ELECTRONIC RESEARCH ANNOUNCEMENTS
OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 7, Pages 28–36 (April 24, 2001)
S 1079-6762(01)00091-9

A STRETCHED EXPONENTIAL BOUND ON THE RATE OF GROWTH OF THE NUMBER OF PERIODIC POINTS FOR PREVALENT DIFFEOMORPHISMS II

VADIM YU. KALOSHIN AND BRIAN R. HUNT

(Communicated by Svetlana Katok)

ABSTRACT. We continue the previous article's discussion of bounds, for prevalent diffeomorphisms of smooth compact manifolds, on the growth of the number of periodic points and the decay of their hyperbolicity as a function of their period n. In that article we reduced the main results to a problem, for certain families of diffeomorphisms, of bounding the measure of parameter values for which the diffeomorphism has (for a given period n) an almost periodic point that is almost nonhyperbolic. We also formulated our results for 1-dimensional endomorphisms on a compact interval. In this article we describe some of the main techniques involved and outline the rest of the proof. To simplify notation, we concentrate primarily on the 1-dimensional case.

1. INTRODUCTION

In the previous article, we described in Section 3 bounds on the growth (as a function of n) of the number of periodic points of period n and the decay of their hyperbolicity for "almost every" $C^{1+\rho}$ diffeomorphism of a finite-dimensional smooth compact manifold. Our definition of "almost every" is based on the measure-theoretic notion of prevalence, described in Section 2 of the previous article. In Section 5 of that article, we reduced the main results to a problem of estimating, within a particular parametrized family of diffeomorphisms, the measure of "bad" parameters — those for which the diffeomorphism has an almost periodic point that is almost nonhyperbolic. The reader should refer to the previous article, primarily Sections 3 and 5, for notation and terminology used below.

2. PERTURBATION OF RECURRENT TRAJECTORIES BY NEWTON INTERPOLATION POLYNOMIALS

The approach we take to estimate the measure of "bad" parameter values in the space of perturbations $HB^N(\mathbf{r})$ 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

Received by the editors December 21, 2000.
2000 *Mathematics Subject Classification.* Primary 37C20, 37C27, 37C35, 34C25, 34C27.
Key words and phrases. Periodic points, prevalence, diffeomorphisms.

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 of 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 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*

$$f_u(x) = f(x) + u\prod_{k=0}^{n-2}(x - x_k). \tag{1}$$

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 is because *one cannot 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 our other 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

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

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

$$\Big||(f_v^n)'(x_0)| - 1\Big| = \left|\left|\prod_{k=0}^{n-1} f'(x_k) + v\prod_{k=0}^{n-2}(x_{n-1} - x_k)^2\prod_{k=0}^{n-2} f'(x_k)\right| - 1\right| > \gamma \tag{3}$$

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

One more time we can see that 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 (1) and (2) 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

$$j^{1,n}f(x_0, \dots, x_{n-1}) = (x_0, \dots, x_{n-1}, f(x_0), \dots, f(x_{n-1}), f'(x_0), \dots, f'(x_{n-1})) . \tag{4}$$

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

$$\mathcal{J}^{1,n}(I, \mathbb{R}) = \underbrace{\mathcal{J}^1(I, \mathbb{R}) \times \cdots \times \mathcal{J}^1(I, \mathbb{R})}_{n \text{ times}}, \tag{5}$$

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

$$\Delta_n(I) = \big\{ \{x_0, \dots, x_{n-1}\} \times \mathbb{R}^{2n} \subset \mathcal{J}^{1,n}(I, \mathbb{R}) : \exists\, i \neq j \text{ such that } x_i = x_j \big\} \tag{6}$$

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 $\Delta_n(I)$ 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

$$\prod_{k=0}^{n-2} |x_{n-1} - x_k| \tag{7}$$

measures how close $\{x_k\}_{k=0}^{n-1}$ is to the diagonal $\Delta_n(I)$, or how independently one can perturb points of a trajectory. One can also say that (7) is a quantitative characteristic of how recurrent a trajectory of length n is. Introduction of this *product of distances along a trajectory* into the analysis of recurrent trajectories is a new point of our paper.

3. 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 $2n - 1$,

$$f_\varepsilon(x) = f(x) + \phi_\varepsilon(x), \qquad \phi_\varepsilon(x) = \sum_{k=0}^{2n-1} \varepsilon_k x^k, \tag{8}$$

where $\varepsilon = (\varepsilon_0, \dots, \varepsilon_{2n-1}) \in \mathbb{R}^{2n}$. The perturbation vector ε consists of coordinates from the Hilbert brick $HB^1(\mathbf{r})$ of analytic perturbations defined in Section 3 of the previous article. 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

$$\mathcal{J}^{1,n}(x_0, \dots, x_{n-1}, \varepsilon) = j^{1,n}\phi_\varepsilon(x_0, \dots, x_{n-1}). \tag{9}$$

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

$$\phi_\varepsilon(x) = \sum_{k=0}^{2n-1} \varepsilon_k x^k = \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j (\mathrm{mod}\ n)}). \tag{10}$$

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 $\mathcal{J}^{1,n}(I, \mathbb{R})$.

$$\begin{tikzcd}
 & \mathcal{DD}^{1,n}(I,\mathbb{R}) = \underbrace{I \times \cdots \times I}_{n \text{ times}} \times \mathbb{R}^{2n} \arrow[d, "\pi^{1,n}"] \\
\underbrace{I \times \cdots \times I}_{n \text{ times}} \times \mathbb{R}^{2n} \arrow[ur, "\mathcal{D}^{1,n}"] \arrow[r, "\mathcal{J}^{1,n}"] & \mathcal{J}^{1,n}(I,\mathbb{R}) = \underbrace{I \times \cdots \times I}_{n \text{ times}} \times \mathbb{R}^{2n}
\end{tikzcd}$$

FIGURE 1. An 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

$\pi^{1,n} : \mathcal{DD}^{1,n}(I,\mathbb{R}) \to \mathcal{J}^{1,n}(I,\mathbb{R})$:

$$
\begin{aligned}
\phi_\varepsilon(x_0) &= u_0,\\
\phi_\varepsilon(x_1) &= u_0 + u_1(x_1 - x_0),\\
\phi_\varepsilon(x_2) &= u_0 + u_1(x_2 - x_0) + u_2(x_2 - x_0)(x_2 - x_1),\\
&\ \vdots\\
\phi_\varepsilon(x_{n-1}) &= u_0 + u_1(x_{n-1} - x_0) + \cdots + u_{n-1}(x_{n-1} - x_0)\cdots(x_{n-1} - x_{n-2}),\\
\phi'_\varepsilon(x_0) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1}(x - x_{j(\mathrm{mod}\ n)})\right)\Big|_{x=x_0},\\
&\ \vdots\\
\phi'_\varepsilon(x_{n-1}) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1}(x - x_{j(\mathrm{mod}\ n)})\right)\Big|_{x=x_{n-1}}.
\end{aligned}
\tag{11}
$$

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 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 $0 \le k \le n-1$, if pairwise 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 the derivative $f'_\varepsilon(x_0)$ at x_0 depends only and nontrivially on u_n. Likewise for $0 \le k \le n-1$, if distinct points $\{x_j\}_{j=0}^{n}$ 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 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 3.1. The *first order divided difference* of g is defined as

$$\Delta g(x_0, x_1) = \frac{g(x_1) - g(x_0)}{x_1 - x_0} \tag{12}$$

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 mth order for $2 \le m \le r$,

$$\Delta^m g(x_0, \ldots, x_m) = \frac{\Delta^{m-1} g(x_0, \ldots, x_{m-2}, x_m) - \Delta^{m-1} g(x_0, \ldots, x_{m-2}, x_{m-1})}{x_m - x_{m-1}} \tag{13}$$

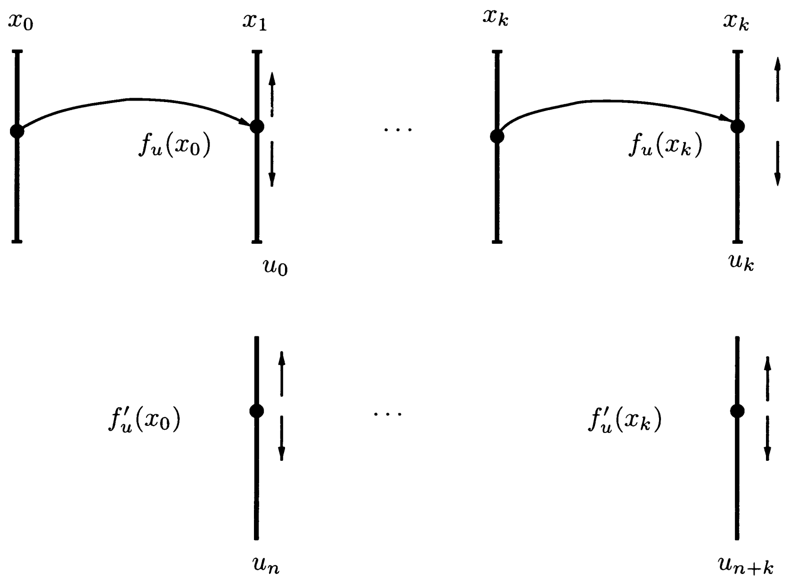

FIGURE 2. Newton coefficients and their action.

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

$$\Delta^m g(x_0, \dots, x_m) = \sum_{i=0}^{m} \frac{g(x_i)}{\prod_{j \neq i}(x_i - x_j)}. \tag{14}$$

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

$$\Delta^m x^k(x_0, \dots, x_m) = p_{k,m}(x_0, \dots, x_m), \tag{15}$$

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

$$p_{k,m}(x_0, \dots, x_m) = \sum_{r_0+\cdots+r_m=k-m} \prod_{j=0}^{m} x_j^{r_j}. \tag{16}$$

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

$$\begin{aligned} g(x) = {} & \Delta^0 g(x_0) + \Delta^1 g(x_0, x_1)(x - x_0) + \cdots \\ & + \Delta^{n-1} g(x_0, \dots, x_{n-1})(x - x_0) \cdots (x - x_{n-2}) \\ & + \Delta^n g(x_0, \dots, x_{n-1}, x)(x - x_0) \cdots (x - x_{n-1}) \end{aligned} \tag{17}$$

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 (17) 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 (10). 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 (10) 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

$$\begin{aligned} u_m &= \Delta^m \left(\sum_{k=0}^{2n-1} \varepsilon_k x^k \right) (x_0, \dots, x_{m(\text{mod } n)}) \\ &= \varepsilon_m + \sum_{k=m+1}^{2n-1} \varepsilon_k p_{k,m}(x_0, \dots, x_{m(\text{mod } n)}) \end{aligned} \tag{18}$$

for $m = 0, \dots, 2n-1$. We call the transformation given by (18) the *Newton map*. Notice that for fixed $\{x_k\}_{k=0}^{2n-1}$, the Newton map is linear and given by an upper triangular matrix with units on the diagonal. Hence it is Lebesgue volume-preserving and invertible, whether or not $\{x_k\}_{k=0}^{2n-1}$ lies on the diagonal $\Delta_n(I)$.

We call the basis of monomials

$$\prod_{j=0}^{k} (x - x_{j(\text{mod } n)}) \quad \text{for} \quad k = 0, \dots, 2n-1 \tag{19}$$

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 $\mu_n(C, \delta, \rho, M_{1+\rho})$ of "bad" parameter values $\vec{\varepsilon} \in HB^N(\mathbf{r})$.

4. Discretization method

The fundamental problem with using the Newton basis to estimate the measure of "bad" parameter values, those for which there is an almost periodic point of period n that is not sufficiently hyperbolic, is that the Newton basis depends on the almost periodic n-tuple $\{x_k\}_{k=0}^{n-1}$. For a particular "bad" parameter value we can fix this n-tuple and the corresponding Newton basis, then estimate the measure of the set of parameters for which a nearby n-tuple is both almost periodic and not sufficiently hyperbolic. But there are a continuum of possible n-tuples, so how can we account for all of the possible cells of "bad" parameter values $\vec{\varepsilon}$ within our parameter brick $HB^N(\mathbf{r})$? At the beginning of Section 2, we indicated that for a particular initial condition x_0 we would obtain an estimate on the measure of "bad" parameter values corresponding to an almost periodic point in a neighborhood of x_0, and thus need only to consider a discrete set of initial conditions. But as the parameter vector $\vec{\varepsilon}$ varies over $HB^N(\mathbf{r})$, there is (for large n at least) a wide range of possible length-n trajectories starting from a particular x_0, so there is no hope of using a single Newton basis to estimate even the measure of "bad" parameter values corresponding to a single x_0.

The solution to this problem is to discretize the entire space of n-tuples $\{x_k\}_{k=0}^{n-1}$, considering only those that lie on a particular grid. If we choose the grid spacing small enough, then every almost periodic orbit of period n that is not sufficiently hyperbolic will have a corresponding pseudotrajectory of length-n on the grid that also has small hyperbolicity. In this way we reduce the problem to bounding the measure of a set of "bad" parameter values corresponding to a particular length-n pseudotrajectory, and then summing the bounds over all possible length-n pseudotrajectories on the chosen grid.

Returning to the general case of $C^{1+\rho}$ diffeomorphisms on B^N, where we assume $0 < \rho \le 1$, the grid spacing we use at stage n is $\tilde{\gamma_n}(C, \delta, \rho) = N^{-1}(M_{1+\rho}^{-2n}\gamma_n(C, \delta))^{1/\rho}$, where $M_{1+\rho} > 1$ is a bound on the $C^{1+\rho}$ norm of the diffeomorphisms $f_{\vec{\varepsilon}}$ corresponding to parameters $\vec{\varepsilon} \in HB^N(\mathbf{r})$. This ensures that when rounded off to the nearest grid points $\{x_k\}_{k=0}^{n-1}$, an almost periodic orbit of length n becomes an $N(M_{1+\rho} + 1)\tilde{\gamma_n}(C, \delta, \rho)$-pseudotrajectory, meaning that $|f_{\vec{\varepsilon}}(x_j) - x_{j+1}| \le N(M_{1+\rho} + 1)\tilde{\gamma_n}(C, \delta, \rho)$ for $j = 0, 1, \dots, n-2$. It also ensures that when rounding, the derivative $df_{\vec{\varepsilon}}$ changes by at most $M_{1+\rho}^{1-2n}\gamma_n(C, \delta)$, which in turn implies that the change in hyperbolicity over all n points is small compared with $\gamma_n(C, \delta)$. (Recall that $\gamma_n(C, \delta)$ is our tolerance for hyperbolicity at stage n.)

Roughly speaking, in the case $N = 1$ our estimate on the measure of "bad" parameter values for a particular n-tuple $\{x_k\}_{k=0}^{n-1}$ is then proportional to $(\tilde{\gamma_n}(C, \delta, \rho))^n \cdot \gamma_n(C, \delta)$, whereas the number of possible n-tuples is proportional to $(\tilde{\gamma_n}(C, \delta, \rho))^{-n}$, making our bound $\mu_n(C, \delta, \rho, \mathbf{r}, M_{1+\rho})$ on the total measure of "bad" parameter values at stage N proportional to $\gamma_n(C, \delta)$. The remaining problem then is to show that for maps satisfying the Inductive Hypothesis of order $n - 1$, we can bound the proportionality factor in such a way that $\mu_n(C, \delta, \rho, \mathbf{r}, M_{1+\rho})$ meets the conditions prescribed in Section 5 of the previous article, namely that it is summable over n and that the sum approaches 0 as $C \to \infty$. (Notice that the sequence $\gamma_n(C, \delta)$ meets these conditions.) The proportionality factor depends on the product of distances described in Section 2, and in [K4] we proceed as follows. At the nth stage we split length-n trajectories of diffeomorphisms satisfying the Inductive Hypothesis into three groups. One group consists of what we call "simple" trajectories for which the product of distances is not too small. For nonsimple trajectories we show that either the trajectory is sufficiently hyperbolic by the Inductive Hypothesis (second group) or the trajectory returns very close to itself before the nth iteration and is simple (not recurrent) up to this point (third group). In the latter case, perturbation by Newton polynomials of order lower than n at the point of a close return can control the behavior of that trajectory up to length n.

Notice that in the preceding paragraph, even if the product of distances is not small, the proportionality factor in our estimate on the measure of "bad" parameter values for a given n-tuple $\{x_k\}_{k=0}^{n-1}$ is large because the parameter measure is normalized to be 1 on a brick $HB^1(\mathbf{r})$ whose sides decay rapidly; the normalization increases the measure by a factor of $r_0 r_1 \cdots r_{n-1} r_{2n-1}$. However, we are able to show that when considering only diffeomorphisms $f_{\vec{\varepsilon}}$ with $\vec{\varepsilon} \in HB^1(\mathbf{r})$, the number of n-tuples we must regard as possible pseudotrajectories of $f_{\vec{\varepsilon}}$ is reduced by the factor $r_0 r_1 \cdots r_{n-2}$. Due to our definition of an admissible sequence $\mathbf{r}$, the remaining factor $r_{n-1} r_{2n-1}$ does not affect the necessary summability properties for the bounds $\mu_n(C, \delta, \rho, \mathbf{r}, M_{1+\rho})$. There is an additional distortion of our estimates that is exponential in n, due to the fact that an image of a finite-dimensional brick of ε-parameters under the Newton map is a parallelepiped of u-parameters, but no longer a brick. This exponential factor is also not problematic, because our bound $\mu_n(C, \delta, \rho, \mathbf{r}, M_{1+\rho})$ decays superexponentially in n.

5. Conclusion

In this and the previous article we have only been able to outline some of the fundamental tools that are needed for the proof of the main result, which will appear in [K5] and [KH]. Here we list some of major difficulties appearing in the proof.

• We must handle almost periodic trajectories of length n that have a close return after $k < n$ iterates, so that, as discussed above, the product of distances along the trajectory is small. The precise definition of a close return is a major problem here. It must not be too restrictive, because we must also show that a trajectory without close returns is simple (the product of distances is not too small).[1]

• In dimension $N > 1$, the Newton interpolation polynomials involve products of differences of coordinates of points, which may be small even though the points themselves are not close. Thus we must be careful about how we construct the Newton basis for a given n-tuple of points $x_0, \dots, x_{n-1} \subset B^N$ and how to incorporate this into the general framework of the space of Newton interpolation polynomials.

• At nth stage of the induction we need to deal with the $(2n)^N$-dimensional space $W_{\leq 2n-1,N}$ of polynomials of degree $2n-1$ in N variables and handle the distortion properties of the Newton map. In a space of such a large dimension, even the ratio of volumes of the unit ball and the unit cube is of order $(2n)^{N(2n)^N}$ [San].

In [K5], [KH], based on [K4], we first prove the main 1-dimensional result for the case $N = \rho = 1$, discussed in Section 4 of the previous article and Sections 2 and 3 of this article, and then, using additional tools and ideas, complete the proof in the general case.

References

[GG] M. Golubitsky and V. Guillemin, *Stable mappings and their singularities*, Springer-Verlag, 1973. MR **49**:6269

[GY] A. Grigoriev, S. Yakovenko, Topology of generic multijet preimages and blow-up via Newton interpolation, J. Diff. Equations **150** (1998), 349–362. MR **99m**:58028

[K4] V. Yu. Kaloshin, Ph.D. thesis, Princeton University, 2001.

[K5] V. Kaloshin, Stretched exponential bound on growth of the number of periodic points for prevalent diffeomorphisms, part 1, in preparation.

[KH] V. Kaloshin, B. Hunt, Stretched exponential bound on growth of the number of periodic points for prevalent diffeomorphisms, part 2, in preparation.

[San] L. Santalo, Integral geometry and geometric probability, Encycl. of Math. and its Appl., Vol. 1, Addison-Wesley, Reading, MA–London–Amsterdam, 1976. MR **55**:6340

Fine Hall, Princeton University, Princeton, NJ 08544

E-mail address: `kaloshin@math.princeton.edu`

Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742

E-mail address: `bhunt@ipst.umd.edu`

[1] This is exactly the place in the proof where we need to impose superexponential decay of our bounds on hyperbolicity and periodicity.