Newton Interpolation Polynomials, Discretization Method, and Certain Prevalent Properties in Dynamical Systems

Anton Gorodetski, Brian Hunt, Vadim Kaloshin*

Mathematics Subject Classification (2000). 37C05, 37C50, 37D25, 37C29.

Keywords. Discretization method, Newton Interpolation Polynomials, prevalence, pseudotrajectory, growth of number of periodic points, Newhouse phenomenon.

Abstract. We describe a general method of studying prevalent properties of diffeomorphisms of a compact manifold M, where by *prevalent* we mean true for Lebesgue almost every parameter ε in a generic finite-parameter family $\{f_\varepsilon\}$ of diffeomorphisms on M.

Usually a dynamical property $\mathcal{P}$ can be formulated in terms of properties $\mathcal{P}_n$ of trajectories of finite length n. Let $\mathcal{P}$ be such a dynamical property that can be expressed only in terms of periodic trajectories. *The first idea* of the method is to *discretize* M and split the set of all possible periodic trajectories of length n for the entire family $\{f_\varepsilon\}$ into a *finite number* of approximating periodic pseudotrajectories. Then for each such pseudotrajectory, we estimate the measure of parameters for which it fails $\mathcal{P}_n$. This bounds the total parameter measure for which $\mathcal{P}_n$ fails by a finite sum over the periodic pseudotrajectories of length n. Application of Newton Interpolation Polynomials to estimate the measure of parameters that fail $\mathcal{P}_n$ for a given periodic pseudotrajectory of length n is *the second idea*.

We outline application of these ideas to two quite different problems:

- Growth of number of periodic points for prevalent diffeomorphisms (Kaloshin-Hunt).
- Palis' conjecture on finititude of number of "localized" sinks for prevalent surface diffeomorphisms (Gorodetski-Kaloshin).

1. Introduction

A classical problem in dynamics, geometry, and topology is the description of generic behavior. Given a set of objects what are the properties of a generic element of the set? This question applies to diffeomorphisms, Riemannian metrics, linear operators, and vector fields, just to give several examples. The traditional approach is based on the category theorem of Baire. A countable intersection of open, dense sets is called a *residual,* or *topologically generic*, set. The Baire category theorem says that topologically generic sets of a complete metric space (or, more generally, Baire space) are dense. The book of Oxtoby [O] provides a rich

*Supported by an Alfred Sloan Research Fellowship and NSF Grant No. DMS-0300229

variety of topologically generic mathematical objects. However, in many different areas of mathematics examples of "wild behavior" of topologically generic objects have been detected (see [HSY], [Ka2], [OY], [Si] and references there). In this paper we are concerned with generic properties in dynamics, particularly those that are not generic topologically but are generic in a measure-theoretic sense.

In the 1960s two main theories in dynamical systems were developed, one of which was designed for conservative systems and called *KAM* for *Kolmogorov–Arnold–Moser* and the other was constructed for general dynamical systems (non-conservative, dissipative) and called *hyperbolic*.

Kolmogorov [Ko], in his plenary talk of ICM 1954, point out that a different notion of genericity may be appropriate: "In order to obtain negative results concerning insignificant or exceptional character of some phenomenon we shall apply the following, somewhat haphazard, technique: if in a class K of functions $f(x)$ one can introduce a finite number of functionals

$$F_1(f),\, F_2(f), \dots,\, F_r(f),$$

which in some sense can naturally be considered as taking "arbitrary" values in general

$$F_1(f) = C_1,\, F_2(f) = C_2, \dots, F_r(f) = C_r$$

from some domain of the r-dimensional space of points $C = (C_1, \dots, C_r)$, then any phenomenon that takes place only if C belongs to a set of zero r-dimensional measure will be regarded exceptional and subject to "neglect"."

A somewhat similar way to define a measure-theoretic genericity, often called *prevalence*, is the following: *We call a property $\mathcal{P}$ prevalent if for a generic*[1] *finite–parameter family $\{f_\varepsilon\}_{\varepsilon \in B}$ for Lebesgue almost every parameter ε the corresponding f_ε satisfies $\mathcal{P}$. If complement of a property is prevalent such a property is called shy.* We shall discuss prevalence further in section 9.

There are many examples when topological genericity and measure-theoretic genericity do not coincide. We just mention a few of them (see [HSY], [Ka2], [OY] for many more).

• *Diophantine numbers* form a set of full measure on the line $\mathbb{R}$, but are topologically negligible (that is the complement of the set is topologically generic).

• For a topologically generic, even open dense, set of circle maps preserving orientation there is a finite number of attracting and repelling periodic orbits. All other orbits accumulate to these orbits both forward or backward in time. However, as the famous example of Arnold, called *Arnold tongues*, shows in the family $f_{\alpha,\varepsilon} : \theta \mapsto \theta + \alpha + \varepsilon \sin \theta$ that the smaller ε is, the smaller is the measure of α values such that $f_{\alpha,\varepsilon}$ has this property. Moreover, the main result of KAM theory says that for consevative systems close to integrable most, in a measure–theoretic sense, motions are *quasiperiodic*.

• In general dynamical systems a dream of the 1960s was to prove that a generic dynamical system is structurally stable. However, this dream evaporated by the end of that decade. One of the beautiful counterexamples is due to Newhouse

[1] we give a rigorous definition in section 9

[N1, N2]. He shows that there is an open set in the space of diffemorphisms of a compact manifold such that a generic diffeomorphism in this open set has *infinitely many coexisting sinks* (attracting periodic orbits). Below we show in some weak sense this phenomenon is shy (see section 7). This phenomenon is closely related to Palis' program [Pa] which is discussed next.

Let $\mathrm{Diff}^r(M)$ be the space of C^r diffeomorphisms of a smooth compact manifold M with the uniform C^r-topology, where $\dim M \geq 2$, and let $f \in \mathrm{Diff}^r(M)$. The main focus of the present paper is *the space of general (nonconservative) diffeomorphisms* $\mathrm{Diff}^r(M)$. The authors believe that the method presented here also applies to conservative systems.

While examples such as Newhouse's show that on open subsets of $\mathrm{Diff}^r(M)$, "wild" phenomena that are not structurally stable can be topologically generic, a measure-theoretic point of view may be more appropriate to describe the dynamical behavior that would typically be observed by a scientist. In the influential paper J. Palis [Pa] proposed a new global view of generic dynamics based on measure theory. He stated the following conjectures on finititude of attractors and their metric stability:
(I) *Denseness of finititude of attractors — there is C^r ($r \geq 1$) dense set D of diffeomorphisms in* $\mathrm{Diff}^r(M)$ *such that each element of D has finitely many attractors, the union of whose basins of attraction has full measure;*
(II) *Existence of physical (SRB) mesure — each attractor of an element of D supports a physical measure that represents the limiting distribution for Lebesgue almost every initial condition in its basin;*
(III) *Metric stability of basins of attraction — for each element in D and each of its attractors, for almost all small C^r perturbations in generic k–parameter families of diffeomorphisms in* $\mathrm{Diff}^r(M)$, $k \in \mathbb{N}$, *there are finitely many attractors whose union of basins is nearly equal in the sense of Lebesgue measure to the basin of the initial attractor; such perturbed attractors support a physical measure.*

Such results have been established for certain examples of dynamical systems. Lyubich [Ly] for the quadratic family of 1–dimensional maps and Avila–Lyubich–de Melo [ALM] for a generic family of analytic unimodal 1–dimensional maps showed that for almost all parameters the attractors are either periodic sinks or carry an absolutely continuous invariant measure. For the 1–dimensional Schrödinger cocycles Avila–Krikorian [AK] showed that for all analytic or C^∞ potentials and almost all rotation numbers the corresponding cocycle is either non–uniformly hyperbolic or reducible.

In this paper we discuss *two* important topologically negligible dynamical properties that are in fact prevalent. One property is (stretched) exponential growth of the number of periodic points and the other is finiteness of number of coexisting "localized" sinks for surface diffeomorphisms.

We hope that the method, outlined in this article, brings a better understanding of prevalent properties of $\mathrm{Diff}^r(M)$ in the direction of Palis' conjectures and other important dynamical properties.

2. Elementary events and a sample result

Here we expose ideas in a general setting. Consider a family of diffeomorphisms $\{f_\varepsilon\}_{\varepsilon\in B} \subset \mathrm{Diff}^r(M)$ of a compact manifold with a probability measure μ supported on the set of parameters B. To avoid distracting details we postpone specification of μ and B.

Let's fix a certain property $\mathcal{P}$ of periodic points of period n. In both cases that we will consider $\mathcal{P}$ is some form of quantitative hyperbolicity. We split the problem into two parts.

• Estimate the measure of the set

$$B_n = \{\varepsilon \in B : \ f_\varepsilon \text{ has a periodic orbit that does not satisfy } \mathcal{P}\} \subset B, \ \mu(B_n) \le \mu_n.$$

• Derive some dynamically interesting properties from this estimate.

The second part essentially depends on a problem. As for the first part, application of Discretization method and Newton Interpolation Polynomials give a uniform approach to get a required estimate. First, we discuss of growth of the number of periodic points (see Theorem 2.1 below).

For $\gamma > 0$ we say that $x = f^n(x)$ is (n,γ)*–hyperbolic* if all eigenvalues of the linearization $df^n(x)$ are at least γ–away from the unit circle[2]. For $\gamma > 0$ it is a weak analog of Kupka–Smale property. Fix some $c > 0$ and a decaying to zero sequence of positive numbers $c\Gamma = \{c\gamma_n\}_{n\in\mathbb{Z}_+}$.

We say that the map f_ε satisfies *the Inductive Hypothesis of order* n *with constants* $c\Gamma$, denoted $f_\varepsilon \in IH(n, c\Gamma)$, if for all $k \le n$ all periodic orbits of period k are $(k, c\gamma_k)$-hyperbolic. Consider a sequence of "bad" sets in the parameter space

$$B_n(c\Gamma) = \{\varepsilon \in B : \ f_\varepsilon \in IH(n-1, c\Gamma), \quad \text{but} \quad f_\varepsilon \notin IH(n, c\Gamma)\}. \tag{1}$$

In other words, $B_n(c\Gamma)$ is the set of "bad" parameter values $\varepsilon \in B$ for which all periodic points with period strictly less than n are sufficiently hyperbolic, but there is a periodic point of period n that is not $(n, c\gamma_n)$-hyperbolic.

Our goal is to find an upper bound

$$\mu\{B_n(c\Gamma)\} \le \mu_n(c\Gamma) \tag{2}$$

for the measure of the set of "bad" parameter values. Then the sum over n of (2) gives an upper bound $\mu\{\cup_n B_n(c\Gamma)\} \le \sum_{n\ge1} \mu_n(c\Gamma)$ on the set of all parameters ε for which f_ε has a periodic point of some period n that is not $(n, c\gamma_n)$–hyperbolic. If the sum converges and $\sum_{n\ge1} \mu_n(c\Gamma) = \mu(c) \to 0$ as $c \to 0$, then for μ–almost every ε there is $c > 0$ such that for every n every periodic point of period n is $(n, c\gamma_n)$–hyperbolic.

This statement (almost) implies that all periodic points of period n are at least $\approx \ c\gamma_n$–apart and, therefore, the number of periodic points is bounded by $\approx (c\gamma_n)^{-\dim M}$ (see [KH1], Prop. 1.1.6). Thus, the key to prove a statement that a certain property is prevalent, i.e. holds for almost every parameter value, is *an estimate of probability (2) of a "bad" event.* One could replace a property of

[2] In [KH1] we use a stronger property of hyperbolicity of periodic points (see sect. 2).

hyperbolicity of periodic points by another property and still the key is to get an estimate of probability to fail a certain dynamical property.

Our goal is to outline the proof of the following result:

Theorem 2.1. *[KH1, Ka3, Ka4] For a prevalent set of diffeomorphisms $f \in$ Diff$^r(M)$, with $1 < r < \infty$ and for any $\delta > 0$ there exists $C = C(\delta)$ such that*

$$P_n(f) := \#\{ \text{ isolated } x \in M : \ f^n(x) = x\} \leq \exp(Cn^{1+\delta}).$$

Density of diffeomorphisms with this property is the classical result of Artin–Mazur [AM] (see also [Ka2] for a simple proof). In [Ka1], using [GST], it is shown that diffeomorphisms having an arbitrary ahead given growth along a subsequence are topologically generic.

In section 7 we briefly describe application of the method of the paper to Newhouse phenomenon from [GK].

3. Strategy to estimate probability of a "bad" event: Discretization Method

The goal of this section is to outline how one can get estimate (2). Usually we do not know where is a "bad" trajectory, which fails $\mathcal{P}$, and what is a dynamics in its neighborhood. So our analysis will be *implicit*. More exactly, we shall consider all possible trajectories in the family $\{f_\varepsilon\}_{\varepsilon \in B}$ and the worst case scenario for each of them.

In order to fail Inductive Hypothesis of order n with constants $c\Gamma$ a diffeomorphism f_ε should have a periodic, but not $(n, c\gamma_n)$–hyperbolic point $x = f^n_\varepsilon(x)$. There is a continuum of possible n–tuples $\{x_k\}_{0\leq k\leq n}$ such that for some $\varepsilon \in B$ we have $f(x_k) = x_{k+1 \pmod n}$ and x_0 is not $(n, c\gamma_n)$–hyperbolic. Instead of looking at the continuum of n–tuples we discretize this space and consider only those n–tuples $\{x_k\}_{0\leq k\leq n}$ that lie on a particular grid, denoted $I_{\widetilde{\gamma_n}}$, and replace trajectories by $\widetilde{\gamma_n}$-pseudotrajectories. If we choose the grid spacing $\widetilde{\gamma_n}$ small enough, then for every (almost) periodic point of period n that is not sufficiently hyperbolic will have a corresponding $\widetilde{\gamma_n}$-pseudotrajectory of length-n on the grid that also has small hyperbolicity. In this way we reduce the problem of bounding the measure of a set of "bad" parameters corresponding to a particular length-n $\widetilde{\gamma_n}$-pseudotrajectory on the chosen grid.

Thus, the basic requirement for the grid size $\widetilde{\gamma_n}$ is that every real periodic trajectory $\{x_k = f^k_\varepsilon(x_0)\}_{0\leq k\leq n}$ of length-n can be approximated by a $\widetilde{\gamma_n}$-pseudotrajectory $\{\tilde{x}_k\}_{0\leq k\leq n}$ so that if x_0 is periodic not $(n, c\gamma_n)$–hyperbolic, then the n–tuple $\{\tilde{x}_k\}_{0\leq k\leq n}$ is not $(n, c\gamma_n/2)$-hyperbolic (see [KH1] sect. 3.2 and [GK] sect. 8 for various definitions).

We call an n–tuple $\{x_k\}_{k=0}^{n-1} \subset I^n_{\widetilde{\gamma_n}}$ *a $\widetilde{\gamma_n}$–pseudotrajectory associated to some ε (or to the map f_ε)* if for each $k = 0, \dots, n-1$ we have $dist(f_\varepsilon(x_{k-1}), x_k) \leq \widetilde{\gamma_n}$ and we call it *a $\widetilde{\gamma_n}$–pseudotrajectory associated to B (or the family $\{f_\varepsilon\}_{\varepsilon\in B}$)* if it is associated to some $\varepsilon \in B$.

The naive idea of estimate (2) consists of two steps:

Step 1. Estimate the number of different $\widetilde{\gamma_n}$–pseudotrajectories $\#_n(\widetilde{\gamma_n})$ associated to B;

Step 2. For an n–tuple $\{x_k\}_{0\le k\le n-1} \subset I^n_{\widetilde{\gamma_n}}$ estimate the measure

$$\begin{aligned}\mu\{\varepsilon \in B :\ \{x_k\}_{0\le k\le n-1} \text{ is a } \widetilde{\gamma_n}\text{–pseudotrajectory associated to } \varepsilon \\ \text{which is } \widetilde{\gamma_n}\text{–periodic, not } (n, c\gamma_n/2)\text{–hyperbolic}\} \le \mu_n(c\gamma_n, \widetilde{\gamma_n}).\end{aligned} \tag{3}$$

Then the product of two numbers $\#_n(\widetilde{\gamma_n})$ and $\mu_n(c\gamma_n, \widetilde{\gamma_n})$ that are obtained in Steps 1 and 2 gives the required estimate. In fact, this simpleminded scheme requires modifications discussed at the end of the next section (see (10–13)).

We start with the second step. For simplicity we shall discuss 1-dimensional maps (see [KH1], sect. 3). In [KH1] sect. 4.2 we discuss difficulties arising to extend this method to multidimensional maps. See also [GK] sect. 10 (resp. [Ka4] sect. 7–8), where 2-dimensional (resp. N-dimensional) case is considered. To treat multidimensional case one use very similar ideas, however, technical difficulties arising due to multidimensionality are fairly involved. Now we show how to estimate probability (3) within a particular polynomial family and then show how to do Step 1 and incorporate the method into the global framework.

4. Newton Interpolation Polynomials and an estimate of probability of a $\widetilde{\gamma_n}$-periodic, not $(n, c\gamma_n/2)$-hyperbolic $\widetilde{\gamma_n}$-pseudotrajectory of length-n

Let M be an interval $[-1, 1]$ and $I_{\widetilde{\gamma_n}} \subset [-1, 1]$ be a $\widetilde{\gamma_n}$-grid. Fix an n-tuple of points $\{x_k\}_{k=0}^{n-1} \subset I_{\widetilde{\gamma_n}}$. Consider the following $2n$–parameter family of maps:

$$f_u(x) = f(x) + \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j\ (mod\ n)}).$$

This family is nothing but Newton Interpolation Polynomial associated to the n–tuple $\{x_k\}_{k=0}^{n-1}$. Denote $\phi_u(x) = \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1}(x - x_{j\ (mod\ n)})$. Notice that

$$
\begin{aligned}
\phi_u(x_0) =& u_0, \\
\phi_u(x_1) =& u_0 + u_1(x_1 - x_0), \\
\phi_u(x_2) =& u_0 + u_1(x_2 - x_0) + u_2(x_2 - x_0)(x_2 - x_1), \\
&\vdots \\
\phi_u(x_{n-1}) =& u_0 + u_1(x_{n-1} - x_0) + \dots \\
&+ u_{n-1}(x_{n-1} - x_0)\dots(x_{n-1} - x_{n-2}), \\
\phi_u'(x_0) =& \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k}(x - x_{j(\mathrm{mod}\ n)})\right)\Big|_{x=x_0}, \\
&\vdots \\
\phi_u'(x_{n-1}) =& \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k}(x - x_{j(\mathrm{mod}\ n)})\right)\Big|_{x=x_{n-1}}.
\end{aligned}
\tag{4}
$$

These formulas are very useful for dynamics. For a given map f and initial point x_0, the image $f_u(x_0) = f(x_0) + \phi_u(x_0)$ 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 $f_u(x_1)$ 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 $f_u(x_k)$ of x_k depends only and nontrivially on u_k.

Suppose now that an n-tuple of pairwise distinct points $\{x_j\}_{j=0}^{n-1}$ and Newton coefficients $\{u_j\}_{j=0}^{n-1}$ are fixed. Then derivative $f_u'(x_0)$ 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 $f_u'(x_k)$ at x_k depends only and nontrivially on u_{n+k}.

As Figure 1 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 $f_u = f + \phi_u$ with such a trajectory.

Using these properties of Newton Interpolation Polynomials we can easily estimate probability (3). Let's split this compound dynamic event into simple ones

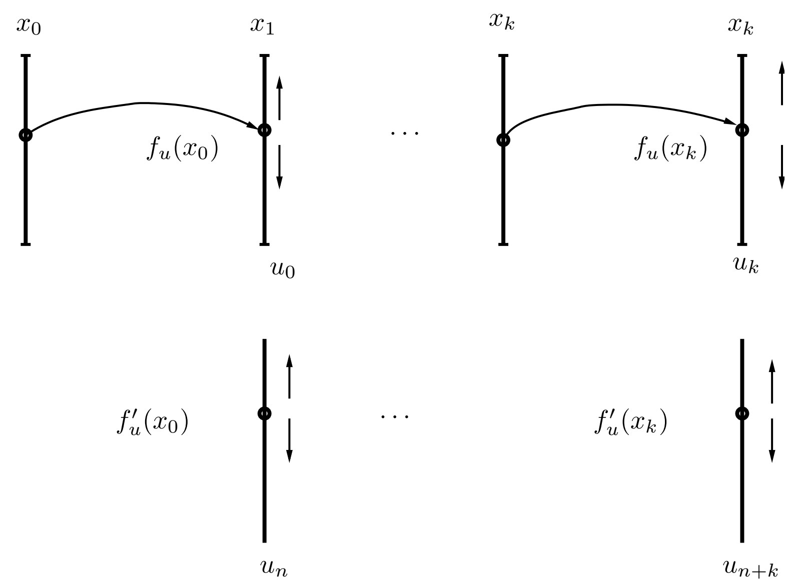

Figure 1. Newton Coefficients and their action

and use the above properties:

$$
\begin{array}{ll}
1. & |f_\varepsilon(x_0) - x_1| \le \widetilde{\gamma_n}; \\
2. & |f_\varepsilon(x_1) - x_2| \le \widetilde{\gamma_n}; \\
 & \cdots; \\
n. & |f_\varepsilon(x_{n-1}) - x_0| \le \widetilde{\gamma_n}; \\
n+1. & \left| \prod_{j=0}^{n-1} |f'_\varepsilon(x_j)| - 1 \right| \le c\gamma_n/2.
\end{array} \tag{5}
$$

First, we find probabilities of these events with respect to u-parameters (see [KH1] sect 3.3 for more details). It turns out that the map relating ε–parameters and u–parameters is one-to-one, linear, and volume-preserving (see sect 5).

Notice that in (4) and Figure 1 the image $f_u(x_0)$ of x_0 is independent of u_k for all $k > 0$. Therefore, the position of $f_u(x_0)$ *depends only on* u_0. For the 1-dimensional Lebesgue measure of u_0's we have

$$
Leb\ \{u_0 :\ |f_u(x_0) - x_1| = |f(x_0) + u_0 - x_1| \le \widetilde{\gamma_n}\} \le 2\widetilde{\gamma_n}
$$

Fix u_0. Similarly, the position of $f_u(x_1)$ *depends only on* u_1 (see (4) and Fig. 1. Thus, we have

$$
Leb\big\{u_1 :\ |f_u(x_1) - x_2| = |f(x_1) + u_0 + u_1(x_1 - x_0) - x_2| \le \widetilde{\gamma_n}\big\} \le \frac{2\widetilde{\gamma_n}}{|x_1 - x_0|}
$$

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

$$Leb\Big\{u_k :\ |f_u(x_k) - x_{k+1}| = \left| f(x_k) + \sum_{m=0}^{k} u_m \prod_{j=0}^{m-1}(x_k - x_j) - x_0 \right| \le \widetilde{\gamma_n} \Big\} \le \frac{2\widetilde{\gamma_n}}{\prod_{j=0}^{k-1} |x_k - x_j|},$$

and for $k = n-1$ we have

$$Leb\left\{u_{n-1} :\ |f_u(x_{n-1}) - x_0| \le \widetilde{\gamma_n}\right\} \le \frac{2\widetilde{\gamma_n}}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|}. \tag{6}$$

In particular, the parameter u_{n-1} *is responsible* for (n, γ_n)-periodicity of the n-tuple $\{x_k\}_{0\le k\le n}$. This formula estimates the *"measure of periodicity"*.

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

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

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

$$\prod_{j=0}^{n-1} (g_{u_{2n-1}})'(x_j) = \left(g'(x_{n-1}) + u_{2n-1} \prod_{j=0}^{n-2} |x_{n-1} - x_j|^2 \right) \prod_{j=0}^{n-2} g'(x_j). \tag{7}$$

To get the final estimate, we use the fact that we are interested only in maps from the family $\{f_u\}_u$. Suppose $|f_u'(x_{n-1})|$ is uniformly bounded by some M_1. For condition $(n+1)$ of (5) to hold, $\left|\prod_{j=0}^{n-1} f_u'(x_j)\right|$ must lie in $[1 - c\gamma_n, 1 + c\gamma_n]$. If this occurs for any u_{2n-1}, then $\left|\prod_{j=0}^{n-2} f_u'(x_j)\right| \ge (1 - c\gamma_n)/M_1$ for all u_{2n-1}, because this product does not depend on u_{2n-1}. Using (7) and the fact that $1 - c\gamma_n \ge 1/2$, we get

$$Leb \left\{ u_{2n-1} :\ \left| \prod_{j=0}^{n-1} |f_u'(x_j)| - 1 \right| \le \frac{c\gamma_n}{2} \right\} \le M_1 \ \frac{4c\gamma_n}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|^2}. \tag{8}$$

This formula estimates the *"measure of hyperbolicity"*.

We can combine all these estimates and get

$$Leb^{n+1} \left\{(u_0, \dots, u_{n-1}, u_{2n-1}) : f_u \text{ satisfies conditions (5) and } \|f_u\|_{C^1} \le M_1\right\} \le \le \frac{4M_1 c\gamma_n}{\prod_{j=0}^{n-2} |x_{n-1} - x_j|^2} \prod_{m=1}^{n-1} \frac{2\widetilde{\gamma_n}}{\prod_{j=0}^{m-1} |x_m - x_j|}. \tag{9}$$

This completes Step 2, but leaves many open questions which we shall discuss while treating Step 1. The estimate of Step 1 then breaks down as follows:

$$\#_n(\widetilde{\gamma_n}) \approx \boxed{\begin{array}{c}\#\text{ of initial}\\ \text{points of } I_{\widetilde{\gamma_n}}\end{array}} \times \boxed{\begin{array}{c}\#\text{ of } \widetilde{\gamma_n}\text{-pseudotrajectories}\\ \text{per initial point}\end{array}} \tag{10}$$

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

$$\mu_n(c\gamma_n, \widetilde{\gamma_n}) \approx \frac{\boxed{\begin{array}{c}\text{Measure of}\\ \text{periodicity (6)}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{hyperbolicity (8)}\end{array}}}{\boxed{\begin{array}{c}\#\text{ of } \widetilde{\gamma_n}\text{-pseudotrajectories}\\ \text{per initial point}\end{array}}}$$

(Roughly speaking, the terms in the numerator represent respectively the measure of parameters for which a given initial point will be $(n, \widetilde{\gamma_n})$-periodic and the measure of parameters for which a given n-tuple is $(n, c\gamma_n)$-hyperbolic; they correspond to estimates (6) and (8) in the next section.) Thus after cancellation, the estimate of the measure of "bad" set $B_n(c\Gamma)$ associated to almost periodic, not sufficiently hyperbolic trajectories becomes:

$$\boxed{\begin{array}{c}\text{Measure of bad}\\ \text{parameters}\end{array}} \leq \boxed{\begin{array}{c}\#\text{ of initial}\\ \text{points of } I_{\widetilde{\gamma_n}}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{periodicity (6)}\end{array}} \times \boxed{\begin{array}{c}\text{Measure of}\\ \text{hyperbolicity (8)}\end{array}} \tag{11}$$

Consider only pseudotrajectories having $\prod_{j=0}^{n-2} |x_{n-1} - x_j| \geq (c\gamma_n)^{1/4}$ *and suppose* $\widetilde{\gamma_n} = M_1^{-n} c\gamma_n$. Then the first term on the right hand side of (11) is of order $(c\gamma_n)^{-1}$. The second term has an upper bound of order $(c\gamma_n)^{3/4}$, and the third term is at most of order $(c\gamma_n)^{1/2}$, so that the product on the right-hand side of (11) is of order at most $(c\gamma_n)^{1/4}$ (up to an exponential function in n). If $c\gamma_n$ is exponentially small with a large exponent in n, then $\mu_n(c\gamma_n, \widetilde{\gamma_n})$ is at most exponentially small. To prove these heuristic formulas we use Newton interpolation polynomials discussed in the next section. This discussion motivates the following

Definition 4.1. *A trajectory* $x_0, \ldots, x_{n-1}$ *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*

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

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. This product of distances is quantitative characteristic of recurrent properties of a trajectory. If it is small enough, then there are close returns of it to x_0 before time n.

Even though most of properties of periodic orbits do not depend on a starting point, it turns out that for the above product, even asymptotically, it *does* matter, where to choose the starting point. A good example to look at is periodic trajectories in a neighborhood of a planar homoclinic tangency (see [KH1] sect. 2.4 for more). It motivates the following

Definition 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.*

In (11) we consider only $(n, c\gamma_n)$–simple pseudotrajectories. To study nonsimple pseudotrajectories we look for their simple almost periodic parts. More exactly, *for each non-$(n, c\gamma_n)$-simple pseudotrajectory we find such a close return, say x_k, that $\{x_j\}_{j=0}^{n-1}$ is almost equal to n/k copies of $\{x_j\}_{j=0}^{k-1}$ and $\{x_j\}_{j=0}^{k-1}$ is $(k, c\gamma_k)$-simple.* Due to closeness sufficient hyperbolicity of $\{x_j\}_{j=0}^{k-1}$ implies sufficient hyperbolicity $\{x_j\}_{j=0}^{n-1}$. Then investigation of the measure of nonhyperbolicity of nonsimple pseudotrajectory reduces to the measure of nonhyperbolicity of its simple almost periodic parts. Thus to obtain $\mu_n(c\gamma_n, \widetilde{\gamma_n})$ from (2) we arrive to the following scheme:

[Measure of bad parameters associated to periodic nonhyperbolic orbits] = (12)

[Measure of bad parameters associated to simple periodic nonhyperbolic orbits (1)] + [Measure of bad parameters associated to nonsimple periodic nonhyperbolic orbits (2)]

[Measure of bad parameters associated to nonsimple periodic nonhyperbolic orbits] = (13)

[Partition of nonsimple periodic orbits into simple almost periodic parts (2.1)] & [Measure of bad parameters associated to short non-simple almost periodic nonhyperbolic orbits (2.2)]

This diagram summarizes the problems we face in the proof.

• Part (1): how to estimate the measure of parameter values (11) associated with simple periodic nonhyperbolic orbits;

• Part (2.1): how to partition a nonsimple periodic orbit into almost periodic parts so that hyperbolicity of an almost periodic part implies hyperbolicity of the whole orbit;

The part (2.2) (how to estimate the measure associated with (11) simple periodic nonhyperbolic shorter orbits) can be treated in the same way as part (1), even though the actual details are usually quite involved (see [KH1], sect. 3.5–3.6).

5. How to collect all simple (almost) periodic pseudotrajectories: the Distortion and Collection Lemmas

In this section for the model family we show how one can justify heuristic estimates (10 – 11). The model family is the family of perturbations of a C^2 map $f : I \to I,\ I = [-1, 1]$ such that $f(I)$ strictly belongs to I

$$f_\varepsilon(x) = f(x) + \sum_{k=0}^{2n-1} \varepsilon_k x^k, \quad \varepsilon = (\varepsilon_0, \ldots, \varepsilon_{2n-1}). \tag{14}$$

This is a $2n$–parameter family. Assume that parameters belong to a brick, called *the brick of standard thickness* with width τ (see [KH1] sect. 3.1 in the 1-dimensional case, [KH1] sect. 4.3, [Ka4] sect. 8.3 in the N-dimensional case, [GK] sect. 2.3 and 11.2 for modified definitions in the 2-dimensional case applicable to the problem of finiteness of localized sinks)

$$HB^{\mathrm{st}}_{<2n}(\tau) = \Big\{ \{\varepsilon_k\}_{k=0}^{2n-1} :\ \forall\ 0 \le k < 2n,\ |\varepsilon_k| < \frac{\tau}{k!} \Big\}.$$

For small enough τ the map $f_\varepsilon : I \to I$ is well defined for all $\varepsilon \in HB^{\mathrm{st}}_{<2n}(\tau)$. Since we are interested in the measure 0 or 1 events, one could chop a brick of another shape into smaller bricks of standard thickness and use the same proof. Suppose $\sup_{\varepsilon \in HB^{\mathrm{st}}_{<2n}(\tau)} \|f_\varepsilon\| < M_1$ for some M_1.

Define the Lebesgue product probability measure, denoted by $\mu^{\mathrm{st}}_{<2n,\tau}$, on the Hilbert Brick of parameters $HB^{\mathrm{st}}_{<2n}(\tau)$ by normalizing the 1-dimensional Lebesgue measure along each component to the 1-dimensional Lebesgue probability measure

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

By definition of $\mu^{\mathrm{st}}_{<2n,\tau}$ we have that $\varepsilon_0, \ldots, \varepsilon_{2n-1}$ are independent uniformly distributed random variables.

How to get from this family to a "generic finite-parameter family" is a tedious two step procedure based on Fubini theorem. The first step is from finite–parameter polynomial families to families of analytic perturbations is discussed in [KH1] sect. 2.3, see also [GK] sect. 3.2. The second step is from analytic perturbations to prevalent finite–parameter families is discussed in [KH1] Appendix C.

Consider an ordered n-tuple of points $\mathbf{X}_n = \{x_k\}_{k=0}^{n-1} \in I^n$. One can define an linear map $\mathcal{L}^1_{\mathbf{X}_n} : \mathbb{R}^{2n}_\varepsilon \to \mathbb{R}^{2n}_u$ given implicitly by the following formulas

$$\sum_{k=0}^{2n-1} \varepsilon_k x^k = \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j(\text{mod } n)}), \tag{15}$$

where $\mathcal{L}^1_{\mathbf{X}_n}(\varepsilon_0, \dots, \varepsilon_{2n-1}) = (u_0, \dots, u_{2n-1})$. In [KH1] sect 2.2 we give an explicit definition of this map using so-called *divided differences*, and call it *Newton map*. It provides relation between ε-coordinates and u-coordinates. It turns out that $\mathcal{L}^1_{\mathbf{X}_n}$ is *volume-preserving and $\mu^{\text{st}}_{<2n,\tau}$-preserving* ([KH1] Lm.2.2.2). *Therefore, estimate (9) in u–space and ε-space are the same.*

We now estimate the distortion of the Newton map $\mathcal{L}^1_{\mathbf{X}_n}$ as a map from the standard basis $\{\varepsilon_k\}_{k=0}^{2n-1}$ in the space of polynomials of degree $< 2n$ to the Newton basis $\{u_k\}_{k=0}^{2n-1}$. It helps to have in mind the following picture characterizing the distortion of the Newton map.

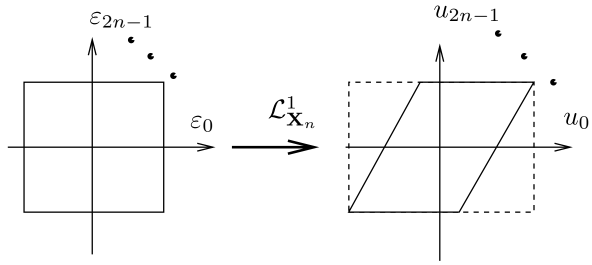

Figure 2. Distortion by the Newton map

The Distortion Lemma. ([KH1] sect 3.4) *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 $\mathcal{L}^1_{\mathbf{X}_n} : \mathbb{R}^{2n}_\varepsilon \to \mathbb{R}^{2n}_u$ be the Newton map, defined implicitly by (15). Then the image of the Brick of standard thickness $HB^{\text{st}}_{<2n}(\tau)$ with width $\tau > 0$ is contained in the Brick of standard thickness $HB^{\text{st}}_{<2n}(3\tau)$ with width 3τ:*

$$\mathcal{L}^1_{\mathbf{X}_n}(HB^{\text{st}}_{<2n}(\tau)) \subset HB^{\text{st}}_{<2n}(3\tau) \subset \mathbb{R}^{2n}_u.$$

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

The proof is simple, provided the Newton map is explicitly defined (see [KH1] sect 2.2).

In the N-dimensional case the statement of the necessary Distortion Lemma is somewhat involved. Even to define the N-dimesional Newton map one has incorporates many multiindices (see [KH1] sect. 4.2–4.3, [Ka4] sect. 8.2–8.3) For the statement and the proof of a modified Distortion Lemma applicable to the problem of finiteness of localized sinks see [GK] sect. 11.4.

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

$$\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau) := \mathcal{L}^1_{\mathbf{X}_n}(HB^{\rm st}_{<2n}(\tau)) \subset \mathbb{R}^{2n}_u$$

is the set of parameters $(u_0,\dots,u_{2n-1})$ that correspond to parameters $(\varepsilon_0,\dots,\varepsilon_{2n-1}) \in HB^{\rm st}_{<2n}(\tau)$. In other words, these are the Newton parameters *allowed by the family* (14) for the n-tuple $\mathbf{X}_n$. It turns out that $\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau)$ has the same volume as $HB^{\rm st}_{<2n}(\tau)$, but the Distortion Lemma tells us in addition that the projection of $\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau)$ onto any coordinate axis is at most a factor of 3 longer than the projection of $HB^{\rm st}_{<2n}(\tau)$.

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 (14) 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 $\mathbf{X}_n \in I^n_{\widetilde{\gamma_n}}$, we consider only the (finite number of) possibilities $x_m,\dots,x_{n-1} \subset I_{\widetilde{\gamma_n}}$. Let

$$\pi^{u,\mathbf{X}_n}_{<2n,\le m} : \mathbb{R}^{2n}_u \to \mathbb{R}^m_u \quad \text{and} \quad \pi^{u,\mathbf{X}_n}_{<2n,m} : \mathbb{R}^{2n}_u \to \mathbb{R}_{u_m}$$

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

$$\begin{aligned} \mathcal{P}^{\rm st}_{<2n,\le m,\mathbf{X}_m}(\tau) &= \cup_{x_m,\dots,x_{n-1}\in I_{\widetilde{\gamma_n}}}\ \pi^{u,\mathbf{X}_n}_{<2n,\le m}(\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau)) \subset \mathbb{R}^m_u \\ \mathcal{P}^{\rm st}_{<2n,m,\mathbf{X}_m}(\tau) &= \cup_{x_m,\dots,x_{n-1}\in I_{\widetilde{\gamma_n}}}\ \pi^{u,\mathbf{X}_n}_{<2n,m}(\mathcal{P}^{\rm st}_{<2n,\mathbf{X}_n}(\tau)) \subset \mathbb{R}_{u_m}. \end{aligned}$$

For each $m < n$, the set $\mathcal{P}^{\rm st}_{<2n,\le m,\mathbf{X}_m}(\tau)$ is a polyhedron and $\mathcal{P}^{\rm st}_{<2n,m,\mathbf{X}_m}(\tau)$ 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 $\mathcal{P}^{\rm st}_{<2n,\le m,\mathbf{X}_m}(\tau)$ consists of all Newton parameters $\{u_j\}_{j=0}^m \in \mathbb{R}^m_u$ that are allowed by the family (14) for the m-tuple $\mathbf{X}_m$.

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

$$f_{u(m),\mathbf{X}_m}(x) = f(x) + \sum_{s=0}^{m} u_s \prod_{j=0}^{s-1}(x - x_j), \tag{16}$$

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

$$\begin{aligned} T^{1,\widetilde{\gamma_n}}_{<2n,\le m,\tau}(f;x_0,\dots,x_{m-1},x_m,x_{m+1}) = \Big\{ & u(m) \in \mathcal{P}^{\rm st}_{<2n,\le m,\mathbf{X}_m}(\tau) \subset \mathbb{R}^m_u : \\ & |f_{u(m),\mathbf{X}_m}(x_{j-1}) - x_j| \le \widetilde{\gamma_n} \text{ for } j = 1,\dots,m+1 \Big\}. \end{aligned}$$

The set $T^{1,\widetilde{\gamma_n}}_{<2n,\le m,\tau}(f;x_0,\dots,x_{m-1},x_m,x_{m+1})$ represents the set of Newton parameters $u(m)=\big(u_0,\dots,u_m\big)$ allowed by the family (14) for which $x_0,\dots,x_{m+1}$ is a $\widetilde{\gamma_n}$-pseudotrajectory of $f_{u(m),\mathbf{X}_m}$ (and hence of $f_{u,\mathbf{X}_n}$ for all valid extensions u and $\mathbf{X}_n$ of $u(m)$ and $\mathbf{X}_m$).

In the following lemma, we collect all possible $\widetilde{\gamma_n}$-pseudotrajectories and estimates of "bad" measure corresponding to those $\widetilde{\gamma_n}$-pseudotrajectories. The idea of the proof of this lemma is the following. Let m be some number $0\le m<n$. Suppose an $(m+1)$-tuple $x_0,\dots,x_m\subset I_{\widetilde{\gamma_n}}$ is fixed and we are interested in the number of possible continuations $x_{m+1}\in I_{\widetilde{\gamma_n}}$ so that $x_0,\dots,x_{m+1}$ is associated to the family (14). Consider the family (16), where $u_0,\dots,u_{m-1}$ are fixed. By Distortion Lemma we have $|u_m|\le\dfrac{3\tau}{m!}$. Rewrite this family

$$f_{u(m),\mathbf{X}_m}(x_m)=f(x_m)+\sum_{s=0}^{m-1}u_s\prod_{j=0}^{s-1}(x_m-x_j)+u_m\prod_{j=0}^{m-1}(x_m-x_j).$$

Since all u_k's except u_m are fixed, the range of x_{m+1} associated to the family (14) is bounded by $3\tau\prod_{j=0}^{m-1}(x_m-x_j)/m!$.

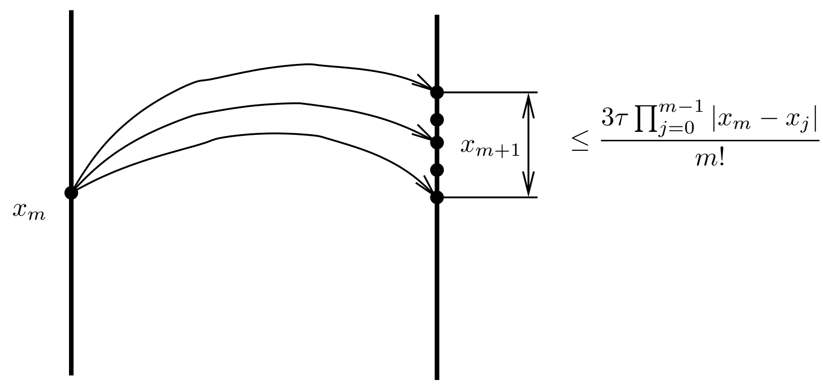

Figure 3. Collection of pseudotrajectories

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

$$\begin{aligned}&\mu^{\rm st}_{<2n,\tau}\big\{\varepsilon:\ \text{there is a } \widetilde{\gamma_n}\text{-periodic } \widetilde{\gamma_n}\text{-pseudotrajectory from } I^n_{\widetilde{\gamma_n}} \text{ starting at } x_0,\ \text{which is}\\ &\text{not } (n,M_1^{3n}c\gamma_n)\text{-hyperbolic}\big\}\le 6^{2n}M_1^{4n+1}\ \frac{(n-1)!}{\tau}\ \frac{(2n-1)!}{\tau}\ c^{1/4}\ \widetilde{\gamma_n}\ \gamma_n^{1/4}.\end{aligned}\qquad(17)$$

Corollary 5.1. *With the notations above the measure of the "bad" parameters satisfies*

$$\begin{aligned}&\mu^{\rm st}_{<2n,\tau}\big\{\varepsilon:\ \text{there is a } \widetilde{\gamma_n}\text{-periodic } \widetilde{\gamma_n}\text{-pseudotrajectory from } I^n_{\widetilde{\gamma_n}},\\ &\text{which is not } (n,M_1^{3n}c\gamma_n)\text{-hyperbolic}\big\}\le 2\cdot 6^{2n}M_1^{4n+1}\ \frac{2(n-1)!}{\tau}\ \frac{(2n-1)!}{\tau}\ c^{1/4}\ \gamma_n^{1/4}.\end{aligned}$$

For the modified Collection Lemma for the N-dimensional case see [Ka4] sect 9.3 and for the problem of finiteness of localized sinks in the 2-dimensional case see [GK] sect. 11.6 respectively.

Since there are $2/\widetilde{\gamma_n}$-grid points of $I_{\widetilde{\gamma_n}} \subset [-1,1]$, this Corollary follows directly from the Collection Lemma. Suppose that $\widetilde{\gamma_n}$-discretization is fine enough to be able to approximate "real" trajectories by $\widetilde{\gamma_n}$-pseudotrajectories well enough (see [KH1] Prop. 3.1.2, its proof, and (3.17) in the 1-dimensional case, [Ka4] sect. 9 in the N-dimensional case, and [GK] sect. 8 for the problem of finiteness of localized sinks). Then up to the error term $6^{2n} M_1^{4n+1}$ this proves (11).

Proof of the Collection Lemma: We prove by backward induction on m that for $x_0, \dots, x_m \subset I_{\widetilde{\gamma_n}}$,

$$\begin{aligned}\mu^{\mathrm{st}}_{<2n,\tau}\big\{&\text{there is a } \widetilde{\gamma_n}\text{-periodic } \widetilde{\gamma_n}\text{-pseudotrajectory from } I^n_{\widetilde{\gamma_n}} \text{ starting with} \\ &x_0,\dots,x_m \text{ which is not } (n, M_1^{3n} c\gamma_n)\text{-hyperbolic}\big\} \le \\ \le 6^{2n-m} M_1^{4n+1}\, &\frac{(n-1)!}{\tau}\, \frac{(2n-1)!}{\tau}\, \mu^{\mathrm{st}}_{<m,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le m-1,\tau}(f; x_0,\dots,x_m)\big\}\; c^{1/4}\; \widetilde{\gamma_n}\; \gamma_n^{1/4},\end{aligned} \tag{18}$$

resulting when $m=0$ in (17).

Consider the case $m = n-1$. Fix an $(n, c\gamma_n/2)$-simple n-tuple $\mathbf{X}_n = \{x_j\}_{j=0}^{n-1} \in I^n_{\widetilde{\gamma_n}}$. Using formulas (6) and (8), we get

$$\begin{aligned}&\mu^{\mathrm{st}}_{n-1,\tau}\{u_{n-1} : |f_{u,\mathbf{X}_n}(x_{n-1}) - x_0| \le \widetilde{\gamma_n}\} \le \\ &\frac{(n-1)!}{\tau} \frac{\widetilde{\gamma_n}}{\prod_{m=0}^{n-2}|x_{n-1}-x_m|} \le \frac{2^{1/4}(n-1)!}{\tau}\; c^{-1/4}\; \widetilde{\gamma_n}\; \gamma_n^{-1/4}\end{aligned}$$

and

$$\begin{aligned}&\mu^{\mathrm{st}}_{2n-1,\tau}\Big\{u_{2n-1} : \Big|\prod_{j=0}^{n-1} |(f_{u,\mathbf{X}_n})'(x_j)| - 1\Big| \le M_1^{3n} c\gamma_n\Big\} \le \\ &4M_1 \frac{(2n-1)!}{\tau}\; \frac{4M_1^{3n} c\gamma_n}{\prod_{m=0}^{n-2}|x_{n-1}-x_m|^2} \le \frac{2^{9/2} M_1^{3n+1}(2n-1)!}{\tau}\; c^{1/2}\; \gamma_n^{1/2}.\end{aligned}$$

The Fubini Theorem, preservation of generalized volume by the Newton map (see [KH1] Lm 2.2.2), and the definition of the product measure $\mu^{\mathrm{st}}_{<2n,\tau}$ imply that

$$\begin{aligned}\mu^{\mathrm{st}}_{<2n,\tau}\big\{&\text{there is a } \widetilde{\gamma_n}\text{-periodic } \widetilde{\gamma_n}\text{-pseudotrajectory from } I^n_{\widetilde{\gamma_n}} \text{ starting with} \\ &x_0,\dots,x_{n-1} \text{ which is not } (n, M_1^{3n} c\gamma_n)\text{-hyperbolic}\big\} \le \\ &\mu^{\mathrm{st}}_{<n-1,\tau}\Big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le n-2,\tau}(f; x_0,\dots,x_{n-1})\Big\} \times \\ &\mu^{\mathrm{st}}_{n-1,\tau}\{u_{n-1} : |f_{u,\mathbf{X}_n}(x_{n-1}) - x_0| \le \tilde{\gamma}_n\} \times \prod_{s=n}^{2n-2} \mu^{\mathrm{st}}_{s,\tau}\{\mathcal{P}^{\mathrm{st}}_{<2n,s,\mathbf{X}_n}(\tau)\} \times \\ &\mu^{\mathrm{st}}_{2n-1,\tau}\Big\{u_{2n-1} : \Big|\prod_{j=0}^{n-1} |(f_{u,\mathbf{X}_n})'(x_j)| - 1\Big| \le M_1^{3n} c\gamma_n\Big\} \le \\ \le 2^{11/4} 3^{n-1} M_1^{4n+1} &\frac{(n-1)!}{\tau} \frac{(2n-1)!}{\tau}\; \mu^{\mathrm{st}}_{<n-1,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le n-2,\tau}(f; x_0,\dots,x_{n-1})\big\}\; c^{1/4}\; \widetilde{\gamma_n}\; \gamma_n^{1/4}\end{aligned}$$

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

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

Since $2^{11/4}3^{n-1} < 6^{n+1}$, this yields the required estimate (18) for $m = n-1$.

Suppose now that for $m+1$, (18) is true and we would like to prove it for m. Denote by $G^{1,\widetilde{\gamma_n}}_{<2n,m,\tau}(f, u(m-1); x_0, \dots, x_m) \subset I_{\widetilde{\gamma_n}}$ the set of points x_{m+1} of the $2\widetilde{\gamma_n}$-grid $I_{\widetilde{\gamma_n}}$ such that the $(m+2)$-tuple $x_0, \dots, x_{m+1}$ is a $\widetilde{\gamma_n}$-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, $G^{1,\widetilde{\gamma_n}}_{<2n,m,\tau}(f, u(m-1); x_0, \dots, x_m)$ is the set of all possible continuations of the $\widetilde{\gamma_n}$-pseudotrajectory $x_0, \dots, x_m$ using all possible Newton parameters u_m allowed by the family (14).

Now if $x_0, \dots, x_m$ is a $\widetilde{\gamma_n}$-pseudotrajectory associated to $u(m) = (u_0, \dots, u_m)$, then at most 2 values of $x_{m+1} \in I_{\widetilde{\gamma_n}}$ are within $\widetilde{\gamma_n}$ of $f_{u(m),\mathbf{X}_m}(x_m)$. 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 $G^{1,\widetilde{\gamma_n}}_{<2n,m,\tau}(f, u(m-1); x_0, \dots, x_m)$. It follows that

$$\sum_{x_{m+1}\in G^{1,\widetilde{\gamma_n}}_{<2n,m,\tau}(f,u(m-1);x_0,\dots,x_m)} \mu^{\rm st}_{\le m,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le m,\tau}(f; x_0, \dots, x_{m+1})\big\} \le$$
$$2\ \mu^{\rm st}_{m,\tau}\{\mathcal{P}^{\rm st}_{<2n,m,\mathbf{X}_n}(\tau)\}\ \mu^{\rm st}_{\le m-1,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le m-1,\tau}(f; x_0, \dots, x_m)\big\}.$$

The Distortion Lemma then implies that

$$\sum_{x_{m+1}\in G^{1,\widetilde{\gamma_n}}_{<2n,m,\tau}(f,u(m-1);x_0,\dots,x_m)} \mu^{\rm st}_{\le m,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le m,\tau}(f; x_0, \dots, x_{m+1})\big\} \le$$
$$6\ \mu^{\rm st}_{\le m-1,\tau}\big\{T^{1,\widetilde{\gamma_n}}_{<2n,\le m-1,\tau}(f; x_0, \dots, x_m)\big\}.$$

Inductive application of this formula completes the proof of the Collection Lemma.□

This completes an outline of treatment of part (1) of (12) for the 1-dimensional case. To carry out part (1) of (12) in the N-dimensional case ($N > 1$) we use the same ideas, but have to overcome additional difficulties. We discuss them in details in [KH1] sect. 4.1 (detailed discussion of the 2-dimensional case can be found in [GK] sect. 10–11 and of the N-dimensional case in [Ka4] sect. 8–9) and just briefly mention them here.

- *(nonuniqueness)* It turns out that there are many ways to write Newton Interpolation Polynomial in N variables;
- *(dynamically essential coordinates)* Among many N-dimensional Newton monomials we need to choose those effective for perturbation (see [KH1] (4.6–4.7), [GK] sect.10.1, and [Ka4] sect. 8.2);
- *(the multidimensional Distortion Lemma)* The 1–dimensional Distortion Lemma leads to an exponential factor 6^{2n} coming from dimension of the space of polynomials of degree $< 2n$ in 1-variable. The space of polynomials of degree $< 2n$ in N-variables is $\sim (2n)^N$. This forces us to find a better multidimensional Distortion Lemma (see [KH1] sect. 4.3, [Ka4] sect. 8.3, and [GK] sect. 11.4).

To treat part (2.1) of (13) we need to analyze nonsimple (recurrent) periodic trajectories of period n knowing that all periodic trajectories of period $< n$ are sufficiently hyperbolic (see (1) of (12)).

6. Partition of nonsimple periodic trajectories into simple almost periodic parts

Analysis of nonsimple periodic trajectories of multidimensional diffeomorphisms, performed in [KH1] and [Ka4] occupies sect. 2.4 and 3.5 in [KH1] and section 5 in [Ka4]. The goal for each nonsimple periodic trajectory $\{x_j = f^j(x_0)\}_{j=0}^{n-1}$ of period n find a close return, say x_k, so that $\{x_j\}_{j=0}^{n-1}$ nearly repeats $\{x_j\}_{j=0}^{k-1}$ exactly n/k times and $\{x_j\}_{j=0}^{k-1}$ is simple. This, in particular, means that hyperbolicity of $\{x_j\}_{j=0}^{k-1}$ and $\{x_j\}_{j=0}^{n-1}$ are closely related. Here we just summarize the strategy to obtain such a partition. This is exactly the step, where *we can't handle a sequence of* $\{c\gamma_n\}_{n\geq 1}$ *which decay slower than a stretched exponential* $\exp(-n^{1+\delta})$ *(*$\delta > 0$*).* In other words, if γ_n decays not too fast, say exponentially, we are unable to find a close return with the above properties (see [KH1], Appendix D for further discussion).

The following definitions are the key elements of the mechanism to find a close return. They quantitatively characterize close returns.

Definition 6.1. *Let g be a diffeomorphism and let D be large positive. A point x_0 (resp. a trajectory $x_0, \dots, x_{n-1} = g^{n-1}(x_0)$ of length n) has a weak (D, n)-gap at a point $x_k = g^k(x_0)$ if*

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

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

This definition characterizes a close return at x_0. For the proof we need a modification of this definition (see [KH1] def. 3.5.3). See the Shift Theorem [KH1] sect 3.5 and [Ka4] sect. 5 for all the details. Recall that $\{c\gamma_n\}_{n\geq 1}$ is the sequence tracking hyperbolicity of periodic trajectories of period n introduced in the beginning of section 2.

Definition 6.2. *Let g be a C^2-smooth diffeomorphism. Let also $c > 0$ and $k < n$ be positive integers. We say that a point x_0 has a (k, n, c)-leading saddle if $|x_0 - x_k| \leq n^{-1}(c\gamma_k)^2$. Also if x_0 is $(n, \widetilde{\gamma_n})$-periodic, we say that x_0 has no (n, c)-leading saddles if for all $k < n$ we have that x_0 has no (k, n, c)-leading saddles.*

Now start with a diffeomophism f satisfying Inductive Hypothesis of order $n-1$ with constants $c\Gamma$, i.e. for any $k < n$ all periodic trajectories of period k are $(k, c\gamma_k)$-hyperbolic. In particular, it means that all periodic trajectories of period $k < n$ are either sinks, or sources, or saddles.

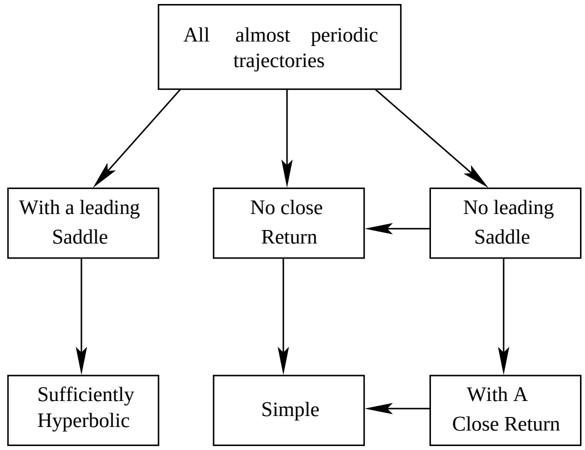

Figure 4. Various types of almost periodic periodic trajectories

• The definition of a leading saddle is designed in such a way that if x_0 has a (k, n, c)-leading saddle, then there is a periodic point $x^* = f^k(x^*)$ close to x_0 (see [Ka4], Lm. 5.2.3). If $x_k, x_{2k}, \dots$ all stay close to x^*, then $x_0 = f^n(x_0)$ inherits hyperbolicity of x^* (see [Ka4], Lm. 5.2.1).

• Suppose x_0 has a (k, n, c)-leading saddle, but for some $p < n/k$ the corresponding x_{pk} leaves a small neighborhood of x^*. Then one can show that x_{pk} has no (n, c)-leading saddles (see [Ka4], Lm. 5.2.4).

• Suppose $\tilde{x}_0 = x_{pk}$ has no (n, c)-leading saddles. It turns out that $\tilde{x}_0$ can have at most $\dim M$ weak (D, n)-gaps at some $\tilde{x}_{k_1}, \dots, \tilde{x}_{k_s}$, $s \le \dim M$. The reason is that each weak (D, n)-gap $\tilde{x}_{k_j}$ after the first one at k_1 implies that the linearization $df^{k_1}(\tilde{x}_0)$ has an almost eigenvalue that is a k_j/k_1-root of unity, and the same is true for $k_{s+1} = n$ (see [Ka4], Thm. 5.1.4).

• Suppose $\tilde{x}_0$ has no (n, c)-leading saddles and has $s(< \dim M)$ weak (D, n)-gaps. Then we can show that it is $(n, c\gamma_n)$-simple (see [Ka4], Thm. 5.3.1 and its extension necessary for the proof Thm. 5.4.1).

This scheme is summarized in the diagram (see Fig. 4).

7. Finititude of number of localized coexisting sinks

In this section we give a short exposition of a result from [GK] concerning Newhouse phenomenon of infinitely many sinks. The primary goal of [GK] is to analyze trajectories *localized in a neighborhood of a fixed HT*. A sink is the simplest attractor. We now introduce notions of an unfolding of a homoclinic tangency and localized trajectories of finite complexity associated to that homoclinic tangency.

Consider a 1-parameter family of perturbations $\{f_\varepsilon\}_{\varepsilon \in I}$, $I = [-\varepsilon_0, \varepsilon_0]$ of a 2-dimensional diffeomorphism $f = f_0 \in \mathrm{Diff}^r(M)$ with homoclinic tangency, where

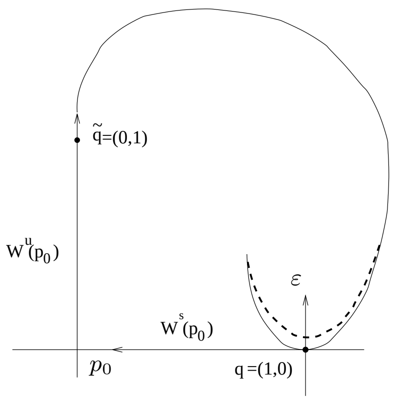

Figure 5. Homoclinic Tangency

ε_0 is small (see Figure 5). Roughly speaking, ε parameterizes oriented distance of the top tip of the unstable manifold to the stable manifold. Such a family is called an *unfolding of an HT*.

Robinson [R], adapting Newhouse's ideas [N1, N2], showed that for such an unfolding there is a sequence of open intervals converging to zero such that for a generic parameter from those intervals the corresponding diffeomorphism f_ε has infinitely many coexisting sinks.

Assume that f has a fixed saddle point $p_0 = f(p_0)$ and that the eigenvalues λ, μ of the linearization $Df(p_0),\ 0 < \lambda < 1 < \mu,$ belong to the open dense set of pairs of eigenvalues for which Sternberg's linearization theorem holds. Then in a small neighborhood $\widetilde{V}$ of p_0 there is a C^r smooth normal coordinate system $(x, y) \in \widetilde{V} \subset \mathbb{R}^2$ such that $f(x, y) = (\lambda x, \mu y)$. Suppose q is the point of homoclinic tangency of $W^s(p_0)$ and $W^u(p_0)$ away from $\widetilde{V}$, and let $\tilde{q} = f^{-1}(q)$ be its preimage.

Extend the coordinate neighborhood $\tilde{V}$ by iterating forward and backward until first it contains $\tilde{q}$ and $f(q)$, respectively. Decreasing $\widetilde{V}$ if necessary we can assume that there are no overlaps. Denote such a neighborhood by V and call it *a normal neighborhood*. By definition V does not contain q (see Figure 6). Consider a neighborhood U (resp. $\widetilde{U} \subset \widehat{U}$) of q (resp. $\tilde{q}$) such that $f(U) \cap U = \emptyset$ (resp. $f^{-1}(\widehat{U}) \cap \widehat{U} = \emptyset$), $f(\tilde{U}) \supset U$, and $f(\widehat{U}) \cap V = \emptyset$. By rescaling coordinate axis one could set q to have coordinates (1,0) and $\tilde{q}$ to have (0,1). Set $\mathcal{V} = V \cup U$. In what follows we *fix* a neighborhood $\mathcal{V}$ once and for all.

Definition 7.1. *We call an invariant set of points* $\mathcal{V}$***-localized*** *if it belongs to* $\mathcal{V}$. *In particular, any invariant set contained in*

$$\Lambda_{\mathcal{V}} = \cap_{n\in\mathbb{Z}} f^n(\mathcal{V})$$

is $\mathcal{V}$*-localized. A periodic point* $f^n(p) = p,\ n \in \mathbb{N}$, *is called* $\mathcal{V}$*-localized if it belongs to* $\Lambda_{\mathcal{V}}$ *and is called* $(\mathcal{V}, s)$*-localized if its trajectory* $\mathfrak{P} = \{f^k(p)\}_{k=1}^n$ *visits* U *exactly* s *times. Call* $s = s(\mathfrak{P})$ *the* ***cyclicity*** *of a* $\mathcal{V}$*-localized periodic orbit.*

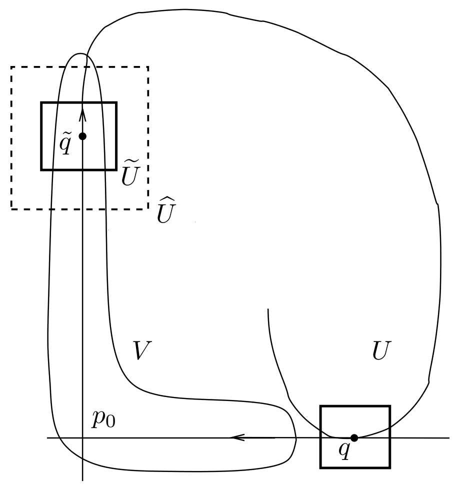

Figure 6. Localization for Homoclinic Tangency

The zoo of $\mathcal{V}$-localized invariant sets is incredibly rich. Below we just mention the authors favorit animals: *Smale's horseshoe, infinitely many coexisting $\mathcal{V}$-localized sinks*[3]*, strange attractor (Benedicks-Carleson [BC], Mora-Viana [MV], Young-Wang [WY]), arbitrarily degenerate periodic points of arbitrary high periods (Gonchenko-Shilnikov-Turaev [GST1]), uniformly and nonuniformly hyperbolic horseshoes as maximal invariant sets $\Lambda_\mathcal{V}$ (Newhouse-Palis [NP], Palis-Takens [PT], Palis-Yoccoz [PY1, PY2], Rios [Ri]).*

The main result of [GK] is the following

Theorem 7.1. *With the above notations, for a generic*[4] *1-parameter family $\{f_\varepsilon\}_{\varepsilon\in I}$ that unfolds an HT at q there is a sequence of numbers $\{\mathbf{N}_s\}_{s\in\mathbb{N}}$ such that for almost every parameter ε and any $D\in\mathbb{N}$ the corresponding f_ε has only finitely many $\mathcal{V}$-localized sinks $\{\mathfrak{P}_j\}_{j\in J}$ whose cyclicity is bounded by D or period exceeds $\mathbf{N}_{s_j}$, where $s_j = s(\mathfrak{P}_j) > D$ is cyclicity of a corresponding sink $\mathfrak{P}_j$. In other words, for almost every parameter ε if there are infinitely many coexisting $\mathcal{V}$-localized sinks $\{\mathfrak{P}_j\}_{j\in J}$, then all but finitely many have cyclicity $s_j = s(\mathfrak{P}_j) > D$ and period $< \mathbf{N}_{s_j}$.*

Remark 7.1. *For 1-loop periodic sinks a similar result is obtained by Tedeschini-Lalli–Yorke [LY]. Dynamical properties of periodic and homoclinic orbits of low cyclicity ($s = 1, 2, 3$) were studied in [GST1], [GStT]. In particular, Gonchenko–Shilnikov found the relation between existence of the infinite number of 2-loop sinks and numerical properties of the invariants of smooth conjugacy [GoS].*

[3]Actually Newhouse [N2] (see also Palis-Takens [PT] for a simplified proof) proved that for a Baire generic set of diffeomorphisms in a Newhouse domain there are infinitely many coexisting sinks. However one can construct infinitely many of those as $\mathcal{V}$-localized.

[4]meaning of "generic" is in the sense of prevalence in the space of 1-parameter families see Section 9 for a definition.

Remark 7.2. *We can choose* $\mathbf{N}_s = s^{5s^2}$.

Palis-Takens [PT] and Palis-Yoccoz [PY1, PY2] investigated generic unfolding of an HT not only for saddle periodic points but also for horseshoes. They investigated parameters *outside* of Newhouse domains. We obtain less sharp properties of the dynamics, but we treat parameters *inside* Newhouse domains too.

8. Discussion of the proof of Theorem 7.1

To prove Theorem 7.1 we follow very similar strategy as to prove Theorem 2.1. First we introduce several notions:

Trajectory type, hyperbolic and parabolic maps. Any $(\mathcal{V}, s)$-localized periodic orbit, by definition, visits U exactly s times and spends $n_1, n_2, \dots, n_s$ consecutive iterates in V, $n = n_1 + n_2 + \dots + n_s + s$. We call an ordered sequence $(n_1, \dots, n_s)$ *type* of a periodic orbit. For a given periodic orbit denote the points of intersection with U by $\mathbf{p}_0,\ \mathbf{p}_1, \dots, \mathbf{p}_{s-1}$ and the corresponding points in $\widetilde{U}$ by $\widetilde{\mathbf{p}}_0 = f^{n_1}(\mathbf{p}_0),\ \widetilde{\mathbf{p}}_1 = f^{n_2}(\mathbf{p}_1), \dots, \widetilde{\mathbf{p}}_{s-1} = f^{n_s}(\mathbf{p}_{s-1})$.

Recall that f is linear in $V \setminus \widetilde{U}$ with eigenvalues $\lambda < 1 < \mu$, $f|_{V\setminus \widetilde{U}}(x, y) = (\lambda x, \mu y)$. Call this linear map *hyperbolic*, denoted L, and $f|_{\widetilde{U}}$ *parabolic*, denoted $\mathcal{P}$.

We replace hyperbolicity of periodic points from (1) by *cone condition*.

Cone condition. To estimate the measure of parameters for which a periodic orbit of a given type is not a sink and even has exponentially large linearization, we introduce the following cone condition. Define at every point $p \in U$ a cone

$$K_A(p) = \{v = (v_x, v_y) \in T_p\mathcal{V} \simeq \mathbb{R}^2 :\ |v_y| \geq \mu^{-A}|v_x|\}.$$

To show that the periodic point $\mathbf{p}_0$ is hyperbolic it turns out that it suffices to find $0 < \alpha \ll 1$ independent of n such that

$$Df^n_\varepsilon(K_{\alpha n}(\mathbf{p}_0)) \subset K_{\alpha n}(\mathbf{p}_0). \tag{19}$$

To verify this condition directly does not seem possible in general. Our plan is to verify that for most parameters this cone condition holds after each loop:

$$Df^{n_i+1}_\varepsilon(K_{\alpha n}(\mathbf{p}_i)) \ \subset\ K_{\alpha n}(\mathbf{p}_{i+1 \pmod s}) \ \text{ for each } i = 0, \dots, s-1. \tag{20}$$

This condition clearly implies (19), because the image of the first cone $K_{\alpha n}(\mathbf{p}_0)$ belongs to the second cone $K_{\alpha n}(\mathbf{p}_1)$. The image of the second one belongs to the third one and so on.

Fix $0 < \alpha \ll 1$. Notice that if all loops are *long*: $n_i > 3\alpha n$, then $L^{n_i}K_{\alpha n}(\mathbf{p}_i)$ is the cone of width angle $< 2\mu^{-\alpha n}$. Fix $1 \leq j \leq s$. To satisfy condition (20) for j we need to avoid the intersection of the cone $Df_{\varepsilon, \widetilde{\mathbf{p}}_j}(L^{n_j}K_{\alpha n}(\mathbf{p}_j))$ and a complement

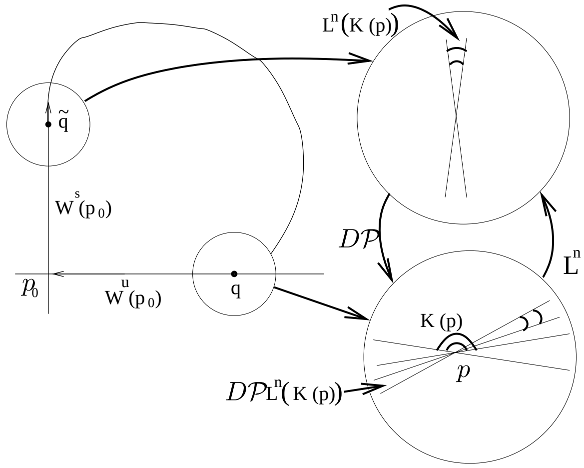

Figure 7. Evolution of cones

to $K_{\alpha n}(\widetilde{\mathbf{p}}_{j+1})$ (see Fig. 7 for $p = \mathbf{p}_{j+1}$). Assume that we can perturb $Df_{\varepsilon,\widetilde{\mathbf{p}}_j}$ by composing with rotation and angle of rotation is a parameter. Then we need to avoid a phenomenon that has "probability" $\sim \mu^{-\alpha n}$. Taking the union over all types $\mathcal{N}_s$, $|\mathcal{N}_s| = n$ we get that probability to fail (20) for some $1 \le i \le s$ is $\sim n^s \mu^{-\alpha n}$. We avoid saying explicitly probability in what space, just assume that it is proportional to angle of rotation, and postpone the exact definition for further discussion.

However, it might happen that one of n_i's is significantly smaller than αn, e.g. $n_s \le \ln n$. In this case, the above argument fails. Indeed, let $n_s = [\ln n]$, $n \gg 1$. Consider the image of the cone $K_{\alpha n}(\mathbf{p}_{s-1})$ after the last loop $L^{n_s} K_{\alpha n}(\mathbf{p}_{s-1})$. It is the cone, whose width angle is of order 1. Taking into account possibility that $Df_{\varepsilon,\widetilde{\mathbf{p}}_{s-1}}$ rotates a vertical vector by $\frac{\pi}{2}$ it is certainly not possible to fulfill (20) by a small perturbation. The natural idea is *to avoid looking at condition (20) after "short" loops.* This leads to combinatorial analysis of type $\mathcal{N}_s$ of trajectories.

Combinatorial analysis of type $\mathcal{N}_s$ of s-loop trajectories Below we don't pay attention to dynamics of a trajectory of type $\mathcal{N}_s$ under consideration. We investigate only properties of the type $\mathcal{N}_s$ itself.

- *Short and long loops* ([GK], sect. 5.1). We shall divide an s-tuple $\mathcal{N}_s = (n_1, \dots, n_s)$ into two groups of *long* and *short* n_i's, because they correspond to loops of a trajectory. After such a division long n_i's should be *much longer* than short n_i's. Denote by t (resp. $s - t$) the number of long (resp. short) loops.
- *Generalized loops and essential returns* ([GK], sect.5.2). Since we can't fulfill (20) after short loops, we combine all loops into groups, called *generalized loops.* Each generalized loop starts with a long loop and is completed by all short loops following afterwards. Therefore, the number of generalized loops equals

the number of long loops. Then we *verify (20) not after each loop, but after each generalized loop*. Denote by $P_0, \dots, P_{t-1}, P_t = P_0 \subset U$ starting points of generalized loops, by $\tilde{P}_0, \dots, \tilde{P}_{t-1}, \tilde{P}_t = \tilde{P}_0$, prestarting points of generalized loops, i.e. $f(\widetilde{P}_i) = P_{i+1},\ i = 0, \dots, t-1$, and by $N_1, \dots, N_t$ their lengths respectively. Then we modify (20) to

$$Df^{N_{i+1}}(K_{\alpha n}(P_i)) \subset K_{\alpha n}(P_{i+1}) \quad \text{for each} \quad i = 0, \dots, t-1. \tag{21}$$

Now the idea presented above has a chance to work. Indeed, let n_j be a long loop and $n_{j+1}, \dots, n_{j+j'}$ be short ones from the corresponding generalized loop. Consider the image of the corresponding cone $K_{\alpha n}(P_j)$ after the generalized loop. Notice that after the long loop n_j the cone $L^{n_j} K_{\alpha n}(P_j)$ is the cone of width angle $< 2\mu^{-\alpha n}$. Since long n_j is so much longer than short loops $n_{j+1}, \dots, n_{j+j'}$ respectively the cone

$$(Df(\tilde{\mathbf{p}}_{j+j'}) \circ L^{n_{j+j'}} \circ \dots \circ Df(\tilde{\mathbf{p}}_{j+1}) \circ L^{n_{j+1}}) \circ (Df(\tilde{\mathbf{p}}_j) \circ L^{n_j} K_{\alpha n}(\mathbf{p}_j))$$

has width angle $< 3\mu^{-\alpha n}$. To satisfy condition (20) for $j + j'$ we need to avoid an interval of rotations (i.e. of parameters) of length $< 5\mu^{-\alpha n}$. This phenomenon still has "probability" $\sim \mu^{-\alpha n}$.

After this combinatorial analysis we face the next difficulty. *We can't perturb $Df(\tilde{\mathbf{p}})$ and $Df(\tilde{\mathbf{p}}')$ independently at nearby points $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{p}}'$.*

Dynamical analysis of trajectories. Assume for a moment that we are interested in properties of *scattered* periodic orbits, that is, such orbits that $P_0, \dots, P_{t-1}$ in U are pairwise well spaced. In particular, it is always the case for 1-loop orbits. In this case the difficulty of nearby points is removed. Using Discretization method and the cone condition (21) one can prove that for most parameters all but a finite number of the periodic orbits are hyperbolic saddles. Moreover, consider for $0 < \gamma' = \mu^{-\alpha' n} \ll \gamma'' = \mu^{-\alpha'' n}$ parameters for which a periodic not enough hyperbolic γ''-scattered γ'-pseudo-orbit of period n exists. In fact, we can show that the measure of these parameters is small[5]. Now we are going to explain how this can be used to treat all periodic orbits, not necessarily scattered. Consider the 2-loop case for illustration. If starting points of loops $\mathbf{p}_0$ and $\mathbf{p}_1$ are far enough from each other, one can perturb differential of parabolic map at their preimages independently, and above arguments allow to estimate the measure of "bad" parameters. Otherwise a periodic orbit can be decomposed into a union of two 1-loop periodic pseudo-orbits, which have nearby endpoints in U. The cone condition (21) for each of these pseudo-orbits holds for most parameters, which implies (19).

Another illustration can be given by the case $t = 1$, i.e. we have one loop which is much longer than all the others. In this case the image of the cone $K_{\alpha n}(\mathbf{p}_0)$ after the application of differential of the map along the orbit has width angle $< 2\mu^{-\alpha n}$, as explained above. Point $\widetilde{\mathbf{p}}_{s-1} = \widetilde{P}_0 = f^{n-1}(\mathbf{p}_0) = f^{-1}(\mathbf{p}_0)$ can not be too close to points $\widetilde{\mathbf{p}}_0, \widetilde{\mathbf{p}}_1, \dots, \widetilde{\mathbf{p}}_{s-2}$. Indeed, the distance between $\mathbf{p}_i$ and x-axis is

[5] The Discretization method in this case, compare to the one described in sections 4–5, requires certain modifications (see [GK] sections 9–11).

$(\mathbf{p}_i)_y \sim \mu^{-n_{i+1}}$. Since $n_1 \gg n_i$ we have $\mu^{-n_1} \ll \mu^{-n_i}$. Therefore the point $\mathbf{p}_0$ can not be too close to points $\mathbf{p}_1, \ldots, \mathbf{p}_{s-1}$, and we can perturb $\phi(\widetilde{\mathbf{p}}_{s-1}) = \phi(f^{-1}(\mathbf{p}_0))$ independently of $\phi(\widetilde{\mathbf{p}}_0), \ldots, \phi(\widetilde{\mathbf{p}}_{s-2})$. This allows to estimate the measure of "bad" parameters.

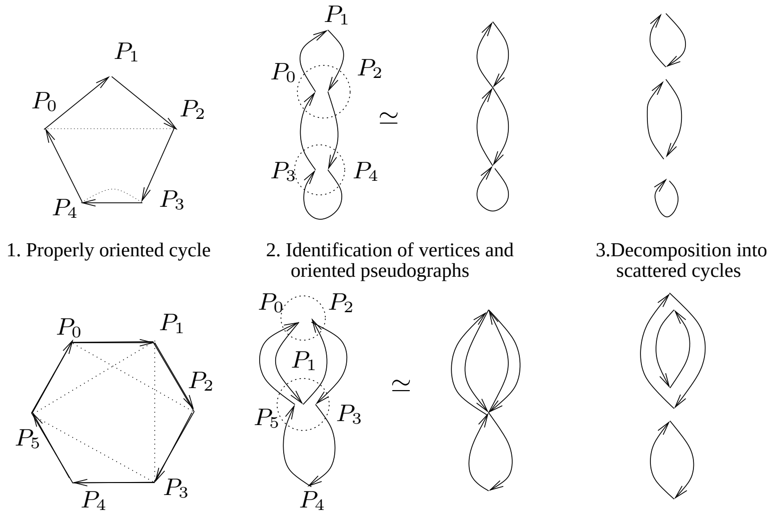

Figure 8. Graph surgery

To consider the general case we represent a periodic orbit as an oriented cyclic graph. Starting points of generalized loops are vertices of this graph, and vertices corresponding to subsequent generalized loops are connected by an oriented edge (see Figure 8, picture 1). It turns out that for some $\gamma' \ll \gamma''$ for any pair of points (P_i, P_j) either $\text{dist}(P_i, P_j) > \gamma''$ or $\text{dist}(P_i, P_j) < \gamma'$ (see [GK] sect. 7). Therefore every pair of vertices is either γ'-close or γ''-far apart (see Figure 8, picture 2). Now all the vertices can be divided into "clouds" or "clusters". Let us identify the vertices in the same cloud of nearby points, as shown on Figure 8, picture 2. The initial cycle is transformed now into oriented pseudograph (see [GK] Def. 20) with the same number of ingoing and outgoing edges at each vertex. Such a pseudograph can be decomposed into the union of oriented cycles (Figure 8, picture 3 and also see [GK] Lm. 7). Each of cycles from this decomposition represents a γ''-scattered γ'-pseudo-orbit. Application of the arguments above to these pseudo-orbits gives inclusion (21) for most values of parameters and implies the cone condition (19) for the initial periodic orbit.

9. Prevalence

Our definition of prevalence for a space $\text{Diff}^r(M)$ of C^r diffeomorphisms on a smooth manifold M is based on the following definition from [HSY] for a complete metric linear space V.

Definition 9.1 (Linear Prevalence). *A Borel set $S \subset V$ is called* shy *if there is a compactly supported Borel probability measure μ on V such that $\mu(S - v) = 0$ for all $v \in V$. More generally, a subset of V is called shy if it is contained in a shy Borel set. A subset of V is called* prevalent *if its complement is shy.*

(Shy sets were previously called "Haar null sets" by Christensen [Chr].) Some important properties of prevalence, proved in [HSY], are:

1. A prevalent set is dense.
2. A countable intersection of prevalent sets is prevalent.
3. A subset of $\mathbb{R}^m$ is prevalent if and only if its complement has Lebesgue measure zero.

Properties 2 and 3 above follow from the Fubini-Tonelli theorem, along with the Tychnoff theorem in the case of Property 2. Property 1 follows from the observation that a transverse measure μ can be localized in the following sense. By compactness of the support of μ, there are arbitrarily small balls with positive measure. Every translation of P must intersect these balls, or equivalently every translation of one of these balls must intersect P.

Along these lines, it is useful to think of a transverse measure for a prevalent set P as a probability space of perturbations, such that at each point v in the space V, choosing a random perturbation and adding it to v yields a point in P with probability one. Often the perturbations can be chosen from a finite dimensional space of parameters, using normalized Lebesgue measure on a bounded subset of parameter space. In this case, we say that P is "finite-dimensionally prevalent".

In other cases, one needs an infinite number of parameters; for example, a property about periodic orbits might be finite-dimensionally prevalent for each fixed period, but higher periods require more parameters. One may be able to choose the parameters from a "Hilbert brick" $HB = J_1 \times J_2 \times \cdots$, where each J_k is an interval of real numbers ε_k, the perturbation corresponding to $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots)$ is $\varepsilon_1 v_1 + \varepsilon_2 v_2 + \cdots$ for some vectors $v_1, v_2, \ldots \in V$, and the probability measure on HB is the infinite product of normalized Lebesgue measure on each interval. We call this measure the uniform measure on HB. A property is then prevalent if for each $v \in V$, the property is true for $v + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \cdots$ for almost every ε with respect to the uniform measure on HB.

The notion of prevalence that we use in $\mathrm{Diff}^r(M)$ is based on this idea of choosing perturbations from a Hilbert brick. Though we cannot add perturbations in this nonlinear space, by embedding M in a Euclidean space $\mathbb{R}^N$, we can perturb elements of $\mathrm{Diff}^r(M)$ in a natural way by means of additive perturbations in the space $C^r(T, \mathbb{R}^N)$ of C^r functions from T to $\mathbb{R}^N$, where T is a neighborhood of the embedded image of M in $\mathbb{R}^N$. The details of this construction are given in Appendix C of [KH1]; here we provide a brief outline.

For N sufficiently large, we can embed M into $\mathbb{R}^N$ by the Whitney embedding theorem; choose an embedding and think of M then as a subset of $\mathbb{R}^N$ (that is, identify it with its image). Choose a neighborhood T of M sufficiently small that the orthogonal projection $\pi : T \to M$ is well-defined. Extend each diffeomorphism $f \in \mathrm{Diff}^r(M)$ to a diffeomorphism F on T, in such a way that F is strongly contracting toward M. We then consider the family of perturbations

$$F_\varepsilon = F + \varepsilon_1 F_1 + \varepsilon_2 F_2 + \cdots .$$

for some functions $F_1, F_2, \ldots \in C^r(T, \mathbb{R}^N)$ and ε in an appropriate Hilbert brick. For the results presented in this paper, $F_1, F_2, \ldots$ are a basis for the polynomials on $\mathbb{R}^N$, but in general they could be any functions that are chosen independently of F.

Next we associate to each F_ε a diffeomorphism $f_\varepsilon \in \mathrm{Diff}^r(M)$. By Fenichel's theorem [Fen], for ε sufficiently small, F_ε has an invariant manifold M_ε close to M, such that $\pi_\varepsilon = \pi|_{M_\varepsilon}$ is invertible. (To be precise, Fenichel's theorem is for flows, and we apply it by considering the suspension flow associated with f.) Furthermore, F_ε is strongly contracting toward M_ε, so that all of its periodic orbits (indeed, all of its nonwandering points) are on M_ε. We then let $f_\varepsilon = \pi_\varepsilon \circ F_\varepsilon \circ \pi_\varepsilon^{-1}$. Because of this smooth conjugacy, we can prove many properties of f_ε by proving them about F_ε.

Given this construction, we make the following definition.

Definition 9.2 (Nonlinear Prevalence). *A subset $P \subset \mathrm{Diff}^r(M)$ is prevalent if for some functions $F_1, F_2, \ldots \in C^r(T, \mathbb{R}^N)$ and a sufficiently small Hilbert brick HB such that the construction above works for every $\varepsilon \in HB$, we have that for each $f \in \mathrm{Diff}^r(M)$, the diffeomorphism f_ε constructed above belongs to P for almost every ε with respect to the uniform measure on HB.*

Of course, this definition depends on the choices made in our construction – the particular embedding of M and the means of extending a diffeomorphism on M to a neighborhood of its embedded image. We emphasize that the results in this paper and any results proved by a similar technique are independent of the details of the construction; the family of polynomial perturbations works regardless of the choices of embedding and extension. In this sense, we do not construct just a single family of perturbations for which our results are true with probability one, but rather an entire class of parametrized families that establish prevalence.

References

[AM] Artin, M., Mazur, B., On periodic orbits, *Ann. Math.* **81** (1965), 82–99.

[AK] Avila, A., Krikorian, R., Reducibility or non–uniform hyperbolicity for quasiperiodic Schrödinger cocycles, to appear in *Ann of Math.*

[ALM] Avila, A., Lyubich, M., de Melo, W., Regular or stochastic dynamics in real analytic families of unimodal maps, *Invent. Math.* 154 (2003), no. 3, 451–550.

[BC] Benedicks, M., Carleson, L., The dynamics of the Hénon map, *Ann. of Math.* (2) 133 (1991), no. 1, 73–169.

[Chr] Christensen, J. P. R., On sets of Haar measure zero in abelian Polish groups, *Israel J. Math.* 13 (1972), 255–260.

[C] Colli, E., Infinitely many coexisting strange attractors, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 15 (1998), no. 5, 539–579.

[GaS] Gavrilov, N., Shilnikov, L., On the three dimensional system close to a system with a structurally unstable homoclinic curve. I. *Math. USSR Sbornik* 17, pp. 467-485, 1972; II. *Math. USSR Sbornik* 19, 139-156, 1973.

[GoS] Gonchenko, S., Shil'nikov, L., On dynamical systems with structurally unstable homoclinic curves, *Soviet Math. Dokl.*, 33(1), 1986, 234-238.

[GST] Gonchenko, S., Shil'nikov, L., Turaev, D., Homoclinic tangencies of an arbitrary order in Newhouse regions (Russian), Dynamical systems, *Itogi Nauki Tekh. Ser. Sovrem. Mat. Prilozh. Temat. Obz., Vseross. Inst. Nauchn. i Tekhn. Inform. (VINITI)*, 67, Moscow, (1999), 69–128.

[GST1] Gonchenko, S., Shil'nikov, L., Tuvaev, D., On models with non-rough Poincare homoclinic curves, *Phisica D* 62 (1-4), 1993, 1-14.

[GStT] Gonchenko, S., Sten'kin, O., Tuvaev, D., Complexity of homoclinic bifurcations and Ω-moduli, *Intern. Journal of Bifur. and Chaos*, vol.6, no. 6, (1996), 969-989.

[GK] Gorodetski, A., Kaloshin, V., How often surface diffeomorphisms have infinitely many sinks and hyperbolicity of periodic points near a homoclinic tangency, preprint, (2004), 99pp, www.its.caltech.edu/∼ kaloshin/research/sinks.pdf.

[Fen] Fenichel, N., Persistence and smoothness of invariant manifolds for flows, *Indiana Univ. Math. J.* 21 (1971), 193–226.

[HSY] Hunt, B. R., Sauer, T., Yorke, J. A., 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.

[Ka1] Kaloshin, V., Generic diffeomorphisms with superexponential growth of the number of periodic orbits, *Comm in Math Phys*, **211**, (2000), no.1, 253-271.

[Ka2] Kaloshin, V., An extension of the Artin-Mazur theorem, *Ann. Math.* **150** (1999), 729–741.

[Ka3] Kaloshin, V., Ph.D. thesis, Princeton University, 2001.

[Ka4] Kaloshin, V., Stretched exponential bound on growth of the number of periodic points for prevalent diffeomorphisms, part II, preprint (2001), 86pp, www.its.caltech.edu/∼ kaloshin/research/Per-pts2.pdf.

[Ka5] Kaloshin, V., Some prevalent properties of smooth dynamical systems, *Proc. Steklov Math. Inst.*, no. 3 **213** (1997), 115–140.

[KH1] Kaloshin, V., Hunt, B., A stretched exponential bound on growth of the number of periodic points for prevalent diffeomorphisms, part I, 95pp, to appear in *Ann of Math.*, www.its.caltech.edu/∼ kaloshin/research/Per-pts1.pdf.

[KH2] Kaloshin, V., Hunt, B., A stretched exponential bound on the rate of growth of the number of periodic points for prevalent diffeomorphisms, *Electr Research Ann of AMS,* **7**, 2001, part I, 17–27, part II, 28–36.

[Ko] Kolmogorov, A.N., Talk at IMC 1956.

[Ly] Lyubich, M., Almost every real quadratic map is either regular or stochastic, *Ann of Math* (2) **156** (2002), no. 1, 1–78.

[MV] Mora, L., Viana, M., Abundance of strange attractors, *Acta Math.* 171, (1993), no. 1, 1–71.

[N1] Newhouse, S., Diffeomorphisms with infinitely many sinks. *Topology* 13, (1974), 9–18.

[N2] Newhouse, S., The abundance of wild hyperbolic sets and nonsmooth stable sets of diffeomorphisms, *Publ. Math. I.H.E.S.*, 50, (1979), 101–151.

[NP] Newhouse, S., Palis, J., Cycles and bifurcation theory, *Asterisque*, 31 (1976), 44–140.

[OY] Ott, W., Yorke, J., Prevalence, *Bull. Amer. Math. Soc.* **42** no. 3 (1993), 263–290.

[O] Oxtoby, J. C., Measure and Category, Springer-Verlag (1971).

[Pa] Palis, J., A Global view of dynamics and a conjecture on the denseness of finititude of attractors, *Astérisque*, **261**, 2000, 335–347.

[PT] Palis, J., Takens, F., Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations. Cambridge University Press, 1993.

[PY1] Palis, J., Yoccoz, J.-Ch., Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension. *Acta Math.* 172 (1994), no. 1, 91–136.

[PY2] Palis, J., Yoccoz, J.-Ch., Fers cheval non uniformment hyperboliques engendrs par une bifurcation homocline et densit nulle des attracteurs. (French) [Nonuniformly hyperbolic horseshoes generated by homoclinic bifurcations and zero density of attractors] *C. R. Acad. Sci. Paris Sér. I Math.* 333 (2001), no. 9, 867–871.

[R] Robinson, C., Bifurcations to infinitely many sinks, *Comm. Math. Phys.* 90, (1986), no. 3, 433–459.

[Ri] Rios, I., Unfolding homoclinic tangencies inside horseshoes: hyperbolicity, fractal dimensions and persistent tangencies, *Nonlinearity* 14 (2001), no. 3, 431–462.

[Si] Simon, B., Operators with singular continuous spectrum. I. General operators, *Ann of Math* (2) **141** (1995) no. 1, 131–145.

[LY] Tedeschini-Lalli, L., Yorke, J., How often do simple dynamical processes have infinitely many coexisting sinks? *Comm. Math. Phys.* 106, (1986), no. 4, 635–657.

[WY] Wang, Q., Young, L.-S., Strange attractors with one direction of instability. *Comm. Math. Phys.* 218 (2001), no. 1, 1–97.

California Institute of Technology
Mathematics 253-37
Pasadena, CA 91125, USA
E-mail: asgor@caltech.edu

University of Maryland
Department of Mathematics
College Park, MD 20742-3281 USA
E-mail: bhunt@ipst.umd.edu

Penn State University
Mathematics Department
University Park, State College, PA 16802, USA
E-mail: kaloshin@math.psu.edu