Delta_{f,k}(x_0)\Big)^\sigma\right\} \\ G_{f,n,k} &= M_1^{2n} \frac{g_{f,k}}{\min \mathbf{X}_k} \end{aligned} \tag{6.22}$$

Now we formulate auxiliary lemmas.

Lemma 6.3.1. *With notations of Theorem 6.1.1 suppose that a point x_0 in B^N has a (D, n, r_2, σ)–gap at x_k, i.e.*

$$|x_0 - x_k| \leq \min\left\{D^{-n\sigma} \min_{1\leq j\leq k-1} |x_0 - x_j|,\ r_2^{4(N^2+N)},\ \Big(\Delta_{f,k}(x_0)\Big)^\sigma\right\} \tag{6.23}$$

and $\Delta_{f,k}(x_0) \geq \max_{1\leq s\leq k-1} \Delta_{f,k}(x_s)$. Then for any $s \neq j = 0, \dots, k-1$ and any $\tilde{x}_j \in B^N$ such that $|\tilde{x}_j - x_j| \leq M_1^n |x_0 - x_k|$ we have

$$\frac{|\tilde{x}_j - x_j|}{|\tilde{x}_j - x_s|} \leq G_{f,n,k} \quad \text{and} \quad \left|1 - \frac{|\tilde{x}_j - x_s|}{|x_j - x_s|}\right| \leq G_{f,n,k}. \tag{6.24}$$

Moreover, for any $j = 0, \dots, k-2$

$$\left|\frac{\tilde{\Delta}_{f,k}(\tilde{x}_j, x_0)}{\Delta_{f,k}(x_0)}\right| \leq N^k \quad \text{and} \quad \left|1 - \tilde{\Delta}_{f,k}(\tilde{x}_{k-1}, x_0)\Delta_{f,k}^{-1}(x_0)\right| \leq 2nNG_{f,n,k}. \tag{6.25}$$

Proof: By the triangle inequality , formula (6.17), and Sublemma

$$|\tilde{x}_j - x_j| \geq |x_j - x_s| - |\tilde{x}_j - x_j| \geq M_1^{-k}|x_0 - x_{|j-s|}| - M_1^n|x_0 - x_k|. \tag{6.26}$$

By definition of a (D, n, r_2, σ)–gap at x_k we have

$$|x_0 - x_k| \leq \frac{g_{f,k}}{\min \mathbf{X}_k}|x_0 - x_{|j-s|}| \quad \text{and} \quad \min \mathbf{X}_k \geq M_1^{2n} g_{f,k} \tag{6.27}$$

which implies that

$$|\tilde{x}_j - x_s| \geq |x_0 - x_k| \left(M_1^{-k} \frac{\min \mathbf{X}_k}{g_{f,k}} - M_1^n \right) \geq M_1^{-n} \frac{\min \mathbf{X}_k}{g_{f,k}} |x_0 - x_k|. \tag{6.28}$$

The condition $|\tilde{x}_j - x_j| \leq M_1^n|x_0 - x_k|$ applied to the last inequality gives what is required in the left-hand part of (6.24). The right-hand part of (6.24) has nearly the same proof.

To prove (6.25) the right-hand part we need

$$\left| \frac{\tilde{\Delta}_{f,k}(\tilde{x}_j, x_0)}{\Delta_{f,k}(x_0)} \right| \leq \left| \prod_{s=0, s\neq j}^{k-2} \frac{(\tilde{x}_j - x_s)_{\sigma_s(\mathbf{X}_k)}}{(x_{k-1} - x_s)_{\sigma_s(X_k)}} \frac{(\tilde{x}_j - x_{k-1})_{\sigma_{k-1}(\mathbf{X}_k)}}{(x_{k-1} - x_j)_{\sigma_j(\mathbf{X}_k)}} \right| \tag{6.29}$$

Multipling and dividing by $\Delta_{f,k}(x_j)$ and using condition $\Delta_{f,k}(x_0) \geq \Delta_{f,k}(x_j)$ we can the last quantity from above by

$$\left| \prod_{s=0, s\neq j}^{k-1} \frac{|\tilde{x}_j - x_s|}{(x_j - x_s)_{\sigma_s(X_k)}} \right| \leq N^{k/2} \left| \prod_{s=0, s\neq j}^{k-1} \frac{|\tilde{x}_j - x_s|}{|x_j - x_s|} \right|. \tag{6.30}$$

Now we can apply the right-hand part of (6.24) and prove the left-hand part of (6.25). To prove the right-hand part of (6.25) we can apply the right-hand part of (6.24). This completes the proof of the Lemma. Q.E.D.

Lemma 6.3.2. *In the above notations for every $i, m = 1, \dots, N$ the im–th entry of the linearization matrix $d\tilde{f}_{U_2}(\tilde{x}_j)$ we have the following form: for $j = k-1$*

$$\begin{aligned} \partial_m \tilde{f}_{U_2,i}(\tilde{x}_{k-1}) =& \partial_m \tilde{f}_i(\tilde{x}_{k-1}) + \\ (\Delta_{f,k}(x_0))^2 \Big\{ & U_{im}\rho(\mathbf{X}_k, \tilde{x}_{k-1}) + 2kG_{f,n,k}(U^i, v(\mathbf{X}_k, \tilde{x}_{k-1})) \Big\}, \end{aligned} \tag{6.31}$$

where $|1 - \rho(\mathbf{X}_k, \tilde{x}_{k-1})| \leq 2kNM_1^{2n} g_{f,k}/\min \mathbf{X}_k$ and $v(\mathbf{X}_k, \tilde{x}_{k-1})$ is a vector in $\mathbb{R}^N$ of at most unit length. For $j = 0, \dots, k-2$ we have

$$\begin{aligned} \partial_m \tilde{f}_{U_2,i}(\tilde{x}_j) =& \partial_m \tilde{f}_i(\tilde{x}_j) + \\ (\Delta_{f,k}(x_0))^2 \Big(& 2G_{f,n,k}\ (U^i, \tilde{e}_j)\rho(\mathbf{X}_k, \tilde{x}_j) + 2kG_{f,n,k}^2(U^i, v(\mathbf{X}_k, \tilde{x}_j)) \Big) \end{aligned} \tag{6.32}$$

where $|\rho(\mathbf{X}_k, \tilde{x}_j)| \leq 1$ and $v(\mathbf{X}_k, \tilde{x}_{k-1})$ is also a vector in $\mathbb{R}^N$ of at most unit length.

Proof: First we prove the first formula. We calculate explicity the entries of the linearization matrix (6.20).

$$\partial_m \tilde{f}_{U_2,i}(\tilde{x}_{k-1}) = \partial_m \tilde{f}_i(\tilde{x}_{k-1}) + (\tilde{\Delta}_{f,k}(\tilde{x}_{k-1}, x_0))^2 \left\{ U_{im} + 2\sum_{s=0}^{k-2} \chi(m=\sigma_s(\mathbf{X}_k)) \frac{(U^i, (\tilde{x}_{k-1} - x_{k-1}))}{(\tilde{x}_{k-1} - x_s)_{\sigma_s(\mathbf{X}_k)}} \right\}, \tag{6.33}$$

provided that $(\tilde{x}_{k-1} - x_s)_{\sigma_s(\mathbf{X}_k)} \neq 0$, otherwise $\tilde{\Delta}_{f,k}(\tilde{x}_{k-1}, x_0)$ vanishes and can be eliminated from the above sum. Now we can apply Lemma 6.3.1 to terms $|\tilde{x}_{k-1} - x_{k-1}|/|(\tilde{x}_{k-1} - x_s)_{\sigma_s(\mathbf{X}_k)}|$ with $s = 0, \ldots, k-2$ and rewrite the left-hand side in the form:

$$\partial_m \tilde{f}_i(\tilde{x}_{k-1}) + (\Delta_{f,k}(x_0))^2 \Big\{ U_{im}\ \rho(\mathbf{X}_k, \tilde{x}_{k-1}) + 2k\ G_{f,n,k}\ (U^i, v(\mathbf{X}_k, \tilde{x}_{k-1})) \Big\},$$

which is the required equality.

Calculations for $j = 0, \ldots, k-2$ a little bit more involved.

$$\partial_m \tilde{f}_{U_2,i}(\tilde{x}_j) = \partial_m \tilde{f}_i(\tilde{x}_j) + (\tilde{\Delta}_{f,k}(\tilde{x}_j, x_0))^2\ \left\{ U_{im} + 2\sum_{s=0}^{k-2} \chi(m=\sigma_s(\mathbf{X}_k)) \frac{(U^i, (\tilde{x}_j - x_{k-1}))}{(\tilde{x}_j - x_s)_{\sigma_s(\mathbf{X}_k)}} \right\}, \tag{6.34}$$

provided that $(\tilde{x}_j - x_s)_{\sigma_s(\mathbf{X}_k)} \neq 0$, otherwise $\tilde{\Delta}_{f,k}(\tilde{x}_j, x_0)$ vanishes and can be eliminated from the above sum. Now we can apply Lemma 6.3.1 to terms $(\tilde{x}_j - x_j)/(\tilde{x}_j - x_s)_{\sigma_s(\mathbf{X}_k)}$ with $s = 0, \ldots, k-2$ and rewrite the left-hand side in the form:

$$\begin{aligned} \partial_m \tilde{f}_i(\tilde{x}_{k-1}) + \Big\{ & 2\chi(m=\sigma_j(\mathbf{X}_k)) \frac{(U^i, \tilde{e}_j)}{|\tilde{x}_j - x_{k-1}|} (\tilde{x}_j - x_j)_{\sigma_j(\mathbf{X}_k)} \tilde{\rho}_j(\mathbf{X}_k, \tilde{x}_j) + \\ & 2\sum_{s=0, s\neq j}^{k-2} \chi(m=\sigma_s(\mathbf{X}_k)) \frac{(\tilde{x}_j - x_j)^2_{\sigma_j(\mathbf{X}_k)}}{|\tilde{x}_j - x_{k-1}|\ |\tilde{x}_j - x_s|} \tilde{\rho}_s(\mathbf{X}_k, \tilde{x}_j) + \\ & U_{im} \frac{(\tilde{x}_j - x_j)^2_{\sigma_j(\mathbf{X}_k)}}{|\tilde{x}_j - x_{k-1}|^2} \tilde{\rho}_{k-1}(\mathbf{X}_k, \tilde{x}_j) \Big\}\ (\Delta_{f,k}(x_0))^2, \end{aligned} \tag{6.35}$$

where $|\tilde{\rho}_s(\mathbf{X}_k, \tilde{x}_j)| \leq 1$ for $s = 0, \ldots, k-1$. To get from (6.34) to (6.35) we replace $(\tilde{x}_j - x_j)_{\sigma_j(\mathbf{X}_k)}$ by $|\tilde{x}_j - x_j|$ and also multiply and divide by corresponding factors. Now we can use estimates from Lemma 6.3.1 and estimate from above

$$|(\tilde{x}_j - x_j)_{\sigma_j(\mathbf{X}_k)}|/|\tilde{x}_j - x_s| \leq M_1^{2n} g_{f,k}/\min \mathbf{X}_k = G_{f,n,k}. \tag{6.36}$$

Direct calculation completes the proof of the Lemma. Q.E.D.

6.4 Comptetion of approximation calculation and the proof of Theorem 6.1.1.

In this Section we derive formulas (6.14) and (6.15) and using an exact expression for an error term in (6.14) we complete the proof of of Theorem 6.1.1. First, we derive an error term of approximation of $d\tilde{f}^k_{U_2}(x_{jk})$ by $d\tilde{f}^k_{U_2}(x_0)$.

Lemma 6.4.1. *With notations of Theorem 6.1.1 for any $j = 0, \dots, p-1$ we have*

$$d\tilde{f}^k_{U_2}(x_{jk}) = d\tilde{f}^k_{U_2}(x_0) + 4\ k\ N\ M_1^n\ G_{f,n,k}\ (\Delta_{f,k}(x_0))^2\ U_2\ V(\mathbf{X}_k, \mathbf{X}_{jk}), \qquad (6.37)$$

where $\mathbf{X}_{jk} = \{x_{jk-k}, x_{jk-k+1}, \dots, x_{jk-1}\}$ is a k–tuple and $V(\mathbf{X}_k, \mathbf{X}_{jk})$ is an $N \times N$ matrix whose entries bounded in absolute value by one.

Proof: Apply Lemma 6.3.2 to each term of the decomposition corresponding to $d\tilde{f}^k_{U_2}(x_{jk})$ into $d\tilde{f}_{U_2}(x_{jk+s})$:

$$d\tilde{f}^k_{U_2}(x_{jk})\ = d\tilde{f}_{U_2}(x_{jk+k-1}) \circ \dots \circ d\tilde{f}_{U_2}(x_{jk}) =$$
$$\Big[d\tilde{f}_{U_2}(x_{k-1}) + \ \Delta^2_{f,k}(x_0)\{U_2(\rho(\mathbf{X}_k, x_{jk+k-1}) - 1) + kG_{f,n,k}U_2V_2(\mathbf{X}_k, x_{jk+k-1})\}\Big] \circ$$
$$\Big[d\tilde{f}_{U_2}(x_{k-2}) + \ 2N\Delta^2_{f,k}(x_0)G_{f,n,k}\{U_2V_1(\mathbf{X}_k, x_{jk+k-2}) + kG_{f,n,k}U_2V_2(\mathbf{X}_k, x_{jk+k-2})\}\Big]$$
$$\circ \dots \circ \Big[d\tilde{f}_{U_2}\ (x_0) + 2N\ \Delta^2_{f,k}(x_0)G_{f,n,k}\{U_2V_1(\mathbf{X}_k, x_{jk}) + kG_{f,n,k}U_2V_2(\mathbf{X}_k, x_{jk})\}\Big],$$

where $\{V_i(\mathbf{X}_k, x_{jk+s})\}_{ij}$ are $N \times N$ matrices whose entries are bounded in absolute value by one. Computation of error terms using the fact that $\|d\tilde{f}_{U_2}(x)\| \le M_1$ completes the proof of the Lemma. Q.E.D.

Lemma 6.4.2. *With notations of Theorem 6.1.1 suppose a point x_0 in B^N has a (D, n, r_2, σ) -gap at x_k and $n = kp$, then we have*

$$d\tilde{f}^n_{U_2}(x_0) = \left(d\tilde{f}^k_{U_2}(x_0)\right)^p + 8\ n\ N\ M_1^{2n}\ G_{f,n,k}\ (\Delta_{f,k}(x_0))^2\ U_2\ V(\mathbf{X}_k, \mathbf{X}_n), \quad (6.38)$$

where $\mathbf{X}_n = \{x_0, x_{jk+1}, \dots, x_{n-1}\}$ is an n–tuple and $V(\mathbf{X}_k, \mathbf{X}_{jk})$ is an $N \times N$ matrix whose entries bounded in absolute value by one.

Proof: Apply Lemma 6.4.1 to each term of the decomposition corresponding to $d\tilde{f}^n_{U_2}(x_0)$ into $d\tilde{f}^k_{U_2}(x_{jk})$:

$$d\tilde{f}^n_{U_2}(x_0) = d\tilde{f}^k_{U_2}(x_{(p-1)k}) \circ \dots \circ d\tilde{f}^k_{U_2}(x_0) =$$
$$\Big[d\tilde{f}^k_{U_2}(x_0) + \ 4\ k\ N\ M_1^n\ G_{f,n,k}\ (\Delta_{f,k}(x_0))^2\ U_2\ V(\mathbf{X}_k, \mathbf{X}_{(p-1)k})\Big] \circ \dots$$
$$\circ \Big[d\tilde{f}^k_{U_2}(x_0) + \ 4\ k\ N\ M_1^n\ G_{f,n,k}\ (\Delta_{f,k}(x_0))^2\ U_2\ V(\mathbf{X}_k, \mathbf{X}_k)\Big] \circ d\tilde{f}^k_{U_2}(x_0).$$

Computation of error terms using the fact that $\|d\tilde{f}_{U_2}(x)\| \leq M_1$ completes the proof of the Lemma. Q.E.D.

Sublemma. *With notations of Theorem 6.1.1 suppose that a point x_0 has a (D, n, r_2, σ) -gap at x_k. Then*

$$\mu_{r_2,N_2}\{U \in C(r_2, N^2) : \ x_0 \quad \text{is not} \quad (n, M_1^{8n} g^{\rho}_{f,k})\text{-hyperbolic}\} \\ \mu_{r_2,N_2}\{U \in C(r_2, N^2) : \ x_0 \quad \text{is not} \quad (k, 2M_1^{8n} g^{\rho}_{f,k})\text{-hyperbolic}\}. \tag{6.39}$$

Proof: By Lemma A.2 if $d\tilde{f}^k_{U_2}(x_0)$ is $2M_1^{8n} g^{\rho}_{f,k}$–hyperbolic, then $\left(d\tilde{f}^k_{U_2}(x_0)\right)^p$ is $2M_1^{8n} g^{\rho}_{f,k}$–hyperbolic. If $\left(d\tilde{f}^k_{U_2}(x_0)\right)^p$ is $2M_1^{8n} g^{\rho}_{f,k}$–hyperbolic, then by Lemmas A.3 and 6.4.2 $d\tilde{f}^n_{U_2}(x_0)$ is $(2M_1^{8n} g^{\rho}_{f,k} - 8nN^3 r_2 M_1^{2n} (\Delta_{f,k}(x_0))^2 G_{f,n,k})$–hyperbolic. Direct calculation using definition of $\Delta_{f,k}(x_0)$ and $G_{f,n,k})$ shows that the last quantity exceeds $M_1^{8n} g^{\rho}_{f,k}$. This proves the Sublemma. Q.E.D.

By the Sublemma the required estimate on the measure of U_2's for which x_0 is not $(n, M_1^{8n} g^{\rho}_{f,k})$–hyperbolic follows from Proposition A.5. This completes the proof of Theorem 6.1.1. Q.E.D.

Chapter 7

The Multijet Space $\mathcal{J}^{1,n}(B^N, \mathbb{R}^N)$ and the Divided Differences Space $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$

In this Chapter we generalize formulas of the 1–dimensional Newton interpolation polynomials, described in Chapter 2, to the N–dimensional Newton interpolation polynomials. We shall reproduce some of comments from Chapters 2 and 3 for the 1–dimensional case to make the presentation of this Chapter selfcontained. In the next Chapter we shall investigate distortion properties of the corresponding Newton map (cf the 1-dimensional Newton map, defined in (2.24)).

Given a positive integer n and a C^1 map $f : B^N \to \mathbb{R}^N$ we define an associated function

$$\begin{aligned} &\mathcal{J}^{1,n} f : \underbrace{B^N \times \cdots \times B^N}_{n \text{ times}} \to \mathcal{J}^{1,n}(B^N, \mathbb{R}^N), \\ &\mathcal{J}^{1,n} f : (x_0, .., x_{n-1}) \to (x_0, .., x_{n-1}; f(x_0), .., f(x_{n-1}); df(x_0), .., df(x_{n-1})), \end{aligned} \tag{7.1}$$

where $df(y)$ defines the linearization of f at point y. In singularity theory this function is called the n–tuple 1–jet of f. The ordinary 1–jet, usually denoted by $j^1 f(x) = (x, f(x), df(x))$, maps the unit ball B^N to the space of 1–jets $\mathcal{J}^1(B^N, \mathbb{R}^N) \simeq B^N \times \mathbb{R}^N \times M_N$, where M_N be the space of $N \times N$ matrices with real coefficients. The product of n copies of $\mathcal{J}^1(B^N, \mathbb{R}^N)$ w is denote by

$$\mathcal{J}^{1,n}(B^N, \mathbb{R}^N) = \underbrace{\mathcal{J}^1(B^N, \mathbb{R}^N) \times \cdots \times \mathcal{J}^1(B^N, \mathbb{R}^N)}_{n \text{ times}}. \tag{7.2}$$

We call $\mathcal{J}^{1,n}(B^N,\mathbb{R}^N)$ the space of *complete n–tuple 1–jets* (or simply *the complete multijet space).*

Notice that $\mathcal{J}^{1,n}(B^N,\mathbb{R}^N)$ is isomorphic to $\left(B^N\right)^n \times \left(\mathbb{R}^N\right)^n \times \left(\mathbb{R}^{N\times N}\right)^n$ after rearranging coordinates. The n–tuple 1–jet of f associates with each n–tuple of points in B^N all the information necessary to determine how close the n–tuple to being a periodic orbit, and if so, how close it is to being nonhyperbolic.

The set

$$\begin{aligned}\Delta_n(B^N) = \Big\{ & \{x_0,\dots,x_{n-1}\} \times \left(\mathbb{R}^N\right)^n \times \left(\mathbb{R}^{N\times N}\right)^n \in \left(B^N\right)^n \times \\ & \left(\mathbb{R}^N\right)^n \times \left(\mathbb{R}^{N\times N}\right)^n \subset \mathcal{J}^{1,n}(B^N,\mathbb{R}^N) : \exists i \neq j \text{ such that } x_i = x_j \Big\}.\end{aligned} \tag{7.3}$$

is called the *diagonal* (or sometimes the *generalized diagonal*). In singularity theory the space of multijets is defined outside of the diagonal $\Delta_n(B^N)$ and is usually denoted by $\mathcal{J}^1_n(B^N,\mathbb{R}^N) = \mathcal{J}^{1,n}(B^N,\mathbb{R}^N) \setminus \Delta_n(B^N)$ (see [GG]). This is the reason we call $\mathcal{J}^{1,n}(B^N,\mathbb{R}^N)$ the *complete* space of multijets. It is easy to see that *a recurrent trajectory $\{x_k\}_{k\in\mathbb{Z}_+}$ of a map $f : B^N \to B^N$ is located in a neighborhood of the diagonal $\Delta_n(B^N) \subset \mathcal{J}^{1,n}(B^N,\mathbb{R}^N)$ for a sufficiently large n.* Below we shall generalize the 1–dimensional blow-up along the diagonal $\Delta_n(I)$ for 1–dimensional maps from Chapter 1 to the N–dimensional blow-up along the diagonal $\Delta_n(B^N)$ for N–dimensional maps using the Grigoriev-Yakovenko construction [GY].

7.1 Newton interpolation and blow-up along the diagonal in the multijet space

Let $\alpha = (\alpha_1,\dots,\alpha_N) \in \mathbb{Z}^N_+$ be a multiindex, $x = (x_1,\dots,x_N) \in \mathbb{R}^N$ be a point, $x^\alpha = x_1^{\alpha_1}\dots x_N^{\alpha_N}$, and $|\alpha| = \sum_{j=1}^N \alpha_j$. Choose a positive n and consider the family of perturbations of a C^1 map $f : B^N \hookrightarrow B^N$ by polynomials of degree $2n-1$

$$f_{\vec\varepsilon}(x) = f(x) + \phi_{\vec\varepsilon}(x), \qquad \phi_{\vec\varepsilon}(x) = \sum_{|\alpha|\le 2n-1} \vec\varepsilon_\alpha x^\alpha, \tag{7.4}$$

where $\vec\varepsilon = \left(\{\vec\varepsilon_\alpha\}_{|\alpha|=0},\dots,\{\vec\varepsilon_\alpha\}_{|\alpha|=2n-1}\right) \in W_{0,N}\times\dots\times W_{2n-1,N}$. Recall that $\dim W_{k,N} = \nu(k,N)$ and $\dim W_{\le k,N} = \eta(k,N) = \sum_{j=0}^k \nu(k,N)$. The perturbation vector $\vec\varepsilon$ consists of coordinates from the Hilbert brick $HB^N(\vec{\mathbf{r}})$ of analytic perturbations defined in Section 1.3. Our goal now is to describe how such perturbations affect the n–tuple 1–jet of f, and since the operator $j^{1,n}$ is linear in f, for the time being we consider only the perturbations $\phi_{\vec{\varepsilon}}$ and their n–tuple 1–jets. For each n–tuple $\{x_k\}_{k=0}^{n-1}$ there is a natural transformation from ε–coordinates to jet-coordinates, which we write

$$\begin{aligned}\mathcal{J}^{1,n} : \left(B^N\right)^n \times \mathbb{R}^{\eta(2n-1,N)} &\to \mathcal{J}^{1,n}(B^N, \mathbb{R}^N)\\ \mathcal{J}^{1,n}(x_0, \dots, x_{n-1}, \vec{\varepsilon}) &= j^{1,n}\phi_{\vec{\varepsilon}}(x_0, \dots, x_{n-1}).\end{aligned} \tag{7.5}$$

For each multiindex $\alpha \in \mathbb{Z}_+^N$ introduce the following notation

$$\begin{aligned}(x; x_0, \dots, x_{|\alpha|-1(mod\ n)})^\alpha = \prod_{i_1=0}^{\alpha_1-1}(x - x_{i_1})_1\\ \prod_{i_2=0}^{\alpha_2-1}(x - x_{\alpha^1+i_1(mod\ n)})_2 \cdots \prod_{i_N=0}^{\alpha_N-1}(x - x_{\alpha^{N-1}+i_1(mod\ n)})_N\end{aligned} \tag{7.6}$$

Instead of working directly with the transformation $\mathcal{J}^{1,n}$, we introduce intermediate u–coordinates based on Newton interpolation polynomials. The relation between ε–coordinates and u–coordinates is given implicitly by

$$\phi_{\vec{\varepsilon}}(x) = \sum_{|\alpha| \le 2n-1} \vec{\varepsilon}_\alpha x^\alpha = \sum_{|\alpha| \le 2n-1} \vec{u}_\alpha (x; x_0, \dots, x_{|\alpha|-1(\text{mod}\ n)})^\alpha. \tag{7.7}$$

Based on this identity, we will define functions

$$\begin{aligned}\mathcal{D}^{1,n} : \left(B^N\right)^n \times \mathbb{R}^{\eta(2n-1,N)} &\to \left(\mathbb{R}^N\right)^n \times \mathbb{R}^{\eta(2n-1,N)}\\ \mathcal{D}^{1,n} : (x_0, \dots, x_{n-1}, \{\vec{\varepsilon}_\alpha\}_{|\alpha| \le 2n-1}) &\to (x_0, \dots, x_{n-1}, \{\vec{u}_\alpha\}_{|\alpha| \le 2n-1})\end{aligned} \tag{7.8}$$

and

$$\begin{aligned}\pi^{1,n} : \left(\mathbb{R}^N\right)^n \times \mathbb{R}^{\eta(2n-1,N)} &\to \mathcal{J}^{1,n}(B^N, \mathbb{R}^N)\\ \pi^{1,n} : (x_0, \dots, x_{n-1}, \{\vec{u}_\alpha\}_{|\alpha| \le 2n-1}) &\to (x_0, \dots, x_{n-1};\\ &\qquad \phi_{\vec{\varepsilon}}(x_0), \dots, \phi_{\vec{\varepsilon}}(x_{n-1}); d\phi_{\vec{\varepsilon}}(x_0), \dots, d\phi_{\vec{\varepsilon}}(x_{n-1})),\end{aligned} \tag{7.9}$$

where $\vec{\varepsilon} = \{\vec{\varepsilon}_\alpha\}$ is defined as the preimage of $\left(\mathcal{D}^{1,n}\right)^{-1}(x_0, \dots, x_{n-1}, \{\vec{u}_\alpha\}_{|\alpha| \le 2n-1})$. So it is clear from the definition that $\mathcal{J}^{1,n} = \pi^{1,n} \circ \mathcal{D}^{1,n}$, or in other words the diagram in Figure 3 commutes.

The intermediate space, which we denote by $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$, is called *the space of divided differences* and consists of n–tuples of points $\{x_k\}_{k=0}^{n-1}$ and $\eta(2n-1, N)/N$ real vector coefficients each N–dimensional vector $\{\vec{u}_\alpha\}_{|\alpha| \le 2n-1}$. According to the construction the space $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$ has the following skew-product structure. Denote by $W^u_{k,N}(x_0, \dots, x_{k-1})$ the space of N–vector polynomials of degree k in N variables vanishing at $x_0, \ldots, x_{k-1}$ with the basis $\{(x; x_0, \ldots, x_{k-1})^\alpha\}_{|\alpha|=k}$. If some of x_j's occurs in $x_0, \ldots, x_{k-1}$ more than once, then polynomials from $W^u_{k,N}(x_0, \ldots, x_{k-1)})$ has to vanish up to the order of the number of repetitions. Denote by

$$\begin{aligned} &W^u_{\le k,N}(x_0, \ldots, x_{k-1}) = \\ &W^u_{0,N}(x_0) \times W^u_{1,N}(x_0, x_1) \times \cdots \times W^u_{k,N}(x_0, \ldots, x_{k-1}). \end{aligned} \tag{7.10}$$

Then

$$\begin{aligned} &\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N) = \{(x_0, \ldots, x_{n-1}, \{\vec{u}_\alpha\}_{|\alpha| \le 2n-1}) \in \\ &\underbrace{B^N \times \cdots \times B^N}_{n \quad \text{times}} \times W^u_{\le 2n-1,N}(x_0, \ldots, x_{2n-1(mod\ n)})\} \end{aligned} \tag{7.11}$$

We will show later that $\mathcal{D}^{1,n}$ is invertible, while in contrast to the 1-dimensional case $\pi^{1,n}$ is not invertible. However, it still defines a blow-up along the diagonal $\Delta_n(B^N)$ in the space of multijets $\mathcal{J}^{1,n}(B^N, \mathbb{R}^N)$ in the sense that codimension of the preimage $(\pi^{1,n})^{-1}(\Delta_n(B^N))$ of the diagonal in the preimage $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$ exceeds codimension of the diagonal $\Delta_n(B^N)$ in the image $\mathcal{J}^{1,n}(B^N, \mathbb{R}^N)$.

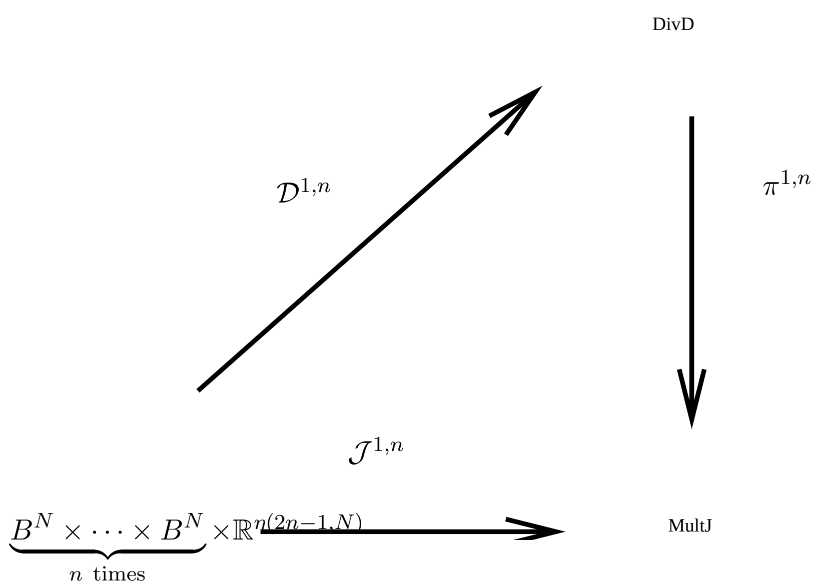

Figure 7.1: Algebraic N–dimensional Blow-up along the Diagonal $\Delta_n(B^N)$

Here are explicit coordinate-by-coordinate formulas giving natural relations for $\pi^{1,n} : \mathcal{DD}^{1,n}(B^N, \mathbb{R}^N) \to \mathcal{J}^{1,n}(B^N, \mathbb{R}^N)$ This mapping is given by

$$\begin{aligned} &\pi^{1,n}(x_0, \ldots, x_{n-1}, u_0, \ldots, u_{2n-1}) = \\ &\quad \Big(x_0, \ldots, x_{n-1}, \phi_\varepsilon(x_0), \ldots, \phi_\varepsilon(x_{n-1}), d\phi_\varepsilon(x_0), \ldots, d\phi_\varepsilon(x_{n-1})\Big), \end{aligned} \tag{7.12}$$

where

$$\begin{aligned}
\phi_{\vec{\varepsilon}}(x_0) &= \vec{u}_0,\\
\phi_{\vec{\varepsilon}}(x_1) &= \vec{u}_0 + \sum_{|\alpha|=1} \vec{u}_\alpha (x_1; x_0)^\alpha,\\
\phi_{\vec{\varepsilon}}(x_2) &= \vec{u}_0 + \sum_{|\alpha|=1} \vec{u}_\alpha (x_2; x_0)^\alpha + \sum_{|\alpha|=2} \vec{u}_\alpha (x_2; x_0, x_1)^\alpha,\\
&\vdots\\
\phi_{\vec{\varepsilon}}(x_{n-1}) &= \vec{u}_0 + \sum_{|\alpha|=1} \vec{u}_\alpha (x_{n-1}; x_0)^\alpha +\\
&\cdots + \sum_{|\alpha|=n-1} \vec{u}_\alpha (x_{n-1}; x_0, \dots, x_{n-2})^\alpha,\\
\frac{\partial}{\partial x}\phi_{\vec{\varepsilon}}(x_0) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} \sum_{|\alpha|=k} \vec{u}_\alpha (x; x_0, \dots, x_{k-1 (\mathrm{mod}\ n)})^\alpha\right)\Big|_{x=x_0},\\
&\vdots\\
\frac{\partial}{\partial x}\phi_{\vec{\varepsilon}}(x_{n-1}) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} \sum_{|\alpha|=k} \vec{u}_\alpha (x; x_0, \dots, x_{k-1 (\mathrm{mod}\ n)})^\alpha\right)\Big|_{x=x_{n-1}},
\end{aligned} \tag{7.13}$$

Similarly to the 1–dimensional case these formulas are very useful for dynamics. For a given base map f and an initial point x_0, the image $f_{\vec{\varepsilon}}(x_0) = f(x_0) + \phi_{\vec{\varepsilon}}(x_0)$ of x_0 depends only on $\vec{u}_0$. Furthermore the image can be set to any desired point by choosing $\vec{u}_0$ appropriately — we say then that it depends non trivially only on $\vec{u}_0$. Suppose x_0, x_1, and $\vec{u}_0$ are fixed, then the image $f_{\vec{\varepsilon}}(x_1)$ of x_1 depends only on $\{\vec{u}_\alpha\}_{|\alpha|=1}$, and as long as $x_0 \neq x_1$ it depends non trivially on $\{\vec{u}_\alpha\}_{|\alpha|=1}$. More generally for $0 \leq k \leq n-1$, if pairwise distinct points $\{x_j\}_{j=0}^k$ and coefficients $\{\vec{u}_\alpha\}_{|\alpha|\leq k-1}$ are fixed, then the image $f_{\vec{\varepsilon}}(x_k)$ of x_k depends only and non trivially on $\{\vec{u}_\alpha\}_{|\alpha|=k}$. However, in contrast with the 1–dimensional case the image $f_{\vec{\varepsilon}}(x_k)$ does not define $\{\vec{u}_\alpha\}_{|\alpha|=k}$ in a unique way, whereas in the 1–dimensional case it does [1].

Suppose now that an n–tuple of pairwise distinct points $\{x_j\}_{j=0}^{n-1}$ not on the diagonal $\Delta_n(B^N)$ and Newton coefficients $\{\vec{u}_\alpha\}_{|\alpha|\leq n-1}$ are fixed. Then the linearization $df_{\vec{\varepsilon}}(x_0)$ at x_0 depends only and non trivially on $\{\vec{u}_\alpha\}_{|\alpha|=n}$. Likewise for $0 \leq k \leq n-1$,

[1] This gives multidimensional Newton interpolation polynomials additional flexibility compare to 1–dimensional ones. In particular, one can try to use lower dimensional Newton coefficients $\{\vec{u}_\alpha\}_{|\alpha|=k}$ to perturb an n–th point of a trajectory under consideration for $k < n$.

if distinct points $\{x_j\}_{j=0}^n$ and Newton coefficients $\{\vec{u}_\alpha\}_{|\alpha|\le n+k-1}$ are fixed, then the linearization $df_{\vec{\varepsilon}}(x_k)$ at x_k depends only and non trivially on $\{\vec{u}_\alpha\}_{|\alpha|=n+k}$.

As Figure 2 illustrates, these considerations show that for any map f and any desired trajectory of distinct points with any given collection of linearizations along it, one can choose Newton coefficients $\{\vec{u}_\alpha\}_{|\alpha|\le 2n-1}$ and explicitly construct a map $f_{\vec{\varepsilon}} = f + \phi_{\vec{\varepsilon}}$ with such a trajectory.

7.2 Language of divided differences and the Newton interpolation formula

In this Section a construction of the space of divided differences $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$ the corresponding map $\mathcal{D}^{1,n}$, and the polynomial $\pi^{1,n}$ described above on figure 1 is presented. We are following presentation from [GY].

Let $f : \mathbb{R}^N \to \mathbb{R}$ be a smooth function in N real variables $x_1, \dots, x_N$.

Definition 7.2.1. *The first order divided difference of f in the variable x_k is the function of $N+1$ variables $x_1, \dots, x_{k-1}, x'_k, x''_k, \dots, x_N$ defined as*

$$\Delta_{x_k} f(x_1, \dots, x_{k-1}, x'_k, x''_k, \dots, x_N) = \frac{f(x_1, \dots, x_{k-1}, x'_k, \dots, x_N) - f(x_1, \dots, x_{k-1}, x''_k, \dots, x_N)}{x'_k - x''_k} \tag{7.14}$$

for $x'_k \ne x''_k$ and extended by its limit value as $\frac{\partial f}{\partial x_k}(x_1, \dots, x_{k-1}, x'_k, \dots, x_N)$ for $x'_k = x''_k = x_k$. Clearly, if f is C^r function, then (e.g., by the Hadamard Lemma), $\Delta_{x_k} f$ is at least C^{r-1}–smooth function of its arguments.

It turns out that iterating this construction is possible [GY] which leads to

Definition 7.2.2. *Let $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}_+^N$ be a multiindex, let f be as above. Then $\Delta_x^\alpha f = \Delta_{x_1}^{\alpha_1} \dots \Delta_{x_N}^{\alpha_N} f$ is called the mixed divided difference of order $|\alpha| = \alpha_1 + \dots + \alpha_N$. This divided difference is a smooth function of $N + |\alpha|$ arguments subdivided into N groups of $\alpha_1 + 1, \dots, \alpha_N + 1$ variables, symmetric with respect to permutations of variables within the same groups.*

It is easy to see that the operators Δ_{x_j} and Δ_{x_k} commute for $k \ne j$, and, therefore, we can use the multiindex notation for divided differences.

The Newton interpolation formula in one variable is discussed in Section 2.2.

7.3 The Newton interpolation formula in multivariables

Consider an n–tuple of points $x_0, \dots, x_{n-1} \in \mathbb{R}^N$ each given by N coordinates. Rearrange coordinates of these points in the following form

$$X^1 = (x_0^1, \dots, x_{n-1}^1) \subset \mathbb{R}, \dots, X^N = (x_0^N, \dots, x_{n-1}^N) \subset \mathbb{R}. \tag{7.15}$$

Each set X^j consists of the same number of points belonging to the corresponding j–th coordinate axis of $\mathbb{R}^N$. Then, given a multiindex $\alpha \in \mathbb{Z}_+^N$ and a smooth function $g(x) = g(x^1, \dots, x^N)$ in N variables we can form the divided difference $\Delta_x^\alpha f(X^1, \dots, X^N)$.

In terms of the divided differences one can write the Newton interpolation polynomial as follows:

$$\begin{gathered} \mathcal{P}(t^1, \dots, t^N) = \\ \sum_{|\alpha| \leq n-1} \Delta_x^\alpha f(X^1, \dots, X^N) \prod_{i_1=0}^{\alpha_1 - 1} (t^1 - x_{i_1}^1) \cdots \prod_{i_N=0}^{\alpha_N - 1} (t^N - x_{i_N}^N). \end{gathered} \tag{7.16}$$

The polynomial $\mathcal{P}(t^1, \dots, t^N)$ has degree $\leq n-1$ in variables $t = (t^1, \dots, t^N)$. The Newton interpolation formula implies that the difference $g(t) - \mathcal{P}(t^1, \dots, t^N)$ vanishes at all points of the Cartesian product grid $\mathbf{X} = X^1 \times \cdots \times X^N \subset \mathbb{R}^N$. Moreover, to obtain interpolation of the 1–st jet of f we replace each X^j by $\mathrm{diag}^2(X^j)$, defined in Section 2.2. The degree of interpolating polynomial will be $\leq 2n-1$.

In the case of a multivariate function $f : \mathbb{R}^N \to \mathbb{R}^s$ interpolating polynomial $\mathcal{P}(t^1, \dots, t^N)$ becomes s–dimensional vector and is interpolating by coordinate functions of f. In this paper we are interested in the case when $s = N$ so we shall concentrate only on this case.

Definition 7.3.1. *Let $g : B^N \to \mathbb{R}^N$ be a sufficiently smooth map and $\mathcal{DD}^{1,n}(B^N, \mathbb{R}^N)$ be the collection of all divided differences with 2 repetitions, $\{\Delta_x^\alpha g(diag^2(\mathbf{X}_{n+k})), \dots, diag^2(\mathbf{X}_{n+k}))\}_\alpha$, where $|\alpha| \leq 2n-1, \mathbf{X}_{n+k} \subset \mathbf{X}_n^2$. This is a linear space naturally equipped with the coordinates $\{x_i, \vec{u}_\alpha :\ 0 \leq i \leq n,\ |\alpha| \leq 2n-1\}$, where x_i and $\vec{u}_\alpha$ are vectors from $\mathbb{R}^N$. The dimension of this space is equal to $nN + \eta(2n-1, N)$.*

The map $\mathcal{D}^{1,n} g$ is defined by

$$\begin{aligned} \mathcal{D}^{1,n} g : (x_0, \dots, x_{n-1}) &\to (x_0, \dots, x_{n-1}, \{\vec{u}_\alpha\}_\alpha), \\ \text{where } \vec{u}_\alpha &= \Delta_x^\alpha g,\ |\alpha| \leq 2n-1. \end{aligned} \tag{7.17}$$

The multivariate interpolation formula together with its derivatives in t_j evaluated at the points of the grid, can be interpreted as a polynomial map restoring multijets from divided differences.

Newton Interpolation on $\mathbb{R}^N$ (abstract version) *The multivariate Newton interpolation formula*

$$\pi^{1,n}(x) = \sum_{|\alpha|\le 2n-1} \vec{u}_\alpha \prod_{i_1=0}^{\alpha_1-1}(x^1 - x^1_{i_1}) \cdots \prod_{i_N=0}^{\alpha_N-1}(x^N - x^N_{i_N}), \tag{7.18}$$

where $x = (x^1, \dots, x^N) \in \mathbb{R}^N$, along with (7.16) defines a polynomial interpolation map $\pi^{1,n} : \mathcal{DD}^{1,n}(B^N, \mathbb{R}^N) \to \mathcal{J}^{1,n}(B^N, \mathbb{R}^N)$ such that $\mathcal{J}^{1,n}g = \pi^{1,n} \circ \mathcal{D}^{1,n}g$ for any sufficiently smooth map $g : B^N \to \mathbb{R}^N$. The degrees of all the components of $\pi^{1,n}$ do not exceed $2n-1$.

Consider the restriction of the map $\mathcal{D}^{1,n}$ to a fixed n–tuple $x_0, \dots, x_{n-1}$

$$\mathcal{L}^N_{x_0,\dots,x_{n-1}} = \mathcal{D}^{1,n}\Big|_{x_0,\dots,x_{n-1}} : W_{\le 2n-1,N} \to \quad W^u_{\le 2n-1,N}(x_0, \dots, x_{2n-1(mod\ n)}), \tag{7.19}$$

then this map is linear and we call this *the Newton map*. In the next Section we shall investigate properties of the Newton map $\mathcal{L}^N_{x_0,\dots,x_{n-1}}$.

7.4 Basic properties of the Newton map

Let $\alpha, \beta \in \mathbb{Z}^N_+$. Define partial ordering $\succ$ for multiindices by $\beta \succ \alpha$ if for each $i = 1, \dots, N$ we have $\beta_i \ge \alpha_i$. Recall that in the introduction for two positive integers m, k and $m+1$ points $x_0, \dots, x_m \in \mathbb{R}$ for $k \ge m$ we denote by $p_{k,m}(x_0, \dots, x_m)$ the homogeneous polynomial of all degree $k - m$ monomials in $x_0, \dots, x_m$ with unit coefficients,

$$p_{k,m}(x_0, \dots, x_m) = \sum_{r_0+\cdots+r_m=k-m} \prod_{j=0}^{m} x_j^{r_j} \tag{7.20}$$

and if $k < m$, then $p_{k,m}(x_0, \dots, x_m)$ is identically zero.

Proposition 7.4.1. *Let $\alpha, \beta \in \mathbb{Z}^N_+$ be multiindices with $|\alpha| = m$ for some $m \in \mathbb{Z}_+$ and $x_0, \dots, x_{m-1} \in \mathbb{R}^N$ be an m–tuple of points given by its coordinates in $\mathbb{R}^N$. Rearrange coordinates of these points in the form (7.15). Then the divided difference $\Delta^\alpha x^\beta(x_0, \dots, x_{m-1})$ is given by the quasihomogeneous polynomial of the form*

$$\Delta^\alpha x^\beta(x_0, \dots, x_{m-1}) = \prod_{i=1}^{N} p_{|\beta_i|,|\alpha_i|}(x_i^0, \dots, x_i^{|\alpha_i|-1}) \tag{7.21}$$

of degree $\sum_{i=0}^{N}(\beta_i-\alpha_i)$ *in the case* $\alpha \succ \beta$ *and by identical zero in the other case. The number of monomials of* $\Delta^\alpha x^\beta(x_0,\dots,x_{m-1})$ *equals*

$$\frac{\beta_1!}{(\beta_1-\alpha_1)!\alpha_1!}\cdots\frac{\beta_N!}{(\beta_N-\alpha_N)!\alpha_N!}, \tag{7.22}$$

Proof: The case $N=1$ is considered in the introduction. The only we need to notice is that the number of monomials of $p_{k,m}(x_0,\dots,x_m)$ equals $\frac{k!}{m!(k-m)!}$ as simple combinatorial arguments show.

To generalize it to multivariable case notice that operators Δ_{x_i} and Δ_{x_j} commute and act independently. Therefore, the number of monomials of $\Delta^\alpha x^\beta(x_0,\dots,x_m)$ equals the product of the number of monomials after an application $\Delta_{x_i}^{\alpha_i}$, for $i=1,\dots,N$. This completes the proof of the Proposition. Q.E.D.

Consider the alphabetic ordering on the space of multiindices from $\mathbb{Z}_+^N$.

Corollary 7.4.2. *With respect to the alphabetic ordering for any* n*–tuple of points* $x_0,\dots,x_{n-1}\in\mathbb{R}^N$ *the Newton map* $\mathcal{L}^N_{x_0,\dots,x_{n-1}}$ *has the following block-form*

$$\begin{pmatrix} Id & 0 \\ A & UT \end{pmatrix} \tag{7.23}$$

where Id *is the* $nN\times nN$ *identity matrix,* UT *is an* $\eta(2n-1,N)\times\eta(2n-1,N)$ *upper triangular matrix with units on the diagonal, "A" is the* $\eta(2n-1,N)\times nN$*–block consisting of all divided differences order up to* $\alpha\in\mathbb{Z}_+^N$ *with* $|\alpha|\le 2n-1$*. Therefore,* $\det\mathcal{L}^N_{x_0,\dots,x_{n-1}}\equiv 1$ *and* $\mathcal{L}^N_{x_0,\dots,x_{n-1}}$ *preserves Lebesgue volume. Moreover,* $\mathcal{L}^N_{x_0,\dots,x_{n-1}}$ *preserves measures from the class* $\mu^N_{\mathbf{r}}$ *for any sequence of positive* r_k*'s.*

Lemma 7.4.3. *Let* $g:[0,1]\to\mathbb{R}$ *be a* C^n *smooth function,* $x_0,\dots,x_n$ *be a set of points in* $[0,1]$*. Define the set of polynomials*

$$\begin{aligned} p_0(x_0)=x_0,\quad p_1(u_1,x_0,x_1)=&(x_1-x_0)u_1+p_0(x_0),\dots,\\ p_{n-1}(u_1,\dots,u_{n-1},x_0,\dots,x_{n-1})=&(x_{n-1}-x_{n-2})u_{n-1}+\\ &p_{n-2}(u_1,\dots,u_{n-2},x_0,\dots,x_{n-2}). \end{aligned} \tag{7.24}$$

Then the n*–th order divided difference*

$$\begin{aligned} &\Delta^n g(x_0,\dots,x_n)=\\ &\int_0^1\int_0^{u_1}\cdots\int_0^{u_{n-1}} g^{(n)}(p_n(u_1,\dots,u_{n-2},t,x_0,\dots,x_n))\,dt\,du_{n-1}\dots du_1, \end{aligned} \tag{7.25}$$

where $p_n(u_1,\dots,u_{n-2},t,x_0,\dots,x_n)$ *is defined as in (7.24).*

Let $x_{\min} = \min x_j,\ x_{\max} = \max x_j$. Then by the mean value Theorem there exists a point $\xi = \xi(x_0, \dots, x_n) \in [x_{\min}, x_{\max}]$ such that

$$\Delta^n g(x_0, \dots, x_n) = \frac{1}{n!} g^{(n)}(\xi). \tag{7.26}$$

Proof: We shall prove (7.25) by induction in n.
The basis of induction $k = 1$.

$$\Delta g(x_0, x_1) = \frac{g(x_1) - g(x_0)}{x_1 - x_0} = \frac{1}{x_1 - x_0} \int_{x_0}^{x_1} g'(t)\ dt$$
$$= \int_0^1 g'((x_1 - x_0)t + x_0)\ dt.$$

Theorem on intermediate value proves the basis step. The inductive step. Suppose that formula (7.25) is already proven for $k = n$. Let's prove it for $k = n + 1$. Notice that

$$\Delta^{n+1} g(x_0, \dots, x_{n-1}) = \frac{\Delta^n g(x_0, \dots, x_{n-2}, x_n) - \Delta^n g(x_0, \dots, x_{n-2}, x_{n-1})}{x_n - x_{n-1}} =$$
$$\frac{1}{x_n - x_{n-1}} \int_0^1 \int_0^{u_1} \dots \int_0^{u_{n-2}} \Big[g^{(n-1)}(p_{n-1}(u_1, \dots, u_{n-2}, t, x_0, \dots, x_{n-2}, x_n)) \tag{7.27}$$
$$- g^{(n-1)}(p_{n-1}(u_1, \dots, u_{n-2}, t, x_0, \dots, x_{n-2}, x_{n-1})) \Big]\ dt\ du_{n-2} \dots du_1 =$$
$$\int_0^1 \int_0^{u_1} \dots \int_0^{u_{n-1}} g^{(n)}(p_n(u_1, \dots, u_{n-2}, t, x_0, \dots, x_n))\ dt\ du_{n-1} \dots du_1.$$

This completes the proof of the Proposition. Q.E.D.

Corollary 7.4.4. *Let $f : B^N \to \mathbb{R}$ be a C^r smooth function, $\alpha \in \mathbb{Z}_+^N$ be a multiindex, $|\alpha| \le n$, $(x_1^0, \dots, x_1^{\alpha_1}), \dots, (x_N^0, \dots, x_N^{\alpha_N})$ be N groups of points such that for each $k = 1, \dots, N$ $0 \le i_k \le \alpha_k$, we have $(x_1^{i_1}, \dots, x_N^{i_N}) \in [-1, 1]$. Then there exists a point $\xi \in B^N$ depending on all x_i^j's such that the $|\alpha|$–th order divided difference*

$$\Delta^\alpha F(x_1^0, \dots, x_1^{\alpha_1}, \dots, x_N^0, \dots, x_N^{\alpha_N}) = \frac{1}{\alpha_1! \dots \alpha_N!}\ \frac{\partial^\alpha F(\xi)}{\partial x^\alpha}. \tag{7.28}$$

The function $\Delta^\alpha F(x_1^0, \dots, x_1^{\alpha_1}, \dots, x_N^0, \dots, x_N^{\alpha_N})$ is $C^{r-|\alpha|}$–smooth.

Proof: Inductive application of formula (7.25) and the mean value Theorem.

Chapter 8

Geometry of a Newton map and the Distortion Lemma

8.1 Motivation for the N–dimensional Distortion Lemma

The main goal of this Chapter to get estimates on distortion of a Newton map which are necessary ingredient for the Discretization Method. In the 1-dimensional case, described in details in Chapter 3, investigation of distortion properties becomes a trivial problem (see Distortion Lemma from section 3.4). The main reason for that is the dimension of the space $W_{\leq 2n-1,1}$ of polynomials of degree $2n-1$ in 1 variable is $2n$. So, even if distortion along each direction is bounded by 3, then total distortion of volumes is 3^{2n} and is an exponential function in n. However, the dimension of the space $W_{\leq 2n-1,N}$ of polynomials of degree $2n-1$ in N variables for $N>1$ is of order $(2n)^N$. Similarly to the 1–dimensional case one can show that distortion along each direction is bounded by 3^N. So, if we follow the "naive" 1–dimensional approach, then distortion is of order $3^{(2n)^N}$ and is a superexponential function in n. This is an unaffordable large for us distortion.

Sometimes when one needs to estimate the measure of a bad set, in our case the set of parameters corresponding to almost periodic nonsufficiently hyperbolic periodic points, this "bad" set is covered by balls of a certain size. If we know the number of those balls we get a required estimate on the measure. Notice now that in the $(2n)^N$–dimensional space ratio of volumes of a ball of diameter 1 and a cube with a side 1 is of order $\left\{(2n)^N\right\}^{(2n)^N}$ [San]. This is again exceeds an exponential function in n by far. To get admittable exponential in n distortion we need to study structure of the Newton map and the space of Divided Differences more closely then in the 1–dimensional case.

In what follows in this Chapter it is useful to have in mind the 1–dimensional discretization method and Distortion Lemma from chapter 3, because the goal of this Chapter is to find an appropriate extension of this Lemma.

Recall that in Section 4.4 for each integer n at the n–th stage of the induction we reduce investigation of the infinite-dimensional Hilbert Brick $HB^N(\vec{\mathbf{r}})$ to the finite-dimensional Cubic Brick of at most standard thickness

$$\begin{gathered} CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0)=\Big\{\{\vec{\varepsilon}_\alpha\}_{|\alpha|\le 2n-1}:\ \forall|\alpha|\le 2n-1,\\ |\vec{\varepsilon}_\alpha|<\lambda_{|\alpha|}\ \text{ and }\ \forall k\le 2n-2\ \ \lambda_k\ge (k+1)^{4N}\lambda_{k+1}\Big\}=\\ C^{\nu(0,N)}(\lambda_0)\times C^{\nu(1,N)}(\lambda_1)\times\cdots\times C^{\nu(2n-1,N)}(\lambda_{2n-1})\\ \subset W_{0,N}\times W_{1,N}\times\cdots\times W_{2n-1,N}, \end{gathered} \tag{8.1}$$

where $C^m(\lambda)$ is the m–dimensional cube with edge λ and $\nu(k,N)$ is dimension of the space of homogeneous N–vector polynomials $W_{k,N}$ of degree k in N variables.

On the space of parameters $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0)$ we have the product measure $\mu^N_{n,\vec{\mathbf{r}}_{2n-1}}$, defined as the Cartesian product of normalized Lebesgue measure μ^N_{k,r_k} induced by the scalar product (1.12)

$$\mu^N_{\le 2n-1,\vec{\mathbf{r}}_{2n-1}}=\times_{k=0}^{2n-1}\mu^N_{k,r_k}, \tag{8.2}$$

where μ^N_{k,r_k}, induced by the scalar product (1.12), on r_k–ball $B^N_k(r_k)$, defined in (1.13) in $W_{j,N}$.

Fix $C>0$, $\delta>0$, and $n\in\mathbb{Z}_+$. For a $C^{1+\rho}$–smooth diffeomorphism $f\in \mathrm{Diff}^{1+\rho}(B^N)$ with $\rho>0$ recall that as in the 1–dimensional case we need to fix a grid size $\tilde{\gamma_n}(C,\delta,\rho)=M_{1+\rho}^{-2n/\rho}\gamma_n^{1/\rho}(C,\delta)$ and consider the $\tilde{\gamma_n}(C,\delta,\rho)$–grid

$$\begin{gathered} \mathbb{Z}^N_{\tilde{\gamma_n}(C,\delta,\rho)}=\{(x_1,\dots,x_N)\in\mathbb{R}^N:\ \exists(k_1,\dots k_N)\in\mathbb{Z}^N\\ \text{such that }\ x_i=k_i\tilde{\gamma_n}(C,\delta,\rho)\ \text{ for each }\ i=1,\dots,N\} \end{gathered} \tag{8.3}$$

in the unit ball B^N. Then we consider all possible pairwise different n–tuples of points $x_0,\dots,x_{n-1}\subset\mathbb{Z}_{\tilde{\gamma_n}(C,\delta,\rho)}$. The key to the Discretization Method, as in the

1–dimensional case, is to define the following set of parameters

$$
\begin{aligned}
&CB_n^{N,\tilde{\gamma}_n(C,\delta,\rho)}(C,\delta,\rho,\tilde{f},\vec{\lambda}_{2n-1};x_0,\dots,x_{n-1}) = \{\vec{\varepsilon}_{2n-1} \in CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0):\\
&x_0,\dots,x_{n-1} \quad \text{is}\\
&a)\ \tilde{\gamma}_n(C,\delta,\rho)\text{-pseudotrajectory associated to } \ \vec{\varepsilon}_{2n-1},\ \text{i.e.}\\
&|(\tilde{f}_{\vec{\varepsilon}_{2n-1}}(x_j)-x_{j+1})_k| \le \tilde{\gamma}_n(C,\delta,\rho) \text{ for each } \ j=0,\dots,n-2;\\
&k=1,\dots,N;\\
&b)\ \ (n,2\gamma_n^{1/\rho}(C,\delta))\text{-periodic, i.e. } \ |(\tilde{f}_{\vec{\varepsilon}_{2n-1}}(x_{n-1})-x_0)_k| \le 2\gamma_n^{1/\rho}(C,\delta);\\
&c)\ \ \text{non-}(n,\gamma_n(C,\delta))\text{-hyperbolic, i.e. } \ d\tilde{f}_{\vec{\varepsilon}_{2n-1}}(x_{n-1})\circ\dots\circ d\tilde{f}_{\vec{\varepsilon}_{2n-1}}(x_0)\\
&\text{is non-}(n,2\gamma_n(C,\delta))\text{-hyperbolic.}\}
\end{aligned}
\tag{8.4}
$$

The goal of the rest of the Chapter is to get an upper estimate for the measure

$$
\mu^N_{\le 2n-1,\vec{\mathbf{r}}_{2n-1}}\left(CB_n^{N,\tilde{\gamma}_n(C,\delta,\rho)}(C,\delta,\rho,\tilde{f},\vec{\lambda}_{2n-1};x_0,\dots,x_{n-1})\right), \tag{8.5}
$$

which is in 1–dimensional case is relatively trivial (see Section 3.3) and obtain bounds on distortion properties of the Newton map, so that we shall be able to extract from the estimate (8.5) an estimate of the measure of all "bad" parameters with simple, almost period, nonsufficiently hyperbolic trajectories of length n.

In the 1–dimensional case we estimated the measure of type (8.5) for all possible n–tuples and then roughly speaking computed the number of all possible n–tuples associated with the corresponding Brick of parameters. Adding all the estimates we get an estimate for the measure of the "bad" parameters. In a view of the 1–dimensional Distortion Lemma there are a lot of cancellations (see Collection Lemma from Section 3.4) in this case. In the N–dimensional case the idea is similar, but realization has to be different in a view of a large dimensionality of appropriate spaces $W_{k,N}$'s and as the result a huge distortion. For simplicity we propose the reader to think that we shall estimate (8.5) for one n–tuple and compute the number of all possible n–tuples associated to the corresponding Brick of parameters. Summation of all those estimates gives an estimate for the measure of "bad" parameters.

Consider now the "Newton" family of polynomial perturbations

$$
\begin{aligned}
f_U(x) &= f(x) + \vec{u}_0 + \vec{u}_1 l_{1,0}(x-x_0) + \vec{u}_2 \prod_{j=0}^{1} l_{2,j}(x-x_j) + \dots\\
&\vec{u}_{n-1}\prod_{j=0}^{n-2} l_{n-1,j}(x-x_j) + U_{2n-1}(x-x_{n-1})\prod_{j=1}^{n-2}(l_{n-1,j}(x-x_j))^2,
\end{aligned}
\tag{8.6}
$$

where $\{l_{k,j}\}_{0\le j<k\le n-1}$ is a collection of linear functionals vanishing at 0, U_{2n-1} is an $N\times N$ matrix and $\{\vec{u}_s=(u_s^1,\dots,u_s^N)\}_{s=0}^{n-1}\subset\mathbb{R}^N$ is a set of N–dimensional vectors. Certainly this family can be rewritten is terms of an explicitly computable $(nN+N^2)$–parameter family of some standard polynomial perturbations. We shall specify this family and these functionals later. For simplicity the reader might think that these functionals measure the first coordinate of corresponding vectors from $\mathbb{R}^N$.

In Chapter 3 we estimate the measure of $U=(\vec{u}_0,\vec{u}_1,\dots,\vec{u}_{n-1},U_{2n-1})$'s corresponding to the event $CB_n^{N,\tilde{\gamma_n}(C,\delta,\rho)}(C,\delta,\rho,f,\vec{\lambda}_{2n-1};x_0,\dots x_{n-1})$ in the 1–dimensional case is shown. It is relatively easy (see Section 3.3 and figure 2). Let's briefly explain this procedure even though it is similar to the 1–dimensional, but notations are quite different.

Step 1. Determine the range and the measure of values of $\vec{u}_0$ for which x_0,x_1 is $\tilde{\gamma_n}(C,\delta,\rho)$–pseudotrajectory. The fact x_0,x_1 being $\tilde{\gamma_n}(C,\delta,\rho)$–pseudotrajectory requires the image of x_0 belong to $\tilde{\gamma_n}(C,\delta,\rho)$–cube centered at x_1, i.e.

$$|(f_U(x_0)-x_1)_k|=|(f(x_0)-x_1)_k+u_0^k|\le\tilde{\gamma_n}(C,\delta,\rho),\tag{8.7}$$

for each $k=1,\dots,N$. Since $f_U(x_0)$ is independent of $\{\vec{u}_s\}_{s=1}^{n-1}$ and U_{2n-1}, (8.7) determines the range of values of $\vec{u}_0$'s being within a certain cube in $\mathbb{R}^N$ with edge $2\tilde{\gamma_n}(C,\delta,\rho)$ and makes estimates of the corresponding measure of $\vec{u}_0$'s trivial:

$$\mu_{r_0}\{\vec{u}_0\in\mathbb{R}^N:\ \vec{u}_0\ \text{ satisfies (8.7)}\}\le\left(\frac{\tilde{\gamma_n}(C,\delta,\rho)}{r_0}\right)^N,\tag{8.8}$$

where $\mu_{r_0}=\frac{Leb_N}{2r_0}$ is the normalized by $2r_0$ Lebesgue measure on $\mathbb{R}^N$.

Step 2. Determine the range and the measure of values of $\vec{u}_1$ for which x_1,x_2 is $\tilde{\gamma_n}(C,\delta,\rho)$–pseudotrajectory. The fact x_1,x_2 being $\tilde{\gamma_n}(C,\delta,\rho)$–pseudotrajectory requires the image of x_1 belong to the $\tilde{\gamma_n}(C,\delta,\rho)$–cube centered at x_2, i.e.

$$\begin{aligned}&|(f_U(x_1)-x_2)_k|=\\&|(f(x_1)-x_2)_k+u_0^k+u_1^k l_{1,0}(x_1-x_0)|\le\tilde{\gamma_n}(C,\delta,\rho),\end{aligned}\tag{8.9}$$

for each $k=1,\dots,N$. Since $f_U(x_1)$ is independent of $\{\vec{u}_s\}_{s=2}^{n-1}$, U_{2n-1} and $\vec{u}_0$ is fixed at Step 1, (8.9) determines the range of values of $\vec{u}_1$'s being within a certain cube in $\mathbb{R}^N$ with edge $2\tilde{\gamma_n}(C,\delta,\rho)/|l_{1,0}(x_1-x_0)|$ and makes estimates of the corresponding measure of $\vec{u}_1$'s trivial:

$$\mu_{r_1}\{\vec{u}_1\in\mathbb{R}^N:\ \vec{u}_1\ \text{ satisfies (8.9)}\}\le\left(\frac{\tilde{\gamma_n}(C,\delta,\rho)}{r_1|l_{1,0}(x_1-x_0)|}\right)^N,\tag{8.10}$$

where $\mu_{r_1} = \frac{Leb_N}{2r_1}$ is the normalized Lebesgue measure on $\mathbb{R}^N$.

Steps 3 through $n-1$ goes along the same line.

Step n (Almost periodicity). Determine the range and the measure of values of $\vec{u}_{n-1}$ for which the n–tuple $x_0, x_1, \ldots, x_{n-1}$ is $\gamma_n^{1/\rho}(C,\delta)$–periodic $\tilde{\gamma}_n(C,\delta,\rho)$–pseudotrajectory. This fact requires the image of x_{n-1} belong to $\gamma_n^{1/\rho}(C,\delta)$–cube centered at x_0, i.e.

$$\begin{aligned} |(f_U(x_{n-1}) - x_0)_k| = |(f(x_{n-1}) - x_0)_k + \sum_{s=0}^{n-2} u_s^k \prod_{j=0}^{s-1} l_{s,j}(x_{n-1} - x_j) \\ + u_{n-1}^k \prod_{j=0}^{n-2} l_{n-1,j}(x_{n-1} - x_j)| \le 2\gamma_n^{1/\rho}(C,\delta), \end{aligned} \tag{8.11}$$

for each $k = 1, \ldots, N$. All the parameters $\vec{u}_0, \ldots, \vec{u}_{n-2}$ are fixed in the previous steps. Therefore, one can determine the range of values of $\vec{u}_{n-1}$'s and makes estimates of the corresponding measure of $\vec{u}_{n-1}$'s trivial:

$$\begin{aligned} \mu_{r_{n-1}}\{\vec{u}_{n-1} \in \mathbb{R}^N : \ \vec{u}_1 \ \text{ satisfies } (8.11)\} \le \\ \left(\frac{2\gamma_n^{1/\rho}(C,\delta)}{r_{n-1} \prod_{j=0}^{n-2} |l_{n-1,j}(x_{n-1} - x_j)|} \right)^N, \end{aligned} \tag{8.12}$$

where $\mu_{r_{n-1}} = \frac{Leb_N}{2r_{n-1}}$ is the normalized Lebesgue measure on $\mathbb{R}^N$.

Step n + 1 (Nonsufficient hyperbolicity). Determine the measure of values of U_{2n-1} for which the n–tuple $x_0, x_1, \ldots, x_{n-1}$ is non-$(n, \gamma_n(C,\delta))$–hyperbolic. This step is a little bit different form all the previous. Denote $U = (\vec{U}_{n-1}; U_{2n-1}) = (\vec{u}_0, \vec{u}_1, \ldots, \vec{u}_{n-1}; U_{2n-1})$. We need to consider how the linearization of f_U along this n–tuple changes with $\vec{U}_{n-1}$ fixed and U_{2n-1} varying

$$\begin{aligned} df_U(x_{n-1}) \circ df_U(x_{n-2}) \circ \cdots \circ df_U(x_0) = \\ (df_{(\vec{U}_{n-1};0)}(x_{n-1}) + U_{2n-1} \prod_{j=0}^{n-2} l_{n-1,j}(x_{n-1} - x_j)) \circ \\ \circ df_{(\vec{U}_{n-1},0)}(x_{n-1}) \circ \cdots \circ df_{(\vec{U}_{n-1},0)}(x_0). \end{aligned} \tag{8.13}$$

The last equality uses the fact that at all points $x_0, \ldots, x_{n-2}$ except the last one the "Newton" family $\{f_{(\vec{U}_{n-1},U_{2n-1})}\}_{U_{2n-1} \in [-r_{2n-1}, r_{2n-1}]^{N^2}}$, defined by (8.6), has zeroes of the second order, though $df_U(x_j)$ is independent of U_{2n-1} for each $j = 0, \ldots, n-2$.

Recall that we denote the space of $N \times N$ matrices with real coefficients by M_N. The goal now is to estimate the measure of U_{2n-1}'s from the r_{2n-1}–cube $C^{N^2}(r_{2n-1})$

centered at the origin of M_N, for which the operator from (8.13) is non-$\gamma_n(C,\delta)$–hyperbolic. Notice that the composition

$$df_{(\vec{U}_{n-1},0)}(x_{n-1}) \circ \dots \circ df_{(\vec{U}_{n-1},0)}(x_0) \tag{8.14}$$

is the composition of the linearizations of the diffeomorphism $f_{(\vec{U}_{n-1},0)}$. So, the norms of linear operators $df_{(\vec{U}_{n-1},0)}(x)$ and $df^{-1}_{(\vec{U}_{n-1},0)}(x)$ are bounded from below and above by M_1 and M_1^{-1} respectively. Therefore, the family of compositions from right-hand side of (**??**) covers $M_1^{-n} r_{2n-1}$–neighborhood in M_N and is covered by $M_1^n r_{2n-1}$–neighborhood of the composition (8.13) for $U_{2n-1} = 0$. So if amount of required hyperbolicity $\gamma_n(C,\delta)$ is significantly smaller then $M^{-n} r_{2n-1}$, then it is quite reasonable to expect that an estimate on the measure of non-$\gamma_n(C,\delta)$–hyperbolic is bounded by $2M_1^{N^2 n} r_{2n-1}^{N^2} \gamma_n(C,\delta)$ (see Lemma A.5), which is by admissibility condition 1.3.1 part C) is a very small number. This is a starting point of our construction.

Denote by $\mu_{r_0,\dots,r_{n-1},r_{2n-1};\vec{U}_{n-1},U_{2n-1}}$ the product of the normalized Lebesgue probability measures $\{\mu_{r_j}\}_{j=0}^{n-1}$ defined above. Combining estimates (8.7-8.13) we get

$$\begin{gathered} \mu_{r_0,\dots,r_{n-1},r_{2n-1};\vec{U}_{n-1},U_{2n-1}} \Big\{ \Big(\vec{U}_{n-1}, U_{2n-1}\Big) : f_{\vec{U}_{n-1},U_{2n-1}} \equiv f_{\vec{\varepsilon}} \\ \text{and } \vec{\varepsilon} \in CB_n^{N,\tilde{\gamma}_n(C,\delta,\rho)}(C,\delta,\rho,\tilde{f},\vec{\lambda}_{2n-1};x_0,\dots,x_{n-1}) \Big\} \le M_1^{4N^2} \times \\ \Big(\frac{\tilde{\gamma}_n(C,\delta,\rho)}{r_0} \; \frac{\tilde{\gamma}_n(C,\delta,\rho)}{r_1 |l_{1,0}(x_1 - x_0)|} \cdots \frac{\tilde{\gamma}_n(C,\delta,\rho)}{r_{n-2} \prod_{j=0}^{n-3} |l_{n-2,j}(x_{n-2} - x_j)|} \Big)^N \\ \Big(\frac{2\gamma^{1/\rho}(C,\delta)}{r_{n-1} \prod_{j=0}^{n-2} |l_{n-1,j}(x_{n-1} - x_j)|} \Big)^N \frac{2\gamma(C,\delta)}{r_{2n-1} \prod_{j=0}^{n-2} |l_{n-1,j}(x_{n-1} - x_j)|^2} \end{gathered} \tag{8.15}$$

Denote by $p_{l_k;x_0,\dots,x_{k-1}}(x) = \prod_{j=0}^{k-1} l_{k,j}(x - x_j)$ for $k = 1,\dots,n$ and $p^s_{x_0,\dots,x_{n-1}}(x) = (x - x_{n-1})_s \prod_{j=0}^{n-1} (l_{k,j}(x - x_j))^2$ for $s = 1,\dots,N$.

The very first problem is that a choice of functionals $\{l_{k,j}\}_{1 \le j < k \le n-1}$ for the Newton family (8.6) is a subtle question. Such a choice should satisfy two major requirements:

• for each pair of j and k with $1 \le j < k \le n-1$ the functional $l_{k,j}$ does not annihilate or nearly annihilate the vector $x_k - x_j$. More exactly, $|l_{k,j}(x_k - x_j)| \ge 1/T|x_k - x_j|$ for some constant $T > 0$.

• monomials $\big\{\{p_{l_k;x_0,\dots,x_{k-1}}\}_{k=1}^n, \{p^s_{x_0,\dots,x_{n-1}}\}_{s=1}^N\big\}$ generate the $N(n+N)$–dimensional subspace in the space $W_{\le 2n-1,N}$ of N–vector polynomials of degree $2n-1$. This subspace should be analyzable on the subject of distortion properties of the Newton map $\mathcal{L}^N_{x_0,\dots,x_{n-1}}$.

8.2 A Choice of Newton functionals

The idea of choosing functionals for the Newton family (8.6) is very natural. Let $v = (v_1, \ldots, v_N) \in \mathbb{R}^N \setminus \{0\}$ be a nonzero vector. We always define the norm in $\mathbb{R}^N$ by $|v| = \sqrt{\sum_{i=1}^N v_i^2}$. Therefore, we have the following estimate

$$\sqrt{N} \max_{1\le i\le N} |v_i| \ge |v| \ge \max_{1\le i\le N} |v_i| \tag{8.16}$$

and see that the maximal component of v differs from $|v|$ at most by a factor of $\sqrt{N}$.

According to the 1–dimensional strategy (see figure 2) we are interested in values of the product $p_{l_k;x_0,\ldots,x_{k-1}}(x) = \prod_{j=0}^{k-1} l_{k,j}(x - x_j)$ only when $x = x_k$. Define the map

$$\begin{gathered} i : \mathbb{R}^N \setminus \{0\} \to \{1, \ldots, N\} \\ i(v) = i(v_1, \ldots, v_N) \text{ if } \ |v_j| = \max_{1\le i\le N} |v_i| \ \text{ and } \ |v_j| > \max_{1\le i<j} |v_i| \end{gathered} \tag{8.17}$$

which determines the minimal index with the maximal in absolute value coordinate of v. In the generic case this simple the index of the maximal in absolute value coordinate. Denote by $m_{k,j}$ the corresponding index $i(x_k - x_j)$ for $0 \le j < k \le n-1$. It is well-defined for an n–tuple pairwise different points.

We define the *dynamical Newton* functionals $\{l^{dyn}_{k,j}\}_{k,j}$ by $l^{dyn}_{k,j}(x - x_j) = (x - x_j)_{m_{k,j}}$ for $0 \le j < k \le n-1$ and call the following monomials

$$p^{dyn}_{x_0,\ldots,x_{k-1}}(x) = \prod_{j=0}^{k-1} l^{dyn}_{k,j}(x - x_j) \quad \text{for} \quad k = 1, \ldots, n \tag{8.18}$$

$$p^{dyn,s}_{x_0,\ldots,x_{n-1}}(x) = (x - x_{n-1})_s \prod_{j=0}^{n-1} (l^{dyn}_{k,j}(x - x_j))^2 \tag{8.19}$$

by *dynamical Newton monomials*. For each $k = 1, \ldots, n-1$ one can compare absolute values of product of distances $\prod_{j=0}^{k-1} |x - x_j|$ and value of the corresponding dynamical Newton monomial $|p^{dyn}_{x_0,\ldots,x_{k-1}}(x)|$ at x_k as follows

$$\prod_{j=0}^{k-1} |x_k - x_j| \le N^{k/2} |p^{dyn}_{x_0,\ldots,x_{k-1}}(x_k)|. \tag{8.20}$$

Now we can define the Newton family which we shall use

$$\begin{aligned} f_U(x) &= f(x) + \vec{u}_0 + \vec{u}_1(x-x_0)_{m_{1,0}} + \vec{u}_2 \prod_{j=0}^{1}(x-x_j)_{m_{2,j}} + \dots \\ &\vec{u}_{n-1}\prod_{j=0}^{n-2}(x-x_j)_{m_{n-1,j}} + U_{2n-1}(x-x_{n-1})\prod_{j=1}^{n-2}\left((x-x_j)_{m_{n-1,j}}\right)^2. \end{aligned} \tag{8.21}$$

This family seems to be the most natural candidate for the Newton family in the sense that $\vec{u}_k$ moves the image of x_k nearly with the maximal speed.

The next problem in that the monomials we have chosen above are not always in the space of divided differences defined in the standard way.

8.3 A Trajectory respecting basis in the space of divided differences and dynamical Newton map.

In Section 7.1 we introduce the following basis in the space of divided differences is defined as follows: Put $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}_+^N$, $\alpha^j = \sum_{i=1}^{j} \alpha_i$. Then the monomials $\{(x; x_0, \dots, x_{|\alpha|-1})^\alpha\}_{|\alpha| \le 2n-1}$ form a basis of the space $W_{\le 2n-1,N}$ of N–vector polynomials of degree $2n-1$. We call this basis *the standard Newton basis.* Here we face the problem: this basis might be inefficient for perturbing trajectories.

Remark 8.3.1. *Let $N = n = 2$, and $x_0, x_1, x_2 \in \mathbb{R}^2$ are such that $(x_2 - x_0)_2 = 0$ and $(x_2 - x_1)_1 = 0$. If $(x, x_0, x_1)^\alpha$ vanishes at x_0 and x_1, then it also vanishes at x_2. Therefore, we are unable to move the image of x_2 within the family $\sum_{|\alpha|=2} u_\alpha (x; x_0, x_1)^\alpha$.*

This is the reason we define a new Newton basis in $W_{\le 2n-1,N}$. For an n–tuple of pairwise different points $x_0, \dots, x_{n-1}$ and $k = 1, \dots, n$ and let $\mathbf{X}_k = (x_0, \dots, x_{k-1})$ be a k–tuple of first k points of the n–tuple. Define an integer-valued vector

$$\mathcal{I}_{\mathbf{X}_k} = (i(x_k - x_0), i(x_k - x_1), \dots, i(x_k - x_{k-1})) \in \{1, \dots, N\}^{k-1} \tag{8.22}$$

and a permutation $\sigma(\mathbf{X}_k) \in S_{k-1}$, which orders components of the vector $\mathcal{I}_{\mathbf{X}_k}$ in an increasing order as follows: put

$$s_{j,\mathbf{X}_k} = \{0 \le i \le k-2 : \ p(x_{k-1} - x_i) = j\} \ \text{ for } \ j = 1, \dots, N \tag{8.23}$$

and denote by $S_{j,\mathbf{X}_k}$ the set $s_{j,\mathbf{X}_k}$ permuted in an increasing order. Then

$$\sigma(\mathbf{X}_k)\big(0,1,\ldots,k-2\big) \to \big(S_{1,\mathbf{X}_k}, S_{2,\mathbf{X}_k},\ldots,S_{N,\mathbf{X}_k}\big). \tag{8.24}$$

Now we define *a trajectory respecting Newton coordinate system* in the space $W_{\leq 2n-1,N}$ by defining the basis as follows: let $\alpha \in \mathbb{Z}_+$, $|\alpha| \leq 2n-1$, and $\alpha^j = \sum_{i=1}^{j}\alpha_i$ Then

$$\begin{aligned}(x;x_0,\ldots,x_{|\alpha|-1(\text{mod }n)})^{\alpha}_{\sigma\left(\mathbf{X}_\alpha\right)} &= \prod_{i_1=0}^{\alpha_1-1}\left(x - x_{\sigma(\mathbf{X}_\alpha)\left(i_1\right)}\right)_1 \\ \prod_{i_2=0}^{\alpha_2-1}\left(x - x_{\sigma(\mathbf{X}_\alpha)\left(i_2\right)}\right)_2 &\cdots \prod_{i_N=0}^{\alpha_N-1}\left(x - x_{\sigma(\mathbf{X}_\alpha)\left(i_N\right)}\right)_N.\end{aligned} \tag{8.25}$$

The space $W_{\leq 2n-1,N}$ with this basis is denoted by $W^{u,dyn}_{\leq 2n-1,N}(x_0,\ldots,x_{2n-1(\text{mod }n)})$ and is called *the dynamical Newton* space with *the trajectory respecting coordinate system.*

We call the Newton map

$$\mathcal{L}^{N,dyn}_{x_0,\ldots,x_{n-1}} : W_{\leq 2n-1,N} \to W^{u,dyn}_{\leq 2n-1,N}(x_0,\ldots,x_{2n-1(\text{mod }n)}) \tag{8.26}$$

implicitly defined by

$$\begin{aligned}\sum_{|\alpha|\leq 2n-1}\vec{\varepsilon}_\alpha x^\alpha &= \sum_{|\alpha|\leq 2n-1}\vec{u}_\alpha(x;x_0,\ldots,x_{|\alpha|-1})^{\alpha}_{\sigma\mathbf{X}_\alpha} \\ \mathcal{L}^{N,dyn}_{x_0,\ldots,x_{n-1}}(\{\vec{\varepsilon}_\alpha\}_{|\alpha|\leq 2n-1}) &= \{\vec{u}_\alpha\}_{|\alpha|\leq 2n-1}\end{aligned} \tag{8.27}$$

the dynamical Newton map. In the trajectory respecting coordinate system for each $k=0,\ldots,n-2$ the dynamical Newton monomial $\prod_{j=0}^{k-1}(x-x_j)_{m_{k,j}}$ from the Newton family (8.21) has a unique multiindex, denoted by $\alpha(\mathbf{X}_k)$, such that

$$\prod_{j=0}^{k-1}(x-x_j)_{m_{k,j}} = (x;x_0,\ldots,x_{k-1})^{\alpha(\mathbf{X}_k)}_{\sigma(\mathbf{X}_k)} = p^{dyn}_{x_0,\ldots,x_{k-1}}(x). \tag{8.28}$$

The last vector-monomial $(x-x_{n-1})\prod_{j=0}^{n-1}(x-x_j)^2_{m_{n-1,j}}$ from (8.21) has N multiindices corresponding to it. Denote by $\alpha_1(\mathbf{X}_n), \alpha_2(\mathbf{X}_n),\ldots,\alpha_N(\mathbf{X}_n)$ multiindices corresponding to $(x-x_{n-1})_1\prod_{j=0}^{n-1}(x-x_j)^2_{m_{n-1,j}}$, $(x-x_{n-1})_2\prod_{j=0}^{n-1}(x-x_j)^2_{m_{n-1,j}},\ldots,(x-x_{n-1})\prod_{j=0}^{n-1}(x-x_j)^2_{m_{n-1,j}}$.

Definition 8.3.2. *Multiindices $\alpha(\mathbf{X}_0),\ldots,\alpha(\mathbf{X}_{n-1}),\{\alpha_i(\mathbf{X}_n)\}_{i=1}^{N}$, defined above, are called dynamically essential multiindices and the $(nN+N^2)$–dimensional plane generated by the corresponding monomials in the space $W^{dyn,u}_{\leq 2n-1}(x_0,\ldots,x_{2n-1(\text{mod }n)})$ is* *called the plane of dynamically essential coordinates and is denoted by* $D^{ess}_{x_0,\dots,x_{n-1}}$*. The orthogonal complement is called a plane of nonessential coordinates and is denoted by* $D^{non}_{x_0,\dots,x_{n-1}}$*. Therefore, we have*

$$W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{n-1}) = D^{ess,dyn}_{x_0,\dots,x_{n-1}} \bigoplus D^{non,dyn}_{x_0,\dots,x_{n-1}}. \tag{8.29}$$

Denote by

$$\begin{aligned} \pi^{ess,dyn}_{x_0,\dots,x_{n-1}} &: W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{n-1}) \to D^{ess,dyn}_{x_0,\dots,x_{n-1}} \\ \pi^{non,dyn}_{x_0,\dots,x_{n-1}} &: W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{n-1}) \to D^{non,dyn}_{x_0,\dots,x_{n-1}} \end{aligned} \tag{8.30}$$

the natural projections along the complement.

For any n–tuple of points $\{x_k\}_{k=0}^{n-1}$ there is a natural one-to-one correspondence between the standard basis $\{x^\alpha\}_{|\alpha|\le 2n-1}$ in the space of N–vector polynomials $W_{\le 2n-1,N}$ of degree $2n-1$ and the dynamical Newton basis $\{(x; x_0,\dots,x_{|\alpha|-1 (\text{mod } n)})^\alpha_\sigma\}_{|\alpha|\le 2n-1}$ in the same space, but denoted by $W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{2n-1(\text{mod } n)})$

$$x^\alpha \quad \leftrightarrow \quad (x; x_0,\dots,x_{|\alpha|-1(\text{mod } n)})^\alpha_{\sigma(\mathbf{X}_{|\alpha|})}\}_{|\alpha|\le 2n-1}. \tag{8.31}$$

Therefore, for any sequence of positive numbers $\vec{\mathbf{r}} = \{r_k\}_{k=0}^\infty$ one can define the product measure $\mu^{N,dyn}_{\le 2n-1,\vec{\mathbf{r}}}$ on $W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{n-1})$ using the definition of the measure $\mu^N_{\le 2n-1,\vec{\mathbf{r}}}$ on $W_{\le 2n-1,N}$ by identifying coordinates with the same multiindex α. Denote by $\mu^{N,dyn,ess}_{\le 2n-1,\vec{\mathbf{r}}}$ and $\mu^{N,dyn,non}_{\le 2n-1,\vec{\mathbf{r}}}$ the measures induced by $\mu^{N,dyn}_{\le 2n-1,\vec{\mathbf{r}}}$ and the projections $\pi^{ess,dyn}_{x_0,\dots,x_{n-1}}$ and $\pi^{non,dyn}_{x_0,\dots,x_{n-1}}$.

Consider the Cubic Brick of at most standard thickness $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1}, 0)$ associated to the initial Brick $HB^N_{\le 2n-1}(\vec{\mathbf{r}})$ as it is defined in (**??**). Consider the set of family dynamically allowed parameters

$$\begin{aligned} \mathcal{P}^{N,st,dyn}_{\le 2n-1,\mathbf{X}_n}(\vec{\lambda}_{2n-1}, 0) = \mathcal{L}^{N,dyn}_{x_0,\dots,x_{n-1}} &\left(CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1}, 0) \right) \\ &\subset W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{2n-1(\text{mod } n)}) \end{aligned} \tag{8.32}$$

associated to the Cubic Brick $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1}, 0)$ and the n–tuple $x_0,\dots,x_{n-1}$ defined in (4.13). As a matter of fact the set $\mathcal{P}^{N,st,dyn}_{\le 2n-1,\mathbf{X}_n}(\vec{\lambda}_{2n-1}, 0)$ is the same Cubic Brick $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1}, 0)$ considered in a different coordinate system. However, the set of family allowed parameters

$$\mathcal{P}^{N,st,dyn}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1}, 0) \subset W^{u,dyn}_{\le k,N}(x_0,\dots,x_k), \tag{8.33}$$

defined as a projection of the parallelepiped $\mathcal{P}^{N,st,dyn}_{\le 2n-1,\mathbf{X}_n}(\vec{\lambda}_{2n-1},0)$ along the complement $W^{u,dyn}_{\le k,N}(x_0,\dots,x_k)$ naturally embedded in $W^{u,dyn}_{\le 2n-1,N}(x_0,\dots,x_{2n-1(\text{mod } n)})$ onto $W^{u,dyn}_{\le k,N}(x_0,\dots,x_k)$ is a parallelepiped, which is not necessarily rectangular.

Define the sets of essential and nonessential allowed parameters by projecting the set $\mathcal{P}^{N,st,dyn}_{\le 2n-1,\mathbf{X}_k}(\vec{\lambda}_{2n-1},0)$ onto planes $\pi^{dyn}_{\le 2n-1,k}\circ D^{ess,dyn}_{x_0,\dots,x_{n-1}}$ and $\pi^{dyn}_{\le 2n-1,k}\circ D^{non,dyn}_{x_0,\dots,x_{n-1}}$ of essential and nonessential parameters inside the space of parameters of order k, i.e.

$$\begin{aligned}\mathcal{P}^{N,st,dyn,ess}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1},0) &= \pi^{ess,dyn}_{x_0,\dots,x_{n-1}}\left(\mathcal{P}^{N,st,dyn}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1},0)\right)\\ \mathcal{P}^{N,st,dyn,non}_{\le 2n-1,\mathbf{X}_k}(\vec{\lambda}_{2n-1},0) &= \pi^{non,dyn}_{x_0,\dots,x_{n-1}}\left(\mathcal{P}^{N,st,dyn}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1},0)\right)\end{aligned} \tag{8.34}$$

The N–dimensional Distortion Lemma. *Let $\{x_0,\dots,x_{n-1}\}\subset B^N$ be an n–tuple of points in the unit ball B^N and $\mathcal{L}^{N,dyn}_{\mathbf{X}_n}: W_{\le 2n-1,N}\to W^u_{\le 2n-1,N}$ be the dynamical Newton map. Choose $k=0,\dots,2n-1$. Let $\vec{\lambda}_{2n-1}=(\lambda_0,\dots,\lambda_{2n-1})\in\mathbb{R}^{2n}_+$ be a vector with strictly positive components such that it defines the Cubic Brick of $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0)$ of at most standard thickness. Then with the notations above for each $k=0,\dots,2n-1$ we have*

$$\begin{aligned}\frac{\mu^{N,dyn,ess}_{\le 2n-1,\vec{\mathbf{r}}}\left(\mathcal{P}^{N,st,dyn,ess}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1})\right)}{\mu^{N,dyn,ess}_{\le 2n-1,\vec{\mathbf{r}}}\left[\pi^{ess,dyn}_{\le 2n-1,k}\circ\pi^{ess,dyn}_{x_0,\dots,x_{n-1}}\circ CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0)\right]} &\le 3^{N^2}\\ \frac{\mu^{N,dyn,non}_{\le 2n-1,\vec{\mathbf{r}}}\left(\mathcal{P}^{N,st,dyn,non}_{\le 2n-1,\le k,\mathbf{X}_k}(\vec{\lambda}_{2n-1})\right)}{\mu^{N,dyn,non}_{\le 2n-1,\vec{\mathbf{r}}}\left[\pi^{ess,dyn}_{\le 2n-1,k}\circ\pi^{non,dyn}_{x_0,\dots,x_{n-1}}\circ CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1},0)\right]} &\le 3^{N^3}\end{aligned} \tag{8.35}$$

Application of the N–dimensional Distortion Lemma provides an important ingredient for the Discretization Method:

Corollary 8.3.3. *With the notation introduced above the measure of the set of "bad" parameters $CB^{N,\tilde{\gamma}_n(C,\delta,\rho)}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};x_0,\dots,x_{n-1})\subset CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)$ corresponding to a fixed $\tilde{\gamma}_n(C,\delta,\rho)$–pseudotrajectory $x_0,\dots,x_{n-1}\subset\mathbb{Z}^N_{\tilde{\gamma}_n(C,\delta,\rho)}$ admits the following estimate*

$$\begin{aligned}\frac{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\left\{CB^{N,\tilde{\gamma}_n(C,\delta,\rho)}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f});x_0,\dots,x_{n-1})\right\}}{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\{CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)\}} &\le M_1^{N^2}\\ \left(\tilde{\gamma}_n(C,\delta,\rho)\ \frac{\tilde{\gamma}_n(C,\delta,\rho)}{|p^{dyn}_{x_0}(x_1)|}\cdots\frac{\tilde{\gamma}_n(C,\delta,\rho)}{|p^{dyn}_{x_0,\dots,x_{n-3}}(x_{n-2})|}\right)^N&\\ N^{2nN}\ 3^{3nN^4}\frac{\gamma_n^{N/\rho+1}(C,\delta)}{\prod_{j=0}^{n-2}|x_{n-1}-x_j|^{2+N}},&\end{aligned} \tag{8.36}$$

where $\{p^{dyn}_{x_0,\ldots,x_{k-1}}(x)\}_{k=1}^{n-2}$ *are dynamical Newton monomials, defined in (8.18).*

This Corollary is an indispensable part of the Discretization Method.

Chapter 9

The Discretization Method for N–dimensional Diffeomorphisms

In this chapter we unite efforts of all the previous chapters and complete the proof of the Main Result (Theorem 1.3.7). Recall our set up. We start with a $C^{1+\rho}$–diffeomorphism $f : B^N \to B^N$ of the closed unit ball B^N into its interior for some positive ρ. Consider a Hilbert Brick $HB^N(\vec{\mathbf{r}})$ an admissible size $\vec{\mathbf{r}}$ with respect to f (see (1.14) and definition 1.3.1). This Hilbert Brick corresponds to a family of analytic functions on the closed unit ball B^N. We choose $HB^N(\vec{\mathbf{r}})$ so that we have a family of diffeomosprisms

$$f_{\vec{\varepsilon}}(x) = f(x) + \phi_{\vec{\varepsilon}}(x) = f(x) + \sum_{\alpha \in \mathbb{Z}_+^N} \vec{\varepsilon}_\alpha x^\alpha, \quad \vec{\varepsilon} = \{\vec{\varepsilon}_\alpha\}_{\alpha \in \mathbb{Z}_+^N} \in HB^N(\vec{\mathbf{r}}). \tag{9.1}$$

Define C^1 and $C^{1+\rho}$–norms M_1 and $M_{1+\rho}$ as in (2.2) respectively for this family. Let $C > \frac{8}{\rho} \ln M_{1+\rho}$ and $D = M_{1+\rho}^{8/\rho}$. Fix a positive δ and denote $\gamma_n(C, \delta) = \exp(-Cn^{1+\delta})$ for $n \in \mathbb{Z}_+$. Our goal is to estimate the measure of "bad" parameters $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f) \subset HB^N(\vec{\mathbf{r}})$, defined in (2.1). Recall that $f \in \text{Diff}^{1+\rho}(B^N)$ satisfies the Inductive Hypothesis of order n denoted $f \in IH(n, C, \delta, \rho)$, if for all$k \leq n$, every $(k, \gamma_k^{1/\rho}(C, \delta))$–periodic point is $(k, \gamma_k(C, \delta))$–hyperbolic. So $B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)$ consists of parameters $\vec{\varepsilon}$ with $f_{\vec{\varepsilon}}$ in $IH(n-1, C, \delta, \rho)$, but not in $IH(n-1, C, \delta, \rho)$.

As we pointed out in Chapter 2 the key to the Proof is to find an upper estimate $\mu_n(C, \delta, \rho, \vec{\mathbf{r}}, M_{1+\rho})$ for the measure of the "bad" set $\mu_{\vec{\mathbf{r}}}^N \{B_n(C, \delta, \rho, \vec{\mathbf{r}}, f)\}$. In order to deal with a finite-dimensional set of parameters in section 2.3 we use simple Fubini/Tonelli arguments and reduce consideration to the finite-parameter family

$HB^N_{\le 2n-1}(\vec{\mathbf{r}})$

$$f_{\vec{\varepsilon}}(x) = f(x) + \phi_{\vec{\varepsilon}_{\le 2n-1}}(x) + \phi_{\vec{\varepsilon}_{>2n-1}}(x) = \\ f(x) + \sum_{|\alpha|\le 2n-1} \vec{\varepsilon}_\alpha x^\alpha + \sum_{|\alpha|>2n-1} \vec{\varepsilon}_\alpha x^\alpha, \tag{9.2}$$

where $\varepsilon_{\le 2n-1} \in HB^N_{\le 2n-1}(\vec{\mathbf{r}})$ varies and $\varepsilon_{>2n-1} \in HB^N_{>2n-1}(\vec{\mathbf{r}})$ is fixed. At this point we need to estimate the measure of "bad" paratemers $B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{>2n-1})$ inside a finite-parameter slice $HB^N_{\le 2n-1}(\vec{\mathbf{r}}) \times \{\varepsilon_{>2n-1}\}$ uniformly over $\varepsilon_{>2n-1} \in HB^N_{>2n-1}(\vec{\mathbf{r}})$. For convenience of notations we use in (2.33) redenote $B_n(C,\delta,\rho,\vec{\mathbf{r}},f,\vec{\varepsilon}_{>2n-1})$ by $B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f})$.

Now to estimate the measure of "bad" parameters $B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f})$ we need to apply the Discretization Method. However, we can not apply this method to an arbitrary Brick $HB^N_{\le 2n-1}(\vec{\mathbf{r}})$ of parameter. So as it is described in section 4.4 to be able to apply the Discretization Method and the N–dimensional Distortion Lemma from Chapter 8, in particular, we cover the Brick $HB^N_{\le 2n-1}(\vec{\mathbf{r}})$ by Cubic Bricks $\{CHB^{N,st}_{\le 2n-1}(\vec{\varepsilon}^{\,j}_{\le 2n-1},\vec{\lambda}(\vec{\mathbf{r}}))\}_{j\in J(\vec{\mathbf{r}})}$ of at most standard thickness. Then estimate the measure of "bad" parameters $B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f})$ located inside of each of the Cubic Bricks $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1})$ uniformly in $j \in J(\vec{\mathbf{r}})$ and combine all the estimates.

Now fix $j \in J(\vec{\mathbf{r}})$ and let $\tilde{f}_j(x) = \tilde{f}(x) + \sum_{|\alpha|\le 2n-1} \varepsilon^j_\alpha x^\alpha$, where $\vec{\varepsilon}^{\,j}_{\le 2n-1} = \{\vec{\varepsilon}^{\,j}_\alpha\}_{|\alpha|\le 2n-1}$. Denote by

$$CB^N_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) = B_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}) \cap CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1}). \tag{9.3}$$

At this point the goal of the proof is reduced to estimating ratio of the measures of the "bad" set inside the Cubic Brick $\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\Big\{CB_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)\Big\}$ and the Cubic Brick itself $\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\Big\{CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1}),\Big\}$ uniformly over $j \in J(\vec{\mathbf{r}})$. Then the sum over all $j \in J(\vec{\mathbf{r}})$ and application of the Covering Lemma from 4.4 completes the proof (see (4.20)).

To estimate the measure of "bad" parameters $CB_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$ we decompose it into two parts simple and nonsimple as in section 2.4 formulas (2.35-2.37). Denote the two sets of "bad" parameters we obtain

$$\begin{aligned} CB^{N,sim}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) &= B^{sim}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) \cap CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1}) \\ CB^{N,non}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) &= B^{non}_n(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) \cap CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1}) \end{aligned} \tag{9.4}$$

with nonhyperbolic either simple or nonsimple trajectories respectively.

Now we apply the Shift Theorem proved in Chapter 5. Let $D = M_{1+\rho}^{8/\rho}$ and $\sigma \geq 12$. Then define for each $k|n$ sets of "bad" parameters

$$\begin{aligned} CB_n^{N,gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma) = \\ B_n^{gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma) \cap CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1}) \end{aligned} \tag{9.5}$$

where $B_n^{gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma)$ is defined in (3.70) and we have an analog of (3.71)

$$CB_n^{N,non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) \subseteq \cup_{k|n} CB_n^{N,gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma). \tag{9.6}$$

Now we follow the strategy of the 1-dimensional case developed in Chapter 3. We fix a cubic brick $CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1})$ or an inxed $j \in J(\vec{\mathbf{r}})$. First, we estimate the measure of "simple bad" parameters $CB_n^{N,sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$. Second, we estimate the measure of "nonsimple bad" parameters $CB_n^{N,non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$, which would complete the proof.

9.1 Estimates of the measure of "bad" parameters for "one" trajectory.

Similarly to the 1-dimensional case choose $\gamma_n(C,\delta,\rho) = M_{1+\rho}^{-2n/\rho}\gamma_n^{1/\rho}(C,\delta)$ and consider $\gamma_n(C,\delta,\rho)$–grid $\mathbb{Z}^N_{\gamma_n(C,\delta,\rho)} \subset \mathbb{R}^N \supset B^N$ with respect to the orthogonal coordinate system fixed in the introduction, i.e.

$$\begin{aligned} \mathbb{Z}^N_{\gamma_n(C,\delta,\rho)} = \{x = (x_1,\dots,x_N) \in \mathbb{R}^N :\ \exists\ (k_1,\dots,k_N) \in \mathbb{Z}^N \\ \text{such that}\ \ x_j = k_j\gamma_n(C,\delta,\rho)\ \ \text{for each}\ \ j = 1,\dots,N\}. \end{aligned} \tag{9.7}$$

The grid $\mathbb{Z}^N_{\gamma_n(C,\delta,\rho)}$ should be small enough to recover from hyperbolicity and periodicity properties of $\gamma_n(C,\delta,\rho)$–pseudotrajectories similar properties of approximated trajectories. For brevity denote $\gamma_n(C,\delta,\rho)$ by $\tilde{\gamma}_n$.

Definition 9.1.1. *We call a k–tuple $x_0,\dots,x_{k-1} \subset \mathbb{Z}^N_{\tilde{\gamma}_n}$ a $\tilde{\gamma}_n$–pseudotrajectory associated to $\vec{\varepsilon}$ (or to the map $\tilde{f}_{\vec{\varepsilon}}$) if for each $j = 1,\dots,k-1$ we have $|\tilde{f}_{\vec{\varepsilon}}(x_{j-1}) - x_j| < \tilde{\gamma}_n$ and a $\tilde{\gamma}_n$–pseudotrajectory associated to $CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1})$ (or to the family $\{\tilde{f}_{\vec{\varepsilon}}\}_{\vec{\varepsilon} \in CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1})}$) if it is associated to some $\varepsilon \in CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1})$.*

A $\tilde{\gamma}_n$–pseudotrajectory $x_0,\dots,x_{n-1}$ of length n associated to some parameter $\vec{\varepsilon} \in CHB_{\leq 2n-1}^{N,st}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\leq 2n-1})$ and for some $\gamma > 0$ is

- *(n,γ)–periodic if $|\tilde{f}_{\vec{\varepsilon}}(x_{n-1}) - x_0| \leq \gamma$,*
- *(n,γ)–simple if $\prod_{j=0}^{n-2}|x_{n-1} - x_j| \geq \gamma^{1/4N}$,*
- *(n,γ)–hyperbolic if $\gamma\big(\prod_{j=0}^{n-1}\ d\tilde{f}_{\vec{\varepsilon}}(x_j)\big) \geq \gamma$.*

We would like to decompose the set of "bad" parameters $CB_n^{N,sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$ into a finite collection of subsets each of "bad" parameters corresponding to a single $\tilde{\gamma}_n$–pseudotrajectory.

$$\begin{aligned} CB_n^{N,sim,\tilde{\gamma}_n}(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{n-1}) = \\ \{\vec{\varepsilon}\in CHB^{N,st}_{\le 2n-1}(\vec{\lambda}_{2n-1}(\vec{\mathbf{r}}),0): \\ \{x_k\}_{k=0}^{n-1}\in\mathbb{Z}^N_{\tilde{\gamma}_n} \text{ is a } \tilde{\gamma}_n\text{–pseudotrajectory associated to } \vec{\varepsilon} \\ \text{and is } \left(n,\frac{\gamma_n}{2}\right)\text{–simple and } (n,2\gamma_n^{1/\rho})\text{–periodic,} \\ \text{but not } (n,2\gamma_n))\text{-hyperbolic}\} \end{aligned} \tag{9.8}$$

By Corollary 8.3.3 from the N–dimensional Distortion Lemma we obtain bound on ratio the $\mu^N_{\le 2n-1,\vec{\mathbf{r}}}$–measures of $CB_n^{N,\tilde{\gamma}_n}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;x_0,\dots,x_{n-1})$ and $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1})$ given by (8.36).

9.2 Decomposition into pseudotrajectories

In this Section as in the 1–dimensional case (Section 3.2) we decompose the set of all "bad" parameters $CB_n^{N,sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$ for which there exists a simple, almost periodic, but not sufficiently hyperbolic trajectory into a finite union of sets of "bad" parameters associated with a particular simple, almost periodic, but not sufficiently hyperbolic pseudotrajectory and then using this decomposition we extend the estimate (8.36) of the measure of "bad" parameters associated with a particular trajectory $\{x_k\}_{k=0}^{n-1}$ to "bad" parameters associated with all possible simple trajectories and get estimate on ratio

$$\frac{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\left\{CB_n^{N,\tilde{\gamma}_n}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)\right\}}{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\left\{CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),\vec{\varepsilon}^{\,j}_{\le 2n-1})\right\}} \le \mu_n^{sim}(C,\delta,\rho,\vec{\mathbf{r}},M_{1+\rho}). \tag{9.9}$$

The goal of this Section is to prove the following

Theorem 9.2.1. *In notations above the following estimate holds true*

$$\mu_n^{sim}(C,\delta,\rho,\vec{\mathbf{r}},M_{1+\rho}) \le \frac{M_1^{N^2}M_{1+\rho}^{2Nn/\rho}3^{4nN}\gamma_n^{(3N-2)/(4N)}}{r_{n-1}^N r_{2n-1}^{N^2+1}}. \tag{9.10}$$

Introduce the union of all "bad" sets associated with $\tilde{\gamma}_n$–pseudotrajectories

$$\begin{array}{c} CB_n^{N,sim,\tilde{\gamma}_n}(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho}) = \\ \cup_{\{x_0,\dots,x_{n-1}\}\subset\mathbb{Z}^N_{\tilde{\gamma}_n}} CB_n^{N,sim,\tilde{\gamma}_n}(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{n-1}). \end{array} \tag{9.11}$$

Most of the sets in the right-hand side are empty and one of our goals is to determine and estimate the number of nonempty ones. Since we increase periodicity and hyperbolicity for pseudotrajectories and decrease simplicity we shall be able to prove that

$$CB_n^{N,sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j) \subset CB_n^{N,sim,\tilde{\gamma}_n}(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho}). \tag{9.12}$$

Intuitively this is true because each trajectory of length n can approximated sufficiently well by a pseudotrajectory of length n which has almost the same periodicity, simplicity, and hyperbolicity as the original one.

"Roughly speaking" ratio (9.9) will be estimated in two steps as in the 1–dimensional case:

Step 1. Estimate the number of different $\tilde{\gamma}_n$–pseudotrajectories $\#_n(\tilde{\gamma}_n,,\vec{\lambda}(\vec{\mathbf{r}}))$ associated to $CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)$;

Step 2. Estimate (8.36) for ratio corresponding to the measure for one pseudotrajectory uniformly over all $(n,\gamma_n(C,\delta)/2)$–simple $\tilde{\gamma}_n$–pseudotrajectories $x_0,\dots,x_{n-1}\subset \mathbb{Z}^N_{\tilde{\gamma}_n}$ or simply make sure that (8.36) is indeed uniform.

Then the product of two numbers $\#_n(\tilde{\gamma}_n,\vec{\lambda}(\vec{\mathbf{r}}))$ and(8.36) which are obtained in Steps 1 and 2 gives the required estimate (9.9). Even though arguments below look very similar to the 1–dimensional case some special case has to be taken, while choosing dynamically essential parameters.

It turns out that in N–dimensional case up similarly to the 1–dimensional case we have that up to an exponential in n factor there is the estimate (3.24) of the measure of "bad" set associated to simple, almost periodic, not sufficiently hyperbolic trajectory. In (3.24) we need to replace $\mathbb{Z}_{\tilde{\gamma}_n}$ by $\mathbb{Z}^N_{\tilde{\gamma}_n}$.

Heuristic arguments are similar to the one in the 1–dimensional case. We shall see that the first term is of order $\tilde{\gamma}_n^{-N}$, the second term is of order $r_{n-1}^{-N}\frac{\gamma_n(C,\delta)}{\prod_{j=0}^{n-2}|x_{n-1}-x_j|}\le r_{n-1}^{-N}\gamma_n^{3/4}(C,\delta)$, and the third ones of order $r_{2n-1}^{-N^2}\frac{\gamma_n}{\prod_{j=0}^{n-2}|x_{n-1}-x_j|^2}\le r_{2n-1}^{-N^2}\gamma_n^{1/2}(C,\delta)$ (compare with second and third with (3.30) and (3.33) respectively). So the product up to an exponential in n factor is of order $\gamma_n^{1/4}(C,\delta)r_{n-1}^{-N}r_{2n-1}^{-N^2}$ and is superexponentially small.

9.3 Newton interpolation polynomials and the N–dimensional Collection Lemma

In this Section we formulate an auxiliary Lemma about the dynamic Newton map $\mathcal{L}^{N,dyn}_{x_0,\dots,x_{n-1}}$ and complete estimation of the measure of all "bad" parameters with a simple, almost periodic, but not sufficiently hyperbolic trajectory (9.9) by collecting all possible "bad" pseudotrajectories (see the N–dimensional Collection Lemma below).

To collect all possible within the family $\{\tilde{f}_{\vec{\varepsilon}^{j}_{\le 2n-1}+\vec{\varepsilon}}\}_{\vec{\varepsilon}\in CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)}$ as in the 1–dimensional case we need to have distortion estimates for the Newton map. Recall that we use not the Newton map, but the dynamical Newton map, defined in (8.26).

$$
\begin{gathered}
f_{\vec{\varepsilon}^{j}_{\le 2n-1}+\vec{\varepsilon}}(x) = f_{\vec{\varepsilon}^{j}_{\le 2n-1}}(x) + \sum_{|\alpha|\le 2n-1} \vec{\varepsilon}_\alpha x^\alpha = \\
f_{\vec{\varepsilon}^{j}_{\le 2n-1}}(x) + \sum_{|\alpha|\le 2n-1} \vec{u}_\alpha \; (x; x_0, \dots, x_{|\alpha|-1})^{\alpha}_{\sigma(\mathbf{X}_\alpha)} \qquad (9.13) \\
\mathcal{L}^{N,dyn}_{x_0,\dots,x_{n-1}}(\{\vec{\varepsilon}_\alpha\}_{|\alpha|\le 2n-1}) = \{\vec{u}_\alpha\}_{|\alpha|\le 2n-1} \quad \text{and} \quad \vec{\varepsilon}\in CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0).
\end{gathered}
$$

For each $m<n$ using the trajectory respecting coordinate system (8.25) in the dynamical Newton space and essential monomials (8.28) we introduce the family of diffeomorphisms $g_{\vec{u}^{ess}_{\le 2n-1,\le m,\mathbf{X}_m}}$

$$
g_{\vec{u}^{ess}_{\le 2n-1,\le m,\mathbf{X}_m}}(x) = g(x) + \sum_{s=0}^{m} \vec{u}_{\alpha(\mathbf{X}_s)} \; (x; x_0, \dots, x_{s-1})^{\alpha(\mathbf{X}_s)}_{\sigma(\mathbf{X}_s)} \qquad (9.14)
$$

where $\mathbf{X}_s=(x_0,\dots,x_{s-1})$ is the initial segment of $\mathbf{X}_n$ oflength s and $\vec{u}^{ess}_{\le 2n-1,\le m,\mathbf{X}_m} = \{\vec{u}_{\alpha(\mathbf{X}_s)}\}_{s=0}^{m}$.

To define more precisely where $\vec{u}^{ess}_{\le 2n-1,\le m,\mathbf{X}_m}$ and $\vec{u}^{non}_{\le 2n-1,\le m,\mathbf{X}_m}$ belong to consider first the case $m=n-1$. In notations of the N–dimensional Distortion Lemma we have $\vec{u}^{ess}_{\le 2n-1,\le n-1,\mathbf{X}_{n-1}} \in D^{ess,dyn}_{x_0,\dots,x_{n-1}}$. It is also natural to denote by $\vec{u}^{non}_{\le 2n-1,\le n-1,\mathbf{X}_{n-1}} \in D^{non,dyn}_{x_0,\dots,x_{n-1}}$ the set of all dynamically nonessential monomials. Recall that the sets of parameters $\mathcal{P}^{N,st,dyn,ess}_{\le 2n-1,\mathbf{X}_n}(\vec{\lambda}(\vec{\mathbf{r}}),0) \subset D^{ess,dyn}_{x_0,\dots,x_{n-1}}$ and $\mathcal{P}^{N,st,dyn,non}_{\le 2n-1,\mathbf{X}_n}(\vec{\lambda}(\vec{\mathbf{r}}),0) \subset D^{non,dyn}_{x_0,\dots,x_{n-1}}$, defined in (8.32), describe the range of family allowed values for $\vec{u}^{ess}_{\le 2n-1,\le n-1,\mathbf{X}_{n-1}}$ and $\vec{u}^{non}_{\le 2n-1,\le n-1,\mathbf{X}_{n-1}}$ within the family (9.13) respectively.

In the case $m<n-1$ situation a bit more complicated. One needs to recall the sets of parameters $\mathcal{P}^{N,st,dyn,ess}_{\le 2n-1,\le m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0)$ and $\mathcal{P}^{N,st,dyn,non}_{\le 2n-1,\le m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0)$, defined by

(8.34). Notice that these sets depends only on $\mathbf{X}_m$ and are independent of $\mathbf{X}_{n-1} \setminus \mathbf{X}_m$. Then we have

$$\begin{aligned} \vec{u}^{ess}_{\leq 2n-1,\leq m,\mathbf{X}_m} &\in \mathcal{P}^{N,st,dyn,ess}_{\leq 2n-1,\leq m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0) \\ \vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m} &\in \mathcal{P}^{N,st,dyn,non}_{\leq 2n-1,\leq m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0). \end{aligned} \tag{9.15}$$

Let $\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m} \in \mathcal{P}^{N,st,dyn,non}_{\leq 2n-1,\leq m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0)$ be the set of nonessential parameters in the sense defined above. Then we introduce the set of the *family (9.13) essential allowed* Newton parameters $\vec{u}^{ess}_{\leq 2n-1,\leq m,\mathbf{X}_m} = \big(\vec{u}_{\alpha(\mathbf{X}_0)},\dots,\vec{u}_{\alpha(\mathbf{X}_m)}\big)$ for which the corresponding diffeomorphism $g_{\vec{u}^{ess}_{\leq 2n-1,\leq m,\mathbf{X}_m},\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m}}$ has a certain $\tilde{\gamma}_n$–pseudotrajectory $x_0,\dots,x_{m+1}$. Namely,

$$\begin{aligned} T^{N,\tilde{\gamma}_n}_{\leq 2n-1,\leq m,\vec{\mathbf{r}}}(\tilde{f}_j;x_0,\dots,x_{m-1},x_m,x_{m+1};\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m}) = \\ \Big\{\vec{u}^{ess}_{\leq 2n-1,\leq m,\mathbf{X}_m} \in \mathcal{P}^{st}_{\leq 2n-1,\leq m,\mathbf{X}_m}(\vec{\lambda}(\vec{\mathbf{r}}),0) \subset W^{u,dyn}_{\leq m,N}: \\ |\tilde{f}_{j,\vec{u}^{ess}_{\leq 2n-1,\leq m,\mathbf{X}_m},\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m}}(x_{j-1})-x_j| \leq \tilde{\gamma}_n \text{ for } j=1,\dots,m+1\Big\}. \end{aligned} \tag{9.16}$$

Notice that the set of parameters $T^{N,\tilde{\gamma}_n}_{\leq 2n-1,\leq m,\vec{\mathbf{r}}}(\tilde{f}_j;x_0,...,x_{m-1},x_m,x_{m+1};\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m})$ does not depend on u_α's with $|\alpha|>m$, because any $(x;x_0,\dots,x_{|\alpha|-1})^\alpha$ has to vanish at all x_j's for $j=0,\dots,m-1$ so the image of each such x_j stays the same.

Based on the definition of the set $CB^{N,sim,\tilde{\gamma}_n}_n(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{n-1})$ of parameters $\vec{\varepsilon}$ for which the corresponding diffeomorphism $f_{\vec{\varepsilon}}$ has a prescribed $\tilde{\gamma}_n$–pseudotrajectory $x_0,\dots,x_{n-1} \subset \mathbb{Z}_{\tilde{\gamma}_n}$ that is almost periodic and not sufficiently hyperbolic, define a set of parameters $\vec{\varepsilon}$ for which the diffeomorphism $f_{\vec{\varepsilon}}$ has only a part of $\tilde{\gamma}_n$–pseudotrajectory $x_0,\dots,x_{m-1} \subset \mathbb{Z}^N_{\tilde{\gamma}_n}$ prescribed

$$\begin{aligned} CB^{N,sim,\tilde{\gamma}_n}_n(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{m-1}) = \{\vec{\varepsilon} \in CHB^{N,st}_{\leq 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)\}: \\ \text{there is a } \tilde{\gamma}_n - \text{pseudotrajectory } \{x_j\}_{j=0}^{m-1} \cup \{x_j\}_{j=m}^{n-1} \subset \mathbb{Z}^N_{\tilde{\gamma}_n} \\ \text{such that } \vec{\varepsilon} \in CB^{N,sim,\tilde{\gamma}_n}_n(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{n-1})\}. \end{aligned} \tag{9.17}$$

The N–dimensional Collection Lemma. *With the notations above for any $m=0,\dots,n-2$ ratio the $\mu^N_{\leq 2n-1,\vec{\mathbf{r}}}$–measure of the discretized "bad" parameters $CB^{N,sim,\tilde{\gamma}_n}_n(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{m-1})$ and $CHB^{N,st}_{\leq 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)$ satisfies*

$$\begin{aligned} &\frac{\mu^N_{\leq 2n-1,\vec{\mathbf{r}}}\big\{CB^{N,sim,\tilde{\gamma}_n}_n(\gamma_n,\vec{\mathbf{r}},\tilde{f}_j,M_{1+\rho};x_0,\dots,x_{m-1})\big\}}{\mu^N_{\leq 2n-1,\vec{\mathbf{r}}}\big\{CHB^{N,st}_{\leq 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)\big\}} \\ &\leq 3^{3nN^4+N^3(n-m-1)} M_1^{N^2}\ r_{n-1}^{-N}\ r_{2n-1}^{-N^2-1} \gamma_n^{\frac{3N-2}{4N}+\frac{N}{\rho}}(C,\delta) \\ &\mu^{N,ess}_{\leq m,\vec{\mathbf{r}}}\big\{T^{N,\tilde{\gamma}_n}_{\leq 2n-1,\leq m,\vec{\mathbf{r}}}(\tilde{f}_j;x_0,\dots,x_{m-1},x_m,x_{m+1};\vec{u}^{non}_{\leq 2n-1,\leq m,\mathbf{X}_m})\big\} \end{aligned} \tag{9.18}$$

and for $m=-1$

$$\frac{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\{CB^{N,sim,\tilde\gamma_n}_n(\gamma_n,\tilde f_j,M_{1+\rho};x_0)\}}{\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\{CHB^{N,st}_{\le 2n-1}(\vec\lambda(\vec{\mathbf{r}}),0)\}} \le 3^{3nN^4+nN^3}\ M_1^{N^2}\ r_{n-1}^{-N}\ r_{2n-1}^{-N^2-1}\ \gamma_n^{\frac{3N-2}{4N}+\frac{N}{\rho}}(C,\delta) \tag{9.19}$$

Proof of the N–dimensional Collection Lemma: In the proof we collect all possible $\tilde\gamma_n$–pseudotrajectories and estimates of "bad" measure corresponding to those $\tilde\gamma_n$–pseudotrajectories. The proof uses backward induction in $(n-m)$.

Consider the case $m=n-2$. Fix an n–tuple $\mathbf{X}_n=(x_0,\dots,x_{n-1})$, apply Corollary 9.1, and notice that

$$\begin{aligned}&\mu^{N,ess}_{\le m,\vec{\mathbf{r}}}\{T^{N,\tilde\gamma_n}_{\le 2n-1,\le m,\vec{\mathbf{r}}}(\tilde f_j;x_0,\dots,x_{m-1},x_m,x_{m+1};\vec u^{non}_{\le 2n-1,\le m,\mathbf{X}_m})\} \le \\ &\Big(\tilde\gamma_n(C,\delta,\rho)\ \frac{\tilde\gamma_n(C,\delta,\rho)}{|p^{dyn}_{x_0}(x_1)|}\cdots\frac{\tilde\gamma_n(C,\delta,\rho)}{|p^{dyn}_{x_0,\dots,x_{n-3}}(x_{n-2})|}\Big)^N \\ &\mu^N_{\le 2n-1,\vec{\mathbf{r}}}\{CHB^{N,st}_{\le 2n-1}(\vec\lambda(\vec{\mathbf{r}}),0)\}.\end{aligned} \tag{9.20}$$

This is the required estimate for $m=n-2$.

Suppose now that for $n-m=p$ estimate (3.43) is provenaand we would like to prove it for $n-(m-1)=p+1$. Denote by $G^{N,\tilde\gamma_n}_{\le 2n-1,\le m-1,\vec{\mathbf{r}}}(\tilde f_j;x_0,\dots,x_m;\vec u^{non}_{\le 2n-1,\le m,\mathbf{X}_m}) \subset \mathbb{Z}^N_{\tilde\gamma_n}$ the set of points x_{m+1} of the $\tilde\gamma_n$–grid $\mathbb{Z}^N_{\tilde\gamma_n}$ so that the $(m+2)$–tuple $x_0,\dots,x_{m+1}$ is associated to $CHB^N_{\le 2n-1}(\vec\lambda(\vec{\mathbf{r}}))$ that is, the set of possible continuations of the $\tilde\gamma_n$-pseudotrajectory $x_0,\dots,x_m$. Notice that the Distortion Lemma implies that

$$\begin{aligned}\sum_{x_{m+1}\in G^{N,\tilde\gamma_n}_{\le 2n-1,\le m-1,\vec{\mathbf{r}}}(\tilde f_j;x_0,\dots,x_m;\vec u^{non}_{\le 2n-1,\le m,\mathbf{X}_m})} &\mu^{N,ess}_{\le m,\vec{\mathbf{r}}}\{T^{N,\tilde\gamma_n}_{\le 2n-1,\le m,\vec{\mathbf{r}}}(\tilde f_j;x_0,\dots,x_{m+1})\} \\ &\le\ 3^N\ \mu^{N,ess}_{\le m-1,\vec{\mathbf{r}}}\{T^{N,\tilde\gamma_n}_{\le 2n-1,\le m-1,\vec{\mathbf{r}}}(\tilde f_j;x_0,\dots,x_m)\}.\end{aligned} \tag{9.21}$$

Inductive application of this formula completes the proof of the Lemma. Q.E.D.

Proof of Theorem 9.2.1: To simplify notations within the proof denote $\gamma_n(C,\delta)$ by γ_n and $M^{-2n/\rho}_{1+\rho}\gamma_n$ by $\tilde\gamma_n$. Assume also that $M_{1+\rho}\ge 2^{1/\rho}$. Let's show first that

$$CB^{N,sim,\tilde\gamma_n}_{n,\vec{\mathbf{r}}}(\gamma_n,\tilde f_j)\subset \cup_{\tilde x_0\in B^N\cap Z^N_{\tilde\gamma_n}} CB^{N,sim,\tilde\gamma_n}_n(\gamma_n,\tilde f_j,M_{1+\rho};x_0) \tag{9.22}$$

If $\vec\varepsilon\in CB^{N,sim,\tilde\gamma_n}_{n,\vec{\mathbf{r}}}(\gamma_n,\tilde f_j)$, then $\vec\varepsilon$ is associated with an essentially (n,γ_n)–simple, (n,γ_n)–periodic, and not (n,γ_n)–hyperbolic trajectory $\{x_k=f^k_{\vec\varepsilon}(x_0)\}_{k=0}^{n-1}$. Recall that the set of parameters $CB^{N,sim,\tilde\gamma_n}_n(\vec{\mathbf{r}},\tilde f_j)$, defined in (9.11), is independent of C,δ, and ρ. The bound on the C^1–norm M_1 implies that for any $x,y\in I$ we have

$|f(x)-f(y)| < M_1|x-y|$. Essential simplicity of the trajectory $\{x_m\}_{m=0}^{n-1}$ implies that for some $j < n$ the shifted trajectory $\{x_{j+k} = f_{\vec{\varepsilon}}^{j+m}(x_0)\}_{m=0}^{n-1}$ is (n, γ_n)–simple and $(n, M_1^n\gamma_n)$–periodic. Let's approximate the shifted trajectory $\{x_{j+m}\}_{m=0}^{n-1}$ by a $\tilde{\gamma}_n$–pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1} \subset \mathbb{Z}_{\tilde{\gamma}_n}^N$ associated to the fixed above parameter $\vec{\varepsilon}$. Consider the $\tilde{\gamma}_n$–pseudotrajectory $\{\tilde{x}_{j+m}\}_{m=0}^{n-1} \subset \mathbb{Z}_{\tilde{\gamma}_n}^N$ starting at $\tilde{x}_j$. Since $|x_j - \tilde{x}_j| \le \tilde{\gamma}_n$, we have

$$|x_{j+1} - \tilde{x}_{j+1}| \le |\tilde{f}_{\vec{\varepsilon}}(x_j) - \tilde{f}_{\vec{\varepsilon}}(\tilde{x}_j)| + |\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_j) - \tilde{x}_{j+1}| \le \frac{M_1^2 - 1}{M_1 - 1}\tilde{\gamma}_n \tag{9.23}$$

By induction in m we have

$$\begin{aligned}|x_{j+m+1} - \tilde{x}_{j+m+1}| &\le |\tilde{f}_{\vec{\varepsilon}}(x_{j+m}) - \tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{j+m})| \\ &+ |\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{j+m}) - \tilde{x}_{j+m+1}| \le \frac{M_1^{m+2} - 1}{M_1 - 1}\tilde{\gamma}_n\end{aligned} \tag{9.24}$$

Therefore, we have

$$\begin{aligned}|\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{j+n-1}) - \tilde{x}_j| &\le |\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{j+n-1}) - \tilde{f}_{\vec{\varepsilon}}(x_{j+n-1})| + \\ &|\tilde{f}_{\vec{\varepsilon}}(x_{j+n-1}) - x_j| + |x_j - \tilde{x}_j| \le M_1^n \gamma_n^{1/\rho}.\end{aligned} \tag{9.25}$$

So, the $\tilde{\gamma}_n$–pseudotrajectory $\{\tilde{x}_{j+k}\}_{k=0}^{n-1}$ is $(n, M_1^n\gamma_n^{1/\rho})$–periodic. Consider the difference of derivatives

$$\left|\prod_{m=0}^{n-j-1} d\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{j+m}) \prod_{m=0}^{j-1} d\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{n+m}) - \prod_{m=0}^{n-1} d\tilde{f}_{\vec{\varepsilon}}(x_m)\right|. \tag{9.26}$$

Since

$$\begin{aligned}\max\{|d\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_m) - d\tilde{f}_{\vec{\varepsilon}}(x_m)|,\ |d\tilde{f}_{\vec{\varepsilon}}(\tilde{x}_{n+m}) - d\tilde{f}_{\vec{\varepsilon}}(x_m)|\} \le \\ M_1\ \max\{|\tilde{x}_m - x_m|,\ |\tilde{x}_{n+m} - x_m|\}^{1/\rho} \le M_1^{1+n\rho} M_{1+\rho}\gamma_n\end{aligned} \tag{9.27}$$

and $|d\tilde{f}_{\vec{\varepsilon}}(x)| \le M_1$ we get that the difference of derivatives (9.26) is bounded by $2nM_{1+\rho}^{2n}\gamma_n$, provided that $M_1^{2n} \gg \gamma_n$. This is certainly the case, because γ_n super-exponentially small in n. Therefore, if the initial exact trajectory $\{x_k\}_{j=0}^{n-1}$ is non–(n, γ_n)—hyperbolic, then the pseudotrajectory is non–$(n, M_{1+\rho}^{2n}\gamma_n)$—hyperbolic and the parameter $\vec{\varepsilon}$ should belong to $CB_{\le 2n-1}^{N,sim,\tilde{\gamma}_n}(M_{1+\rho}^{2n}\gamma_n, \vec{\mathbf{r}}, \tilde{f}_j, M_{1+\rho}; \tilde{x}_0)$. Similarly, one can show that if $\{x_{j+m}\}_{m=0}^{n-1}$ is (n, γ_n)—simple, then its approximating pseudotrajectory $\{\tilde{x}_k\}_{j=0}^{n-1}$ is $(n, \frac{\gamma_n}{2})$–simple, because of grid size and triangle inequality.

The number of starting point $\tilde{x}_0 \subset B^N \cup Z_{\tilde{\gamma}_n}^N$ for a $\tilde{\gamma}_n$–pseudotrajectory approximately equals $1/\tilde{\gamma}_n^N$. Therefore, multiplying the estimate (9.19) by $1/\tilde{\gamma}_n$ and plugging in the values of $\tilde{\gamma}_n$ and γ_n we get (9.9). This completes the proof of Theorem 9.2.1. Q.E.D.

9.4 The N–dimensional Discretization Method for trajectories with a gap

In previous Sections we developed the N-dimensional Discretization Method and estimated the measure of the set $B_n^{N,sim}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)$ associated with essentially simple almost periodic trajectories (9.9). In this and coming Sections we consider "bad" parameters $B_n^{N,non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j,\gamma_n)$ associated with essentially nonsimple almost periodic trajectories (9.4) and get an estimate on the measure of this set.

Theorem 9.4.1. *Let $\{\tilde{f}_{\vec{\varepsilon}}\}_{\varepsilon\in CHB^{N,st}_{\le 2n-1}(\vec{\lambda}(\vec{\mathbf{r}}),0)}$ be the family of polynomial perturbations (9.13), defined inside of a Cubic Brick of at most standard thickness (4.8), and $M_{1+\rho}$ be an upper bound on the $C^{1/\rho}$–norm of this family. Then with the above notations for any $C>0,\ \delta>0,\ \rho>0$ and a sufficiently small positive γ_n, e.g. $\gamma_n\le\gamma_n(C,\delta)$, the following estimate on the measure of parameters associated with maps $\tilde{f}_{\vec{\varepsilon}}$'s with an $(n,\gamma_n^{1/\rho})$–periodic, essentially non-(n,γ_n)–simple, and non-(n,γ_n)–hyperbolic point holds*

$$\mu^{N,st}_{\le 2n-1,\vec{\mathbf{r}}}\{CB_n^{N,ess}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j)\}\le\frac{C}{\ln M_{1+\rho}}\ n^{1+\delta}\ 3^{2n+2}\ M_{1+\rho}^{10n+1}\ D^{-n},\qquad(9.28)$$

where $D=M_{1+\rho}^{8/\rho}$.

Using definition (3.70) and the Shift Theorem from Chapter 5 we have decomposition (3.71) which in our notations has the form

$$CB_n^{N,non}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f},\gamma_n)\subseteq\ \cup_{k|n}CB_n^{N,\mathrm{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f},\gamma_n;D,\sigma).\qquad(9.29)$$

Theorem 9.4.2. *With the conditions of Theorem 9.4.1 let $\sigma\ge 12$ and k be some integer which divides n. Then for $D=M_{1+\rho}^{8/\rho}$ and a sufficiently small positive γ_n, e.g. $\gamma_n<\gamma_n(C,\delta)$, we have the following estimate on the measure of diffeomorphisms $\tilde{f}_{\vec{\varepsilon}}$'s associated with an (n,γ_n)–periodic, but not (n,γ_n)–hyperbolic point that has a (D,n,τ,σ)–gap at the k–th point of its trajectory.*

$$\mu^{st}_{\le 2n-1,\vec{\mathbf{r}}}\{CB_n^{N,st,gap(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j,\gamma_n)\}\le\frac{C}{\ln M_{1+\rho}}\ n^{\delta}\ 3^{2n+2}\ M_{1+\rho}^{10n+1}\ D^{-n}.(9.30)$$

Recall definition of (D,n,τ,σ)-gap (3.59) and of the (D,n,τ,σ)–gap number associated to the k–tuple $\{x_0,g(x_0),\dots,g^{k-1}(x_0)\}$ and the Brick $HB^N_{\le 2n-1}(\vec{\mathbf{r}})$. Similarly to (3.76) introduce the following set of "bad" parameters.

Denote $\mathbf{X}_k(x_0, g) = \{x_0, g(x_0), \dots, g^{k-1}(x_0)\}$. Notice that it follows from the definition of the (D, n, τ, σ)–gap number that if x_0 has a (D, n, τ, σ)–gap at x_k, then x_0 is $(k, \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k(x_0, g), D, \sigma))$–periodic. Introduce the set of "bad" parameters

$$\begin{aligned} CB_n^{N,\mathrm{gap}(k),\Delta^{st}}(C,\delta,\rho,\mathbf{r},\tilde{f};D,\sigma) = \{\vec{\varepsilon} \in CHB^{N,st}_{\le 2n-1,\vec{\mathbf{r}}}(\vec{\lambda}(\vec{\mathbf{r}}),0) : \\ \tilde{f}_{\vec{\varepsilon}} \in IH(n-1,C,\delta,\rho),\ \tilde{f}_{\vec{\varepsilon}} \text{ has a point } x_0 \text{ which has a } (D,n,\tau,\sigma)\text{-gap at} \\ x_k = \tilde{f}^k_{\vec{\varepsilon}}(x),\ \text{no } (C,\delta,\rho)\text{-leading saddles and} \\ \text{is not } (k, M^{2n}_{1+\rho}\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k(x_0,\tilde{f}_{\vec{\varepsilon}}),D,\sigma))\text{-hyperbolic}\}. \end{aligned} \tag{9.31}$$

Using standard approximation argument one can prove the following

Lemma 9.4.3. *Let $C > 0,\ \delta > 0,\ \rho,\ \sigma > 1$, and $n \in \mathbb{Z}_+$ be some numbers, and let $k \in \mathbb{Z}_+$ divide n. Then with the above notations for any $\gamma_n \le \gamma_n(C, \delta)$ and $D = M^{8/\rho}_{1+\rho}$ we have*

$$CB_n^{N,\mathrm{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j,\gamma_n;D,\sigma) \subset CB_n^{N,\mathrm{gap}(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma). \tag{9.32}$$

It is useful to recall Remark 3.5.9 as a highlight of the goal. Let's now split the set $CB_n^{N,\mathrm{gap}(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma)$ into a finite union. Denote $(M_{1+\rho})^{4ni}\gamma_n^{1/4\rho}(C,\delta)$ by $\tilde{\gamma}_{n,i}(C,\delta,\rho)$. Definition 3.5.10 introduces points recurrent of i-th order.

Define the set of parameters from $CB_n^{N,\mathrm{gap}(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma,i)$ associated to an i–th order recurrent w.r.t. (C,δ,τ,D,σ) trajectories which satisfies conditions (9.31).

$$\begin{aligned} B_n^{N,\mathrm{gap}(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma,i) = \{\vec{\varepsilon} \in CHB^{N,st}_{\le 2n-1,\vec{\mathbf{r}}}(\vec{\lambda}(\vec{\mathbf{r}}),0) : \\ \tilde{f}_{\vec{\varepsilon}} \text{ has a point which satisfies (9.31) and} \\ \text{is an } i\text{-th recurrent w.r.t. } (C,\delta,\rho,\vec{\mathbf{r}},D,\sigma)\}. \end{aligned} \tag{9.33}$$

Lemma 9.4.4. *With the notations of Lemma 9.4.3 for $M = [Cn^\delta / \ln M_{1+\rho}]$ we have*

$$\begin{aligned} &CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma) \subseteq \\ &\bigcup_{i=0}^{M} CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma,i). \end{aligned} \tag{9.34}$$

Proposition 9.4.5. *Let $C > 0,\ \delta > 0,\ \rho,\ \sigma > 0$, and $n \in \mathbb{Z}_+$ be some numbers, and let $k \in \mathbb{Z}_+$ divide n. Then with the above notations for any $\gamma_n \le \gamma_n(C,\delta)$, $D = M^8_{1+\rho}$, and any $i \in \mathbb{Z}_+$ such that $0 \le i \le M$ we have*

$$\mu^{N,st}_{\le 2n-1}\{CB_n^{gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma,i)\} \le 3^{2n+2}\ M^{10n+1}_{1+\rho}\ D^{-n}. \tag{9.35}$$

Derive Theorem 9.4.1 based on Lemmas 9.4.3 and 9.4.4, and Proposition 9.4.5. Let γ_n be sufficiently small, e.g. $\gamma_n \le \gamma_n(C,\delta)$, and let $\sigma > 0$, and $D = M_{1+\rho}^{8/\rho}$. Then

• By The Shift Theorem the set of all parameters $CB_n^{N,st,\text{non}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j,\gamma_n)$ associated to maps $\tilde{f}_{\vec{\varepsilon}}$'s with an (n,γ_n)–periodic, essentially non-(n,γ_n)–simple, but not (n,γ_n)–hyperbolic point are contained in the union (3.73) over all k dividing n of parameters $CB_n^{N,st,\text{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f},\gamma_n;D,\sigma)$ associated to maps $\tilde{f}_{\vec{\varepsilon}}$'s with an (n,γ_n)–periodic, but not (n,γ_n)–hyperbolic point with a (D,n,r_{2k},σ)–gap at x_k.

• By Lemma 9.4.3 the set of parameters $CB_n^{N,st,\text{gap}(k)}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f};D,\sigma)$ is contained in the union of the set of parameters $CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma)$ associated to maps $\tilde{f}_{\vec{\varepsilon}}$'s which have a non $(k,\Delta^{st}_{k,n,\tau}(\mathbf{X}_k,D,\sigma))$–hyperbolic point x_0 with a (D,n,r_{2k},σ)–gap at x_k.

• By Lemma 9.4.4 the set of parameters $B_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma)$ is contained in the union of the sets of parameters $\{CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma,i)\}_{i=0}^{M}$ such that the i–th set $CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma,i)$ is associated to maps $\tilde{f}_{\vec{\varepsilon}}$'s which have a non $(k,\Delta^{st}_{k,n,\tau}(\mathbf{X}_k,D,\sigma))$–hyperbolic point x_0 with a (D,n,r_{2k},σ)–gap at x_k such that the k–tuple $\{x_j = \tilde{f}^j_{\vec{\varepsilon}}(x_0)\}_{j=0}^{k-1}$ is i–th recurrent.

• Proposition 9.4.5 provides an estimate of the measure of the set $CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\vec{\mathbf{r}},\tilde{f}_j;D,\sigma,i)$ all the sets above. Application of Lemma 9.4.4 completes the proof of Theorem 9.4.1.

Proofs Lemmas 9.4.3 and 9.4.4 are similar to the one for Lemmas 3.5.8 and 3.5.12 and are omitted.

9.5 The measure of maps $\tilde{f}_{\vec{\varepsilon}}$ with i–th recurrent, not sufficiently hyperbolic trajectories with a gap

We shall prove Proposition 9.4.5 in three steps.

Step 1. Reduction to polynomial perturbations of degree $2k-1$.

The measure $\mu^{N,st}_{\le 2n-1,\vec{\mathbf{r}}}$ is the Lebesgue product probability measure and each its component $\mu^{N,st}_{m,\vec{\mathbf{r}}}$ is the Lebesgue probability measure (see (1.15))

$$\mu^{N,st}_{\le 2n-1,\vec{\mathbf{r}}} = \mu^{N,st}_{\le 2k-1,\vec{\mathbf{r}}} \times \left(\times_{m=2k}^{2n-1} \mu^{N,st}_{\le 2n-1,\vec{\mathbf{r}}} \right). \tag{9.36}$$

Therefore, by Fubini Theorem it is sufficient to prove that

$$\mu^{N,st}_{\le 2k-1,\mathbf{r}}\left\{CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\mathbf{r},\tilde{f}_j;D,\sigma,i)\cap\{(\vec{\varepsilon}_{2k},\dots,\vec{\varepsilon}_{2n-1})\}\right\} \le 3^{2n+2}\ M_{1+\rho}^{10n+1}\ D^{-n}.$$

uniformly over

$$\begin{aligned}\{(\vec{\varepsilon}_{2k},\dots,\vec{\varepsilon}_{2n-1})\} &\in CHB^{N,st}_{2k,\mathbf{r}}(\vec{\lambda}(\mathbf{r}),0)\times\\ CHB^{N,st}_{2k+1,\mathbf{r}}(\vec{\lambda}(\mathbf{r}))&\times\dots\times CHB^{N,st}_{2n-1,\mathbf{r}}(\vec{\lambda}(\mathbf{r})).\end{aligned} \tag{9.37}$$

To simplify notations we omit $(\vec{\varepsilon}_{2k},\dots,\vec{\varepsilon}_{2n-1})$ and write as if the set of parameters $CB_n^{N,st,gap(k),\Delta^{st}}(C,\delta,\rho,\mathbf{r},\tilde{f}_j;D,\sigma,i)$ is a subset of the Brick $CHB^{N,st}_{\le 2k-1,\mathbf{r}}(\vec{\lambda}(\mathbf{r}),0)$.

From now on we consider the $2k$–parameter family of polynomial perturbations

$$\left\{\tilde{f}_{\vec{\varepsilon}_{\le 2k-1}}=\tilde{f}_j(x)+\sum_{|\alpha|\le 2k-1}\vec{\varepsilon}_\alpha x^\alpha\right\}, \tag{9.38}$$

where $\vec{\varepsilon}_{\le 2k-1}\in CHB^{N,st}_{\le 2k-1,\mathbf{r}}$. Recall that $CHB^{N,st}_{\le 2k-1,\mathbf{r}}(\vec{\lambda}(\mathbf{r}),0)$ is supplied with the Lebesgue product probability measure $\mu^{N,st}_{\le 2k-1,\mathbf{r}}$.

Step 2. An estimate of parameters associated with a single trajectory $\{x_j\}_{j=0}^{n-1}$ *which has a gap at* x_k*, is* i*–th recurrent, and not sufficiently hyperbolic.*

By analogy with (3.15) and using definition 3.6.1 we define

$$\begin{aligned}CB_n^{N,st,sim,\Delta^{st},\tilde{\gamma}_n}(\vec{\lambda}(\mathbf{r}),\tilde{f},\gamma_n,M_{1+\rho};x_0,\dots,x_{k-1}) = &\{\vec{\varepsilon}\in CHB^{N,st}_{\le 2k-1}(\vec{\lambda}(\mathbf{r}),0):\\ &\{x_m\}_{m=0}^{k-1}\text{ is a }\tilde{\gamma}_n\text{-pseudotrajectory associated to }\vec{\varepsilon}\\ &\text{which is }\big(k,M_{1+\rho}^{2n}\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\ \sigma)\big)\text{-periodic}\\ &\text{and has a }(D,n,r_{2k},\sigma)\text{-gap at }x_k=\tilde{f}^k_{\vec{\varepsilon}}(x)\}\end{aligned} \tag{9.39}$$

To show an analogy with the case simple trajectories consider the following

Problem 9.5.1. *Estimate the measure of* $\vec{\varepsilon}\in CHB^{N,st}_{\le 2k-1}(\vec{\lambda}(\mathbf{r}),0)$ *for which this* n*–tuple* $\{x_j\}_{j=0}^{k-1}$ *is*

$$\begin{aligned}&A)\ \tilde{\gamma}_n\text{-pseudotrajectory, i.e. } |\tilde{f}_{\vec{\varepsilon}}(x_j)-x_{j+1}|\le\tilde{\gamma}_n\quad\text{for}\quad j=0,\dots,k-2;\\ &B)\ (k,\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D))\text{-periodic, i.e. } |\tilde{f}_{\vec{\varepsilon}}(x_{n-1})-x_0|<\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\sigma);\\ &C)\ \text{not }(n,(M_{1+\rho})^{2n}\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\sigma))\text{-hyperbolic, i.e.}\\ &\qquad\left|(\tilde{f}^n_{\vec{\varepsilon}})'(x_0)-1\right|\le M_{1+\rho}^{2n}\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\sigma);\end{aligned} \tag{9.40}$$

For a fixed k–tuple of points $\mathbf{X}_k = (x_0, \dots, x_{k-1})$ we can consider the Newton family of polynomial perturbations decomposed into dynamically essential parameters

$$\begin{aligned}\tilde{f}_{\vec{u}^{ess}_{\le k-1}, \mathbf{X}_k}(x) = \tilde{f}(x) + \sum_{s=0}^{k-1} \vec{u}_{\alpha(\mathbf{X}_s)} \ (x; x_0, \dots, x_{s-1})^{\alpha(X_s)}_{\sigma(\mathbf{X}_s)} \\ + \sum_{i=1}^{N} \vec{u}_{\alpha_i(\mathbf{X}_k)} \ (x; x_0, \dots, x_{k-1})^{\alpha(X_k)}_{\sigma(\mathbf{X}_s)},\end{aligned} \tag{9.41}$$

where multiindices $\{\alpha(X_s)\}_{s=0}^{k-1}$ and $\{\alpha_i(X_k)\}_{i=1}^{N}$ are defined in (8.28). Recall that we notice in (2.17) and fig. 2.2 that for any $m < k$ and $s > 0$ the image $\tilde{f}_{\vec{u}^{ess}_{\le k-1}, \mathbf{X}_k}(x_m)$ (resp. the linearization $d\tilde{f}_{\vec{u}^{ess}_{\le k-1}, \mathbf{X}_k}(x_m)$) of (resp. at) the point x_m is independent of the Newton coefficients $\vec{u}_{\alpha(\mathbf{X}_{m+s})}$'s (resp. $\vec{u}_{\alpha(\mathbf{X}_{m+n+s})}$'s) with larger index. This implies that Newton coefficients $\vec{u}_{\alpha(\mathbf{X}_0)}, \dots, \vec{u}_{\alpha(X_{k-2})}$ can be choisen to satisfy the property (A), the Newton coefficient $\vec{u}_{\alpha(X_{k-1}}$ can be choisen to satisfy the property (B) (almost periodicity), and the Newton coefficients $\vec{u}_{\alpha_i(\mathbf{X}_k)}$ can be choisen to satisfy the property (C) (almost hyperbolicity).

Following formulas (3.27), (3.28), and formulas (3.30) and (3.33) with n replaced by k and γ_n replaced by $\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)$ as the 1-dimensional model we have that the measure of $\vec{u}^{ess}_{\le k-1}$ with conditions (9.40) is bounded by

$$\begin{aligned}&\mu^{N,st,ess}_{\le 2k-1, \vec{\mathbf{r}}} \Big\{ \vec{u}^{ess}_{\le k-1} \in W^u_{\le 2k-1,1} : \tilde{f}_{\vec{u}^{ess}_{\le k-1}, \mathbf{X}_k} \text{ satisfies conditions (9.40)} \Big\} \le \\ &M^{4N^2}_{1+\rho} \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma) \ 3^{N^3(k-1)} \prod_{m=1}^{k-2} \frac{\tilde{\gamma}^N_n}{r^N_{n-1} \ \prod_{j=0}^{k-2} |x_{k-1} - x_j|} \\ &\frac{1}{r^{N^2+1}_{2n-1}} \ \frac{\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)}{\prod_{j=0}^{k-2} |x_{k-1} - x_j|} \ \frac{M^{2n}_{1+\rho} \ \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)}{\prod_{j=0}^{k-2} |x_{k-1} - x_j|^2}.\end{aligned} \tag{9.42}$$

Consider last two terms in the right-hand side product. Definition of the (D, n, r_{2k}, σ)–gap number, $\sigma \ge 12$, and the inequality $min(a, b, c) \le a^{1/2} \ b^{1/4} \ c^{1/4}$ show that these two terms are bounded by

$$\begin{aligned}&M^{2n}_{1+\rho} \, (r_{2k})^2 \, \frac{\left(\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)\right)^2}{(\prod_{j=0}^{m-2} |x_{m-1} - x_j|)^3} \\ &\le 2M^{2n}_{1+\rho} D^{-n/2} \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)\end{aligned} \tag{9.43}$$

Therefore,

$$\begin{aligned}&\mu^{N,st,ess}_{\le 2k-1,\vec{\mathbf{r}}}\left\{\vec{u}^{ess}_{\le k-1}\in W^u_{\le 2k-1,1}:\ \tilde f_{\vec{u}^{ess}_{\le k-1},\mathbf{X}_k}\text{ satisfies conditions (3.85)}\right\}\\ &\le 3^{N^3(k-1)}\ \prod_{m=1}^{k-2}\frac{\tilde\gamma_n}{r^N_{2k}\ \prod_{j=0}^{m-1}|x_m-x_j|}\ M^{2n+1}_{1+\rho}\ D^{-n/2}\ \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\sigma).\end{aligned}\tag{9.44}$$

The next step of the proof of Proposition 3.5.13 is as in the case of simple trajectories (Section 9.2) is to collect all possible "bad" pseudotrajectories and make sure that those pseudotrajectories indeed approximate sufficiently well all "real bad" trajectories.

Step 3. A grid $\mathbb{Z}^N_{\tilde\gamma_{n,i}}$ of i dependent size and collection of all "bad" i–th recurrent not sufficiently hyperbolic trajectories $\{x_j\}_{j=0}^{k-1}$ with a gap at x_k.

In the case of simple trajectories in Section 9.2 we considered only one $\tilde\gamma_n$–grid of a fixed size and collected of all simple "bad" trajectories in the Collection Lemma. In the case of i–th recurrent trajectories with a gap at x_k we define a grid $\mathbb{Z}^N_{\tilde\gamma_{n,i}}$ of i dependent size $\tilde\gamma_{n,i}$. Then we prove that $\tilde\gamma_{n,i}$–pseudotrajectories approximate real i–th recurrent trajectories with a gap at x_k sufficiently well. Finally, we collect all possible i–th recurrent $\tilde\gamma_{n,i}$–pseudotrajectories with a gap at x_k and sum estimates of the measures of "bad" sets. Let's we realize this program.

Put $\tilde\gamma_{n,i}=M^{4ni}_{1+\rho}\gamma_n^{1/4}$ and call

$$\mathbb{Z}_{\tilde\gamma_{n,i}}=\{x\in\mathbb{R}^N:\exists k\in\mathbb{Z}\text{ such that }x=k\tilde\gamma_{n,i}\}\tag{9.45}$$

the grid of i–th order.

Definition 9.5.2. *Let $\{x_j\}_{j=0}^{k-1}\subset\mathbb{Z}^N_{\tilde\gamma_{n,i}}$ be a k–tuple for some $i\in\mathbb{Z}_+$. Then the k–tuple $\{x_j\}_{j=0}^{k-1}$ is called the i–th recurrent w.r.t. $D>0,\sigma>0$, and $M_{1+\rho}>0$ if*

$$M^{-5\sigma n}_{1+\rho}\ \tilde\gamma^\sigma_{n,i}\le\Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k,D,\sigma)\le M^{\sigma n}_{1+\rho}\ \tilde\gamma^\sigma_{n,i}.\tag{9.46}$$

In notations and terminology of our induction if the k–tuple $\{x_j\}_{j=0}^{k-1}$ is $\tilde\gamma_{n,i}$–pseudotrajectory associated to some parameter $\vec\varepsilon\in CHB^{N,st}_{\le 2k-1,\vec{\mathbf{r}}}(\vec\lambda(\vec{\mathbf{r}}),0)$ and is i–th recurrent w.r.t. a gap admissible constant D and the $C^{1+\rho}$–norm $M_{1+\rho}$ of the family (3.83), then we say that $\{x_j\}_{j=0}^{k-1}$ is i–th recurrent pseudotrajectory.

Notice also that one pseudotrajectory could be the i–th and the $(i+1)$–st recurrent at the same time. This just means that we will have an undesirable double counting when we will collect all possible pseudotrajectories.

Define a set of parameters which is a discretized version of the set of parameters $CB^{N,st,gap(k),\Delta^{st}}_n(\vec{\mathbf{r}},\tilde f;D,\sigma,i)$ (3.79) associated with all i–th recurrent trajectories of length k which has a weak (D, n, r_{2k}, σ)–gap at x_k and is not $M_{1+\rho}^{2n} \Delta^{st}_{k,n,r_{2k}} (\mathbf{X}_k(x_0, \tilde{f}_{\vec{\varepsilon}}), D, \sigma)$–hyperbolic.

$$\begin{aligned} CB_n^{N,st,gap(k),\Delta^{st}}(\tilde{f}, \tilde{\gamma}_{n,i}; D, \sigma, i) = \{\vec{\varepsilon} \in CHB^{N,st}_{\leq 2n-1, \vec{\mathbf{r}}} : \tilde{f}_{\vec{\varepsilon}} \in IH(n-1, C, \delta, \rho), \\ \text{there is a } k\text{-tuple } \{x_j\}_{j=0}^{k-1} \subset \mathbb{Z}^N_{\tilde{\gamma}_{n,i}} \text{ associated to } \varepsilon \text{ which is} \\ i\text{-th recurrent, has a } (2D, n, 2r_{2k}, \sigma)\text{-gap at } x_k = \tilde{f}^k_{\vec{\varepsilon}}(x), \\ \text{and not } M_{1+\rho}^{2n} \Delta^{st}_{k,n,2r_{2k}} (2\mathbf{X}_k(x_0, \tilde{f}_{\vec{\varepsilon}}), 2D, \sigma)\text{-hyperbolic}\}, \end{aligned} \tag{9.47}$$

where $2\mathbf{X}_k$ means that in the formula (3.75) for the (D, n, r_{2k}, σ)–gap number in the last term $\prod_{j=0}^{k-1} |x_{k-1} - x_j|$ we multiply each term $|x_{k-1} - x_j|$ by 2.

Similarly to the case of simple trajectories we prove that after discretization all real "bad" trajectories can be sufficiently well approximated by pseudotrajectories of a certain grid so that quantities of periodicity, hyperbolicity, existence of a gap, and product of distances along the trajectory are almost the same. Namely,

Lemma 9.5.3. *With above notations for constants* $\gamma_n \leq M_{1+\rho}^{-8n} \gamma_n(C, \delta)$ *and* $\tilde{\gamma}_{n,i} \leq M_{1+\rho}^{-2n/\rho} \tilde{\gamma}_{n,i}^{1/\rho}(C, \delta)$ *we have (see definition 3.5.10 and (9.45))*

$$CB_n^{N,st,gap(k),\Delta^{st}}(\vec{\mathbf{r}}, \tilde{f}; D, \sigma, i) \subset CB_n^{N,st,gap(k),\Delta^{st}}(\vec{\mathbf{r}}, \tilde{f}, \tilde{\gamma}_{n,i}; D, \sigma, i) \tag{9.48}$$

To complete the proof of Proposition 3.5.13 consider the estimate of "bad" measure (9.44) for one pseudotrajectory and notice that the proof of the Collection Lemma can be applied to collect all pseudotrajectories with the required property. This completes the proof of Proposition 3.5.13 and, because of remarks at the end of the last Section, it also proves Theorems 9.4.1.

Proof of Lemma 9.4.3. Fix a parameter $\vec{\varepsilon} \in CB_n^{N,st,gap(k)}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f}, \gamma_n; D, \sigma)$. By definition (9.31) there is an (n, γ_n)–periodic, not (n, γ_n)–hyperbolic point x_0 with a (D, n, τ, σ)–gap at $x_k = \tilde{f}^k_{\vec{\varepsilon}}(x_0)$. It is sufficient to show that if x_0 is $(k, M_{1+\rho}^{2n} \Delta^{st}_{k,n,r_{2k}} (\mathbf{X}_k, D, \sigma))$–hyperbolic, then x_0 should be (n, γ_n)–hyperbolic.

Since n is divisible by k, we can split the trajectory $\{x_j = \tilde{f}^j_{\vec{\varepsilon}}(x_0)\}_{j=0}^{n-1}$ of length n into $p = n/k$ parts of length k each. Consider the linearization

$$(d\tilde{f}^n_{\vec{\varepsilon}})(x_0) = (d\tilde{f}^k_{\vec{\varepsilon}})(x_{(p-1)k}) \cdot \quad \ldots \quad \cdot (d\tilde{f}^k_{\vec{\varepsilon}})(x_k) \cdot (d\tilde{f}^n_{\vec{\varepsilon}})(x_0). \tag{9.49}$$

Since we are interested in large D, assume that $D = M_{1+\rho}^{8/\rho}$. By definition of a $(D, n, \vec{\mathbf{r}}, \sigma)$–gap at x_k we have

$$|x_0 - x_k| \leq \Delta^{st}_{k,n,r_{2k}} (\mathbf{X}_k, D, \sigma) \leq 2\ D^{-n}. \tag{9.50}$$

Therefore, for every $1 \le j \le p-1$ and $0 \le s < k$

$$|x_s - x_{jk+s}| \le 2\ M_1^n\ D^{-n} \tag{9.51}$$

and is exponentially small in n. This implies that the trajectory $\{x_j\}_{j=0}^{n-1}$ of length n consists of p almost identical parts of length k each. Thus, for each $0 \le m < n$ we have

$$\left| d\tilde{f}_{\vec{\varepsilon}}(x_m) - d\tilde{f}_{\vec{\varepsilon}})(x_{m\ (mod\ k)}) \right| \le M_1^n \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma). \tag{9.52}$$

This implies that

$$\left| d\tilde{f}^n_{\vec{\varepsilon}}(x_0) - \left(d\tilde{f}^k_{\vec{\varepsilon}}(x_0) \right)^p \right| \le n(M_{1+\rho})^{2n} \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma). \tag{9.53}$$

Therefore, if $\left| d\tilde{f}^n_{\vec{\varepsilon}}(x_0) - Id \right| \le \gamma_n \le \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma)$, then

$$\left| d\tilde{f}^k_{\vec{\varepsilon}}(x_0) - Id \right| \le \gamma_n \le (M_{1+\rho})^{3n} \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k, D, \sigma) \tag{9.54}$$

and $\vec{\varepsilon}$ should belong to $B_n^{N,st,gap(k),\Delta^{st}}(C, \delta, \rho, \vec{\mathbf{r}}, \tilde{f}_j; D, \sigma)$. This completes the proof of the Lemma. Q.E.D.

Proof of Lemma 9.4.4. It follows from the Shift Theorem (last two lines) with $N = 1$ that if a point x_0 has no (C, δ)–leading saddles and (D, n, τ, σ)–gap at x_k, then x_0 is $(k, \gamma_n(C, \delta)$–simple. Therefore, for the (D, n, τ, σ)–gap number we have the following bounds

$$\gamma_n^{\sigma/4}(C, \delta) = \exp\left(-C \frac{\sigma}{4} n^{1+\delta} \right) \le \Delta^{st}_{k,n,r_{2k}}(\mathbf{X}_k(x_0, \tilde{f}_{\vec{\varepsilon}}) \le 2D^{-n} \tag{9.55}$$

or each such an x_0 with a (D, n, r_{2k}, σ)–gap at x_k is an i–th recurrent for some $i \ge 0$. This completes the proof of the Lemma. Q.E.D.

Proof of Lemma 9.5.3. The proof is by approximation arguments similar to the proof of Lemma 3.5.8 above. Q.E.D.

Appendix A: Properties of hyperbolicity

In this paper we have quantified the hyperbolicity of periodic points in order to bound from below the distance between periodic points of the same period. Recall the definitions of $\gamma(L)$, $\gamma_n(x, f)$, and $\gamma_n(f)$ from (1.6) and (1.7) and the text between them. In this appendix we will prove Proposition 1.1.6, along with a result that say that the hyperbolicity of a power of a linear operator is no smaller than the hyperbolicity of the operator.

Proposition 1.1.6 follows immediately from the following lemma.

Lemma A.1. *Given the hypotheses of Proposition 1.1.6, for every pair of distinct period* n *points of* f*, say* $x = f^n(x) \neq y = f^n(y)$*, the distance* $|x - y|$ *between them is at least* $(M_{1+\rho}^{-n(1+\rho)}\gamma_n(f))^{1/\rho}$.

Proof: Let $v = (y - x)/|x - y|$. Then

$$\begin{aligned} v = \frac{f^n(y) - f^n(x)}{|x-y|} &= \frac{1}{|x-y|}\int_0^{|x-y|} \frac{d}{d\lambda} f^n(x + \lambda v) d\lambda \\ &= \frac{1}{|x-y|}\int_0^{|x-y|} df^n(x + \lambda v) v d\lambda. \end{aligned} \tag{.56}$$

It follows that

$$\frac{1}{|x-y|}\int_0^{|x-y|} (df^n(x + \lambda v)v - v) d\lambda = 0. \tag{.57}$$

Let $w = df^n(x)v - v$; by hypothesis $|w| \geq \gamma_n(f)$. Also,

$$\begin{aligned} |df^n(x + \lambda v)v - v - w| &= |(df^n(x + \lambda v) - df^n(x))v| \\ &\leq \|df^n(x + \lambda v) - df^n(x)\|. \end{aligned} \tag{.58}$$

Now

$$
\begin{aligned}
&df^n(x+\lambda v) - df^n(x) \\
&\quad = [df(f^{n-1}(x+\lambda v)) - df(f^{n-1}(x))]df^{n-1}(x+\lambda v) \\
&\quad\quad + df(f^{n-1}(x))[df(f^{n-2}(x+\lambda v)) - df(f^{n-2}(x))]df^{n-2}(x+\lambda v) \\
&\quad\quad + \cdots + df^{n-1}(f(x))[df(x+\lambda v) - df(x)].
\end{aligned}
\tag{.59}
$$

Since $M_{1+\rho}$ is an upper bound on the $C^{1+\rho}$ norm of f, it is an upper bound on the norm of $df(z)$ for all $z \in M$. It follows that $|f^k(x+\lambda v) - f^k(x)| \leq M_{1+\rho}^k \lambda$ for $k = 0, 1, \ldots, n-1$, and hence

$$
\begin{aligned}
\|df^n(x+\lambda v) - df^n(x)\| \leq & \sum_{k=0}^{n-1} M_{1+\rho}(M_{1+\rho}^k \lambda)^\rho M_{1+\rho}^{n-1} \\
= & M_{1+\rho}^n \frac{M_{1+\rho}^{n\rho} - 1}{M_{1+\rho}^\rho - 1} \lambda^\rho < M_{1+\rho}^{n(1+\rho)} \lambda^\rho.
\end{aligned}
\tag{.60}
$$

(Recall that we assumed $M_{1+\rho} > 2^{1/\rho}$ in the definition of $M_{1+\rho}$.)

By the results above we then have

$$
\begin{aligned}
0 = & \frac{1}{|x-y|} \int_0^{|x-y|} (df^n(x+\lambda v)v - v) \cdot \frac{w}{|w|} d\lambda \\
\geq & \frac{1}{|x-y|} \int_0^{|x-y|} (|w| - \|df^n(x+\lambda v) - df^n(x)\|) d\lambda \\
> & \frac{1}{|x-y|} \int_0^{|x-y|} (\gamma_n(f) - M_{1+\rho}^{n(1+\rho)} \lambda^\rho) d\lambda \\
> & (\gamma_n(f) - M_{1+\rho}^{n(1+\rho)} |x-y|^\rho).
\end{aligned}
\tag{.61}
$$

From this we get the desired upper bound on $|x-y|$. This completes the proof of the Lemma. Q.E.D.

Notice that the notion of hyperbolicity $\gamma(L)$ of a linear operator L as a lower bound on $|Lv - v|$ for unit vectors v occurs naturally in the proof above. It is not possible to make an analogous estimate with the same power on the period n hyperbolicity of f if the hyperbolicity is defined in the more usual manner, as in [Y], taking the minimum distance of the eigenvalues of L from the unit circle in $\mathbb{C}$. To see this, consider the following C^2 map $f : \mathbb{R}^N \to \mathbb{R}^N$ for small $\gamma > 0$:

$$
\begin{aligned}
&f(x_1, x_2, \ldots, x_N) = \\
&((1+\gamma)x_1 - x_2, (1+\gamma)x_2 - x_3, \ldots, (1+\gamma)x_n - x_{n-1}, (1+\gamma)x_n - x_1^2).
\end{aligned}
\tag{.62}
$$

Notice that f has two nearby fixed points, 0 and $(\gamma^N, \gamma^{N+1}, \ldots, \gamma^{2N-1})$, that are within roughly γ^N of each other. Notice also that $df(0)$ is upper triangular and hence all N of its eigenvalues are $1+\gamma$, so that by the eigenvalue notion of hyperbolicity 0 is an (n,γ)-hyperbolic fixed point of f. (Though df has eigenvalues closer to the unit circle at the other fixed point, they are still much farther away than γ^N for large N.) On the other hand, for $v = (1, \gamma, \ldots, \gamma^{N-1})$ we have $|v|$ slightly larger than 1 while

$$|Lv - v| = |(0, 0, \ldots, \gamma^N)| = \gamma^N, \tag{.63}$$

so that our notion of hyperbolicity is commensurate with the spacing between the fixed points.

To this point, when using the hyperbolicity of a linear operator L, it has only been important that $|Lv - v|$ not be small for unit vectors v. The reason for estimating from below $|Lv - \exp(2\pi i\phi)v|$ for $\phi \in [0,1)$ in (1.6) is that the hyperbolicity of L will be a lower bound on the hyperbolicity of L^k for positive integers k. In terms of diffeomorphisms, this estimate gives a bound on how close points of period nk can lie to a hyperbolic point of period n. For the eigenvalue-based notion of hyperbolicity, the estimate is trivial, but for our notion it must be proved.

Lemma A.2. *For every linear operator $L : \mathbb{C}^N \to \mathbb{C}^N$ and $k \in \mathbb{Z}^+$, we have $\gamma(L^k) \geq \gamma(L)$.*

Proof: Suppose $\gamma(L^k) < \gamma(L)$; then for some $\phi \in [0,1)$ and unit vector $v \in \mathbb{C}^N$ we have $|L^k v - \exp(2\pi i\phi)v| < \gamma(L)$. Without loss of generality we may assume that $\phi = 0$; otherwise replace L with $\exp(-2\pi i\phi/k)L$, so that the $\gamma(L)$ and $\gamma(L^k)$ are unaffected and $|L^k v - v| < \gamma(L)$. Let $\omega = \exp(2\pi i/k)$, and for $j = 0, 1, \ldots, k-1$ let

$$w_j = v + \omega^j Lv + \omega^{2j} L^2 v + \cdots + \omega^{(k-1)j} L^{k-1} v. \tag{.64}$$

Notice that $w_0 + w_1 + \cdots + w_{k-1} = kv$, and since v is a unit vector we must have $|w_j| \geq 1$ for some j. But

$$\begin{aligned} Lw_j - \omega^{-j} w_j &= Lv - \omega^{-j} v + \omega^j L^2 v - Lv + \cdots \omega^{(k-1)j} L^k v - \omega^{(k-2)j} L^{k-1} v \\ &= \omega^{(k-1)j} L^k v - \omega^{-j} v = \omega^{-j}(Lv - v), \end{aligned} \tag{.65}$$

the last step because ω is a k-th root of unity. This yields $|Lw_j - \omega^{-j} w_j| = |Lv - v| < \gamma(L)$, contradicting the definition of $\gamma(L)$. This completes the proof of the lemma. Q.E.D.

The next lemma is a simple estimate on how much a small perturbation of a linear operator can change its hyperbolicity.

Lemma A.3. *For any pair of linear operators L and Δ of $\mathbb{R}^N$ into itself, hyperbolicity satisfies the estimate*

$$\gamma(L+\Delta) \geq \gamma(L) - \|\Delta\|. \tag{.66}$$

Proof: By the definition of hyperbolicity,

$$\gamma(L+\Delta) = \inf_{\phi\in[0,1)} \inf_{\|v\|=1} |(L+\Delta)v - \exp(2\pi i\phi)v|. \tag{.67}$$

By triangle inequality, for all $v \in \mathbb{R}^N$,

$$|(L+\Delta)v - \exp(2\pi i\phi)v| \geq |Lv - \exp(2\pi i\phi)v| - |\Delta v|. \tag{.68}$$

This implies the statement of the lemma. Q.E.D.

The following lemma generalizes the previous two lemmas.

Lemma A.4. *For all $N \times N$ matrices $A, A_1, A_2, \ldots, A_m$,*

$$\gamma(A_m A_{m-1} \cdots A_1) \geq \gamma(A) - \sum_{j=1}^{m} \|A - A_j\|. \tag{.69}$$

Proof: Choose $v_0 \in \mathbb{C}^N$ and $\phi \in \mathbb{R}$ such that

$$|A_m A_{m-1} \cdots A_1 v_0 - e^{i\phi} v_0| = \gamma(A_m A_{m-1} \cdots A_1)|v_0|. \tag{.70}$$

Let $v_1 = A_1 v_0$, $v_2 = A_2 A_1 v_0$, $\ldots$, $v_m = A_m A_{m-1} \cdots A_1 v_0$. For $j = 0, 1, \ldots, m-1$ let

$$\omega_j = e^{i(-\phi+2\pi j)/m} \tag{.71}$$

and

$$u_j = v_0 + \omega_j v_1 + \omega_j^2 v_2 + \cdots + \omega_j^{m-1} v_{m-1}. \tag{.72}$$

Choose ℓ for which $|v_\ell| = \max(|v_0|, |v_1|, \ldots, |v_{m-1}|)$, and notice that

$$\sum_{j=0}^{m-1} \omega_j^{-\ell} u_j = m v_\ell. \tag{.73}$$

Thus there exists j such that

$$|u_j| \geq |v_\ell| = \max(|v_0|, |v_1|, \ldots, |v_{m-1}|). \tag{.74}$$

Then we have

$$
\begin{aligned}
&\gamma(A) \le |Au_j - \omega_j^{-1}u_j|/|u_j| \\
=&\frac{|-\omega_j^{-1}v_0 + Av_0 - v_1 + A\omega_j v_1 - \omega_j v_2 + \cdots + A\omega_j^{m-1}v_{m-1}|}{|u_j|} \\
=&\frac{|\omega_j^{m-1}(v_m - e^{i\phi}v_0) + (A-A_1)v_0 + (A-A_2)\omega_j v_1 + \cdots (A-A_m)\omega_j^{m-1}v_{m-1}|}{|u_j|} \\
\le&\frac{\gamma(A_mA_{m-1}\cdots A_1)|v_0| + \|A-A_1\||v_0| + \|A-A_2\||v_1| + \cdots + \|A-A_m\||v_{m-1}|}{|u_j|} \\
\le&\gamma(A_mA_{m-1}\cdots A_1) + \|A-A_1\| + \|A-A_2\| + \cdots + \|A-A_m\|,
\end{aligned}
$$

which is equivalent to the desired inequality. Q.E.D.

Proposition A.5. *Let $r \le 1 \le K$ be positive numbers and A, B be linear operators of $\mathbb{R}^N$ into itself given by $N \times N$ matrices from $M_N(\mathbb{R})$ with real entries. Consider an N^2-parameter family $\{A_U = A + UB\}_{U \in C^{N^2}(r)}$, where $C^{N^2}(r)$ is the cube in $M_N(\mathbb{R})$ whose entries are bounded in absolute value by r. Suppose that $\|B\|, \|B^{-1}\| \le K$. Then for the Lebesgue product probability measure μ_{r,N^2} on the cube $C^{N^2}(r)$ and all $0 < \gamma \le \min(r,1)$, we have*

$$\mu_{r,N^2}\left\{U \in C^{N^2}(r):\ \gamma(A_U) \le \gamma\right\} \le \frac{C(N)K^{2N^2}\gamma}{r^2}, \tag{.75}$$

where the constant $C(N)$ depends only on N.

Proof: For $0 < \gamma \le 1$ and $\phi \in [0,1)$, define the sets of non-γ-hyperbolic matrices by

$$
\begin{aligned}
NH_N^{\gamma}(\mathbb{R}) &= \{L \in M_N(\mathbb{R}):\ \gamma(L) \le \gamma\}, \\
NH_N^{\gamma,\phi}(\mathbb{R}) &= \{L \in M_N(\mathbb{R}):\ \inf_{|v|=1} |(L - \exp(2\pi i\phi)v| \le \gamma\}.
\end{aligned} \tag{.76}
$$

Then

$$NH_N^{\gamma}(\mathbb{R}) = \cup_{\phi\in[0,1)} NH_N^{\gamma,\phi}(\mathbb{R}). \tag{.77}$$

We claim that

$$NH_N^{\gamma}(\mathbb{R}) \subset \cup_{j=0,\dots,[5/\gamma]-1} NH_N^{2\gamma, j/[5/\gamma]}(\mathbb{R}). \tag{.78}$$

Indeed, suppose that $L \in NH_N^\gamma(\mathbb{R})$. Then for some number $\phi \in [0,1)$ and vector $v \in \mathbb{R}^N$ with $|v| = 1$, we have $|(L - \exp(2\pi i\phi))v| \leq \gamma$. Let j be the nearest integer to $[5/\gamma]\phi$ and let $\phi_\gamma = j/[5/\gamma]$; then $\phi - \phi_\gamma \leq 1/(2(5/\gamma - 1)) < \gamma/(2\pi)$. Thus

$$\begin{aligned} &|(L - \exp(2\pi i\phi_\gamma))v| \leq \\ &\quad |(L - \exp(2\pi i\phi))v| + |\exp(2\pi i\phi) - \exp(2\pi i\phi_\gamma)| \leq 2\gamma \end{aligned} \tag{.79}$$

and $L \in NH_N^{2\gamma, j/[\pi/\gamma+1]}(\mathbb{R})$ as claimed.

Next, we claim that every matrix in $NH_N^{2\gamma,j/[5/\gamma]}(\mathbb{R})$ lies within 2γ of a matrix in $NH_N^{0,j/[5/\gamma]}(\mathbb{R})$, where we use the Euclidean ($\mathbb{R}^{N^2}$) norm on $M_N(\mathbb{R})$ (not the matrix norm). Consider $L \in NH_N^{2\gamma,j/[5/\gamma]}(\mathbb{R})$, $\phi \in [0,1)$, and $v \in \mathbb{R}^N$ with $|v| = 1$ and $|(L - \exp(2\pi ij/[5/\gamma]))v| \leq 2\gamma$. Let $w = (L - \exp(2\pi ij/[5/\gamma]))v$ and let $M \in M_N(\mathbb{R})$ be the matrix whose kth row is $w_k v$, where w_k is the kth coordinate of w. Then the Euclidean norm of M is $|w| \leq 2\gamma$ and $Mv = w$, so that $(L - M - \exp(2\pi ij/[5/\gamma]))v = 0$ and hence $L - M \in NH_N^{0,j/[5/\gamma]}(\mathbb{R})$.

We complete the estimate (.75) by estimating for each j the number of γ-balls needed to cover $NH_N^{0,j/[5/\gamma]}(\mathbb{R})$ within an appropriate bounded domain. It then follows from the previous paragraph that if we inflate each of these balls to the concentric ball of radius 3γ, the collection of larger balls will cover $NH_N^{2\gamma,j/[5/\gamma]}(\mathbb{R})$, and from the paragraph before that the union over j of these covers will then cover $NH_N^\gamma(\mathbb{R})$. To this end, we show that each $NH_N^{0,j/[5/\gamma]}(\mathbb{R})$ is a real algebraic set and compute its codimension[1]. Then we will apply an estimate of Yomdin [Y] on the number of γ-balls necessary to cover a given algebraic set by polynomials of known degree.

Notice that

$$NH_N^{0,\phi}(\mathbb{R}) = \{L \in M_N(\mathbb{R}) : \ \det(L - \exp(2\pi i\phi)Id) = 0\}. \tag{.80}$$

We split into the two cases $\exp(2\pi i\phi) \in \mathbb{R}$ (that is, $\phi = 0$ or $1/2$) and $\exp(2\pi i\phi) \notin \mathbb{R}$. In the first case, the equation $\det(L \pm Id) = 0$ is a polynomial of degree N in the entries of L, so $NH_N^{0,0}(\mathbb{R})$ and $NH_N^{0,1/2}(\mathbb{R})$ are real algebraic sets defined by a single polynomial of degree N.

In the second case, decompose the equation $\det(L - \exp(2\pi i\phi)Id) = 0$ into two parts: $Re[\det(L - \exp(2\pi i\phi)Id)] = 0$ and $Im[\det(L - \exp(2\pi i\phi)Id)] = 0$. Each part is given by a real polynomial of degree N. Furthermore, these two polynomials are algebraically independent, since otherwise $Re[\det(L - \exp(2\pi i\phi)Id)]$ and $Im[\det(L - \exp(2\pi i\phi)Id)]$ would satisfy some polynomial relationship and, thus,

[1]Unfortunately $NH_N^0(\mathbb{R})$, in contrast to $NH_N^{0,j/[5/\gamma]}(\mathbb{R})$, is not algebraic

$\det(L-\exp(2\pi i\phi)Id)$ would take on values only in some real algebraic subset of the complex plane. However, for $N \geq 2$ (which is necessary for complex eigenvalues), by considering real diagonal matrices L we see that the values of $\det(L-\exp(2\pi i\phi)Id)$ contain an open set in $\mathbb{C}$. Therefore, $NH_N^{0,\phi}(\mathbb{R})$ is a real algebraic set given by two algebraically independent polynomials of degree N.

Covering Lemma for Algebraic Sets *([Y], Lemma 4.6) Let $V \subset \mathbb{R}^m$ be an algebraic set given by k algebraically independent polynomials $p_1, \ldots, p_k$ of degrees $d_1, \ldots, d_k$ respectively, i.e. $V = \{x \in \mathbb{R}^m : p_1(x) = 0, \ldots, p_k(x) = 0\}$. Let $C_A^m(s)$ be the cube in $\mathbb{R}^m$ with side $2s$ centered at some point A. Then for $\gamma \leq s$, the number of γ-balls necessary to cover $V \cap C_A^m(s)$ does not exceed $C(D,m)\,(2s/\gamma)^{m-k}$, where the constant $C(D,m)$ depends only on the dimension m and product of degrees $D = \prod_i d_i$.*

Remark A.6. *Some additional arguments based on Bezout's Theorem give an upper estimate of $C(D,m)$ by $2^m D$ for γ sufficiently small.*

To complete the proof of Proposition A.5, we apply the Covering Lemma for Algebraic Sets to each $NH^{0,j/[5/\gamma]}(\mathbb{R})$, where $j = 0, \ldots, [5/\gamma]-1$, with $m = N^2$, $s = Kr$, and A as in the statement of the proposition. (Notice that if $U \in C^{N^2}(r)$ then $A + UB \in C_A^{N^2}(Kr)$, so we need only cover the part of $NH^{0,j/[5/\gamma]}(\mathbb{R})$ lying in the latter set.) In the case that $j/[5/\gamma] = 0$ or $1/2$, we have $k = 1$, $d_1 = N$, and $D = N$, so the number of covering γ-balls is bounded by $C(N, N^2)(2Kr/\gamma)^{N^2-1}$. In the case of other j, we have $k = 2$, $d_1 = d_2 = N$, and $D = N^2$, so the number of covering γ-balls is bounded by $C(N^2, N^2)(2Kr/\gamma)^{N^2-2}$. The number of j's of the second type is less than $5/\gamma$. Combining all these estimates along with (.78) we get that $NH^\gamma(\mathbb{R}) \cap C_A^{N^2}(Kr)$ can be covered by $C(N^2, N^2)(2 + 5/(2Kr))(2Kr/\gamma)^{N^2-1}$ balls of radius 3γ.

Finally, notice that the preimage of a ball of radius 3γ under the map $U \mapsto A + UB$ is contained in a ball of radius $3K\gamma$, whose μ_{r,N^2}-measure is less than $(3K\gamma/r)^{N^2}$. Therefore the total measure of $3K\gamma$-balls needed to cover the set $\{U \in C^{N^2}(r) : \gamma(A + UB) \leq \gamma\}$ is at most $C(N)K^{2N^2}\gamma/r^2$, where the constant $C(N)$ depends only on N. Q.E.D.

Appendix B: Orthogonal transformations of R^N and the spaces of homogeneous polynomials

In this appendix, we prove that the scalar product (1.12) in the space $W_{k,N}$ of homogeneous N-vector polynomials of degree k in N variables is invariant with respect to orthogonal transformations of $\mathbb{R}^N$.

Lemma B.1. *Let $x \in \mathbb{R}^N$ be given by N coordinates $x = (x_1, \ldots, x_N)$. For $k \in \mathbb{Z}_+$, consider homogeneous polynomials $p_k(x) = \sum_{|\alpha|=k} \vec{\varepsilon}_\alpha x^\alpha \in W_{k,N}$ in N variables, where $x^\alpha = x_1^{\alpha_1} \ldots x_N^{\alpha_N}$. Let $O \in SO(N)$ be an orthogonal transformation of $\mathbb{R}^N$. Denote by $x' = (x'_1, \ldots, x'_N)$ the new coordinate system induced by the relation $x = Ox'$. Write $p'_k(x') = p_k(Ox') = \sum_{|\alpha|=k} \vec{\varepsilon}'_\alpha (x')^\alpha$ in the new coordinate system. Then for all $\{\vec{\varepsilon}_\alpha\}_{|\alpha|=k}$ and $\{\vec{\nu}_\alpha\}_{|\alpha|=k}$,*

$$\sum_{|\alpha|=k} \binom{k}{\alpha}^{-1} \langle \vec{\varepsilon}_\alpha, \vec{\nu}_\alpha \rangle = \sum_{|\alpha|=k} \binom{k}{\alpha}^{-1} \langle \vec{\varepsilon}'_\alpha, \vec{\nu}'_\alpha \rangle. \tag{.81}$$

Proof: For this lemma it will be helpful to use a different notation for monomials. Given a k-tuple $\beta = (\beta_1, \ldots, \beta_k) \in \{1, 2, \ldots, N\}^k$, define $x^{(\beta)} = x_{\beta_1} x_{\beta_2} \cdots x_{\beta_k}$. Notice that $x^{(\beta)} = x^\alpha$ where α_i is the number of indices j for which $\beta_j = i$. Write $\alpha(\beta)$ for the multiindex corresponding in this manner to the k-tuple β, and notice that for each multiindex α there are $\binom{k}{\alpha}$ different k-tuples β for which $\alpha(\beta) = \alpha$. Let $|\beta| = k$.

Given a polynomial p_k as in the statement of the lemma, we can write $p_k(x) = \sum_{|\beta|=k} \vec{\varepsilon}_{(\beta)} x^{(\beta)}$, where $\vec{\varepsilon}_{(\beta)} = \binom{k}{\alpha}^{-1} \vec{\varepsilon}_{\alpha(\beta)}$. We can also rewrite the scalar product as follows:

$$\sum_{|\alpha|=k} \binom{k}{\alpha}^{-1} \langle \vec{\varepsilon}_\alpha, \vec{\nu}_\alpha \rangle = \sum_{|\beta|=k} \langle \vec{\varepsilon}_{(\beta)}, \vec{\nu}_{(\beta)} \rangle. \tag{.82}$$

(Remember that for each α, there are $\binom{k}{\alpha}$ corresponding terms on the right-hand side.)

Our goal is then to show that

$$\sum_{|\beta|=k} \langle \vec{\varepsilon}_{(\beta)}, \vec{\nu}_{(\beta)} \rangle = \sum_{|\beta|=k} \langle \vec{\varepsilon}'_{(\beta)}, \vec{\nu}'_{(\beta)} \rangle. \tag{.83}$$

We prove this by induction on k. For $k = 0$, the identity is trivial. Assume that the identity now holds for some $k \geq 0$. Given β with $|\beta| = k$ and $i \in \{1, 2, \ldots, N\}$, let (β, i) be the $(k+1)$-tuple $(\beta_1, \ldots, \beta_k, i)$. Also, let $\vec{\varepsilon}^{\,i}_{(\beta)} = \vec{\varepsilon}_{(\beta,i)}$. The reason for this alternate notation is that we will mean something different below by $\vec{\varepsilon}^{\,i\prime}_{(\beta)}$ and $\vec{\varepsilon}'_{(\beta,i)}$. In the former case, we fix i and apply the coordinate transformation O to the polynomial $\sum_{|\beta|=k} \vec{\varepsilon}^{\,i}_{(\beta)} x^{(\beta)}$ to get the coefficients $\vec{\varepsilon}^{\,i\prime}_{(\beta)}$. In the latter case, we transform the polynomial $\sum_{|(\beta,i)|=k+1} \vec{\varepsilon}_{(\beta,i)} x^{(\beta,i)}$ to get the coefficients $\vec{\varepsilon}'_{(\beta,i)}$.

Next, notice that

$$\sum_{|(\beta,i)|=k+1} \vec{\varepsilon}_{(\beta,i)} x^{(\beta,i)} = \sum_{i=1}^{N} x_i \sum_{|\beta|=k} \vec{\varepsilon}^{\,i}_{(\beta)} x^{(\beta)}. \tag{.84}$$

Applying the coordinate change $x = Ox'$ to both sides, we get

$$\sum_{|(\beta,i)|=k+1} \vec{\varepsilon}'_{(\beta,i)} (x')^{(\beta,i)} = \sum_{i=1}^{N} \sum_{j=1}^{N} O_{ij} x'_j \sum_{|\beta|=k} \vec{\varepsilon}^{\,i\prime}_{(\beta)} (x')^{(\beta)}. \tag{.85}$$

It follows that

$$\vec{\varepsilon}'_{(\beta,j)} = \sum_{i=1}^{N} O_{ij} \vec{\varepsilon}^{\,i\prime}_{(\beta)}. \tag{.86}$$

A similar identity holds with $\vec{\varepsilon}$ replaced by $\vec{\nu}$, whereupon

$$\sum_{|(\beta,j)|=k+1} \langle \vec{\varepsilon}'_{(\beta,j)}, \vec{\nu}'_{(\beta,j)} \rangle = \sum_{j=1}^{N} \sum_{|\beta|=k} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \langle O_{ij} \vec{\varepsilon}^{\,i\prime}_{(\beta)}, O_{\ell j} \vec{\nu}^{\,\ell\prime}_{(\beta)} \rangle. \tag{.87}$$

Since O is an orthogonal matrix, $\sum_{j=1}^{N} O_{ij} O_{\ell j} = \delta_{i\ell}$. Exchanging the order of summation on the right-hand side above, we then have

$$\sum_{|(\beta,j)|=k+1} \langle \vec{\varepsilon}'_{(\beta,j)}, \vec{\nu}'_{(\beta,j)} \rangle = \sum_{i=1}^{N} \sum_{|\beta|=k} \langle \vec{\varepsilon}^{\,i\prime}_{(\beta)}, \vec{\nu}^{\,i\prime}_{(\beta)} \rangle = \sum_{|(\beta,i)|=k+1} \langle \vec{\varepsilon}^{\,i}_{(\beta)}, \vec{\nu}^{\,i}_{(\beta)} \rangle \tag{.88}$$

by the inductive hypothesis. Q.E.D.

Appendix C: Embedding of the space of diffeomorphisms of a compact manifold $\mathbf{Diff}^r(M)$ into the one of the ball $\mathbf{Diff}^r(B^N)$

In this Appendix we describe how to extend a diffemorphism of a compact manifold embedded into a Euclidean space and what conditions we need to ensure that the results of Sacker [Sac] and Fenichel [F] about persistence of invariant manifolds apply.

Recall that M is a smooth (C^∞) compact manifold. First we consider a manifold $\tilde{M} = M \times [0,1]/\sim$, where the equivalence relation is defined by $(x,0) \sim (f(x),1)$ for all $x \in M$. $\tilde{M}$ is as smooth as f is and carries a naturally defined vector field X_f whose time one map, restricted to $M \times \{0\}$ coincides with f. Such a construction is usually called *suspension*. Now embed $\tilde{M}$ into the interior of the closed unit ball $B^N \subset \mathbb{R}^N$. Recall that given a smooth (C^∞) compact manifold $\tilde{M}$ of dimension D, for $N > 2D$ the Whitney Embedding Theorem says that a generic smooth function from $\tilde{M}$ to $\mathbb{R}^N$ is a diffeomorphism between $\tilde{M}$ and its image. For a moment denote this image by $E(\tilde{M})$, where $E : \tilde{M} \to \mathbb{R}^N$ is the embedding. Suppose also that the image $E(\tilde{M})$ is contained in the unit ball $B^N(O)$ whose center has distance 2 from the origin. Now embed $\tilde{M}$ into $\mathbb{R}^{N+1}$ as follows: Points on $\tilde{M}$ have naturally defined coordinates $\{(x,\phi) \in M \times [0,1)\}$ once $M \times \{0\}$ is fixed. Embed $B^N(O) \times \mathbb{S}^1 = \{(y,\phi) \in B^N(O) \times [0,2\pi) \subset \mathbb{R}^{N+1}\}$ by

$$(y,\phi) \mapsto (\sin\phi \|E(y)\|, \cos\phi E(y)) \in \mathbb{R}^{N+1}.$$

Using this formula embed $\tilde{M}$. By construction $M \times \{0\}$ belongs to the hyperplane, where the first component equals zero. Now to simplify notation, we identify $\tilde{M}$ and M with their images, so that $\tilde{M}$ and $M \times \{0\}$ become submanifolds of $\mathbb{R}^{N+1}$.

Let $\tilde{T} \subset \mathbb{R}^{N+1}$ be a closed neighborhood of $\tilde{M}$, chosen sufficiently small that there is a well-defined projection $\tilde{\pi} : \tilde{T} \to \tilde{M}$ for which $\tilde{\pi}(\tilde{x})$ is the closest point in $\tilde{M}$ to $\tilde{x}$. Then for each $\tilde{y} \in \tilde{M}$, $\tilde{\pi}^{-1}(\tilde{y})$ is an $(N-D)$-dimensional disk. Then we can extend each vector field X_f on $\tilde{M}$ to a vector field $\mathcal{X}_f$ on $\tilde{T}$ so that the component tangent to $\tilde{\pi}^{-1}(\tilde{y})$ is strongly dominated by the orthogonal one and is directed toward $\tilde{y}$. Consider now the Poincaré return map of $\mathcal{X}_f$ from $T = \tilde{T} \cap \{B^N(O) \times \{0\}\}$ into $B^N(O) \times \{0\}$, which is well-defined by the construction. Denote this map by F. The vector field $\mathcal{X}_f$ is directed so that F maps T strictly inside itself.

Now the closed neighborhood T of M can be considered as a subset of $\mathbb{R}^N$ and can be chosen sufficiently small that there is a well-defined projection $\pi : T \to M$ for which $\pi(x)$ is the closest point in M to x. Then for each $y \in M$, $\pi^{-1}(y)$ is an $(N-D)$-dimensional disk. For $0 < \rho < 1$ and $y, z \in M$ choose a linear function $g_{\rho,y,z} : \pi^{-1}(y) \to \pi^{-1}(z)$ that maps y to z and contracts distances by a factor of ρ, and such that the dependence of $g_{\rho,y,z}$ on y and z is C^r. Then we can extend each $F \in \mathrm{Diff}^r(M)$ to a function $f \in C^r(T)$ that is a diffeomorphism from T to a subset of its interior by letting

$$f(x) = g_{\rho,\pi(x),F(\pi(x))}(x)$$

where $\rho = \min(\|F^{-1}\|_{C^1}^{-r}, 1)/2$. Then by Fenichel's Theorem [F] every sufficiently small perturbation $f_\varepsilon \in C^r(T)$ of such an f has an invariant manifold $M_\varepsilon \subset T$ for which $\pi|_{M_\varepsilon}$ is a C^r diffeomorphism from M_ε to M. Then to such an f_ε we can associate a diffeomorphism $F_\varepsilon \in \mathrm{Diff}^r(M)$ by letting

$$F_\varepsilon(y) = \pi(f_\varepsilon(\pi|_{M_\varepsilon}^{-1}(y))).$$

Notice that the periodic points of f_ε all lie on M_ε and are in one-to-one correspondence with the periodic points of f_ε. Furthermore, because F_ε and $f_\varepsilon|_{M_\varepsilon}$ are conjugate, the hyperbolicity of each periodic orbit is the same for either map. Thus any estimate on $P_n(f_\varepsilon)$ or $\gamma_n(f_\varepsilon)$ applies also to F_ε.

Appendix D: Pathological examples of decay of product of distances of recurrent trajectories

In this Appendix we present several examples of horseshoe diffeomorphisms which show that with the ideas we have in this paper the estimate $\exp(Cn^{1+\delta})$ on the growth of the number of periodic points (the Main Theorem from section 1.3) for any real positive number δ is asymptotically the optimal one. More exactly, the Shift Theorem, stated in 3.5, is crucial to split all almost periodic trajectories into classes as in (3.12). In section 3.5 we outline the proof of this Theorem and it might be helpful to review the strategy presented there, especially, the last remark right before subsection 3.5.1. Here, for $\gamma_n(C,\delta) = \exp(-Cn \ln n)$, which is the "roughly" inverse of the number of periodic points, we give an example of a periodic orbit which is nonsimple, have no leading saddles, and no close returns (gaps), defined in section 3.5. Thus we can't deal with these kind of trajectories with our methods.

Example 1 Consider the sequence of periodic orbits of a horseshoe map with symbolic dynamics

$$\begin{aligned}
S_0 &= 1\\
S_1 &= 0\\
S_2 &= 01\\
S_3 &= 010\\
S_4 &= 01001\\
S_5 &= 01001010\\
S_6 &= 0100101001001\\
&\vdots \; .
\end{aligned}$$

Each sequence is the concatenation of the previous two sequences; it can also be generated from the previous sequence by the substitution rules $0 \to 01$ and $1 \to 0$. The number of symbols in S_n is the nth Fibonacci number F_n.

Notice also that

$$\begin{aligned} S_n &= S_{n-1}S_{n-2} \\ &= S_{n-2}S_{n-3}S_{n-2} \\ &= S_{n-3}S_{n-4}S_{n-3}S_{n-3}S_{n-4} \\ &= S_{n-4}S_{n-5}S_{n-4}S_{n-4}S_{n-5}S_{n-4}S_{n-5}S_{n-4} \\ &= \ldots. \end{aligned}$$

More formally, the sequence S_n can be generated from S_k for any $0 \leq k \leq n$ by replacing each 0 in S_k by S_{n-k+1} and each 1 by S_{n-k}. We refer below to this decomposition of S_n into copies of S_{n-k} and S_{n-k+1} as "decomposition k"

Every three symbol subsequence of S_n is either 010, 100, 001, or 101. Furthermore, regarding each S_n as a cyclic sequence, each of the four triplets above occurs at least once in S_4, at least once in S_5, at least twice in S_6, and in general at least F_{n-4} times in S_n for $n \geq 4$. The importance of this observation below will be that in decomposition k for $4 \leq k \leq n$, each of the substrings $S_{n-k+1}S_{n-k}S_{n-k+1}$, $S_{n-k}S_{n-k+1}S_{n-k+1}$, $S_{n-k+1}S_{n-k+1}S_{n-k}$, and $S_{n-k}S_{n-k+1}S_{n-k}$ occurs at least F_{k-4} times.

Now let $x_0, x_1, \ldots, x_{p-1}$ be points in the periodic orbit with symbolic dynamics S_n, where $p = F_n$ is the length of S_n No matter where the symbolic sequence of x_0 starts within S_n, we claim that for n sufficiently large,

$$\prod_{j=1}^{p-1} |x_0 - x_j| \leq e^{(c_1 - c_2 n)p} \leq e^{c_1 p - c_3 p \log p}$$

for some positive constants c_1, c_2, c_3 independent of n. The latter inequality follows from the fact that $p \leq 2^n$, so it remains to prove the first inequality.

Assume that the distance between any two points in the nonwandering set is at most 1. Say the symbolic sequence of x_0 starts at the mth symbol of S_n. If for some positive integer q, the block of $2q-1$ symbols centered at the mth symbol is repeated centered at the ℓth symbol, then the distance between the points x_0 and $x_{\ell-m}$ is bounded above by e^{-cq} for an appropriate positive constants c. Here the index $\ell - m$ is taken modulo p.

Now for $4 \leq k \leq n$, in decomposition k the mth symbol in S_n lies in a copy of either S_{n-k} or S_{n-k+1}, which in turn lies in the middle of one of the four substrings $S_{n-k+1}S_{n-k}S_{n-k+1}$, $S_{n-k}S_{n-k+1}S_{n-k+1}$, $S_{n-k+1}S_{n-k+1}S_{n-k}$, and $S_{n-k}S_{n-k+1}S_{n-k}$ described above. Each such substring occurs at least F_{k-4} times in S_n, and all but one of these occurrences contributes a factor of at most $e^{-cF_{n-k}}$ to the product of distances $|x_0 - x_j|$. Therefore for $n \geq 6$,

$$\begin{aligned}
\prod_{j=1}^{p-1} |x_0 - x_j| &\leq \prod_{k=6}^{n} (e^{-cF_{n-k}})^{F_{k-4}-F_{k-5}} \\
&= e^{-c\sum_{k=6}^{n} F_{n-k}F_{k-6}} \\
&\leq e^{-c(n-5)F_n/F_8} \\
&= e^{c(5-n)p/34}.
\end{aligned}$$

(The estimate $F_{n-k}F_{k-6} \geq F_n/F_8$ can be proved by induction, but heuristically this type of estimate follows from the fact that F_n is approximately an exponential function of n.)

Example 2 Consider now the aperiodic nonwandering orbit of the horseshoe map whose symbolic dynamics are given as follows. Given a sequence of positive integers $k_1, k_2, \ldots$, let $S_0 = 0$ and $S_n = 1S_{n-1}S_{n-1}\cdots S_{n-1}1$ where S_{n-1} occurs $2k_n+1$ consecutive times. For example, if $k_n = n$ then

$$\begin{aligned}
S_0 &= 0 \\
S_1 &= 10001 \\
S_2 &= 1100011000110001100011000111 \\
&\vdots \quad .
\end{aligned}$$

Each sequence is symmetric, and for $n \geq 1$, each S_n contains a copy of S_{n-1} at its center. Let L_n be the length of S_n; then $L_0 = 1$ and $L_n = (2k_n + 1)L_{n-1} + 2$ for $n \geq 1$. One can easily check that $k_1k_2\cdots k_n \leq L_n \leq 5^n k_1k_2\cdots k_n$.

Let x_0 be the point on the nonwandering set whose symbolic sequence has middle L_n symbols S_n for each $n \geq 0$. By symmetry, to estimate the product of distances $|x_0 - x_j|$ as j goes from 1 to $L_n - 1$, we can estimate the product as j goes from $-(L_n - 1)$ to $L_n - 1$, excluding $j = 0$, and take the square root of the latter estimate.

As in the previous example, let c be a positive constant such that $|x_0 - x_j| \leq e^{-cq}$, where q is the largest positive integer for which the middle $2q - 1$ symbols of the sequences for x_0 and x_j agree, or $q = 0$ if their middle symbols do not agree. Then for all $n \geq 1$ and $-k_n \leq m \leq k_n$ we have $|x_0 - x_{mL_{n-1}}| \leq e^{-c(k_n-|m|+1/2)L_{n-1}}$. The square root of the product of these upper bounds, excluding $m = 0$, is

$$\prod_{m=1}^{k_n} e^{-c(k_n-m+1/2)L_{n-1}} = e^{-ck_n^2 L_{n-1}/2} \leq e^{-ck_n L_n/10}.$$

Here we used the inequality $5k_n L_{n-1} \geq (2k_n+1)L_{n-1} + 2 = L_n$.

In addition, for n and m as above and all $-k_{n-1} \leq p \leq k_{n-1}$ we have $|x_0 - x_{mL_{n-1}+pL_{n-2}}| \leq e^{-c(k_{n-1}-|p|+1/2)L_{n-2}}$. The square root of the product of these upper bounds, exluding $p = 0$, is

$$\begin{aligned} \prod_{m=-k_n}^{k_n} \prod_{p=1}^{k_{n-1}} e^{-c(k_{n-1}-p+1/2)L_{n-2}} &= e^{-c(2k_n+1)k_{n-1}^2 L_{n-2}/2} \\ &\leq e^{-c(2k_n+1)k_{n-1}(L_{n-1}+1)/12}. \end{aligned}$$

Here we used the inequality $6k_{n-1}L_{n-2} \geq (2k_{n-2}+1)L_{n-2} + 3 = L_{n-1} + 1$. Then in turn we can say $(2k_n+1)(L_{n-1}+1) \geq (2k_n+1)L_{n-1} + 3 = L_n + 1$, so that the bound on the product above can be replaced by $e^{-ck_{n-1}L_n/12}$.

In a similar manner, we can bound above another set of terms contributing to the product of distances $|x_0 - x_j|$ by $e^{-ck_{n-\ell}L_n/12}$ for $\ell = 2, 3, \ldots, n-1$. Multiplying all these bounds together we conclude that

$$\prod_{j=1}^{L_n-1} |x_0 - x_j| \leq e^{-c(k_1+k_2+\cdots+k_n)L_n/12}.$$

Notice that if $k_n = k$ independent of n, then $L_n \sim (2k+1)^n$ and $k_1 + k_2 + \cdots + k_n = nk \sim \log L_n$. Thus we get an estimate similar to Example 1.

If $k_n = n^\alpha$, then $\log L_n \sim n \log n$ and hence $k_1 + k_2 + \cdots + k_n \sim n^{\alpha+1} \sim (\log L_n)^{\alpha+1}$, loosely speaking. The closest returns to x_0 are of the form $-\log|x_0 - x_{L_n}| \sim k_{n+1}L_n \sim L_n(\log L_n)^\alpha$, loosely speaking again. Thus if we attempt to apply the Inductive Hypothesis with $\gamma_j = C^{-j(\log j)^\beta}$, this example with $\alpha = \beta - 1/2$ shows that the product of distances along a hyperbolic trajectory can be smaller than any fixed power of γ_j for arbitrarily large j, despite the fact that the closest return over j iterations is larger than any fixed power of γ_j for j sufficiently large.

References

[A] A. Arnold, Problems of Arnold's seminar, 1989.

[AM] M. Artin & B. Mazur, On periodic orbits, Ann. Math. 81 (1965), 82–99.

[F] N. Fenichel, Persistence and smoothness of invariant manifolds for flows, Indiana Univ. Math. J. 21 (1971), 193–226.

[GG] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*, Springer-Verlag (1973).

[GST1] S. Gonchenko, L. Shil'nikov, D. Turaev, On models with non-rough Poincaré homoclinic curves, Physica D 62 (1993), 1–14.

[GST2] S. Gonchenko, L. Shil'nikov, D. Turaev, Homoclinic tangencies of an arbitrary order in Newhouse regions, (Russian) Dynamical systems (Russian) (Moscow, 1998), 69–128, Itogi Nauki Tekh. Ser. Sovrem. Mat. Prilozh. Temat. Obz., 67, Vseross. Inst. Nauchn. i Tekhn. Inform. (VINITI), Moscow, 1999.

[GK] A. Gorodetski, V. Kaloshin, How often surface diffeomorphisms have infinitely many sinks and hyperbolicity of periodic points near a homoclinic tangency, preprint, 2004, 99pp, www.its.caltech.edu/ tilde kaloshin .

[GY] A. Grigoriev, S. Yakovenko, Topology of Generic Multijet Preimages and Blow-up via Newton Interpolation, J. of Diff. Equations, **150**, 349-362, (1998).

[G] M. Gromov, On entropy of holomorphic maps, preprint.

[HSY] B. Hunt, T. Sauer, J. Yorke, Prevalence: a translation-invariant almost every for infinite dimensional spaces, Bull. Amer. Math. Soc. **27** (1992), 217–238; Prevalence: an addendum, Bull. Amer. Math. Soc. **28** (1993), 306–307.

[K1] V. Kaloshin, An extension of the Artin-Mazur Theorem, Ann. Math. 150 (1999), 729–741.

[K2] V. Kaloshin, Generic diffeomorphisms with superexponential growth of number of periodic orbits, Comm in Math Phys, **211**, (2000), no.1, 253-271.

[K3] V. Kaloshin, Some prevalent properties of smooth dynamical systems, Tr. Math. Inst. Steklova 213 (1997), 123–151.

[K4] V. Kaloshin, Stretched exponential estimate on growth of the number of periodic points for prevalent diffeomorphisms, PhD Thesis, Princeton University, 2001.

[K5] V. Kaloshin, Stretched exponential estimate on growth of the number of periodic points for prevalent diffeomorphisms, part 2, preprint, 86pp, www.its.caltech.edu/ tilde kaloshin.

[KH] V. Kaloshin, B. Hunt, A stretched exponential bound on the rate of growth of the number of periodic points for prevalent diffeomorphisms. I & II. Electron. Res. Announc. Amer. Math. Soc. 7 (2001), 17–27 & 28–36.

[KK] V. Kaloshin, O. Kozlovski, An example of a C^r-unimodal map with an arbitrary fast growth of the number of periodic points, preprint, 5pp.

[MMS] M. Martens, W. de Melo, S. Van Strien, Julia-Fatou-Sullivan theory for real one-dimensional dynamics, Acta Math. **168**, (1992), no.3-4, 273-318.

[PM] J. Palis and W. de Melo, *Geometric Theory of Dynamical Systems: An Introduction*, Springer-Verlag (1982).

[O] J. C. Oxtoby, *Measure and Category*, Springer-Verlag (1971).

[San] L. Santalo, Integral Geometry and Geometric Probability. Encycl of Math and its Appl, Vol. **1**. Addison-Wesley Publ Co., Mass.-London-Amsterdam, 1976.

[Sac] R. J. Sacker, A perturbation Theorem for invariant manifolds and Hölder continuity, J. Math. Mech. 18 (1969), 705–762.

[Sar] A. Sard, The measure of the critical points of differentiable maps, Bull. Amer. Math. Soc. 48 (1942), 883–890.

[Sz] W. Szlenk, An Introduction to the theory of smooth dymanical systems, Wiley & Sons, New York, 1984.

[W] H. Whitney, Differentiable manifolds, Ann. Math. 37 (1936), 645–680.

[Y] Y. Yomdin, A quantitative version of the Kupka-Smale Theorem, Ergod. Th. & Dynam, Sys. (1985), **5**, 449–472.