Ergod. Th. & Dynam. Sys. (1999), **19**, 155–199
Printed in the United Kingdom © *1999 Cambridge University Press*

Regularity of invariant graphs for forced systems

JAROSLAV STARK
Centre for Nonlinear Dynamics and its Applications, University College London, Gower Street, London, WC1E 6BT, UK

(*Received* 17 *July* 1996 *and accepted in revised form* 9 *March* 1998)

Abstract. Many applications of nonlinear dynamics involve forced systems. We consider the case where for a fixed input the driven system is contracting; this is for instance the situation in certain classes of filters, and in the study of synchronization. When this contraction is uniform, it can easily be shown that there exists a globally attracting invariant set which is the graph of a function from the driving state space to the driven state space; this is a special case of the well known concept of an inertial manifold for more general systems. If the driving state space is a manifold and the contraction is sufficiently strong this invariant set is a normally hyperbolic manifold, and hence smooth. The aim of this paper is to extend this result in two directions: firstly, where we only have uniform contraction for a compact invariant set of input states, and secondly where the contraction rates are non-uniform (and hence defined by Liapunov exponents and analogous quantities). In both cases the invariant graph is only defined over closed subsets of the input space, and hence we need to define an appropriate notion of smoothness for such functions. This is done in terms of the Whitney extension theorem: a function is considered Whitney smooth if it satisfies the conditions of this theorem and hence can be extended to a smooth function of the whole input space.

1. *Introduction*

Many dynamical systems of both practical and theoretical importance consist of one system driven by another. The simplest, best known, and most widely studied case of this is that of periodic forcing, but a variety of other examples, involving more complex forms of forcing, can be found throughout engineering and the applied sciences. Where such systems are designed to perform some useful task, the forced system will usually be stable in some sense or another, at least in the absence of any forcing.

A particularly topical class of such systems, which helped to motivate the work in the present paper, is that of *recursive filters*; that is filters which possess some kind of internal dynamics of their own. When used to process signals generated by another deterministic system (for instance, arising from a laboratory experiment, or a piece of mechanical or electrical machinery) these can exhibit interesting interactions between the filter's own dynamics and the dynamics of the input signal (e.g. Badii *et al* 1998, Broomhead *et al*

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

1992, Stark 1993, Stark and Davies 1994, Pecora and Carroll 1996, Hunt *et al* 1996, Stark 1997). Another interesting example occurs in the case of so-called synchronized systems (e.g. Pecora and Carroll 1990, 1991); a number of potential applications of this phenomenon to telecommunications problems has been suggested. See Hunt *et al* (1997) and Stark (1997) for a discussion of synchronization in the present context.

1.1. *Skew products.* The standard approach to studying the dynamics of driven systems is to extend the state space of the driven system to include the driving system and give an enlarged system of the form (in discrete time)

$$x_{n+1} = f(x_n)$$
$$y_{n+1} = g(x_n, y_n)$$

where $x_n \in X$ represents the state of the driving system and $y_n \in Y$ is the state of the driven system. Systems of this kind are often called *skew products*.

The situation that we are interested in here is when g is in some sense contracting in the y direction. This is, for instance, almost always the case with recursive filters (e.g. Stark and Davies 1994), and in a weaker form is a necessary condition for synchronization (e.g. Pecora and Carroll 1990, 1991). When this contraction is uniform, it is easy to see that there exists an invariant attracting set for the overall dynamics, which is the graph of a continuous function $\Phi : X \to Y$ (Hirsch and Pugh 1970, Theorem 6.1a, Hirsch *et al* 1977, Theorem 3.1, or see Theorem 1.1 below). When such a Φ exists it gives a topological conjugacy between the dynamics of (f, g) on the graph of Φ and dynamics of the driving system f. Since the graph of Φ is attracting, all the recurrent behaviour of (f, g) lies in this graph and hence the asymptotic dynamics of (f, g) are identical to the dynamics of f, up to a continuous co-ordinate change. Another way of putting this is that the dynamics of g are slaved to that of f. Furthermore, in this case the dynamics of g are in a certain sense stable: if trajectories are perturbed away from the graph of Φ they return to it at an asymptotically exponential rate. In the context of recursive filters this remark provides a framework for the study of the stability of such filters (Stark and Davies 1994); this is a matter of considerable practical importance (e.g. Ljung and Ljung 1985, Slock 1992, Regalia 1993).

Topological conjugacy is, however, quite a weak notion of equivalence of dynamical systems; for instance, it typically does not preserve invariants such as eigenvalues of periodic orbits, Liapunov exponents or fractal dimensions. It is therefore natural to ask under what circumstances Φ is smooth. A sufficient condition for this is that the strongest contraction in the driving dynamics x is weaker than the weakest contraction in y. This follows from the fact that in this case the graph of Φ is a normally hyperbolic manifold, and it is well known that such manifolds are smooth (e.g. Hirsch *et al* 1977 or Ruelle 1989). In fact, appealing to normal hyperbolicity is slight overkill; the required regularity result is the main conclusion of the C^r section theorem (Hirsch and Pugh 1970, Theorem 6.1c, Hirsch *et al* 1977, Theorem 3.5). This theorem is in fact the main ingredient in proving the smoothness of normally hyperbolic manifolds, and hence the two arguments are essentially the same.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

When the above condition is satisfied and hence Φ is smooth, the dynamics of (f, g) on the graph of Φ are smoothly conjugate to f. In the context of time series analysis, when g represents a recursive filter, this is sufficient to ensure that the filter does not alter the fractal dimension or Liapunov exponents of the input signal (e.g. Badii *et al* 1988, Broomhead *et al* 1992, Stark and Davies 1994, Pecora and Carroll 1996, Hunt *et al* 1996, Stark 1997). As we shall see below (see also Davies and Campbell 1996, Campbell and Davies 1996), somewhat weaker conditions are also sufficient to guarantee this. Furthermore, in the case of a recursive least-squares filter, the graph of Φ corresponds to the solution of the relevant least-squares problem, and hence the regularity of Φ determines the regularity of the dependence of this solution on the input data.

It should also be noted that the graph of Φ is a special case of the concept of an *inertial manifold* (e.g. Temam 1988) which in recent years has been the subject of intense study in the field of partial differential equations. In fact, most proofs of the existence and regularity of inertial manifolds amount to a demonstration of a $\mathcal{C}^r$ section (or Lipschitz section in the worst case) type result, within the particular context of the class of equations under consideration.

1.2. *Whitney smoothness of invariant graphs.* From a practical point of view, the requirement that X is a compact manifold and our estimates of the contraction rates of f and g are valid over the whole of X is a serious drawback to the $\mathcal{C}^r$ section theorem. In many cases this is unrealistically strong: for instance, suppose we wish to study a recursive filter driven by a signal generated by the Hénon map (e.g. Badii *et al* 1988, Broomhead *et al* 1992, Stark 1993, 1997, Stark and Davies 1994, Pecora and Carroll 1996, Hunt *et al* 1996). Then $X = \mathbb{R}^2$, and the maximum contraction in the Hénon map is not bounded as one goes to infinity in $\mathbb{R}^2$. Hence one cannot use the standard $\mathcal{C}^r$ section theorem to deduce the existence of a smooth graph $\Phi : X \to Y$.

Of course, in this example we are not really interested in the existence or regularity of Φ for the whole of X; after all the input sequence will be restricted to states $x \in X$ corresponding to the attractor of the Hénon map. This attractor is a compact f-invariant set $\Lambda \subset X$. This motivates us to restrict our study to graphs of functions which are only defined over such a set. This is a natural approach within the context of hyperbolic dynamics; after all, except in the case of Anosov systems we only ever assume hyperbolicity on a compact invariant basic set (e.g. Ruelle 1989, Katok and Hasselblatt 1995).

A similar issue arises if we do not possess uniform estimates on the contraction rates of f and g, and instead work with average rates defined in terms of the most negative Liapunov exponent of f, and an analogous quantity for g (defined in §1.4 below). In such a case we can control the behaviour of Φ only on a nested sequence of compact sets Λ_k, and hence again we need to define a notion of smoothness for functions $\Phi : \Lambda_k \to Y$.

Restricting to a closed Λ poses no problem as far existence, continuity and Hölder regularity are concerned. Indeed, weaker forms of the section theorem merely require X to be a compact metric space, and hence apply to this case (Hirsch and Pugh 1970, Theorems 6.1a and 6.1b, Hirsch *et al* 1977, Theorems 3.1 and 3.8, and see below). However, when we come to discuss the smoothness of $\Phi : \Lambda \to Y$, we run into the

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

question of what we mean by this. The natural concept to use is that of differentiability in the Whitney sense. This was first suggested in the present context by Davies and Campbell (1996) (see also Campbell and Davies 1996). Since we wish to use this for non-integer degrees of differentiability it is convenient to use Stein's formulation of this. Thus, recall that if $\varphi : \Lambda \to \mathbb{R}^m$ where $\Lambda \subset \mathbb{R}^n$ is a closed set, we say that φ is $\mathcal{C}^{1+\gamma}$ in the Whitney sense if there exists a function $A : \Lambda \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and a constant $C \geq 0$ such that

$$\|\varphi(x) - \varphi(x') - A(x)(x - x')\| \leq Cd(x, x')^{1+\gamma} \tag{1.1}$$

$$\|A(x) - A(x')\| \leq Cd(x, x')^{\gamma} \tag{1.2}$$

for all $x, x' \in \Lambda$. In such a case, by the Stein (1970) version of the Whitney (1934) extension theorem, φ can be extended to a $\mathcal{C}^{1+\gamma}$ function on the whole of $\mathbb{R}^n$, with $D_x\varphi = A(x)$ for all $x \in \Lambda$. We thus think of A as a derivative of φ in the Whitney sense. Higher degrees of differentiability are defined similarly, see Stein (1970) for details.

This definition can be extended to the case of a compact $\Lambda \subset X$ when X and Y are manifolds by using local charts in the usual way. Thus, A is now a function $A : \Lambda \to \mathcal{L}(T_\Lambda X, TY)$ such that $A(x) \in \mathcal{L}(T_x X, T_{\varphi(x)}Y)$ for all $x \in \Lambda$. It is convenient to require φ to be uniformly Lipschitz. Then, since Λ is compact we may choose $r > 0$ sufficiently small such that for all $x \in \Lambda$, $\exp_x^{-1} : N_x(r) \to T_x X$ and $\exp_{\varphi(x)}^{-1} : N_{\varphi(x)}(Kr) \to T_{\varphi(x)}Y$ are embeddings, where $N_x(r) = \{z : d(x, z) \leq r\}$, $N_y(r) = \{y : d(y, y') \leq r\}$ and K is a Lipschitz constant for φ (so that $d(\varphi(x), \varphi(x')) \leq Kd(x, x')$ for all $x, x' \in \Lambda$). Condition (1.1) then becomes

$$\|\exp_{\varphi(x)}^{-1}(\varphi(x)) - \exp_{\varphi(x)}^{-1}(\varphi(x')) - A(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\| \leq Cd(x, x')^{1+\gamma} \tag{1.3}$$

for all $x, x' \in \Lambda$ such that $d(x, x') \leq r$. Condition (1.2) of course makes sense in this new setting. Note that $\exp_x^{-1}(x) = 0$, and $\exp_{\varphi(x)}^{-1}(\varphi(x)) = 0$ so that we could write $\|\exp_{\varphi(x)}^{-1}(\varphi(x')) - A(x)\exp_x^{-1}(x')\|$ for the left-hand side of (1.3), but this makes the relationship to (1.1) much less transparent.

1.3. *Regularity of graphs over invariant sets.* The first main result of this paper is the generalization of the $\mathcal{C}^{1+\gamma}$ section theorem to the case of uniform contraction on a f-invariant compact set Λ. Before we give a precise statement of this, we describe conditions for the existence, continuity and Hölder regularity of Φ in such a case. The theorems given here are minor modifications of those appearing in Hirsch and Pugh (1970), and Hirsch *et al* (1977). Despite this we give our own proofs here. This is for two reasons: firstly, some of the intermediate lemmas are required in the proof of the $\mathcal{C}^{1+\gamma}$ theorem and, secondly, these elementary proofs serve to introduce most of the important ideas both in the $\mathcal{C}^{1+\gamma}$ theorem, and in the proofs of analogous results in the non-uniformly hyperbolic case described in the next section. The latter in particular are extremely messy, and hence readers are urged to first familiarize themselves with the basic ideas presented here.

Note that Hirsch *et al* develop their results using the language of vector bundles over X. However, the first step in their proofs is to reduce to a trivial bundle $X \times Y$, and hence their results are no more general than the ones given here. Conversely, Hirsch *et al* assume that

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

$c = 1$, $c' = 1$ in (1.4)–(1.6) below, but the reduction to this case is elementary. Finally, we allow the fibre Y to be an arbitrary metric space rather than requiring it to possess a vector space structure; this does not affect the proofs in any way and is largely cosmetic.

The fundamental condition required in all these results is that g is uniformly contracting in the Y direction. More precisely we assume that there exists a $\lambda < 0$ and a $c \leq 0$ such that

$$d(g_x^{(n)}(y), g_x^{(n)}(y')) \leq ce^{\lambda n} d(y, y') \tag{1.4}$$

for all $x \in \Lambda$ where $g_x^{(n)} : Y \to Y$ is given by $g_x^{(n+1)}(y) = g(f^n(x), g_x^{(n)}(y))$ with $g_x^{(1)}(y) = g_x(y) = g(x, y)$. Then we have the following.

THEOREM 1.1. (Hirsch and Pugh 1970, Hirsch *et al* 1977) *Suppose X is a metric space, $f : X \to X$ a homeomorphism, $\Lambda \subset X$ a closed f-invariant set, Y a complete metric space and $g : X \times Y \to Y$ is continuous, satisfies (1.4) for all $x \in \Lambda$, and for some $y_0 \in Y$ the function $g(x, y_0) : X \times Y \to Y$ is bounded. Then there exists a bounded continuous function $\Phi : \Lambda \to Y$ such that the graph of Φ is (f, g)-invariant and attracting for all $(x, y) \in \Lambda \times Y$.*

Note that by (1.4), if $g(x, y_0)$ is bounded for one y_0, then it is bounded for all $y_0 \in Y$. Furthermore if Λ is compact, then this condition is automatically satisfied. To obtain any additional regularity for Φ, we need to bound the strongest (local) contraction in f. This is most easily expressed in terms of the expansion in f^{-1}. If X is a manifold and f a diffeomorphism, the natural way to formulate this is

$$\|D_x f^{-n}\| \leq c' e^{\mu n} \tag{1.5}$$

for some $\mu \in \mathbb{R}$ and $c' \geq 0$. Since the conditions we impose on μ are upper bounds we may always take $\mu \geq 0$; in fact for $\mu < 0$, the regularity results below follow trivially from standard stable manifold theory. When considering the Hölder or Lipschitz properties of Φ, there is no need for f to be differentiable. Instead of the above condition we thus require that

$$d(f^{-n}(x), f^{-n}(x')) \leq c' e^{\mu n} d(x, x') \tag{1.6}$$

for all $x, x' \in \Lambda$. Note that if Λ is compact and f is $\mathcal{C}^{1+\alpha}$ and satisfies (1.5), then given $\varepsilon > 0$, there exists a $\delta > 0$ such that for all x, x' with $d(x, x') \leq \delta e^{-(\mu+\varepsilon)n}$ we have $d(f^{-n}(x), f^{-n}(x')) \leq c' e^{(\mu+\varepsilon)n} d(x, x')$. Furthermore, if K is the diameter of Λ, we have $d(f^{-n}(x), f^{-n}(x')) \leq K\delta^{-1} e^{(\mu+\varepsilon)n} d(x, x')$ for all $d(x, x') \geq \delta e^{-(\mu+\varepsilon)n}$. Hence in this case, (1.5) implies (1.6), apart from replacing μ by $\mu + \varepsilon$. This argument also shows that, at least for compact Λ, it is not unreasonable to expect (1.6) to hold globally, rather than just for sufficiently nearby x, x'.

We also need some control on the dependence of g on x. In particular, we shall require $g(x, y)$ to be a Hölder function of x, for a fixed y. Since Y is not necessarily bounded it is not reasonable to require a uniform Hölder condition on the whole of $\Lambda \times Y$. It turns out to be sufficient that there exists an $0 < \alpha \leq 1$ such that for any bounded $W \subset Y$ there exists a constant $K_W \geq 0$ satisfying,

$$d(g(x, y), g(x', y)) \leq K_W d(x, x')^\alpha \tag{1.7}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

for all $x, x' \in \Lambda$, $y \in W$. We make no assumptions on the dependence of K_W on W. We then have the following.

THEOREM 1.2. (Hirsch and Pugh 1970, Hirsch *et al* 1977) *Suppose that in addition to the hypotheses of Theorem 1.1, f satisfies (1.6) and g satisfies (1.7). Then Φ is uniformly Hölder on Λ with exponent γ, for any $0 < \gamma \leq \alpha$, such that $\lambda + \gamma\mu < 0$. In particular if $\lambda < -\mu$ and $\alpha = 1$, then Φ is Lipschitz.*

Note that as Hirsch *et al* (1977) point out, the statement of this in Hirsch and Pugh (1970) is incorrect since it lacks condition (1.7). Hirsch *et al* (1977) go too far the other way: they require that (1.7) holds with $\alpha = 1$, which is unnecessarily strong.

Finally, we come to discuss the smoothness of Φ. To do this we need a differentiable structure on both X and Y and hence we assume that both these spaces are $\mathcal{C}^\infty$ Riemannian manifolds and that f and g are $\mathcal{C}^{1+\alpha}$ maps. Since we need uniformity in the sizes of charts about points in Λ, we assume that Φ is compact. We then have the following.

THEOREM 1.3. *Suppose X and Y are $\mathcal{C}^\infty$ Riemannian manifolds, f is a $\mathcal{C}^{1+\alpha}$ diffeomorphism satisfying (1.5), $\Lambda \subset X$ is a compact f-invariant set, and g is uniformly $\mathcal{C}^{1+\alpha}$ on compact sets, and satisfies (1.4). Suppose that $0 < \gamma \leq \alpha$ is such that $\lambda + (1+\gamma)\mu < 0$. Then Φ is $\mathcal{C}^{1+\gamma}$ in the Whitney sense.*

For clarity of exposition, we give the main part of the proof for Y an open convex subset of a Banach space Y_0 such that $(f, g)(X \times \bar{Y}) \subset X \times Y$, where $\bar{Y}$ is the closure of Y. This of course includes the situation where $Y = Y_0$, which for instance is the relevant case in most applications to recursive filters. A sketch of the modifications required to cover the case of an arbitrary manifold Y is given in §4.4.

When Λ is actually a compact manifold, the above theorem reduces to the standard $\mathcal{C}^r$ section theorem of Hirsch and Pugh (1970) and Hirsch *et al* (1977). The latter, of course, holds for higher degrees of smoothness than $1 + \gamma$, and by analogy we expect the obvious generalization of Theorem 1.3 to hold, namely that if f and g are $\mathcal{C}^{r+\alpha}$ and $\lambda + (r+\gamma)\mu < 0$, then Φ is $\mathcal{C}^{r+\gamma}$. In the case of the standard $\mathcal{C}^r$ section theorem, the passage from $\mathcal{C}^{1+\gamma}$ to $\mathcal{C}^{r+\gamma}$ is straightforward: from the proof of the $\mathcal{C}^{1+\gamma}$ theorem it is easy to see that derivatives of functions converge at a rate $\lambda + \mu$, hence by induction the derivative of Φ is $\mathcal{C}^{r-1+\gamma}$, and hence Φ is $\mathcal{C}^{r+\gamma}$. Unfortunately, this approach no longer works for the Whitney notion of smoothness: we can certainly use it prove that Φ is Whitney $\mathcal{C}^{1+\gamma}$, and the function A in (1.1) is Whitney $\mathcal{C}^{r-1+\gamma}$, but this no longer implies that Φ is $\mathcal{C}^{r+\gamma}$. It should be possible to give a direct non-inductive proof of a $\mathcal{C}^{r+\gamma}$ theorem using the same techniques that we use to prove Theorem 1.3. However, since this would be technically very messy and would introduce no new concepts, we have not included it here. In any case, in applications it is the $\mathcal{C}^{1+\gamma}$ result which is generally the most useful.

Finally, note that one might be tempted to deduce Theorem 1.3 from stable manifold theory; thus under the conditions of the theorem Λ has a centre unstable manifold of precisely the correct smoothness which is locally a graph over a neighbourhood of Λ. The problem with this approach is that this manifold is not necessarily unique, and hence one cannot show that $x' \in \Lambda$ must lie in the local centre unstable manifold of $x \in \Lambda$, even for arbitrarily close x, x'.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

1.4. *Non-uniform convergence rates.* The second direction in which we extend the $\mathcal{C}^r$ section theorem is to relax the requirement of uniform estimates on the contraction rates of f and g. In many applications, such uniform estimates are either unavailable, or too weak for Theorem 1.3 to hold. An example of this occurs in synchronization (Pecora and Carroll 1990, 1991) where the contraction in g is usually specified in terms of Liapunov exponent-like quantities (see Stark 1997 for further discussion). Similarly, returning to the example of forcing by the Hénon map in §1.2, since this map is not uniformly hyperbolic, a uniform estimate of its contraction rate will be an overestimate of its real behaviour and will lead to pessimistic estimates of the smoothness of Φ. Finally, note that recently Ashwin *et al* (1994, 1996) have discovered a variety of interesting phenomena in the neighbourhood of invariant manifolds with non-uniform normal contraction.

The non-uniformity of contraction and expansion rates is familiar in many other branches of dynamics and has motivated the development of Liapunov exponents and Pesin theory. The question that we address here, therefore, is to what extent the results of the previous section carry over to the situation where we only have information about average rates of contraction. It turns out that using standard ideas from Pesin theory we can still say a great deal about the existence and regularity of Φ.

Clearly, to talk about average rates of contraction, we need an invariant measure. We thus assume that we are given an invariant measure σ of f (not necessarily ergodic). In order to have some measure of contraction rates we need to assume that for σ almost every x the maximum expansion in g_x is bounded, and that the bound is an integrable function of x. This is sufficient to ensure (§5.1 below) that

$$\lambda(x) = \lim_{n\to\infty} \frac{1}{n} \sup_{y,y'\in Y} \log \frac{d(g_x^{(n)}(y), g_x^{(n)}(y'))}{d(y, y')}$$

exists for almost every x (and is a measurable function of x). A reasonable notion of non-uniform contraction is then that there exists a $\lambda < 0$ such that

$$\lambda(x) \le \lambda \tag{1.8}$$

for σ almost every x. Note that this implies that given any $\eta > 0$ there exists a measurable function $c : X \to \mathbb{R}^+ = \{z \in \mathbb{R} : z \ge 0\}$ such that

$$d(g_x^{(n)}(y), g_x^{(n)}(y')) \le c(x)e^{(\lambda+\eta)n} d(y, y') \tag{1.9}$$

which makes the relationship to (1.4) much more transparent. We then have the following.

THEOREM 1.4. *Suppose X is a compact metric space, $f : X \to X$ a homeomorphism, σ an invariant measure, Y a complete metric space and $g : X \times Y \to Y$ is continuous and satisfies (1.8). Then there exists an f-invariant set $\Lambda \subset X$ such that $\sigma(\Lambda) = 1$ and a function $\Phi : \Lambda \to Y$ such that the graph of Φ is invariant and attracting under (f, g).*

This theorem has recently been applied by Nicol and Broomhead (1999) to the study of the stability of forced systems to random perturbations. We now turn to the regularity of Φ. First, by analogy to Theorem 1.1, we consider boundedness and continuity. If $c(x)$ is not bounded we cannot expect Φ to be bounded either. The best we can do is to construct a nested sequence of compact sets $\Lambda_k \subset \Lambda_{k+1}$ with $\sigma(\Lambda_k) \to 1$ as $k \to \infty$ such that Φ

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

is bounded on each Λ_k, and we have some control of the norm of $\Phi|_{\Lambda_k}$ as $k \to \infty$. In the absence of a vector space structure on Y, we express this by choosing a distinguished point $y_0 \in Y$, and giving bounds on $d(\Phi(x), y_0)$ for all $x \in \Lambda_k$, for each k. Since the set of singularities of Φ can be dense in Λ, we see that Φ in general need not be continuous in the conventional sense (see §7.2 for a further discussion of this point). Instead, we find that Φ is uniformly continuous on each Λ_k. We thus obtain the following non-uniform analogue of Theorem 1.1.

THEOREM 1.5. *Suppose X is a compact metric space, $f : X \to X$ a homeomorphism, σ an f-invariant measure, Y a complete metric space and $g : X \times Y \to Y$ is continuous and satisfies (1.8). Then given any $\eta > 0$ we can construct a nested sequence $\Lambda_k \subset \Lambda_{k+1}$ of compact sets such that $\sigma(\Lambda_k) \to 1$ as $k \to \infty$, $\Phi|_{\Lambda_k}$ is uniformly continuous and there exists a $c \geq 0$ such that $d(\Phi(x), y_0) \leq ce^{\eta k}$ for all $x \in \Lambda_k$, for all k.*

Of course, if Y is in fact a Banach space, we can describe the bounds on Φ by $\|\Phi|_{\Lambda_k}\| \leq ce^{\eta k}$ where $\|\cdot\|$ is the sup norm on bounded functions $\varphi : \Lambda_k \to Y$. This theorem is essentially a generalization of Theorem 5 of Campbell and Davies (1996) to the case of unbounded Y. Note that it gives Theorem 1.4 as an immediate corollary, but since Theorem 1.4 can be proved by much more elementary means we also give a direct proof (in §6.2 below).

As in §1.3, to obtain any additional regularity for Φ, we need both to bound the strongest contraction in f (for which we need to impose some additional structure on X and f) and to have some control over the growth of g.

Bounds on the contraction of f are rather more straightforward than was the case for g. In all those cases where these are required we may assume that X is a $\mathcal{C}^\infty$ compact Riemannian manifold and f is a $\mathcal{C}^{1+\alpha}$ diffeomorphism. Then by the sub-additive ergodic theorem (Kingman 1968) there exists a set Λ of full measure such that

$$\mu(x) = \lim_{n\to\infty} \frac{1}{n} \log \|D_x f^{-n}\|$$

is a f-invariant measurable function. Note that $\mu(x)$ is the largest Liapunov exponent of f^{-1} with respect to σ (since σ is not ergodic this is a measurable function rather than a constant). Thus $-\mu(x)$ is the most negative exponent of f; in other words $-\mu(x)$ gives the strongest contraction in f. Since X is compact $\mu(x)$ is bounded above and so we may define μ to be the essential supremum of $\mu(x)$, i.e. μ is the largest constant such that

$$\mu(x) \leq \mu \tag{1.10}$$

for σ almost every $x \in X$. Note that given any $\eta > 0$ there exists a measurable function $b : X \to \mathbb{R}^+$ such that

$$\|D_x f^{-n}\| \leq b(x)e^{(\mu+\eta)n}$$

and hence (1.10) is the natural analogue of condition (1.5).

We will make use of two increasingly strong conditions on growth rates of g as $d(y, y_0) \to \infty$. To show that Φ is Hölder continuous with exponent α we shall need $g(x, y)$ to be a Hölder function of x, with some control of the Hölder norm as

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

$d(y, y_0) \to \infty$. In particular, we require that there exist a β and constants K_g, $K_{g,0} \geq 0$ such that

$$d(g(x, y), g(x', y)) \leq (K_{g,0} + K_g d(y, y_0)^\beta) d(x, x')^\alpha. \tag{1.11}$$

To discuss the smoothness of Φ we need a differentiable structure on Y. For convenience we shall assume that Y is an open convex subset of a Banach space Y_0 such that $(f, g)(X \times \bar{Y}) \subset X \times Y$, where $\bar{Y}$ is the closure of Y. This is the same setting in which the main part of the proof of Theorem 1.3 is given. It should be possible to prove analogous results if Y is a manifold as long as we have some control over the differentiable structure as we go out to infinity. We shall require g to be uniformly $\mathcal{C}^{1+\alpha}$ on compact sets and the Hölder norm of Dg to satisfy the following growth condition

$$\|D_{(x,y)}g - D_{(x',y')}g\| \leq (K'_{g,0} + K'_g(m(y, y'))^\beta)(d(x, x')^\alpha + d(y, y')^\alpha) \tag{1.12}$$

for some constants K'_g, $K'_{g,0} \geq 0$, where $m(y, y') = \max\{d(y, y_0), d(y', y_0)\}$ and $\|D_{(x,y)}g\|$ is the induced operator norm derived from the norm on Y_0. This, of course, depends on the choice of charts on X; however, we shall always use charts given by $\exp_x$ and only apply the above estimate when x' is in the same chart as x and hence the left-hand side is defined unambiguously. We then get this analogue of Theorem 1.2.

THEOREM 1.6. *Suppose X is a compact Riemannian manifold, f a $\mathcal{C}^{1+\alpha}$ diffeomorphism satisfying (1.10), σ an f-invariant measure, Y a complete metric space and $g: X \times Y \to Y$ a $\mathcal{C}^{1+\alpha}$ map satisfying (1.8) and (1.11). Then for any $0 < \gamma \leq \alpha$ such that $\lambda + \gamma\mu < 0$ and any sufficiently small $\eta > 0$, we can construct a nested sequence $\Lambda_k \subset \Lambda_{k+1}$ of compact sets such that $\sigma(\Lambda_k) \to 1$ as $k \to \infty$, $\Phi|_{\Lambda_k}$ is Hölder with exponent γ and the Hölder constant of $\Phi|_{\Lambda_k}$ is bounded by $ce^{b\eta k}$, for some $b > 0$, i.e. $d(\Phi(x), \Phi(x')) \leq ce^{b\eta k} d(x, x')^\gamma$ for all $x, x' \in \Lambda_k$.*

In particular, if $\lambda < -\mu$ then Φ is uniformly Lipschitz on sets of measure $1 - \varepsilon$ for any $\varepsilon > 0$. As Davies and Campbell (1996) point out, when σ is ergodic, this implies that the information dimension of $\tilde{\sigma}$ is the same as the information dimension of σ, where $\tilde{\sigma}$ is the (f, g) invariant measure on $X \times Y$ induced from σ by Φ. A weaker version of this theorem was proved by Campbell and Davies (1996) under slightly different conditions. Their approach, however, did not permit them to give an explicit Hölder exponent, and in particular they did not show that if $\lambda < -\mu$ then $\Phi|_{\Lambda_k}$ is uniformly Lipschitz. Finally, it should be pointed out that the Λ_k in this theorem are not necessarily the same as those in Theorem 1.5, since they have to take account of the behaviour of f as well as g.

Finally, we come to discuss the smoothness of Φ. In general, Λ cannot be expected to be an open set, and so as in §1.3 our notion of smoothness is that given by the Whitney extension theorem. Furthermore, as in Theorems 1.5 and 1.6, we cannot hope for uniform smoothness on the whole of Λ. Instead, we show that Φ is smooth in the Whitney sense on each Λ_k, and give bounds on the growth rate of its norm as $k \to \infty$.

THEOREM 1.7. *Suppose X is a compact Riemannian manifold, f a $\mathcal{C}^{1+\alpha}$ diffeomorphism satisfying (1.10), σ an f-invariant measure, Y an open convex subset of a Banach space Y_0 such that $(f, g)(X \times \bar{Y}) \subset X \times Y$ and $g : X \times Y \to Y$ is a $\mathcal{C}^{1+\alpha}$ map satisfying (1.8) and*

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

(1.12). Suppose that $\lambda < -\mu$ and $0 < \gamma \leq \alpha$ is such that $\lambda + (1+\gamma)\mu < 0$. Then for any sufficiently small $\eta > 0$ we can construct a nested sequence $\Lambda_k \subset \Lambda_{k+1}$ of compact sets such that $\sigma(\Lambda_k) \to 1$ as $k \to \infty$, and $\Phi|_{\Lambda_k}$ is $\mathcal{C}^{1+\gamma}$ in the Whitney sense; more precisely there exists an A^* and constants $C \geq 0$, $r > 0$, $b > 0$ such that for all $x, x' \in \Lambda_k$ with $d(x, x') \leq re^{-\eta k}$ we have

$$\|\Phi(x) - \Phi(x') - A^*(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\| \leq Ce^{b\eta k} d(x, x')^{1+\gamma}$$
$$\|A^*(x) - A^*(x')\| \leq Ce^{b\eta k} d(x, x')^{\gamma}$$
$$\|A^*(x)\| \leq Ce^{b\eta k}.$$

As in the case of Theorem 1.3, we expect that it should be possible to give an analogous result for higher degrees of smoothness, but have not done so.

This theorem is a generalization of the $\mathcal{C}^1$ results in Davies and Campbell (1995) for linear recursive filters. For such systems g is of the form $g(x, y) = Ay + \psi(x)$ for a linear map $A : Y \to Y$ and some function $\psi : X \to Y$. It is then obviously possible to write down an explicit analytic expression for Φ, and study its regularity by direct calculation. As far as we are aware, their paper is the first application of Whitney smoothness to invariant manifolds, although they only consider $\mathcal{C}^1$ conditions, rather than $\mathcal{C}^{1+\alpha}$. Campbell and Davies (1996) investigate the smoothness of Φ for general systems with bounded Y, but only derive much weaker results. Note also that both of these papers rather unnecessarily require σ to be ergodic.

2. *Existence and continuity under uniform contraction*

Proof of Theorem 1.1. Let $\mathcal{C}^0$ be the space of bounded continuous functions $\varphi : \Lambda \to Y$. We define the usual metric of uniform convergence on $\mathcal{C}^0$

$$d(\varphi, \varphi') = \sup_{x \in \Lambda} d(\varphi(x), \varphi'(x)),$$

$\mathcal{C}^0$ with this metric is a complete metric space. Define the usual graph transform Γ by

$$\Gamma(\varphi)(x) = g(f^{-1}(x), \varphi(f^{-1}(x))).$$

We claim that $\Gamma(\varphi) \in \mathcal{C}^0$, i.e. that $\Gamma(\varphi)$ is bounded. Thus let y_0 be a point such that $g(x, y_0)$ is bounded; say $d(g(x, y_0), y_0) \leq C$ for some constant C. Since φ is bounded we have $d(\varphi(f^{-1}(x)), y_0) \leq C'$ for some constant C'. Then by (1.4), we have

$$\begin{aligned} d(\Gamma(\varphi)(x), y_0) &\leq d(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), y_0)) + d(g(f^{-1}(x), y_0), y_0) \\ &\leq ce^{\lambda} d(\varphi(f^{-1}(x)), y_0) + C \leq ceC' + C, \end{aligned}$$

as required. Thus Γ is a map $\Gamma : \mathcal{C}^0 \to \mathcal{C}^0$. Now

$$\Gamma^n(\varphi)(x) = g^{(n)}(f^{-n}(x), \varphi(f^{-n}(x)))$$

where $g^{(n)}$ is defined by $g^{(n+1)}(x, y) = g(f^n(x), g^{(n)}(x, y))$ with $g^{(1)} = g$. Thus, by (1.4)

$$\begin{aligned} d(\Gamma^n(\varphi)(x), \Gamma^n(\varphi')(x)) &= d(g^{(n)}(f^{-n}(x), \varphi(f^{-n}(x))), g^{(n)}(f^{-n}(x), \varphi'(f^{-n}(x)))) \\ &\leq ce^{\lambda n} d(\varphi(f^{-n}(x)), \varphi'(f^{-n}(x))) \\ &\leq ce^{\lambda n} d(\varphi, \varphi'). \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Taking the supremum over all $x \in \Lambda$,

$$d(\Gamma^n(\varphi), \Gamma^n(\varphi')) \leq ce^{\lambda n} d(\varphi, \varphi'). \tag{2.1}$$

Since $\lambda < 0$, choose m such that $\kappa = ce^{\lambda m} < 1$, then $d(\Gamma^m(\varphi), \Gamma^m(\varphi')) < \kappa d(\varphi, \varphi')$. Hence, by the contraction mapping theorem, Γ^m has a unique attracting fixed point Φ. Furthermore, $\Gamma^m(\Gamma(\Phi)) = \Gamma(\Gamma^m(\Phi)) = \Gamma(\Phi)$ and hence $\Gamma(\Phi)$ is also a fixed point of Γ^m. Hence $\Gamma(\Phi) = \Phi$. Since any fixed point of Γ is also a fixed point of Γ^m, Φ must be the unique fixed point of Γ, and it is attracting by (2.1).

By definition

$$\begin{aligned}\Phi(f(x)) &= \Gamma(\Phi)(f(x)) \\ &= g(f^{-1}(f(x)), \Phi(f^{-1}(f(x)))) \\ &= g(x, \Phi(x)).\end{aligned}$$

Hence, as required, the graph of Φ is invariant under (f, g). Finally, by induction we have $\Phi(f^n(x)) = g^{(n)}(x, \Phi(x))$ and hence,

$$\begin{aligned}d(\Phi(f^n(x)), g^{(n)}(x, y)) &= d(g^{(n)}(x, \Phi(x)), g^{(n)}(x, y)) \\ &\leq ce^{\lambda n} d(\Phi(x), y).\end{aligned}$$

Thus $d(\Phi(f^n(x)), g^{(n)}(x, y)) \to 0$ as $n \to \infty$, and the graph of Φ is attracting under (f, g). This completes the proof of Theorem 1.1. □

Note that by replacing d by an adapted Liapunov metric on Y we could have assumed that $c = 1$. However, this is unnecessarily complicated, particularly since in this case the adapted metric depends on the base point x. The argument given above is thus much more straightforward.

3. *Hölder and Lipschitz continuity under uniform contraction*

We now aim to prove Theorem 1.2. Note that by the above we see that any graph invariant under $(f, g)^m$ is also invariant (and attracting) under (f, g). Thus, by replacing (f, g) with $(f, g)^m$ for a sufficiently large m if necessary, we may from now on take $c = 1$ in (1.4). Thus, given $C > 0$, if we define

$$\mathcal{C}^0_C = \{\varphi \in \mathcal{C}^0 : d(\varphi, \Phi) \leq C\}$$

then by (2.1) we have $d(\Gamma(\varphi), \Phi) = d(\Gamma(\varphi), \Gamma(\Phi)) < d(\varphi, \Phi)$ and hence $\Gamma(\mathcal{C}^0_C) \subset \mathcal{C}^0_C$. Similarly we may take $c' = 1$ in (1.6), at the cost of replacing μ by $\mu + \varepsilon$, for $\varepsilon > 0$, arbitrarily small.

For $0 < \gamma \leq 1$, let $\mathcal{C}^\gamma \subset \mathcal{C}^0$ be the space of Hölder continuous functions with exponent γ:

$$\mathcal{C}^\gamma = \{\varphi \in \mathcal{C}^0 : d(\varphi(x), \varphi(x')) \leq Cd(x, x')^\gamma \text{ for some } C \geq 0, \text{ for all } x, x' \in \Lambda\}.$$

As usual, in the special case of $\gamma = 1$, we denote this by $\mathcal{C}^{\mathrm{Lip}}$. Note that if $\varphi \in \mathcal{C}^\gamma$ then $\Gamma(\varphi) \in \mathcal{C}^\gamma$. For $\varphi \in \mathcal{C}^\gamma$ define

$$H_\gamma(\varphi) = \sup_{x, x' \in \Lambda} \frac{d(\varphi(x), \varphi(x'))}{d(x, x')^\gamma}.$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Given $C > 0$ define

$$\mathcal{C}^\gamma_C = \{\varphi \in \mathcal{C}^\gamma : H_\gamma(\varphi) \leq C\}.$$

To prove that Φ is Hölder we shall show that for a sufficiently large $C > 0$ we have $\Gamma(\mathcal{C}^\gamma_C) \subset \mathcal{C}^\gamma_C$. The fact that $\Phi \in \mathcal{C}^\gamma_C$ then immediately follows from (2.1), and the following well known lemma.

LEMMA 3.1. *$\mathcal{C}^\gamma_C$ is closed in $\mathcal{C}^0$ (in the $\mathcal{C}^0$ topology) for any $C \geq 0$.*

Proof. Let $\varphi_n \in \mathcal{C}^\gamma$ be a sequence such that $d(\varphi_n, \varphi) \to 0$ as $n \to \infty$ for some $\varphi \in \mathcal{C}^0$ and $H_\gamma(\varphi_n) \leq C$. It suffices to prove that $H_\gamma(\varphi) \leq C$. Given $\varepsilon > 0$ we have

$$d(\varphi(x), \varphi(x')) \leq d(\varphi(x), \varphi_n(x)) + d(\varphi_n(x), \varphi_n(x')) + d(\varphi(x'), \varphi_n(x'))$$

for any $x, x' \in \Lambda$. Choose n sufficiently large such that $d(\varphi_n, \varphi) \leq \varepsilon$. Then $d(\varphi(x), \varphi_n(x)) \leq \varepsilon$, and $d(\varphi(x'), \varphi_n(x')) \leq \varepsilon$. Thus,

$$\begin{aligned} d(\varphi(x), \varphi(x')) &\leq 2\varepsilon + d(\varphi_n(x), \varphi_n(x')) \\ &\leq 2\varepsilon + Cd(x, x')^\gamma. \end{aligned}$$

But this holds for all $\varepsilon > 0$, and hence $d(\varphi(x), \varphi(x')) \leq Cd(x, x')^\gamma$ for all $x, x' \in \Lambda$. □

Let

$$N_\Phi(C) = \{(x, y) \in \Lambda \times Y : d(y, \Phi(x)) \leq C\}.$$

Given $\gamma \leq \alpha$, let $K_C \geq 0$ be such that $d(g(x, y), g(x', y)) \leq K_C d(x, x')^\gamma$ for all $x, x' \in \Lambda$, $y \in N_\Phi(C)$. To prove Theorem 1.2 it thus remains to show the following.

PROPOSITION 3.2. *Suppose $\gamma \leq \alpha$ such that $\lambda + \gamma\mu < 0$, $C > 0$ and $C_\gamma \geq K_C e^{\mu\gamma}(1 - e^{\lambda+\mu\gamma})^{-1}$. Then $\Gamma(\mathcal{C}^0_C \cap \mathcal{C}^\gamma_{C_\gamma}) \subset \mathcal{C}^0_C \cap \mathcal{C}^\gamma_{C_\gamma}$.*

Proof. Suppose $\varphi \in \mathcal{C}^0_C \cap \mathcal{C}^\gamma_{C_\gamma}$. By (2.1) we have $\Gamma(\varphi) \in \mathcal{C}^0_C$. Furthermore,

$$\begin{aligned} d(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) \leq\ & d(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi(f^{-1}(x')))) \\ & + d(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))). \end{aligned}$$

Now

$$\begin{aligned} d(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi(f^{-1}(x')))) &\leq e^\lambda d(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))) \\ &\leq e^\lambda C_\gamma d(f^{-1}(x), f^{-1}(x'))^\gamma \\ &\leq C_\gamma e^{\lambda+\mu\gamma} d(x, x')^\gamma. \end{aligned}$$

Turning to the second term, by (1.6) we have,

$$\begin{aligned} d(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))) &\leq K_C d(f^{-1}(x), f^{-1}(x'))^\gamma \\ &\leq K_C e^{\mu\gamma} d(x, x')^\gamma. \end{aligned}$$

Thus, since $(C_\gamma e^{\lambda+\mu\gamma} + K_C e^{\mu\gamma}) \leq C_\gamma$,

$$\begin{aligned} d(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) &\leq (C_\gamma e^{\lambda+\mu\gamma} + K_C e^{\mu\gamma}) d(x, x')^\gamma \\ &\leq C_\gamma d(x, x')^\gamma \end{aligned}$$

and hence $H_\gamma(\Gamma(\varphi)) \leq C_\gamma$ as required. □

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

COROLLARY 3.3. *$\Phi \in \mathcal{C}^\gamma$ for any $\gamma \le \alpha$ such that $\lambda + \gamma\mu < 0$. In particular, if $\lambda < -\mu$, then Φ is uniformly Lipschitz on Λ.*

Proof. Fix some $y_0 \in Y$ and let $\varphi_0 : \Lambda \to Y$ be the constant function $\varphi_0(x) = y_0$. Let $C = d(\varphi_0, \Phi)$ and $\varphi_n = \Gamma^n(\varphi_0)$. Then $\varphi_n \in \mathcal{C}^0_C$ for all n and $d(\varphi_0, \Phi) \to 0$ as $n \to \infty$. Clearly $\varphi_0 \in \mathcal{C}^\gamma$ and $H_\gamma(\varphi_0) = 0$. Then by induction $H_\gamma(\varphi_n) \le C_\gamma$, and hence $\Phi \in \mathcal{C}^\gamma$ by Lemma 4.1. □

This completes the proof of Theorem 1.2.

4. *Whitney smoothness under uniform contraction*

In this section we prove Theorem 1.3. We begin by first constructing a candidate A^* for the derivative of Φ using a fixed point argument, then we show that (Φ, A^*) is attracting in the space of pairs (φ, A) which in turn implies that (Φ, A^*) satisfies (1.1) and (1.2) by an argument similar to that used in the previous section. From now on assume that $\lambda < -\mu$ and hence Φ is Lipschitz. Recall that we assume $c = 1$ in (1.4) and $c' = 1$ in (1.6), by replacing (f, g) with $(f, g)^m$ for a sufficiently large m if necessary.

We shall first give the proof in the technically simpler case where Y is an open convex subset of a Banach space Y_0 such that $(f, g)(X \times \bar{Y}) \subset X \times Y$, where $\bar{Y}$ is the closure of Y (allowing of course the possibility that $Y = Y_0$). The definition (1.3) of Whitney $\mathcal{C}^{1+\gamma}$ now becomes

$$\|\varphi(x) - \varphi(x') - A(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\| \le C d(x, x')^{1+\gamma} \tag{4.1}$$

for all $x, x' \in \Lambda$ such that $d(x, x') \le r$. We describe the modifications required to extend the proof to Y a manifold in §4.4 below.

Note that Y itself is not necessarily complete, but $\bar{Y}$ is. Hence if $\lambda < -\mu$ we can apply the results of §2 and §3 to $\bar{Y}$ to deduce the existence of an invariant Lipschitz graph Φ in $X \times \bar{Y}$; by invariance this must then lie in $X \times Y$. Denote the norm on Y_0 by $\|\cdot\|$, so that the metric Y is given by $d(y, y') = \|y - y'\|$. We shall usually identify $T_y Y$ with Y_0.

4.1. *Existence of derivative.*

Let $\mathcal{L}(T_x X, Y_0)$ be the space of linear maps from $T_x X$ to Y_0, and denote by $\mathcal{L}(T_\Lambda X, Y_0)$ the vector bundle whose fibre above $x \in \Lambda$ is $\mathcal{L}(T_x X, Y_0)$. Let $\mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ be the space of continuous sections of $\mathcal{L}(T_\Lambda X, Y_0)$, thus if $A \in \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ then $A(x)$ is a linear map from $T_x X$ to Y_0. The graph transform extends to $\mathcal{L}^0 = \mathcal{C}^0 \times \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ in a natural way by

$$\tilde{\mathbf{\Gamma}}(\varphi, A) = (\Gamma(\varphi), \Gamma_{\mathcal{L}}(\varphi, A))$$

where

$$\Gamma_{\mathcal{L}}(\varphi, A)(x) = D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})$$

with $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. We define a norm $\|A\|$ on $\mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ by

$$\|A\| = \sup_{x \in \Lambda} \|A(x)\|$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

where $\|A(x)\|$ is the operator norm on $\mathcal{L}(T_xX, Y_0)$ with respect to the norm on T_xX given by the Riemannian metric. We then construct a metric on $\mathcal{C}^0$ by

$$d((\varphi, A), (\varphi', A')) = \max\{d(\varphi, \varphi'), \|A - A'\|\}.$$

LEMMA 4.1. *Suppose $\varphi \in \mathcal{C}^0$ and $A, A' \in \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$. Then*

$$\|\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi, A')\| \le e^{\lambda+\mu}\|A - A'\|.$$

Proof. We have

$$\begin{aligned}(\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi, A'))(x) = {} & D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1}) \\ & - D_w g.(D_x f^{-1}, A'(f^{-1}(x)).D_x f^{-1})\end{aligned}$$

where $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. Given $x \in \Lambda$ and $v \in T_xX$ we have

$$D_w g.(D_x f^{-1}, A(f^{-1}(x))D_x f^{-1}).v = D_w^{(1)} g.u + D_w^{(2)} g.A(f^{-1}(x)).u$$

where $D_w^{(1)} g$ and $D_w^{(2)} g$ are the partial derivatives of g, and $u = D_x f^{-1}.v$. Note that $D_w^{(2)} g = D_{\varphi(f^{-1}(x))} g_{f^{-1}(x)}$, and hence $\|D_w^{(2)} g\| \le e^{\lambda}$. Thus

$$\begin{aligned}\|(\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi, A'))(x).v\| &= \|D_w^{(2)} g.(A(f^{-1}(x)) - A'(f^{-1}(x))).u\| \\ &\le e^{\lambda}\|(A(f^{-1}(x)) - A'(f^{-1}(x)))\|\,\|D_x f^{-1}.v\| \\ &\le e^{\lambda+\mu}\|A - A'\|\,\|v\|\end{aligned}$$

as required. □

COROLLARY 4.2. *If $\lambda < -\mu$, then $\tilde{\mathbf{\Gamma}}$ has a unique fixed point $(\Phi, A^*) \in \{\Phi\} \times \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$, where Φ is the unique fixed point of Γ, obtained previously.*

4.2. *Convergence in $\mathcal{L}^0$.* We now need to show that (Φ, A^*) is attracting in the whole of $\mathcal{L}^0$, rather than just $\{\Phi\} \times \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$. Since g is uniformly $\mathcal{C}^{1+\alpha}$ on compact sets, given $C > 0$, there exists a constant $K'_C \ge 0$ such that

$$\|D_{(x,y)} g - D_{(x',y')} g\| \le K'_C(d(x, x')^{\alpha} + d(y, y')^{\alpha}) \tag{4.2}$$

for all $x, x' \in \Lambda$ such that $d(x, x') \le r$, and all $y, y' \in Y$ such that $d(y, \Phi(x)) \le C$ and $d(y', \Phi(x)) \le C$. Then we have the following.

LEMMA 4.3. *Suppose $\varphi, \varphi' \in \mathcal{C}^0_C$, $A \in \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$. Then*

$$\|\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi', A)\| \le K'_C e^{\mu}(1 + \|A\|)d(\varphi, \varphi')^{\alpha}.$$

Proof. Suppose $x \in \Lambda$ and $v \in T_xX$. Then

$$(\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi', A))(x).v = (D_w g - D_{w'} g).(u, A(f^{-1}(x)).u)$$

where $u = D_x f^{-1}.v$ and $w = (f^{-1}(x), \varphi(f^{-1}(x)))$, $w' = (f^{-1}(x), \varphi'(f^{-1}(x)))$. Since $d(\varphi(f^{-1}(x)), \Phi(x)) \le C$ and $d(\varphi'(f^{-1}(x)), \Phi(x)) \le C$, by (4.2) we have

$$\begin{aligned}&\|(D_w g - D_{w'} g).(u, A(f^{-1}(x).u))\| \\ &\quad \le K'_C d(\varphi(f^{-1}(x)), \varphi'(f^{-1}(x)))^{\alpha}(\|u\| + \|A(f^{-1}(x)).u\|) \\ &\quad \le K'_C d(\varphi, \varphi')^{\alpha}(1 + \|A\|)e^{\mu}\|v\|\end{aligned}$$

as required. □

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

COROLLARY 4.4. *Suppose* $(\varphi, A) \in \mathcal{C}^0_C \times \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ *and* $\lambda < -\mu$. *Then for any* $\varepsilon > 0$ *such that* $\varepsilon < \max\{-(\lambda+\mu), -\lambda\alpha\}$ *there exists* $c > 0$ *such that*

$$\|A_n - A^*\| \le ce^{-\varepsilon n}$$

where $(\varphi_n, A_n) = \tilde{\mathbf{\Gamma}}^n(\varphi, A)$.

Proof. We have

$$\begin{aligned} A_{n+1} - A^* &= \Gamma_{\mathcal{L}}(\varphi_n, A_n) - \Gamma_{\mathcal{L}}(\Phi, A^*) \\ &= \Gamma_{\mathcal{L}}(\varphi_n, A_n) - \Gamma_{\mathcal{L}}(\varphi_n, A^*) + \Gamma_{\mathcal{L}}(\varphi_n, A^*) - \Gamma_{\mathcal{L}}(\Phi, A^*). \end{aligned}$$

Thus

$$\begin{aligned} \|A_{n+1} - A^*\| &\le e^{\lambda+\mu}\|A_n - A^*\| + K'_C e^{\mu}(1 + \|A\|)d(\varphi_n, \Phi)^\alpha \\ &\le e^{\lambda+\mu}\|A_n - A^*\| + c'e^{\lambda n\alpha} \end{aligned}$$

where $c' = K'_C e^{\mu}(1 + \|A\|)d(\varphi, \Phi)^\alpha$. Thus let $c = \max\{\|A - A^*\|, c'e^{\varepsilon}(1 - e^{\lambda+\mu+\varepsilon})^{-1}\}$. We claim that by induction we have

$$\|A_n - A^*\| \le ce^{-\varepsilon n}.$$

This holds trivially for $n = 0$. Suppose it holds for $n > 0$. Then

$$\begin{aligned} \|A_{n+1} - A^*\| &\le ce^{-\varepsilon n}e^{\lambda+\mu} + c'e^{\lambda n\alpha} \\ &\le ce^{-\varepsilon(n+1)}(e^{\lambda+\mu+\varepsilon} + c'c^{-1}e^{\varepsilon}e^{(\lambda\alpha+\varepsilon)n}) \end{aligned}$$

and hence since $e^{(\lambda\alpha+\varepsilon)n} \le 1$ and $c'c^{-1}e^{\varepsilon} \le (1 - e^{\lambda+\mu+\varepsilon})$ we have

$$\begin{aligned} \|A_{n+1} - A^*\| &\le ce^{-\varepsilon(n+1)}(e^{\lambda+\mu+\varepsilon} + 1 - e^{\lambda+\mu+\varepsilon}) \\ &\le ce^{-\varepsilon(n+1)} \end{aligned}$$

as required. □

COROLLARY 4.5. *If* $\lambda < -\mu$, *then* $d(\tilde{\mathbf{\Gamma}}^n(\varphi, A), (\Phi, A^*)) \to 0$ *as* $n \to \infty$ *for all* $(\varphi, A) \in \mathcal{L}^0$.

4.3. *Whitney smoothness.* Suppose $x, x' \in \Lambda$ and $x' \in N_x$, recall that Ψ_x^{-1} is a chart centred at x. Following (4.1), we define

$$\Delta((\varphi, A), x, x') = \varphi(x) - \varphi(x') - A(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))$$

for $x, x' \in \Lambda$ such that $d(x, x') \le r$, and hence $x' \in N_x$. Our aim is to show that for some A^* there exists a $C \ge 0$ and $0 < \gamma \le 1$ such that

$$\|\Delta((\Phi, A^*), x, x')\| \le Cd(x, x')^{1+\gamma} \tag{4.3}$$

for all $x, x' \in \Lambda$ such that $d(x, x') \le r$. Thus define

$$\begin{aligned} \mathcal{L}^{1+\gamma} = \{(\varphi, A) \in \mathcal{L}^0 : \|\Delta((\varphi, A), x, x')\| \le Cd(x, x')^{1+\gamma} \text{ for some } C \ge 0, \\ \text{for all } x, x' \in \Lambda \text{ such that } d(x, x') \le r\}. \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

For $(\varphi, A) \in \mathcal{L}^{1+\gamma}$ we define

$$H_{1+\gamma}(\varphi, A) = \sup_{x,x' \in \Lambda, x' \in N_x} \frac{\|\Delta((\varphi, A)x, x')\|}{d(x, x')^{1-\gamma}}.$$

Given $C > 0$ define

$$\mathcal{L}_C^{1+\gamma} = \{(\varphi, A) \in \mathcal{L}^{1+\gamma} : H_{1+\gamma}(\varphi, A) \leq C\}.$$

We now proceed very much as in §3. First, we show that $\mathcal{L}_C^{1+\gamma}$ is closed in $\mathcal{L}^0$; this is exactly analogous to Lemma 3.1. We will then go on to prove that if $\lambda < -\mu(1+\gamma)$, then if $(\varphi, A) \in \mathcal{L}_C^{1+\gamma}$ for a sufficiently large $C > 0$ (depending on $d(\varphi, \Phi)$, $\|A - A^*\|$ and $H_{\text{Lip}}(\varphi)$) then $\tilde{\mathbf{\Gamma}}(\varphi, A) \in \mathcal{L}_C^{1+\gamma}$. Using Corollary 4.5, this of course immediately implies that $(\Phi, A^*) \in \mathcal{L}^{1+\gamma}$. Thus we have the following.

LEMMA 4.6. *$\mathcal{L}_C^{1+\gamma}$ is closed in $\mathcal{L}^0$ (in the $\mathcal{L}^0$ topology) for any $C \geq 0$.*

Proof. Let $(\varphi_n, A_n) \in \mathcal{L}^{1+\gamma}$ be a sequence such that $H_{1+\gamma}(\varphi_n, A_n) \leq C$ for some $C \geq 0$ and such that $d((\varphi_n, A_n), (\varphi, A)) \to 0$ as $n \to \infty$ for some $(\varphi, A) \in \mathcal{L}^0$. We need to show that $H_{1+\gamma}(\varphi, A) \leq C$. Suppose $x, x' \in \Lambda$ and $x' \in N_x$. Then given $\varepsilon > 0$ we have

$$\begin{aligned}\|\Delta((\varphi, A), x, x')\| \leq \|\Delta((\varphi_n, A_n), x, x')\| + \|\varphi(x) - \varphi_n(x)\| + \|\varphi(x') - \varphi_n(x')\| \\ + \|(A(x) - A_n(x))(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\|.\end{aligned}$$

Choose n sufficiently large such that $d(\varphi, \varphi') \leq \varepsilon$ and $\|A - A_n\| \leq \varepsilon d(x, x')^{-1}$. Then $\|\varphi(x) - \varphi_n(x)\| = d(\varphi(x), \varphi_n(x)) \leq \varepsilon$ and similarly $\|\varphi(x') - \varphi_n(x')\| \leq \varepsilon$. Since $\|\exp_x^{-1}(x) - \exp_x^{-1}(x')\| = d(x, x')$, we have

$$\begin{aligned}\|(A(x) - A_n(x))(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\| &\leq \|A(x) - A_n(x)\| d(x, x') \\ &\leq \|A - A_n\| d(x, x') \leq \varepsilon.\end{aligned}$$

Thus

$$\begin{aligned}\|\Delta((\varphi, A), x, x')\| &\leq 3\varepsilon + \|\Delta((\varphi_n, A_n), x, x')\| \\ &\leq 3\varepsilon + C d(x, x')^{1+\gamma}.\end{aligned}$$

This holds for arbitrary $\varepsilon > 0$, and hence $\|\Delta((\varphi, A), x, x')\| \leq C d_{\eta,x}(x, x')^{1+\gamma}$, as required. □

Recall that we denote the space of Lipschitz functions (i.e. $\mathcal{C}^\gamma$ with $\gamma = 1$) by $\mathcal{C}^{\text{Lip}}$. By Proposition 3.2, we can find constants C_0 and C_1 such that $\tilde{\mathbf{\Gamma}}(\mathcal{C}_{C_0}^0 \cap \mathcal{C}_{C_1}^{\text{Lip}}) \subset \mathcal{C}_{C_0}^0 \cap \mathcal{C}_{C_1}^{\text{Lip}}$. Fix C_0, C_1 and given $C \geq 0$ let

$$\mathcal{L}_C^0 = \{(\varphi, A) \in \mathcal{L}^0 : \varphi \in \mathcal{C}_{C_0}^0 \cap \mathcal{C}_{C_1}^{\text{Lip}}, \|A\| \leq C\}.$$

Observe that the constant c in Corollary 4.4 depends on (φ, A) only through $d(\varphi, \Phi)$ and $\|A - A^*\|$. Thus, by replacing $\tilde{\mathbf{\Gamma}}$ by $\tilde{\mathbf{\Gamma}}^m$ for a sufficiently large m if necessary, we can choose $C_{\mathcal{L}}$ sufficiently large such that $\tilde{\mathbf{\Gamma}}(\mathcal{L}_{C_{\mathcal{L}}}^0) \subset \mathcal{L}_{C_{\mathcal{L}}}^0$. Since Df^{-1} is Hölder and Λ is compact there exists a constant $K_f \geq 0$ such that

$$\|D_x f^{-1}.\exp_x^{-1}(x') - \exp_{f^{-1}(x)}^{-1}(f^{-1}(x'))\| \leq K_f d(x, x')^{1+\alpha} \tag{4.4}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

for all $x, x' \in \Lambda$ such that $d(x, x') \le r$. Finally let

$$K_g = \sup_{x\in\Lambda,\, y\in N_\Phi(C_0)} \|D_{(x,y)}g\|. \tag{4.5}$$

Let

$$C_{1+\gamma} = \max\{C_0, C_{\mathcal{L}}, r^{-\gamma}e^{\mu\gamma}(C_1 + C_{\mathcal{L}}), \\ (1 - e^{-(\lambda+\mu(1+\gamma))})^{-1}(K'_C(1 + C_1) + K_g K_f + e^{\lambda} C_{\mathcal{L}} K_f)\}.$$

We then have the following.

PROPOSITION 4.7. *Suppose* $\lambda < -\mu$, $\lambda + (1 + \gamma)\mu < 0$, *with* $0 < \gamma \le \alpha$. *Then* $\tilde{\mathbf{\Gamma}}(\mathcal{L}^{1+\gamma}_{C_{1+\gamma}} \cap \mathcal{L}^0_{C_{\mathcal{L}}}) \subset \mathcal{L}^{1+\gamma}_{C_{1+\gamma}} \cap \mathcal{L}^0_{C_{\mathcal{L}}}$.

Proof. We need to prove that if $H_{1+\gamma}(\varphi, A) \le C_{1+\gamma}$, $d(\varphi_n, \Phi) \le C_0$, $H_{\mathrm{Lip}}(\varphi) \le C_1$, and $\|A\| \le C_{\mathcal{L}}$, then $\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\| \le C_{1+\gamma} d(x, x')^{1+\gamma}$ for all $x, x' \in \Lambda$ such that $d(x, x') \le r$. First suppose that $d(x, x') \le re^{-\mu}$ so that $d(f^{-1}(x), f^{-1}(x')) \le r$. Then

$$\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x') = g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) \\ - D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})(\exp_x^{-1}(x) - \exp_x^{-1}(x'))$$

with $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. Let $u_{x'} = -\exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))$ and $v_{x,x'} = \varphi(f^{-1}(x)) - \varphi(f^{-1}(x'))$. Then by (4.2), and using the fact that $d(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))) \le C_1 d(f^{-1}(x), f^{-1}(x'))$, we get

$$\begin{aligned} &\|g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) - D_w g.(u_{x'}, v_{x,x'})\| \\ &\quad \le K'_C(d(f^{-1}(x), f^{-1}(x'))^{1+\alpha} + d(\varphi(f^{-1}(x)), \varphi(f^{-1}(x')))^{1+\alpha}) \\ &\quad \le K'_C(1 + C_1) d(f^{-1}(x), f^{-1}(x'))^{1+\alpha} \\ &\quad \le c_0 d(x, x')^{1+\alpha} \end{aligned}$$

where $c_0 = K'_C(1 + C_1)e^{\mu(1+\alpha)}$. On the other hand, since $\exp_x^{-1}(x) = \mathbf{0}$,

$$\begin{aligned} &D_w g.(u_{x'}, v_{x,x'}) - D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})(\exp_x^{-1}(x) - \exp_x^{-1}(x')) \\ &\quad = D_w g.(D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x')), \mathbf{0}) \\ &\qquad - D_w g.(\mathbf{0}, \varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) - A(f^{-1}(x)).D_x f^{-1}.\exp_x^{-1}(x')). \end{aligned}$$

Using (4.4) and (4.5) we have,

$$\begin{aligned} &\|D_w g.(D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x')), \mathbf{0})\| \\ &\quad \le \|D_w g\|\, \|D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))\| \\ &\quad \le c_1 d(x, x')^{1+\alpha} \end{aligned}$$

where $c_1 = K_g K_f$. Similarly

$$\begin{aligned} &\|D_w g.(\mathbf{0}, A(f^{-1}(x)).(D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))))\| \\ &\quad \le e^{\lambda}\|A(f^{-1}(x))\|\, \|D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))\| \\ &\quad \le c_2 d(x, x')^{1+\alpha} \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

where $c_2 = e^{\lambda} C_{\mathcal{L}} K_f$. Finally

$$\begin{aligned}\|\varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) - A(f^{-1}(x)).\exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))\| \\ &= \|\Delta((\varphi, A), f^{-1}(x), f^{-1}(x'))\| \\ &\leq C_{1+\gamma} d(f^{-1}(x), f^{-1}(x'))^{1+\gamma} \\ &\leq C_{1+\gamma} e^{\mu(1+\gamma)} d(x, x')^{1+\gamma}\end{aligned}$$

and hence

$$\begin{aligned}\|D_w g.(\mathbf{0}, \varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) - A(f^{-1}(x)).\exp^{-1}_{f^{-1}(x)}(f^{-1}(x')))\| \\ \leq C_{1+\gamma} e^{\lambda+\mu(1+\gamma)} d(x, x')^{1+\gamma}.\end{aligned}$$

Combining all these estimates, we have

$$\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\| \leq (c_0 + c_1 + c_2 + C_{1+\gamma} e^{\lambda+\mu(1+\gamma)}) d(x, x')^{1+\gamma}.$$

Since $C_{1+\gamma} \geq (1 - e^{-(\lambda+\mu(1+\gamma))})^{-1}(c_0 + c_1 + c_2)$ we have

$$\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\| \leq C_{1+\gamma} d(x, x')^{1+\gamma}$$

as required. It remains for us to deal with the case $d(x, x') > re^{-\mu}$. By hypothesis, we have

$$\begin{aligned}d(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) &\leq C_1 d(x, x') \\ \|\Gamma_{\mathcal{L}}(\varphi, A)(x)\| &\leq C_{\mathcal{L}}.\end{aligned}$$

Since $\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x') = \Gamma(\varphi)(x) - \Gamma(\varphi)(x') - \Gamma_{\mathcal{L}}(\varphi, A)(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))$, this gives

$$\begin{aligned}\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\| &\leq C_1 d(x, x') + C_{\mathcal{L}} d(x, x') \\ &\leq r^{\gamma_e \mu\gamma} (C_1 + C_{\mathcal{L}}) d(x, x') r^{\gamma} e^{-\mu\gamma} \\ &\leq C_{1+\gamma} d(x, x')^{1+\gamma}\end{aligned}$$

as required. □

COROLLARY 4.8. $(\Phi, A^*) \in \mathcal{L}^{1+\gamma}$.

4.4. *Extension to fibre bundles.* We now briefly sketch the modifications required to extend the above proof to the case of an arbitrary manifold Y, so that $\Lambda \times Y$ is a trivial fibre bundle. This requires essentially no new ideas, and is achieved by transferring to local co-ordinates at several stages in the above argument.

Given two vector bundles E and F over Λ, we denote by $\mathcal{L}(E, F)$ the bundle whose fibre above $x \in \Lambda$ is the space $\mathcal{L}(E_x, F_x)$ of linear maps from E_x to F_x, where E_x and F_x are the fibres above x in E and F, respectively. Now, if $\varphi \in \mathcal{C}^0$, the pullback bundle $\varphi^* TY$ is a vector bundle over Λ (where the fibre above $x \in \Lambda$ is just $T_{\varphi(x)} Y$). Thus we may form the bundle $\mathcal{L}(T_\Lambda X, \varphi^* TY)$ whose fibre over $x \in \Lambda$ is $\mathcal{L}(T_x X, T_{\varphi(x)} Y)$. Let $\mathcal{L}^0(\varphi) = \mathcal{C}^0(\mathcal{L}(T_\Lambda X, \varphi^* TY))$ be the space of continuous sections of this bundle, thus if

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

$\tilde{A} \in \mathcal{L}^0(\varphi)$ then $\tilde{A}(x) \in \mathcal{L}(T_x X, T_{\varphi(x)} Y)$ at each $x \in \Lambda$. Note that we can naturally think of such an $\tilde{A}$ as a continuous, fibre preserving map from $T_\Lambda X$ to TY, linear on each fibre, and such that $\tilde{A}(T_x X) \subset T_{\varphi(x)} Y$; it will thus play the role of (φ, A) in the proof above.

Let $\mathcal{L}^0 = \mathcal{C}^0(\mathcal{L}(T_\Lambda X, \mathcal{C}^{0*} TY))$ be the union of the $\mathcal{L}^0(\varphi)$ over $\varphi \in \mathcal{C}^0$. This space has a natural vector bundle structure over $\mathcal{C}^0$, where the fibre over φ is just $\mathcal{L}^0(\varphi)$. Note that $\mathcal{C}^0$ is actually a smooth Banach manifold (e.g. see Eells 1966, Eliasson 1967), and $\mathcal{L}^0$ is a smooth bundle, but we do not need to make use of this here. The graph transform extends to $\mathcal{L}^0$ in a natural way by

$$\tilde{\mathbf{\Gamma}}(\tilde{A})(x) = T_w g.(T_x f^{-1}, \tilde{A}(f^{-1}(x)).T_x f^{-1})$$

with $w = (f^{-1}(x), \varphi(f^{-1}(x)))$.

Note that $\mathcal{L}^0(\Phi)$ has a linear structure with a norm given by

$$\|\tilde{A}\| = \sup_{x \in \Lambda} \|\tilde{A}(x)\|$$

where $\|\tilde{A}(x)\|$ is the operator norm on $\mathcal{L}(T_x X, T_{\varphi(x)} Y)$. Since $\tilde{\mathbf{\Gamma}}(\mathcal{L}^0(\Phi)) \subset \mathcal{L}^0(\Phi)$, exactly the same argument as in Lemma 4.1 shows that $\tilde{\mathbf{\Gamma}}$ has a fixed point $\tilde{A}^*$ in $\mathcal{L}^0(\Phi)$. In order to show that this is attracting in the whole of $\mathcal{L}^0$ we need to be able to compare $\tilde{A}$ and $\tilde{A}'$ for $\tilde{A} \in \mathcal{L}^0(\varphi)$, $\tilde{A}' \in \mathcal{L}^0(\varphi')$ with $\varphi \neq \varphi'$. We do this by constructing a trivialization of $\mathcal{L}^0$ in a neighbourhood of $\mathcal{L}^0(\Phi)$. Since Λ is compact, there exists an $s > 0$ such that for all $x \in \Lambda$, $\exp_{\Phi(x)}^{-1} : N_{\Phi(x)}(2s) \to T_{\Phi(X)} Y$ is an embedding, where $N_y(s) = \{y : d(y, y') \leq s\}$. Let

$$\mathcal{C}_s^0 = \{\varphi \in \mathcal{C}^0 : d(\varphi, \Phi) \leq s\}, \quad \mathcal{L}_s^0 = \{\tilde{A} \in \mathcal{L}^0(\varphi) : \varphi \in \mathcal{C}_s^0\}.$$

Then if $\tilde{A} \in \mathcal{L}_s^0$ we have $\tilde{A}(x).v \in TN_{\Phi(x)}(s)$ for all $v \in T_x X$, $x \in \Lambda$. Thus $(T \exp_{\Phi(x)}^{-1} \circ \tilde{A}(x)).v \in T(T_{\Phi(x)} Y)$, which we can naturally identify with $T_{\Phi(x)} Y \times T_{\Phi(x)} Y$. Under this identification the first component of $(T \exp_{\Phi(x)}^{-1} \circ \tilde{A}(x)).v$ is just $\exp_{\Phi(x)}^{-1}(\varphi(x))$, and the second can be written as $A(x).v$ where $A(x)$ is a linear map from $T_x X$ to $T_{\Phi(x)} Y$. Thus if we define $\bar{\varphi} = \exp_\Phi^{-1} \circ \varphi$ we have $T \exp_\Phi^{-1} \circ \tilde{A} = (\bar{\phi}, A) \in \mathcal{C}_s^0(\Phi^* TY) \times \mathcal{L}^0(\Phi)$, where $\mathcal{C}_s^0(\Phi^* TY)$ is the space of sections in $\Phi^* TY$ of norm less than s, which we can think of as continuous functions $\bar{\varphi} : \Lambda \to TY$ such that $\bar{\varphi}(x) \in T_{\Phi(x)} Y$, and $\|\bar{\varphi}\| = d(\varphi, \Phi) \leq s$. We thus have an identification between $\mathcal{L}_s^0$ and $\mathcal{C}_s^0(\Phi^* TY) \times \mathcal{L}^0(\Phi)$. We use this to give us a metric on $\mathcal{L}_s^0$ by

$$\begin{aligned} d(\tilde{A}, \tilde{A}') &= d((\bar{\varphi}, A), (\bar{\varphi}', A')) \\ &= \max\{\|\bar{\varphi} - \bar{\varphi}'\|, \|A - A'\|\}. \end{aligned}$$

Note that $\|\bar{\varphi} - \bar{\varphi}\|$ is an equivalent metric to $d(\varphi, \varphi')$, and the norm on $\|A - A'\|$ is just the operator norm on $\mathcal{L}^0(\Phi)$. By (2.1) we have $\Gamma(\mathcal{C}_s^0) \subset \mathcal{C}_s^0$ and hence $\tilde{\mathbf{\Gamma}}(\mathcal{L}_s^0) \subset \mathcal{L}_s^0$. The identification thus induces a transform on $\mathcal{C}_s^0(\Phi^* TY) \times \mathcal{L}^0(\Phi)$ (which we still denote by $\tilde{\mathbf{\Gamma}}$). The space $\mathcal{C}_s^0(\Phi^* TY) \times \mathcal{L}^0(\Phi)$ thus now plays the same role as $\mathcal{C}_s^0 \times \mathcal{C}^0(\mathcal{L}(T_\Lambda X, Y_0))$ in §4.2. In particular, if we write

$$\tilde{\mathbf{\Gamma}}(\bar{\varphi}, A) = (\bar{\mathbf{\Gamma}}(\bar{\varphi}), \bar{\mathbf{\Gamma}}_{\mathcal{L}}(\bar{\varphi}, A))$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

then we have

$$\bar{\mathbf{\Gamma}}(\bar{\varphi})(x) = h_x \circ g(f^{-1}(x), \varphi(f^{-1}(x)))$$
$$\bar{\mathbf{\Gamma}}_{\mathcal{L}}(\bar{\varphi}, A)(x).v = Th_x.T_w g.(u, Th^{-1}_{f^{-1}(x)}.(\varphi(f^{-1}(x)), A(f^{-1}(x).u))$$

with $h_x = \exp^{-1}_{\Phi(x)}$, $\varphi(f^{-1}(x)) = h^{-1}_{f^{-1}(x)}(\bar{\varphi}(f^{-1}(x)))$, $w = (f^{-1}(x), \varphi(f^{-1}(x)))$ and $u = T_x f^{-1}.v$. Then since g is uniformly $\mathcal{C}^{1+\alpha}$ on bounded sets there exists a constant $K'_s \geq 0$ such that

$$\|Th_{f(x)}.T_{(x,y)}g.(v, Th_x^{-1}(h_x(y), v')) - Th_{f(x)}.T_{(x,y')}g.(v, Th_x^{-1}(h_x(y), v'))\|$$
$$\leq K'_s d(y, y')^{\alpha}(\|v\| + \|v'\|)$$

for all $x \in \Lambda$, $v \in T_x X$, and all $y, y', v' \in T_{\Phi(x)}Y$ such that $\|y\| \leq s$ and $\|y'\| \leq s$. Then the same proof as Lemma 4.3 gives

$$\|\bar{\mathbf{\Gamma}}_{\mathcal{L}}(\bar{\varphi}, A) - \bar{\mathbf{\Gamma}}_{\mathcal{L}}(\bar{\varphi}', A)\| \leq K'_s e^{\mu}(1 + \|A\|)\|\bar{\varphi} - \bar{\varphi}'\|^{\alpha}$$

and the same argument as Corollary 4.4 shows that $\tilde{A}^*$ is attracting on the whole of $\mathcal{L}^0$.

It remains to discuss the modifications required to Lemma 4.6 and Propositon 4.7. The first of these is elementary and is left as an exercise. Turning to the latter, observe that $\bar{\mathbf{\Gamma}}(\bar{\varphi}) = \bar{g} \circ (Id, \bar{\varphi}) \circ f^{-1}$, where $\bar{g}$ is given by $\bar{g}(x, y) = h_{f(x)} \circ g(x, h_x^{-1}(y))$. This is in exact analogy to $\Gamma(\varphi) = g \circ (Id, \varphi) \circ f^{-1}$, and hence it is tempting to simply mimic the proof of Proposition 4.7 using $(\bar{\varphi}, A)$ and $\bar{g}$ instead of (φ, A) and g. The problem with this is the implicit dependence of $\bar{g}$ on x through h_x. Since *a priori* Φ is only Lipschitz, this means that $\bar{g}$ is not necessarily even differentiable with respect to x. Furthermore, this means that replacing (φ, A) and g in the definition of $\Gamma_{\mathcal{L}}(\varphi, A)$ by $(\bar{\varphi}, A)$ and $\bar{g}$ does not give $\bar{\mathbf{\Gamma}}_{\mathcal{L}}(\bar{\varphi}, A)$.

This indicates that the correct approach must be to use the same chart on Y for a whole neighbourhood of x. Thus, recall that we restrict ourselves to $\varphi \in \mathcal{C}^{\mathrm{Lip}}_{C_1}$ for some $C_1 \geq 0$. Assume that r is sufficiently small so that $C_1 r \leq s$. Then if $d(x, x') \leq r$, we have $d(\Phi(x), \varphi(x')) \leq d(\Phi(x), \Phi(x')) + d(\Phi(x'), \varphi(x')) \leq C_1 r + s \leq 2s$. Thus $\varphi(x') \in N_{\Phi(x)}(2s)$ and $h_x(\varphi(x'))$ is well defined, and we may define $\hat{\phi} = h_x \circ \varphi : N_x(r) \to N_{\Phi(x)}(2s)$. Similarly, suppose that $d(f(x), f(x')) \leq re^{-\mu}$, so that $d(x, x') \leq r$, and $y' \in Y$ is such that $d(y, \Phi(x)) \leq s$, then $d(g(x', y), \Phi(f(x))) \leq d(g(x', y), \Phi(f(x')))+d(\Phi(f(x')), \Phi(f(x'))) \leq e^{\lambda}s+C_1 re^{-\mu} \leq 2s$ (recall that $\mu \geq 0$). Thus $\hat{g} = h_{f(x)} \circ g \circ (Id, h_x^{-1}) : f^{-1}(N_{f(x)}(re^{-\mu})) \times N_{\Phi(x)}(s) \to N_{\Phi(f(x))}(2s)$ is well defined and as smooth as g is. Furthermore, locally the induced graph transform is given by

$$\tilde{\mathbf{\Gamma}}(\hat{\varphi}, \hat{A}) = (\hat{g} \circ (Id, \hat{\varphi}) \circ f^{-1}, D\hat{g}.(Id, \hat{A} \circ f^{-1}).Df^{-1})$$

where $(\hat{\varphi}, \hat{A}) = Th_x.\tilde{A}$. This is consistent with the action on (φ, A) and we may therefore imitate the proof of Proposition 4.7 using $(\hat{\varphi}, \hat{A})$ and $\hat{g}$. Of course, the metric induced on $N_{x'}(s)$ by h_x is not necessarily the same as that given by $h_{x'}$, and hence for instance $\|A(x')\| \neq \|\hat{A}(x')\|$. However, the two metrics are equivalent, and since Λ is compact and the dependence of h_x on x is continuous, the distortion in passing from $h_{x'}$ to h_x and *vice versa* is uniformly bounded on all Λ. Working in the local co-ordinates given by h_x is therefore justified.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

5. *Existence under non-uniform contraction*

In this section we give conditions for the average contraction of g to be well defined and then give a direct proof of Theorem 1.4.

5.1. *Existence of λ.*

LEMMA 5.1. *Suppose that there exists a $\xi \in L^1(X, \sigma)$ such that*

$$d(g_x(y), g_x(y')) \leq \xi(x) d(y, y')$$

for σ almost every x. Then for all $n \geq 0$,

$$\xi_n(x) = \sup_{y, y' \in Y} \frac{d(g_x^{(n)}(y), g_x^{(n)}(y'))}{d(y, y')}$$

exists for almost every x.

Proof. Define

$$\xi_n(x, y, y') = \frac{d(g_x^{(n)}(y), g_x^{(n)}(y'))}{d(y, y')}.$$

Then

$$\begin{aligned} \xi_n(x, y, y') &= \prod_{i=0}^{n-1} \xi_1(f^i(x), g_x^{(i)}(y), g_x^{(i)}(y')) \\ &\leq \prod_{i=0}^{n-1} \xi(f^i(x)) \end{aligned}$$

and hence $\xi_n(x, y, y')$ is bounded above for almost every x, as required. □

The following will be useful in §6.4.

LEMMA 5.2. *Each ξ_n is lower semi-continuous at all $x \in X$, and hence measurable.*

Proof. Given $\varepsilon > 0$, choose $y, y' \in Y$ such that $d(g_x^{(n)}(y), g_x^{(n)}(y')) \leq (\xi_n(x) - \varepsilon) d(y, y')$. Since $g_x^{(n)}(y)$ is continuous in x, there exists a $\delta > 0$ such that $d(g_{x'}^{(n)}(y), g_{x'}^{(n)}(y')) \leq (\xi_n(x) - 2\varepsilon) d(y, y')$, for all x' such that $d(x, x') \leq \delta$. Hence $\xi_n(x') \geq \xi_n(x) - 2\varepsilon$ for all x' such that $d(x, x') \leq \delta$, as required. □

LEMMA 5.3. *The sequence ξ_n is sub-multiplicative, i.e.*

$$\xi_{n+m}(x) \leq \xi_n(x) \xi_m(f^n(x))$$

for σ almost every x.

Proof. We have

$$\begin{aligned} \xi_{n+m}(x, y, y') &= \xi_n(x, y, y') \xi_m(f^n(x), g_x^{(n)}(y), g_x^{(n)}(y')) \\ &\leq \xi_n(x) \xi_m(f^n(x)) \end{aligned}$$

as required. □

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

It immediately follows that $\log \xi_n$ is sub-additive, that is $\log \xi_{n+m}(x) \leq \log \xi_n(x) + \log \xi_m(f^n(x))$. We can thus apply the sub-additive ergodic theorem (Kingman 1968) to $\log \xi_n$ to deduce that the limit

$$\lambda(x) = \lim_{n\to\infty} \frac{1}{n} \log \xi_{n(x)} \tag{5.1}$$

exists for σ almost every x and is f invariant. Note that the standard sub-additive ergodic theorem requires $\log \xi_n$ to be in $L^1(X, \sigma)$ which is not necessarily the case. However, Arnold 1998 gives an extended version for which it suffices that $(\log \xi_n)^+ = \max\{0, \log \xi_n\}$ is in $L^1(X, \sigma)$, which is the case here. The extension is straightforward and is achieved by applying the standard theorem to the sub-additive sequence $\max\{\log \xi_n(x), Kn\}$ and then allowing the constant K to become arbitrarily negative.

5.2. *Proof of existence of* Φ. We now want to give a direct proof of Theorem 1.4 (that is, not via Theorem 1.5). To do this we need to extend ξ_n to measure contraction backwards along an orbit. Since we do not necessarily assume that g_x is invertible for any x, we proceed as follows. Define

$$\zeta_n(x) = \xi_n(f^{-n}(x)).$$

Note that

$$\begin{aligned}\zeta_{n+m}(x) &= \xi_{n+m}(f^{-(n+m)}(x)) \\ &\leq \xi_n(f^{-(n+m)}(x))\xi_m(f^n(f^{-(n+m)}(x))) \\ &= \xi_n(f^{-n}(f^{-m}(x)))\xi_m(f^{-m}(x)) \\ &= \zeta_n(f^{-m}(x))\zeta_m(x).\end{aligned}$$

Hence ζ_n is sub-multiplicative with respect to f^{-1}. As above, since $(\log \zeta_n)^+ \in L^1(X, \sigma)$ the extended version of the sub-additive ergodic theorem (Arnold 1998) implies that the limit

$$\lambda^-(x) = \lim_{n\to\infty} \frac{1}{n} \log \zeta_n(x) \tag{5.2}$$

exists for σ almost every x and is f^{-1}-invariant. The crucial result is then as follows.

LEMMA 5.4. $\lambda^-(x) = \lambda(x)$ *for* σ *almost every* x.

Proof. The underlying idea used here is taken from Katok and Mendoza (1995). Given $\varepsilon > 0$, by (5.1) and (5.4) there exist measurable functions $c^+(x)$ and $c^-(x)$ such that

$$(\lambda(x) - \varepsilon)n - c^+(x) \leq \log \xi_n(x) \leq (\lambda(x) + \varepsilon)n + c^+(x)$$

and

$$(\lambda^-(x) - \varepsilon)n - c^-(x) \leq \log \zeta_n(x) \leq (\lambda^-(x) + \varepsilon)n + c^-(x).$$

Given $c \geq 0$, let $X_c = \{x : c^+(x) \leq c\}$. Then $\sigma(X_c) \to 1$ as $c \to \infty$ and for almost every $x \in X$ we have $x \in X_c$ for c sufficiently large. By the Poincaré recurrence theorem (e.g.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Katok and Hasselblatt 1995) for almost every $x \in X_c$ there exists a sequence $n_i \to \infty$ such that $f^{-n_i}(x) \in X_c$. Then $\log \zeta_{n_i}(x) = \log \xi_{n_i}(f^{-n_i}(x)) \leq (\lambda(f^{-n_i}(x)) + \varepsilon)n_i + c^+(f^{-n_i}(x)) \leq (\lambda(x) + \varepsilon)n + c$. Thus $(\lambda^-(x) - \varepsilon)n_i - c^-(x) \leq (\lambda(x) + \varepsilon)n_i + c$ and so

$$\lambda^-(x) \leq \lambda(x) + 2\varepsilon + \frac{1}{n_i}(c + c^-(x)).$$

Letting $n_i \to \infty$ we have $\lambda^-(x) \leq \lambda(x) + 2\varepsilon$ for almost all $x \in X_c$. But c and ε are arbitrary and hence $\lambda^-(x) \leq \lambda(x)$ for almost all $x \in X$. Similarly $(\lambda^-(x)+\varepsilon)n_i+c^-(x) \geq \log \zeta_{n_i}(x) \geq (\lambda(f^{-n_i}(x)) - \varepsilon)n_i - c^+(f^{-n_i}(x)) \geq (\lambda(x) - \varepsilon)n - c$ so that

$$\lambda^-(x) \geq \lambda(x) - 2\varepsilon - \frac{1}{n_i}(c + c^-(x))$$

and hence $\lambda^-(x) \geq \lambda(x) - 2\varepsilon$ for almost all $x \in X_c$. Thus $\lambda^-(x) \geq \lambda(x)$ for almost all $x \in X$, and therefore $\lambda^-(x) = \lambda(x)$ for almost all $x \in X$ as required. □

As a corollary of this, we obtain the following.

Proof of Theorem 1.4. Choose $0 < \varepsilon < -\lambda$ and let Λ' be a set of full measure such that for all $x \in \Lambda'$ we have $\lambda^-(x) < \lambda$ and such that there exists a measurable function $c : \Lambda' \to \mathbb{R}^+$ with $\log \zeta_n(x) \leq (\lambda + \varepsilon)n + \log c(x)$ for all $x \in \Lambda'$. Let

$$\Lambda = \bigcap_{n \in \mathbb{Z}} f^n(\Lambda').$$

Then Λ is f-invariant, $\sigma(\Lambda) = 1$ and

$$d(g^{(n)}(f^{-n}(x), y), g^{(n)}(f^{-n}(x), y')) \leq c(x)e^{(\lambda+\varepsilon)n}d(y, y')$$

for all $x \in \Lambda$ and all $y, y' \in Y$, where $g^{(n)}(x, y) = g_x^{(n)}(y)$. Fix some $y_0 \in Y$ and let

$$y_n = g^{(n)}(f^{-n}(x), y_0).$$

Then

$$\begin{aligned} d(y_n, y_{n+1}) &= d(g^{(n)}(f^{-n}(x), y_0), g^{(n)}(f^{-n}(x), y_1)) \\ &\leq c(x)e^{(\lambda+\varepsilon)n}d(y_0, y_1). \end{aligned}$$

Hence

$$\begin{aligned} d(y_n, y_{n+m}) &\leq \sum_{i=n}^{n+m-1} d(y_i, y_{i+1}) \\ &\leq \sum_{i=n}^{n+m-1} c(x)e^{(\lambda+\varepsilon)i}d(y_0, y_1) \\ &\leq e^{(\lambda+\varepsilon)n}\frac{1}{1 - e^{\lambda+\varepsilon}}c(x)d(y_0, y_1). \end{aligned}$$

Hence $\{y_n\}$ is a Cauchy sequence and converges to a limit. Call this limit $\Phi(x)$. It is easy to see that the graph of $\Phi(x)$ is invariant under (f, g). Thus

$$\begin{aligned} \Phi(f(x)) &= \lim_{n\to\infty} g^{(n)}(f^{-n}(f(x)), y_0) \\ &= \lim_{n\to\infty} g(f^{n-1}(f^{-n}(f(x))), g^{(n-1)}(f^{-n}(f(x)), y_0)) \\ &= \lim_{n\to\infty} g(x, g^{n-1}(f^{-(n-1)}(x), y_0)). \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Thus, since g is continuous

$$\begin{aligned}\Phi(f(x)) &= g\Big(x, \lim_{n\to\infty} g^{(n-1)}(f^{-(n-1)}(x), y_0)\Big)\\ &= g(x, \Phi(x))\end{aligned}$$

as required. By induction we have $\Phi(f^n(x)) = g^{(n)}(x, \Phi(x))$ and hence

$$\begin{aligned}d(\Phi(f^n(x)), g^{(n)}(x, y)) &= d(g^{(n)}(x, \Phi(x)), g^{(n)}(x, y))\\ &\le c^+(x)e^{(\lambda+\varepsilon)n}d(\Phi(x), y).\end{aligned}$$

Thus $d(\Phi(f^n(x)), g^{(n)}(x, y)) \to 0$ as $n \to \infty$, and the graph of Φ is attracting. This completes the proof of Theorem 1.4. □

6. *Adapted metrics*

The principal tool which allows us to extend the results of §3 and §4 to non-uniform convergence rates is the construction of adapted metrics and norms. In the case of uniform contraction this is due to Mather (1968), whilst the non-uniform case was independently developed by Pesin (1976, 1977). A helpful account of this can be found in Katok and Mendoza (1995) on which we base much of our exposition. The reader may also find it helpful to consult Fathi *et al* (1983). Unfortunately, whilst this standard approach is perfectly adequate in our context for the construction of an adapted metric for f, there seems to be no satisfactory way of generalizing one crucial step to obtain analogous results for g. We are able to overcome this difficulty using a lemma due to Arnold (1998). This applies equally well to the construction of f and so we give a unified account for both f and g.

6.1. *Construction of the adapted norm.* The basic idea of adapting a norm to the dynamics is simple, and most easily explained in the context of the driving map f. Since in fact we shall only ever require an adapted norm for f^{-1}, we shall present the construction in terms of this, rather than f itself. Thus assume that f is a $\mathcal{C}^{1+\alpha}$ diffeomorphism, is a $\mathcal{C}^\infty$ compact Riemannian manifold X, and that σ is a f-invariant measure (not necessarily ergodic). As already described in the introduction, the largest Liapunov exponent of f^{-1} with respect to σ is given by

$$\mu(x) = \lim_{n\to\infty} \frac{1}{n} \log \|D_x f^{-n}\|. \tag{6.1}$$

Since f is a diffeomorphism and X is compact, $\mu(x)$ is bounded, and so we set μ to be the essential supremum of $\mu(x)$. Thus there exists an invariant set Λ of full measure such that $\mu(x) \le \mu$ for all $x \in \Lambda$. By analogy to §4, we expect the graph transform to contract derivatives by on average $e^{(\lambda+\mu)}$ per time step. In order to establish the smoothness of the invariant graph we require the same estimate to hold every time step rather than just on average. In the case of f^{-1}, this means that we need an estimate like $\|D_x f^{-1}\| \le e^{\mu}$ for all $x \in \Lambda$. Unfortunately, all that we can deduce from (6.1) is that for any $\eta > 0$ there exists a measurable function $b' : \Lambda \to \mathbb{R}^+$ such that

$$\|D_x f^{-n}\| \le b'(x)e^{(\mu+\eta)n}.$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

If the contraction on Λ is not uniform, then b' will be unbounded and we cannot proceed as we did in §§2–4 by replacing f by f^m for a sufficiently large m. To overcome this difficulty, we modify the Riemannian metric at each point so that in the new metric we obtain $\|D_x f^{-1}\| \leq e^{\mu+2\eta}$. Since η can be made as small as we require, the additional factor $e^{2\eta}$ will cause us no difficulties. Such a metric may be obtained by defining an adapted norm $\|v\|_\eta$ on $T_x X$ by

$$\|v\|_\eta = \sum_{n=0}^{\infty} e^{-(\mu+2\eta)n} \|D_x f^{-n}.v\|.$$

We obviously have

$$\|v\| \leq \|v\|_\eta \leq b(x)\|v\|$$

where $b(x) = (1 - e^{-\eta})^{-1} b'(x)$. Thus $\|\cdot\|_\eta$ is well defined and equivalent to $\|\cdot\|$ at each point in Λ (though not uniformly so if b is not bounded). The following simple lemma, applied to $Z = T_\Lambda X$, $T = Df^{-1}$ and $b(v) = \|v\|$ directly yields the required estimate $a(D_x f^{-1}(v)) = \|D_x f^{-1}\|_\eta \leq e^{\mu+2\eta}$.

LEMMA 6.1. *Let Z be an arbitrary space and $T : Z \to Z$ some transformation. Suppose that $b : X \to \mathbb{R}^+$ is a function which satisfies*

$$\limsup_{n\to\infty} \frac{1}{n} \log b(T^n(z)) \leq \mu \tag{6.2}$$

for all $z \in Z$. Then given any $\eta > 0$ define

$$a(z) = \sum_{n=0}^{\infty} e^{-(\mu+2\eta)n} b(T^n(z)).$$

This converges for all $z \in Z$ and satisfies

$$a(T(z)) \leq e^{\mu+2\eta} a(z).$$

Proof. By (6.2) there exists a function b' such that $b(T^n(z)) \leq b'(z)e^{(\mu+\eta)n}$. We then compute directly

$$\begin{aligned} a(T(z)) &= \sum_{n=0}^{\infty} e^{-(\mu+2\eta)n} b(T^{n+1}(z)) \\ &= e^{\mu+\eta} \sum_{n=1}^{\infty} e^{-(\mu+\eta)n} b(T^n(z)) \\ &= e^{\mu+\eta}(a(z) - b(z)). \end{aligned}$$

By definition $b(z) > 0$ and hence $a(z) - b(z) < a(z)$, giving the required result. □

We may then define a sequence of Pesin sets by $\Lambda_k = \{x \in \Lambda : b(x) \leq e^{\eta k}\}$. Since b can be chosen to be lower semi-continuous (see §6.3), the Λ_k can be taken to be closed and hence compact (see Lemma 6.3). The adapted norm is uniformly equivalent to the standard metric on each Λ_k and can be extended to a neighbourhood of Λ_k in X using the exponential mapping (see §6.3). If we restrict ourselves to a particular Λ_k we can thus carry out most of the same computations as in §§2–4.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

6.2. *Growth of bounds along orbits.* The one crucial remaining problem is that the Λ_k are not f invariant. Thus we cannot define a graph transformation just for functions over Λ_k. In particular, to compute $\Gamma(\varphi)$ over Λ_k, we need to know φ over $f^{-1}(\Lambda_k)$. We therefore require to control the relationship between Λ_k and $f^{-1}(\Lambda_k)$. This requires knowledge of the growth of b along orbits. The standard approach to this, due to Pesin (1976, 1977), uses the invertibility of f (or more precisely of Df). There seems to be no satisfactory way of using this technique to obtain the required bounds for the contraction by g. Instead, therefore, we use the following fundamental lemma, due to Arnold (1998) (see also Tanny (1974)).

LEMMA 6.2. *Let $T : Z \to Z$ be a measurable transformation and σ a T-invariant measure. Suppose that $b : Z \to \mathbb{R}^+$ is a measurable function such that $b(z) > 0$ for σ almost every z. Suppose that $(\log b \circ T - \log b)^+ \in L^1(Z, \sigma)$, where $(\log b \circ T - \log b)^+ = \max\{0, \log b \circ T - \log b\}$. Then for σ almost every z*

$$\lim_{n\to\infty} \frac{1}{n} \log b(T^n(z)) = 0.$$

Proof. Let $\theta = \log b \circ T - \log b$. Since $\theta^+ \in L^1(Z, \sigma)$, the extended version of the Birkhoff ergodic theorem (Arnold 1998) implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \theta(T^i(z)) \to \bar{\theta}(z), \quad \text{for } \sigma \text{ a.e. } z$$

where $\bar{\theta} : Z \to \mathbb{R} \cup \{-\infty\}$. However,

$$\sum_{i=0}^{n-1} \theta(T^i(z)) = \log b(T^n(z)) - \log b(z)$$

and hence

$$\lim_{n\to\infty} \frac{1}{n} \log b(T^n(z))$$

exists for σ almost every z.

For a given $m \in \mathbb{N}$, let $U_m = \{z : \log b(z) \le m\}$. Then the union of the U_m has full measure, and for all sufficiently large m we have $\sigma(U_m) > 0$. Fix an m so that this is the case. By the Poincaré recurrence theorem (e.g. Katok and Hasselblatt 1995), almost every $z \in U_m$ returns to U_m infinitely often. Thus there exists a sequence n_i such that $T^{n_i}(z) \in U_m$. Hence $\log b(T^{n_i}(z)) \le m$ and so

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log b(T^n(z)) &= \lim_{n\to\infty} \frac{1}{n} \log b(T^{n_i}(z)) \\ &\le \lim_{n\to\infty} \frac{1}{n} m \\ &= 0. \end{aligned}$$

The same argument applied to $V_m = \{z : \log b(z) \ge -m\}$ shows that

$$\lim_{n\to\infty} \frac{1}{n} \log b(T^n(z)) \ge 0$$

for almost every $z \in V_m$. Since almost every $z \in Z$ lies in all U_m and V_m for m large enough and m can be chosen arbitrarily, we deduce that (6.2) holds for σ almost every $z \in Z$, as required. □

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

In fact, under the conditions of the lemma $\log b$ must be in $L^1(Z,\sigma)$ (Tanny 1974). Also, we can replace the condition $\max\{0, \log b \circ T - \log b\} \in L^1(Z,\sigma)$ by $\min\{0, \log b \circ T - \log b\} \in L^1(Z,\sigma)$.

We can now apply Lemma 6.1 (with $\eta/2$ rather than η) to show that we may replace b by a function which satisfies $a(f^{-1}(x)) \le e^{\eta} a(x)$. Thus define

$$a(x) = \sum_{n=0}^{\infty} e^{-\eta n} b(f^{-n}(x)). \tag{6.3}$$

If we construct the Pesin sets using a instead of b, that is $\Lambda_k = \{x \in \Lambda : a(x) \le e^{\eta k}\}$, we immediately obtain $f^{-1}(\Lambda_k) \subset \Lambda_{k+1}$, as required. Note that $a(x) \le b(x)$ and hence we still have $\|v\|_\eta \le e^{\eta k}\|v\|$ on each Λ_k. It remains to show that the Λ_k are closed. This follows from the following lemma, which uses an argument due to Pesin (1977).

LEMMA 6.3. *Suppose that b is lower semi-continuous and satisfies*

$$\lim_{n\to\infty} \frac{1}{n} \log b(f^{-n}(x)) = 0.$$

Define a by (6.3). Then for any $c > 0$ the set $\Lambda_c = \{x \in \Lambda : a(x) \le c\}$ is closed.

Proof. Suppose that we have a sequence $x_i \to x$ with $x_i \in \Lambda_c$. Since b is lower semi-continuous, for any fixed n we have

$$b(f^{-n}(x)) \le \liminf_{i\to\infty} b(f^{-n}(x_i)).$$

Hence for any n_0

$$\begin{aligned}\sum_{n=0}^{n_0-1} e^{-\eta n} b(f^{-n}(x)) &\le \liminf_{i\to\infty} \sum_{n=0}^{n_0-1} e^{-\eta n} b(f^{-n}(x_i)) \\ &\le \liminf_{i\to\infty} a(x_i) \\ &\le c.\end{aligned}$$

Choose a measurable function $b' : \Lambda \to \mathbb{R}^+$ such that $b(f^{-n}(x)) \le b'(x) e^{\eta n/2}$, and given $\varepsilon > 0$, choose n_0 such that $e^{-\eta n_0/2}(1 - e^{-\eta/2})^{-1} b'(x) \le \varepsilon$. Then

$$\begin{aligned}a(x) &= \sum_{n=0}^{n_0-1} e^{-\eta n} b(f^{-n}(x)) + \sum_{n=n_0}^{\infty} e^{-\eta n} b(f^{-n}(x)) \\ &\le c + \sum_{n=n_0}^{\infty} e^{-\eta n} b'(x) e^{\eta n/2} \\ &\le c + \varepsilon.\end{aligned}$$

This holds for any $\varepsilon > 0$, and hence $x \in \Lambda_c$, as required. □

6.3. *Regular neighbourhoods for f.* In order to complete the construction of the adapted norm for f^{-1} it remains to show that the hypotheses of Lemmas 6.1 and 6.3 are satisfied, and in particular that if we define $b : \Lambda \to \mathbb{R}^+$ by

$$b(x) = \sup_n \{\|D_x f^{-n}\| e^{-(\mu+\eta)n}\}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

then $\max\{0, \log b \circ f^{-1} - \log b\} \in L^1(X, \sigma)$ and b is lower semi-continuous.

To show the former, given $x \in \Lambda$, and an $\varepsilon > 0$, choose n_0 such that $\|D_x f^{-n_0}\| e^{-(\mu+\eta)n_0} \geq b(x) - \varepsilon$. Now, $D_{f(x)} f^{-n_0-1}.D_x f = D_x f^{-n_0}$ and hence $\|D_{f(x)} f^{-n_0-1}\| \, \|D_x f\| \geq \|D_x f^{-n_0}\|$. Thus $\|D_{f(x)} f^{-n_0-1}\| \geq C^{-1} \|D_x f^{-n_0}\|$ where C is the infimum of $\|Df\|$ on the whole of X. This gives $b(f(x)) \geq \|D_{f(x)} f^{-n_0-1}\| e^{-(\mu+\eta)(n_0+1)} \geq C^{-1} \|D_x f^{-n_0}\| e^{-(\mu+\eta)(n_0+1)} \geq C^{-1}(b(x) - \varepsilon) e^{-(\mu+\eta)}$. Since ε was arbitrary, we get $b(f(x)) \geq C^{-1} b(x) e^{-(\mu+\eta)}$. Applying this at $f^{-1}(x)$ gives $\log b(f^{-1}(x)) - \log b(x) \geq \log C + (\mu + \eta)$. Thus in fact $\log b \circ f^{-1} - \log b$ is bounded above, and $\max\{0, \log b - \log b \circ f^{-1}\} \in L^1(X, \sigma)$ as claimed.

To show lower semi-continuity, let $b_n(x) = \|D_x f^{-n}\| e^{-(\mu+\eta)n}$. Each b_n is continuous and $b_n(x) \leq b(x)$ for all n. Suppose we have a sequence $x_i \to x$ with $b(x_i) \leq K$ for all i. Then $b_n(x_i) \leq K$ for all i and all n. Fix n, then $b_n(x_i) \to b_n(x)$ and hence $b_n(x) \leq K$. This holds for all n, and hence $b(x) \leq K$, as required.

The final step is to use the exponential mapping to give an adapted metric on a neighbourhood of each Λ_k in X. This is a standard argument due to Pesin (1977) given here for the convenience of the reader. We largely follow the exposition of Katok and Mendoza (1995).

Let $Q(x) : T_x X \to \mathbb{R}^m$ be the linear map whose pullback takes the standard Euclidean norm on $\mathbb{R}^m$ to the adapted norm on $T_x X$, so that $\|v\|_\eta = \|Q_x.v\|$ for all $v \in T_x X$. Note that $\|Q_x\|$ is measurable and $\|Q_x\| \leq a(x)$. Since $\|v\|_\eta \geq \|v\|$ we have $\|Q_x\| \geq 1$. Define $\Psi_x : B_{r(x)}(\mathbf{0}) \to X$ by $\Psi_x = \exp_x \circ Q_x^{-1}$ where $\exp_x : T_x X \to X$ is the exponential mapping in the standard Riemann metric and $B_{r(x)}(\mathbf{0}) \subset \mathbb{R}^m$ is a ball of radius $r(x)$ about the origin in $\mathbb{R}^m$. Let $U_{x,r}$ be a ball of radius r about the origin in $T_x X$. Since X is compact, there exists a constant r such that $\exp_x$ is an embedding on $U_{x,r}$ for all $x \in X$. Then since $\|Q_x^{-1}.v\| \leq \|v\|$, we see that if we choose $r(x) \leq r$ then Ψ_x is a $\mathcal{C}^\infty$ embedding of $B_{r(X)}(\mathbf{0})$ into X.

Let $N_x = \Psi_x(B_{r(x)}(\mathbf{0}))$. Then Ψ_x^{-1} defines an adapted metric on N_x by

$$d_{\eta,x}(z, z') = \|\Psi_x^{-1}(z) - \Psi_x^{-1}(z')\|$$

for all $z, z' \in N_x$. Note that by definition $\Psi_x^{-1}(x) = \mathbf{0}$, and hence $d_{\eta,x}(x, x') = \|\Psi_x^{-1}(x')\|$. Also since $\exp_x$ is an isometry we have

$$d(z, z') \leq d_{\eta,x}(z, z') \leq b(x) d(z, z')$$

for all $z, z' \in N_x$, where d is the standard Riemannian metric on X. If $r(x)$ is sufficiently small, then $f^{-1}(N_x) \subset \exp f^{-1}(x)(U f^{-1}(x), r)$ and hence

$$f_x^{-1} = [\Psi f^{-1}(x)]^{-1} \circ f^{-1} \circ \Psi_x$$

is well defined as a map $f_x^{-1} : B_{r(x)}(0) \to \mathbb{R}^m$. If f^{-1} is $\mathcal{C}^{1+\alpha}$ and $r(x)$ is sufficiently small (in terms of $1/a(x)$ and the Hölder norm of f^{-1}) then

$$\|D_u f_x^{-1}\| \leq e^{\mu+3\eta}$$

for all $u \in B_{r(x)}(\mathbf{0})$ and hence

$$\|D_z f^{-1}\|_\eta \leq e^{\mu+3\eta} \tag{6.4}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

for all $z \in N_x$ and

$$d_{\eta, f^{-1}(x)}(f^{-1}(z), f^{-1}(z')) \leq e^{\mu+3\eta} d_{\eta,x}(z, z') \tag{6.5}$$

for all $z, z' \in N_x$. The function $r(x)$ is bounded by $c[a(x)]^{-1/\alpha}$ for some $c \geq 0$ (where α is the Hölder norm of Df^{-1}). Thus by using Lemma 6.1 with an appropriate choice of η we may assume that $r(x)$ satisfies $e^{-\eta} r(x) \leq r(f^{-1}(x))$. We may then define a sequence of Pesin sets Λ_k by

$$\Lambda_k = \{x \in \Lambda : r(x) \geq e^{-\eta k}, a(x) \leq e^{\eta k}\}. \tag{6.6}$$

Then $\Lambda_k \subset \Lambda_{k+1}$, $f^{-1}(\Lambda_k) \subset \Lambda_{k+1}$ and

$$\Lambda = \bigcup_{k \in \mathbb{N}} \Lambda_k$$

as required.

6.4. *The adapted metric on Y.* The construction of an adapted metric on Y is similar, although since we have no invariant measure on Y and g_x is not necessarily invertible we cannot directly deduce the required properties from standard results. Instead we apply Lemmas 6.1 and 6.2 above. Recall that

$$\xi_n(x) = \sup_{y, y' \in Y} \frac{d(g_x^{(n)}(y), g_x^{(n)}(y'))}{d(y, y')}$$

and given $\eta > 0$ define

$$c(x) = \sup_n \{\xi_n(x) e^{-(\lambda+\eta)n}\}.$$

Recall that for each n, ξ_n is lower semi-continuous at all $x \in X$ by Lemma 5.2 and hence so is $c_n = e^{-(\lambda+\eta)n} \xi_n$. Thus, as in the previous section c is also lower semi-continuous: if we have a sequence $x_i \to x$ with $c(x_i) \leq K$ for all i then $c_n(x_i) \leq K$ for all i and n and hence $c_n(x) \leq K$ for all n.

We now want to show that

$$\lim_{n \to \infty} \frac{1}{n} \log c(f^{-n}(x)) \leq 0$$

for almost all $x \in \Lambda$. The one slight difficulty is that we do not necessarily have $c(x) > 0$ for almost all $x \in \Lambda$, as is required by Lemma 6.2. However, as we shall see if $c(x) = 0$ then $c(f^{-1}(x)) = 0$, which allows us to apply Lemma 6.2 to the invariant set $\Omega = \{x \in \Lambda : c(f^{-n}(x)) > 0 \text{ for all } n \geq 0\}$. Thus, given $x \in \Lambda$, and $\varepsilon > 0$, choose n_0 such that $c_{n_0}(x) \geq c(x) - \varepsilon$. By Lemma 5.3, ξ_n is sub-multiplicative and hence $\xi_{n_0}(x) \leq \xi_1(x) \xi_{n_0-1}(f(x)) \leq \xi_{n_0}(f^{-1}(x)) \xi_1(x)$. Thus $c(x) - \varepsilon \leq c_{n_0}(x) \leq \xi_1(x) \xi_{n_0-1}(f(x)) e^{-(\lambda+\eta)n_0} \leq \xi_1(x) c_{n_0-1}(f(x)) e^{-(\lambda+\eta)} \leq \xi_1(x) e^{-(\lambda+\eta)} c(f(x))$. Since ε is arbitrary, we get $c(x) \leq \xi_1(x) e^{-(\lambda+\eta)} c(f(x))$. Evaluating this at $f^{-1}(x)$ we get $c(f^{-1}(x)) \leq \xi_1(f^{-1}(x)) e^{-(\lambda+\eta)} c(x)$. Hence as claimed if $c(x) = 0$, then $c(f^{-1}(x)) = 0$. This implies that if $c(f^{-n}(x)) = 0$ for all $m \geq n$ and hence if $x \in \Lambda \backslash \Omega$ we have $c(f^{-n}(x)) = 0$ for all sufficiently large n. Thus

$$\lim_{n \to \infty} \frac{1}{n} \log c(f^{-n}(x)) = -\infty$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

for all $x \in \Lambda\backslash\Omega$. If $\sigma(\Lambda\backslash\Omega) = 1$, then we are done. Otherwise, observe that $f^{-1}(\Omega) \subset \Omega$, and $\log c(f^{-1}(x)) - \log c(x) \leq \log \xi_1(f^{-1}(x)) - (\lambda + \eta)$ on Ω. Since we are assuming (recall Lemma 5.1) that $\xi_1 \in L^1(X, \sigma)$ we see that $(\log c \circ f^{-1} - \log c)^+ \in L^1(\Omega, \sigma)$. Hence by Lemma 6.2 we have

$$\lim_{n\to\infty} \frac{1}{n} \log c(f^{-n}(x)) = 0$$

for almost every $x \in \Omega$, as required. We now proceed as §6.1. Define

$$a'(x) = \sum_{n=0}^{\infty} e^{-\eta n} c(f^{-n}(x)).$$

By Lemma 6.1 we have $a'(f^{-1}(x)) \leq e^{\eta} a'(x)$ for all x. We may thus define the Pesin sets for g by

$$\Lambda_k = \{x : a'(x) \leq (1 - e^{-\eta}) e^{\eta k}\}. \tag{6.7}$$

Then $\Lambda_k \subset \Lambda_{k+1}$ and $f^{-1}(\Lambda_k) \subset \Lambda_{k+1}$. Finally by Lemma 6.3, each Λ_k is closed.

We are now in a position to define the adapted metric on Y. Note that $\xi_n(x) e^{-(\lambda+\eta)n} \leq c(x) \leq a'(x)$ and hence $d(g_x^{(n)}(y), g_x^{(n)}(y')) \leq \xi_n(x) d(y, y') \leq a'(x) e^{(\lambda+\eta)n}$. Thus

$$d_\eta(y, y') = \sum_{n=0}^{\infty} e^{-(\lambda+2\eta)n} d(g_x^{(n)}(y), g_x^{(n)}(y'))$$

is well defined, and $d_\eta(y, y') \leq (1 - e^{-\eta})^{-1} a'(x) d(y, y')$. Since trivially $d(y, y') \leq d_\eta(y, y')$, we see that $d_\eta(y, y')$ is equivalent to $d_\eta(y, y')$ on each Λ_k and

$$d(y, y') \leq d_\eta(y, y') \leq e^{\eta k} d(y, y')$$

on Λ_k. Note that d_η depends on x and hence we shall often write $d_{\eta,x}$ to make this explicit. It will turn out to be convenient to express the adapted nature of d_η in terms of the distance between $g(f^{-1}(x), y)$ and $g(f^{-1}(x), y')$. Since $g_x^{(n)}(g(f^{-1}(x), y)) = g_{f^{-1}(x)}^{(n+1)}(y)$, essentially the same calculation as Lemma 6.1 gives

$$\begin{aligned}
&d_{\eta,x}(g(f^{-1}(x), y), g(f^{-1}(x), y')) \\
&\quad = \sum_{n=0}^{\infty} e^{-(\lambda+2\eta)n} d(g_{f^{-1}(x)}^{(n+1)}(y), g_{f^{-1}(x)}^{(n+1)}(y')) \\
&\quad = e^{\lambda+2\eta} \sum_{n=1}^{\infty} e^{-(\lambda+2\eta)n} d(g_{f^{-1}(x)}^{(n)}(y), g_{f^{-1}(x)}^{(n)}(y')) \\
&\quad = e^{\lambda+2\eta}(d_{n,f^{-1}(x)}(y, y') - d(y, y')) \\
&\quad \leq e^{\lambda+2\eta} d_{n,f^{-1}(x)}(y, y').
\end{aligned}$$

If g is smooth, we can also define a norm $\| \ \|_\eta$ on $T_y Y$ by

$$\|v\|_\eta = \sum_{n=0}^{\infty} e^{-(\lambda+2\eta)n} \|D_y g_x^{(n)}.v\|.$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Then just as above, we have

$$\|v\| \leq \|v\|_\eta \leq e^{\eta k}\|v\|$$

for $x \in \Lambda_k$ and

$$\|D_y g_x\|_\eta \leq e^{\lambda+2\eta}.$$

Note that $\|\ \|_\eta$ depends on y as well as x, and hence we have to be careful in identifying Y with $T_y Y$ when Y is a Banach space. Unfortunately, it does not seem possible to define an adapted norm which is independent of y.

7. *Continuity under non-uniform contraction*

In this section we prove Theorem 1.5 using the same strategy as in §2. Although the existence of Φ is given by Theorem 1.4, this turns out not to be very useful, and it is much simpler to reprove the existence as a by-product of the proof of the bound $d(\Phi(x), y_0) \leq ce^{\eta k}$. Note that the Pesin sets Λ_k in this section are defined purely in terms of g, i.e. they satisfy (6.6).

7.1. *Existence and boundedness of the invariant graph.* Let $\mathcal{C} = \{\varphi : \Lambda \to Y\}$ be the space of functions from Λ to Y as in §2 define the graph transform $\Gamma : \mathcal{C} \to \mathcal{C}$ by

$$\Gamma(\varphi)(x) = g(f^{-1}(x), \varphi(f^{-1}(x))).$$

Let $\mathcal{B}_\eta \subset \mathcal{C} \times \mathcal{C}$ be defined by

$$\mathcal{B}_\eta = \{(\varphi, \varphi') \in \mathcal{C} \times \mathcal{C} : d_\eta(\varphi(x), \varphi'(x)) \leq ce^{\eta k} \text{ for some } c \geq 0, \text{ for all } x \in \Lambda_k, \text{ for all } k\}.$$

For $(\varphi, \varphi') \in \mathcal{B}_\eta$ let

$$d_\eta(\varphi, \varphi') = \inf\{c : d_\eta(\varphi(x), \varphi'(x)) \leq ce^{\eta k} \text{ for all } x \in \Lambda_k, \text{ for all } k\}.$$

Let y_0 be a point $y_0 \in Y$ and $\varphi_0 \in \mathcal{C}$ the constant function $\varphi_0(x) = y_0$. Define $\mathcal{C}_\eta \subset \mathcal{C}$ by

$$\mathcal{C}_\eta = \{\varphi \in \mathcal{C} : (\varphi, \varphi_0) \in \mathcal{B}_\eta\}.$$

Note that this is independent of the choice of y_0 and d_η gives $\mathcal{C}_\eta$ the structure of a complete metric space. This is our analogue of the space of bounded functions. Also, if $\varphi \in \mathcal{C}_\eta$ and $(\varphi, \varphi') \in \mathcal{B}_\eta$, then $\varphi' \in \mathcal{C}_\eta$.

LEMMA 7.1. *Suppose* $(\varphi, \varphi') \in \mathcal{B}_\eta$. *Then* $(\Gamma(\varphi), \Gamma(\varphi')) \in \mathcal{B}_\eta$ *and*

$$d_\eta(\Gamma(\varphi), \Gamma(\varphi')) \leq e^{(\lambda+3\eta)} d_\eta(\varphi, \varphi').$$

Proof. Given $x \in \Lambda_k$ we have

$$\begin{aligned} d_\eta(\Gamma(\varphi)(x), \Gamma(\varphi')(x)) &= d_\eta(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi'(f^{-1}(x)))) \\ &\leq e^{(\lambda+2\eta)} d_\eta(\varphi(f^{-1}(x)), \varphi'(f^{-1}(x))). \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Since $f^{-1}(x) \in \Lambda_{k+1}$ we have

$$d_\eta(\varphi(f^{-1}(x)), \varphi'(f^{-1}(x))) \leq d_\eta(\varphi, \varphi')e^{\eta(k+1)}$$

and hence

$$d_\eta(\Gamma(\varphi)(x), \Gamma(\varphi')(x)) \leq e^{(\lambda+3\eta)}d_\eta(\varphi, \varphi')_\eta e^{\eta k}$$

as required. □

COROLLARY 7.2. $\Gamma(\mathcal{C}_\eta) \subset \mathcal{C}_\eta$.

Proof. Let $\varphi_0(x) = y_0$. Then $\varphi_1(x) = \Gamma(\varphi_0)(x) = g(f^{-1}(x), y_0)$. Since g and f^{-1} are continuous and X is compact, φ_1 is bounded on X; say $d(\varphi_1(x), y_0) \leq c$ for all $x \in X$. Thus $d_\eta(\varphi_1(x), y_0) \leq ce^{\eta k}$ for all $x \in \Lambda_k$, and hence $\varphi_1 \in \mathcal{C}_\eta$. Now suppose $\varphi \in \mathcal{C}_\eta$, then $(\varphi, \varphi_0) \in \mathcal{B}_\eta$ and hence $(\Gamma(\varphi), \varphi_1) \in \mathcal{B}_\eta$, and hence $\Gamma(\varphi) \in \mathcal{C}_\eta$, as required. □

COROLLARY 7.3. *If $0 < 3\eta < -\lambda$ then Γ has a unique fixed point in $\mathcal{C}_\eta$. We shall denote this by Φ.*

For a fixed y_0, denote

$$\mathcal{C}_{\eta,C} \in \{\varphi \in \mathcal{C}_\eta : d_\eta(\varphi, \varphi_0) \leq C\}.$$

COROLLARY 7.4. *If $C_0 \geq (1 - e^{\lambda+3\eta})^{-1}d_\eta(\varphi_1, \varphi_0)$, where $\varphi_1 = \Gamma(\varphi_0)$ as above, then $\Gamma(\mathcal{C}_{\eta,C_0}) \subset \mathcal{C}_{\eta,C_0}$, and hence in particular $d_\eta(\Phi, \varphi_0) \leq (1 - e^{\lambda+3\eta})^{-1}d_\eta(\varphi_1, \varphi_0)$.*

Proof. Suppose $\varphi \in \mathcal{C}_{\eta,C_0}$. Then

$$\begin{aligned} d_\eta(\Gamma(\varphi), \varphi_0) &\leq d_\eta(\Gamma(\varphi), \Gamma(\varphi_0)) + d_\eta(\Gamma(\varphi_0), \varphi_0) \\ &\leq e^{(\lambda+3\eta)}d_\eta(\varphi, \varphi_0) + d_\eta(\varphi_1, \varphi_0) \\ &\leq e^{(\lambda+3\eta)}C_0 + C_0(1 - e^{(\lambda+3\eta)}) \\ &\leq C_0. \end{aligned}$$

□

7.2. *Continuity.* The continuity of Φ is a more or less immediate consequence of the previous section, as long as we express this in terms of d_η. The basic idea is that since $d_\eta(\varphi, \varphi')$ is just a weighted supremum metric on the space $\mathcal{C}_\eta$, then if we take a sequence φ_n of continuous functions with $d_\eta(\varphi_n, \Phi) \to 0$ as $n \to \infty$ then Φ ought to be continuous.

The only issue is the question of what precisely it means for a function to be continuous with respect to d_η; the problem is that this metric depends on the base point x, and hence there is a choice in how we evaluate $d_\eta(\varphi(x), \varphi(x'))$. In view of the construction of regular neighbourhoods above (§6) it might seem natural to define a continuous function as one for which $d_{\eta,x}(\varphi(x_i), \varphi(x)) \to 0$ as $i \to \infty$ for any sequence $x_i \in \Lambda$ with $x_i \to x \in \Lambda$. Unfortunately, it is not necessarily true that if $d_\eta(\varphi_n, \varphi) \to 0$ as $n \to \infty$ and $d_{\eta,x}(\varphi_n(x_i), \varphi_n(x)) \to 0$ as $x_i \to x$ for all n then $d_{\eta,x}(\varphi(x_i), \varphi(x)) \to 0$ as $x_i \to x$. If we were to attempt to apply the standard proof of this type of result we would want to use the following kind of estimate: $d_{\eta,x}(\varphi(x_i), \varphi(x)) \leq d_{\eta,x}(\varphi(x), \varphi_n(x)) + d_{\eta,x}(\varphi_n(x_i), \varphi_n(x)) + d_{\eta,x}(\varphi(x_i), \varphi_n(x_i))$. Whilst the first two terms can be made arbitrarily small, we have no control over the third, although $d_{n,x_i}(\varphi(x_i), \varphi_n(x_i)) \to 0$ as $n \to \infty$,

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

the metrics d_{η,x_i} and $d_{\eta,x}$ are not equivalent, and in particular the best estimate we have is $d_{\eta,x}(y, y') \leq e^{\eta(k_i-k)} d_{\eta,x_i}(y, y')$, where $x_i \in \Lambda_{k_i}$. Thus if $k_i \to \infty$ as $i \to \infty$ sufficiently fast, we cannot guarantee that $d_{\eta,x}(\varphi(x_i), \varphi_n(x_i)) \to 0$.

It is thus reasonable to restrict ourselves to a fixed Λ_k and hence define the following.

Definition 7.5. We say that $\varphi \in \mathcal{C}_\eta$ is η-continuous if $\varphi|_{\Lambda_k}$ is continuous for each k. Since d and d_η are equivalent on Λ_k, the continuity of $\varphi|_{\Lambda_k}$ may be expressed with respect to either metric. Note that since Λ_k is compact, $\varphi|_{\Lambda_k}$ is actually uniformly continuous. We denote the space of such functions as $\mathcal{C}^0_\eta$. This is our analogue of $\mathcal{C}^0$ in §2.

LEMMA 7.6. *$\mathcal{C}^0_\eta$ is closed in $\mathcal{C}_\eta$.*

Proof. Suppose $\varphi_n \in \mathcal{C}^0_\eta$ is a sequence such that $d_\eta(\varphi_n, \varphi) \to 0$ as $n \to \infty$ and $x_i \in \Lambda_k$ is a sequence such that $x_i \to x$ as $i \to \infty$ (since Λ_k is closed $x \in \Lambda_k$). Then $d(\varphi(x_i), \varphi(x)) \leq d(\varphi(x), \varphi_n(x)) + d(\varphi_n(x_i), \varphi_n(x)) + d(\varphi(x_i), \varphi_n(x_i))$. Given $\varepsilon > 0$, choose n sufficiently large such that $d_\eta(\varphi_n, \varphi) \leq \varepsilon e^{-\eta k}$. Then $d(\varphi(x_i), \varphi_n(x_i)) \leq d_\eta(\varphi(x_i), \varphi_n(x_i)) \leq e^{\eta k} d_\eta(\varphi_n, \varphi) \leq \varepsilon$ and similarly for $d(\varphi(x), \varphi_n(x))$. Thus $d(\varphi(x_i), \varphi(x)) \leq d(\varphi_n(x_i), \varphi_n(x)) + 2\varepsilon$. Now choose i_0 such that $d(\varphi_n(x_i), \varphi_n(x)) \leq \varepsilon$ for all $i \geq i_0$. Then $d(\varphi(x_i), \varphi(x)) \leq 3\varepsilon$ for all $i \geq i_0$, and hence φ is continuous on Λ_k as required. □

LEMMA 7.7. $\Gamma(\mathcal{C}^0_\eta) \subset \mathcal{C}^0_\eta$.

Proof. Suppose $\varphi \in \mathcal{C}^0_\eta$. Then $\varphi|_{\Lambda_{k+1}}$ is continuous and hence $\varphi \circ f^{-1}$ is continuous on Λ_k. Thus $\Gamma(\varphi) = g \circ (f^{-1}, \varphi \circ f^{-1})$ is continuous on Λ_k, as required. □

COROLLARY 7.8. $\Phi \in \mathcal{C}^0_\eta$.

Proof. Let $\varphi_0(x) = y_0$, clearly $\varphi_0 \in \mathcal{C}^0_\eta$. Thus $\Gamma^n(\varphi_0) \in \mathcal{C}^0_\eta$, but $d_\eta(\Gamma^n(\varphi_0), \Phi) \to 0$ as $n \to \infty$. □

This completes the proof of Theorem 1.5.

8. *Hölder and Lipschitz continuity under non-uniform contraction*

We now turn to Theorem 1.6, and hence from now on assume that f is a $\mathcal{C}^{1+\alpha}$ diffeomorphism and g satisfies (1.12). The Pesin sets Λ_k are now required to satisfy both (6.6) and (6.7), i.e. they are the intersection of the sets given by each of these definitions.

8.1. *Hölder estimates.*

Just as for continuous functions, we need to define Hölder functions in terms of their restrictions to each Λ_k. However, when we have an explicit modulus of continuity, it makes sense to ask for some control as $k \to \infty$. Thus, we have the following.

Definition 8.1. We say that $\varphi : \Lambda \to Y$ is γ, η, b-Hölder if there exists a $c \geq 0$ such that

$$d(\varphi(x), \varphi(x')) \leq c e^{b\eta k} d(x, x')^\gamma$$

for all $x, x' \in \Lambda_k$.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

As usual it is convenient to work in terms of adapted metrics. Since $d_{\eta,x}(x,x')$ is only defined for $x' \in N_x$, the appropriate function space is given by the following definition.

Definition 8.2.

$$\mathcal{C}^{\gamma}_{\eta,b} = \{\varphi \in \mathcal{C}^0_\eta : d_{\eta,x}(\varphi(x),\varphi(x')) \le c e^{b\eta k} d_{\eta,x}(x,x')^\gamma \text{ for some } c \ge 0, \\ \text{for all } x, x' \in \Lambda_k \text{ with } x' \in N_x, \text{ for all } k\}.$$

For $\gamma = 1$, we write $\mathcal{C}^{\mathrm{Lip}}_{\eta,b}$. The space $\mathcal{C}^{\gamma}_{\eta,b}$ is the appropriate generalization of $\mathcal{C}^\gamma$ in §3 to the present case of non-uniform contraction.

LEMMA 8.3. *If $\varphi \in \mathcal{C}^{\gamma}_{\eta,b}$ then φ is γ, η, b'-Hölder for any $b' \ge \max\{3, b+\gamma\}$.*

Proof. Fix $x, x' \in \Lambda_k$. If $x' \in N_x$, then $d(\varphi(x),\varphi(x')) \le d_{\eta,x}(\varphi(x),\varphi(x')) \le c e^{b\eta k} d_{\eta,x}(x,x')^\gamma \le c e^{(b+\gamma)\eta k} d(x,x')^\gamma$. On the other hand, if $x' \notin N_x$ then since $\varphi \in \mathcal{C}^0_\eta$, we have $d(\varphi(x), y_0) \le d_{\eta,x}(\varphi(x), y_0) \le e^{\eta k} C_0$ and $d(\varphi(x'), y_0) \le d_{\eta,x'}(\varphi(x'), y_0) \le e^{\eta k} C_0$ where $C_0 = d_\eta(\varphi, \varphi_0)$. Furthermore, $d(x,x') \ge [a(x)]^{-1} r(x) \ge e^{-2\eta k}$. Thus

$$\begin{aligned} d(\varphi(x),\varphi(x')) &\le d(\varphi(x), y_0) + d(\varphi(x'), y_0) \\ &\le 2e^{\eta k} C_0 \\ &\le 2C_0 e^{3\eta k} d(x,x') \\ &\le 2C_0 (K_X)^{(1-\gamma)} e^{3\eta k} d(x,x')^\gamma \end{aligned}$$

where $K_X = \max\{d(x,x') : x, x' \in X\}$ (recall that X is compact, and hence bounded). □

For $\varphi \in \mathcal{C}^{\gamma}_{\eta,b}$ we define

$$H_{\gamma,\eta,b}(\varphi) = \sup_{x,x' \in \Lambda_k, x' \in N_x} e^{-b\eta k} \frac{d_{\eta,x}(\varphi(x),\varphi(x'))}{d_{\eta,x}(x,x')^\gamma}.$$

A standard argument gives the following analogue of Lemma 3.1.

LEMMA 8.4. *Let $\varphi_n \in \mathcal{C}^{\gamma}_{\eta,b}$ be a sequence such that $H_{\gamma,\eta,b}(\varphi_n) \le C$ for some $C \ge 0$ and $d_\eta(\varphi_n, \varphi) \to 0$ as $n \to \infty$ for some $\varphi \in \mathcal{C}_\eta$. Then $\varphi \in \mathcal{C}^{\gamma}_{\eta,b}$ and $H_{\gamma,\eta,b}(\varphi) \le C$.*

Proof. Fix $x, x' \in \Lambda_k$, with $x' \in N_x$. Then given $\varepsilon > 0$ we have

$$d_{\eta,x}(\varphi(x),\varphi(x')) \le d_{\eta,x}(\varphi(x),\varphi_n(x)) + d_{\eta,x}(\varphi_n(x),\varphi_n(x')) + d_{\eta,x}(\varphi(x'),\varphi_n(x')).$$

Choose n sufficiently large such that $d_\eta(\varphi_n,\varphi) \le \varepsilon e^{-2\eta k}$. Then $d_{\eta,x}(\varphi(x),\varphi_n(x)) \le e^{\eta k} d_\eta(\varphi_n,\varphi) \le \varepsilon$, and $d_{\eta,x}(\varphi(x'),\varphi_n(x')) \le e^{\eta k} d(\varphi(x'),\varphi_n(x')) \le e^{\eta k} d_{\eta,x'}(\varphi(x'),\varphi_n(x')) \le e^{2\eta k} d_\eta(\varphi_n,\varphi) \le \varepsilon$. Thus

$$\begin{aligned} d_{\eta,x}(\varphi(x),\varphi_n(x')) &\le 2\varepsilon + d_{\eta,x}(\varphi_n(x),\varphi_n(x')) \\ &\le 2\varepsilon + C e^{b\eta k} d_{\eta,x}(x,x')^\gamma. \end{aligned}$$

But this holds for all $\varepsilon > 0$, and hence $d_{\eta,x}(\varphi(x),\varphi(x')) \le C e^{b\eta k} d_{\eta,x}(x,x')^\gamma$, as required. □

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

We now aim to show that $\Phi \in \mathcal{C}^{\gamma}_{\eta,b}$ for some $b > 0$. This follows directly from the following analogue of Proposition 3.2.

PROPOSITION 8.5. *Suppose $\gamma \le \alpha$ such that $\lambda + \gamma\mu < 0$ and $0 < \eta < -(\lambda + \mu\gamma)/(4 + 3\gamma)$. Let $b = \max\{3, \beta + 1\}$ and choose C_0 sufficiently large so that $\Gamma(\mathcal{C}_{\eta,C_0}) \subset \mathcal{C}_{\eta,C_0}$, and $C_\gamma \le \max\{(K_{g,0} + K_g C_0^{\beta} e^{\eta\beta}) e^{(\mu+3\eta)\alpha}(1 - e^{-\eta})^{-1}, 2C_0 e^{\mu+4\eta}\}$. Then if $\varphi \in \mathcal{C}_{\eta,C_0}$ and*

$$d_{\eta,x}(\varphi(x), \varphi(x')) \le C_\gamma e^{b\eta k} d_{\eta,x}(x, x')^{\gamma}$$

for all $x, x' \in \Lambda_k$ and $x' \in N_x$, we have

$$d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) \le C_\gamma e^{b\eta k} d_{\eta,x}(x, x')^{\gamma}.$$

Proof. Note that for $x \in \Lambda_k$, N_x contains a ball of radius $e^{-2k\eta}$, about x. First suppose that $\|\Psi_x^{-1}(x')\| = d_{\eta,x}(x, x') \le e^{-\eta k} e^{-(\mu+4\eta)}$. Then $\|f_x^{-1}(\Psi_x^{-1}(x'))\| \le e^{\mu+3\eta} e^{-\eta k} e^{-(\mu+4\eta)} \le e^{-\eta(k+1)}$ and hence $f_x^{-1}(\Psi_x^{-1}(x')) \in B_{r(f^{-1}(x))}(\mathbf{0}) = [\Psi_{f^{-1}(x)}]^{-1} N_{f^{-1}(x)}$. Thus $f^{-1}(x') = \Psi_{f^{-1}(x)} \circ f_x^{-1} \circ \Psi_x^{-1}(x') \in N_{f^{-1}(x)}$. Then

$$\begin{aligned} d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) &\le d_{\eta,x}(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi(f^{-1}(x')))) \\ &\quad + d_{\eta,x}(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))). \end{aligned}$$

Now

$$\begin{aligned} &d_{\eta,x}(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi(f^{-1}(x')))) \\ &\quad \le e^{\lambda+2\eta} d_{\eta, f^{-1}(x)}(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))). \end{aligned}$$

Since $f^{-1}(x), f^{-1}(x') \in \Lambda_{k+1}$, we have by (6.5)

$$\begin{aligned} d_{\eta, f^{-1}(x)}(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))) &\le C_\gamma e^{b\eta(k+1)} d_{\eta, f^{-1}(x)}(f^{-1}(x), f^{-1}(x'))^{\gamma} \\ &\le C_\gamma e^{b\eta(k+1)} e^{(\mu+3\eta)\gamma} d_{\eta,x}(x, x')^{\gamma}. \end{aligned}$$

Since $(\lambda + 3\eta) + (\mu + 3\eta)\gamma < -\eta$ we get

$$\begin{aligned} &d_{\eta,x}(g(f^{-1}(x), \varphi(f^{-1}(x))), g(f^{-1}(x), \varphi(f^{-1}(x')))) \\ &\quad \le C_\gamma e^{b\eta k} e^{\lambda+2\eta} e^{(\mu+3\eta)\gamma} e^{\eta} d_{\eta,x}(x, x')^{\gamma} \\ &\quad \le e^{-\eta} C_\gamma e^{b\eta k} d_{\eta,x}(x, x')^{\gamma}. \end{aligned}$$

Turning to the second term, we have

$$\begin{aligned} &d(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))) \\ &\quad \le (K_{g,0} + K_g d(\varphi(f^{-1}(x')), y_0)^{\beta}) d(f^{-1}(x), f^{-1}(x'))^{\alpha}. \end{aligned}$$

Since $f^{-1}(x') \in \Lambda_{k+1}$, we have $d(\varphi(f^{-1}(x')), y_0) \le d_{\eta,x}(\varphi(f^{-1}(x')), y_0) \le C_0 e^{\eta(k+1)}$. As above, we have $d(f^{-1}(x), f^{-1}(x')) \le e^{(\mu+3\eta)} d_\eta(x, x')$. Since $x' \in N_x$ we have $d_{\eta,x}(x, x') \le e^{-\eta k} \le 1$ and hence $d_{\eta,x}(x, x')^{\alpha} \le d_{\eta,x}(x, x')^{\gamma}$. Thus $d(f^{-1}(x), f^{-1}(x'))^{\alpha} \le d_{n, f^{-1}(x)}(f^{-1}(x), f^{-1}(x'))^{\alpha} \le e^{(\mu+3\eta)\alpha} d_{\eta,x}(x, x')^{\alpha} \le e^{(\mu+3\eta)\alpha} d_{\eta,x}(x, x')^{\gamma}$

$$\begin{aligned} &d_{\eta,x}(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))) \\ &\quad \le e^{\eta k} d(g(f^{-1}(x), \varphi(f^{-1}(x'))), g(f^{-1}(x'), \varphi(f^{-1}(x')))) \\ &\quad \le e^{\eta k} (K_{g,0} + K_g K^{\beta} e^{\eta\beta(k+1)}) e^{(\mu+3\eta)\alpha} d_{\eta,x}(x, x')^{\gamma} \\ &\quad \le c e^{b\eta k} d_{\eta,x}(x, x')^{\gamma} \end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

where $c = (K_{g,0} + K_g C_0^\beta e^{\eta\beta}) e^{(\mu+3\eta)\alpha}$. Thus, since $C_\gamma \geq c(1-e^{-\eta})^{-1}$ we have

$$\begin{aligned} d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) &\leq (c + e^{-\eta} C_1) e^{b\eta k} d_{\eta,x}(x,x')^\gamma \\ &\leq C_\gamma e^{b\eta k} d_{\eta,x}(x,x')^\gamma \end{aligned}$$

as required.

Now suppose that $d_{\eta,x}(x,x') \geq e^{-\eta k} e^{-(\mu+4\eta)}$. We employ the same argument as in the proof of Lemma 7.1. Since $d(x,x') \geq [a(x)]^{-1} d_{\eta,x}(x,x')$, we have $d(x,x') \geq e^{-2\eta k} e^{-(\mu+4\eta)}$. Since $\Gamma(\varphi) \in \mathcal{C}_{\eta,C_0}$ we have $d_{\eta,x}(\Gamma(\varphi)(x), y_0) \leq C_0 e^{\eta k}$ and $d_{\eta,x'}(\Gamma(\varphi)(x'), y_0) \leq C_0 e^{\eta k}$. Thus $d_{\eta,x}(\Gamma(\varphi)(x'), y_0) \leq e^{\eta k} d(\Gamma(\varphi)(x'), y_0) \leq C_0 e^{2\eta k}$. Note that $x' \in N_x$ and hence we still have $d_\eta(x,x') \leq d_\eta(x,x')^\gamma$. Hence

$$\begin{aligned} d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) &\leq d_{\eta,x}(\Gamma(\varphi)(x), y_0) + d_{\eta,x}(\Gamma(\varphi)(x'), y_0) \\ &\leq C_0 e^{\eta k}(1 + e^{\eta k}) \\ &\leq C_0 e^{\eta k}(1 + e^{\eta k}) e^{\eta k} e^{(\mu+4\eta)} d_{\eta,x}(x,x') \\ &\leq 2C_0 e^{\mu+4\eta} e^{3\eta k} d_{\eta,x}(x,x')^\gamma \\ &\leq C_\gamma e^{b\eta k} d_\eta(x,x')^\gamma \end{aligned}$$

as required. □

COROLLARY 8.6. *$\Phi \in \mathcal{C}^\gamma_{\eta,b}$ for any $\gamma \leq \alpha$ such that $\lambda + \gamma\mu < 0$, any $\eta > 0$ such that $\eta < -(\lambda + \mu\gamma)/(4 + 3\gamma)$ and any $b \geq \max\{3, \beta + 1\}$. In particular, if $\lambda < -\mu$, and η is sufficiently small, then Φ is uniformly Lipschitz on each Λ_k.*

This completes the proof of Theorem 1.6.

9. *Whitney smoothness under non-uniform contraction*

Finally, we come to the proof of Theorem 1.7. This closely parallels that of Theorem 1.3, the main difference is that the various bounds are weighted by factors of $e^{b\eta k}$ for some b. As in §4, since $\lambda < -\mu$, Theorem 1.6 gives us the existence of an invariant graph $\Phi \in \mathcal{C}^{\mathrm{Lip}}_{\eta,b}$ in $X \times \bar{Y}$, and since $(f,g)(X \times \bar{Y}) \subset X \times Y$, this must actually lie in $X \times Y$. As before, denote the norm on Y_0 by $\|\cdot\|$. Recall that we have an adapted norm $\|\cdot\|_\eta$ on $T_y Y$ which was defined at the end of §6.2. As in §4, we identify $T_y Y$ with Y_0, but additional care must be taken here when discussing the adapted norm under this identification, since it depends on y. Note that if $x \in \Lambda_k$ we have

$$\|y - y'\| \leq d_{\eta,x}(y,y') \leq e^{\eta k} \|y - y'\|$$

and since

$$\|y - y'\| \leq \|y - y'\|_{\eta,x,y} \leq e^{\eta k} \|y - y'\|$$

we get

$$e^{-\eta k} \|y - y'\|_{\eta,x,y'} \leq \|y - y'\|_{\eta,x,y} \leq e^{\eta k} \|y - y'\|_{\eta,x,y'} \tag{9.1}$$

and

$$e^{-\eta k} \|y - y'\|_{\eta,x,y} \leq d_{\eta,x}(y,y') \leq e^{\eta k} \|y - y'\|_{\eta,x,y}. \tag{9.2}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Finally, we define

$$K_g'' = \sup_{x \in X} \|D_{(x,y_0)} g\|.$$

9.1. *Existence of derivative.* As in §4.1, let $\mathcal{L}(T_x X, Y_0)$ be the space of linear maps from $T_x X$ to Y_0, and denote by $\mathcal{L}(T_\Lambda X, Y_0)$ the vector bundle whose fibre above $x \in \Lambda$ is $\mathcal{L}(T_x X, Y_0)$. Let $\mathcal{C}_\mathcal{L} = \mathcal{C}(\mathcal{L}(T_\Lambda X, Y_0))$ be the space of sections of $\mathcal{L}(T_\Lambda X, Y_0)$, thus if $A \in \mathcal{C}_\mathcal{L}$ then $A(x)$ is a linear map from $T_x X$ to Y_0. As before, the graph transform extends to $\mathcal{C} \times \mathcal{C}_\mathcal{L}$ by

$$\tilde{\mathbf{\Gamma}}(\varphi, A) = (\Gamma(\varphi), \Gamma_\mathcal{L}(\varphi, A))$$

where

$$\Gamma_\mathcal{L}(\varphi, A)(x) = D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})$$

with $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. Observe that we have a Liapunov norm on $\mathcal{L}(T_x X, Y_0)$ induced from the respective Liapunov norms on $T_x X$ and Y_0. Thus for $A(x) \in \mathcal{L}(T_x X, Y_0)$

$$\|A(x)\|_{\eta,x,y} = \sup\{\|A(x).v\|_{\eta,x,y} : v \in T_x X, \|v\|_{\eta,x} = 1\}.$$

Given $\varphi \in \mathcal{C}_\eta$ we define $\mathcal{C}_{\mathcal{L},\eta,b,\varphi} \subset \mathcal{C}_\mathcal{L}$ by

$$\mathcal{C}_{\mathcal{L},\eta,b,\varphi} = \{A \in \mathcal{C}_\mathcal{L} : \|A(x)\|_{\eta,x,\varphi(x)} \le c e^{\eta b k} \quad \text{for some } c \ge 0, \text{ for all } x \in \Lambda_k, \text{ for all } k\}$$

and define a norm $\|A\|_{\eta,b,\varphi}$ on $\mathcal{C}_{\mathcal{L},\eta,b,\varphi}$ by

$$\|A\|_{\eta,b,\varphi} = \inf\{c : \|A(x)\|_{\eta,x,\varphi(x)} \le c e^{\eta b k} \text{ for all } x \in \Lambda_k, \text{ for all } k\}.$$

Note that if $\varphi, \varphi' \in \mathcal{C}_\eta$ and $A \in \mathcal{C}_{\mathcal{L},\eta,b,\varphi}$, then

$$\begin{aligned} \|A(x).v\|_{\eta,x,\varphi'(x)} &\le e^{\eta k} \|A(x).v\| \\ &\le e^{\eta k} \|A(x).v\|_{\eta,x,\varphi(x)} \\ &\le c e^{\eta(b+1)k} \end{aligned}$$

for all $x \in \Lambda_k$, and hence $A \in \mathcal{C}_{\mathcal{L},\eta,b+1,\varphi'}$ and $\|A\|_{\eta,b+1,\varphi'} \le \|A\|_{\eta,b,\varphi}$. Finally, define

$$\mathcal{L}_{\eta,b} = \{(\varphi, A) : \varphi \in \mathcal{C}_\eta, A \in \mathcal{C}_{\mathcal{L},\eta,b-1,\varphi}\}$$

with a metric

$$d_{\eta,b}((\varphi, A), (\varphi', A')) = \max\{d_\eta(\varphi, \varphi'), \|A - A'\|_{\eta,b,\varphi}, \|A - A'\|_{\eta,b,\varphi'}\}.$$

9.2. *Existence of a fixed point.* We start with an exact analogue of Lemma 4.1.

LEMMA 9.1. *Suppose $\varphi \in \mathcal{C}_\eta$ and $A, A' \in \mathcal{C}_\mathcal{L}$ such that $(A - A') \in \mathcal{C}_{\mathcal{L},\eta,b,\varphi}$ with $b > 0$. Then we have $\Gamma_\mathcal{L}(\varphi, A) - \Gamma_\mathcal{L}(\varphi, A') \in \mathcal{C}_{\mathcal{L},\eta,b,\Gamma(\varphi)}$ and*

$$\|\Gamma_\mathcal{L}(\varphi, A) - \Gamma_\mathcal{L}(\varphi, A')\|_{\eta,b,\Gamma(\varphi)} \le e^{\lambda+\mu+(4+b)\eta} \|A - A'\|_{\eta,b,\varphi}.$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Proof. We have

$$(\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi, A'))(x) = D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1}) - D_w g.(D_x f^{-1}, A'(f^{-1}(x)).D_x f^{-1})$$

where $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. Note that $g(w) = \Gamma(\varphi)(x)$. Fix $x \in \Lambda_k$. For $v \in T_x X$ we have

$$D_w g.(D_x f^{-1}, A(f^{-1}(x))D_x f^{-1}).v = D_w^{(1)} g.u + D_w^{(2)} g.A(f^{-1}(x)).u$$

where $D_w^{(1)} g$ and $D_w^{(2)} g$ are the partial derivatives of g, and $u = D_x f^{-1}.v$. Note that $D_w^{(2)} g = D_{\varphi(f^{-1}(x))} g_{f^{-1}(x)}$, and hence $\|D_w^{(2)} g\|_\eta \le e^{\lambda+2\eta}$, whilst $\|D_x f^{-1}\|_\eta \le e^{\mu+2\eta}$ by (6.4). Thus

$$\begin{aligned}
&\|D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1}).v - D_w g.(D_x f^{-1}, A'(f^{-1}(x)).D_x f^{-1}).v\|_{\eta,x,g(w)} \\
&\quad = \|D_w^{(2)} g.(A(f^{-1}(x)) - A'(f^{-1}(x))).u\|_{\eta,x,g(w)} \\
&\quad \le e^{\lambda+2\eta} \|(A(f^{-1}(x)) - A'(f^{-1}(x)))\|_{\eta,w} \|D_x f^{-1}.v\|_\eta \\
&\quad \le e^{\lambda+2\eta}\, e^{\eta b(k+1)} \|A - A'\|_{\eta,b,\varphi} e^{\mu+2\eta} \|v\|_\eta.
\end{aligned}$$

Hence

$$\begin{aligned}
&\|D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1}) - D_w g.(D_x f^{-1}, A'(f^{-1}(x)).D_x f^{-1})\|_{\eta,x,\Gamma(\varphi)(x)} \\
&\quad \le e^{\eta bk} e^{\lambda+\mu+(4+b)\eta} \|A - A'\|_{\eta,b,\varphi}
\end{aligned}$$

as required. □

COROLLARY 9.2. $\tilde{\mathbf{\Gamma}}(\mathcal{L}_{\eta,b}) \subset \mathcal{L}_{\eta,b}$.

Proof. We proceed exactly as in Corollary 7.2. Suppose $(\varphi, A) \in \mathcal{L}_{\eta,b}$. Set $A_0 = \mathbf{0}$, then $A_1(x) = \Gamma_{\mathcal{L}}(\varphi, A_0)(x) = D_{(f^{-1}(x),\varphi(f^{-1}(x)))} g.(D_x f^{-1}, \mathbf{0})$. Since g, f^{-1} and $D_x f^{-1}$ are continuous and X is compact, A_1 is bounded in the standard metric, and hence $A_1 \in \mathcal{C}_{\mathcal{L},\eta,b,\Gamma(\varphi)}$ for any $b > 0$. Since $A - A_0 = A \in \mathcal{C}_{\mathcal{L},\eta,b,\varphi}$, we have $\Gamma_{\mathcal{L}}(\varphi, A) - A_1 \in \mathcal{C}_{\mathcal{L},\eta,b,\Gamma(\varphi)}$ and hence $\Gamma_{\mathcal{L}}(\varphi, A) \in \mathcal{C}_{\mathcal{L},\eta,b,\Gamma(\varphi)}$. Thus $(\Gamma(\varphi), \Gamma_{\mathcal{L}}(\varphi, A)) \in \mathcal{L}_{\eta,b}$ as required. □

Finally, we obtain the counterpart of Corollary 4.2.

COROLLARY 9.3. *If $\lambda < -\mu$, then $\tilde{\mathbf{\Gamma}}$ has a unique fixed point $(\Phi, A^*) \in \{\Phi\} \times \mathcal{C}_{\mathcal{L},\eta,b,\Phi}$ for all η such that $0 < \eta < (\lambda + \mu)/(4 + b)$, where Φ is the unique fixed point of Γ, obtained previously.*

9.3. *Convergence in $\mathcal{L}_{\eta,b}$.* Exactly as in §4.2, we now aim to show that (Φ, A^*) is attracting in the whole of $\mathcal{L}_{\eta,b}$, rather than just $\{\Phi\} \times \mathcal{C}_{\mathcal{L},\eta,b,\Phi}$. The next lemma thus corresponds to Lemma 4.3.

LEMMA 9.4. *Suppose $\varphi, \varphi' \in \mathcal{C}_\eta$, $A \in \mathcal{C}_{\mathcal{L},\eta,b',\varphi}$ for some $b' > 0$, and $b \ge \alpha + \beta + b' + 1$. Then $\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi', A) \in \mathcal{C}_{\mathcal{L},\eta,b,\Gamma(\varphi)}$ and*

$$\|\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi', A)\|_{\eta,b,\Gamma(\varphi)} \le K_b \hat{m}(\varphi, \varphi')^\beta \|A\|_{\eta,b',\varphi} d_\eta(\varphi, \varphi')^\alpha$$

where $\hat{m}(\varphi, \varphi') = \max\{d_\eta(\varphi, \varphi_0), d_\eta(\varphi', \varphi_0)\}$ and $K_b = 2(K'_{g,0} + K'_g) e^{\mu+(\alpha+\beta+b'+3)}$.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Proof. Suppose $x \in \Lambda_k$ and $v \in T_x X$. Then

$$(\Gamma_{\mathcal{L}}(\varphi, A) - \Gamma_{\mathcal{L}}(\varphi', A))(x).v = (D_w g - D_{w'} g).(u, A(f^{-1}(x)).u)$$

where $u = D_x f^{-1}.v$ and $w = (f^{-1}(x), \varphi(f^{-1}(x)))$, $w' = (f^{-1}(x), \varphi'(f^{-1}(x)))$. Note that by (6.4), $\|v\|_\eta \le e^{\mu+2\eta}\|v\|_\eta$. Thus, using (1.12), we have

$$\begin{aligned}
&\|(D_w g - D_{w'} g).(u, A(f^{-1}(x).u)\| \\
&\quad \le (K'_{g,0} + K'_g m(\varphi(f^{-1}(x)), \varphi'(f^{-1}(x)))^\beta) d(\varphi(f^{-1}(x)), \varphi'(f^{-1}(x)))^\alpha \\
&\qquad \times (\|u\| + \|A(f^{-1}(x)).u\|) \\
&\quad \le (K'_{g,0} + K'_g e^{\eta\beta(k+1)} \hat{m}(\varphi, \varphi')^\beta) e^{\eta\alpha(k+1)} d_\eta(\varphi, \varphi')^\alpha \\
&\qquad \times (1 + e^{b'\eta(k+1)}\|A\|_{\eta,b',\varphi}) e^{\mu+2\eta}\|v\|_\eta.
\end{aligned}$$

Hence

$$\begin{aligned}
&\|(D_w g - D_{w'} g).(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})\|_{\eta,b,\Gamma(\varphi)} \\
&\quad \le 2(K'_{g,0} + K'_g) e^{\mu+(\alpha+\beta+b'+3)\eta} e^{\eta(\alpha+\beta+b'+1)k} \hat{m}(\varphi, \varphi')^\beta \|A\|_{\eta,b',\varphi} d_\eta(\varphi, \varphi')^\alpha
\end{aligned}$$

as required. □

COROLLARY 9.5. *If $\lambda < -\mu$, then $d_{\eta,b}(\tilde{\mathbf{\Gamma}}^n(\varphi, A), (\Phi, A^*)) \to 0$ as $n \to \infty$ for all $(\varphi, A) \in \mathcal{L}_{\eta,b}$, for any $b \ge \alpha + \beta + 4$, and η such that $0 < \eta < (\lambda + \mu)/(3 + b)$.*

Proof. This is almost identical to the proof of Corollary 4.4. Let $(\varphi_n, A_n) = \tilde{\mathbf{\Gamma}}^n(\varphi, A)$ for $n \ge 1$ (we still use φ_0 to denote the constant function $\varphi_0(x) = y_0$). By Lemma 7.1 we have $d_\eta(\varphi_n, \Phi) \le e^{(\lambda+3\eta)n} d_\eta(\varphi, \Phi)$. Since $\|A_n - A^*\|_{\eta,b,\Phi} \le \|A_n - A^*\|_{\eta,b-1,\varphi_n}$, it remains to prove that $\|A_n - A^*\|_{\eta,b-1,\varphi_n} \to 0$. Observe that

$$\begin{aligned}
A_{n+1} - A^* &= \Gamma_{\mathcal{L}}(\varphi_n, A_n) - \Gamma_{\mathcal{L}}(\Phi, A^*) \\
&= \Gamma_{\mathcal{L}}(\varphi_n, A_n) - \Gamma_{\mathcal{L}}(\varphi_n, A^*) + \Gamma_{\mathcal{L}}(\varphi_n, A^*) - \Gamma_{\mathcal{L}}(\Phi, A^*).
\end{aligned}$$

Note that $d_\eta(\varphi_n, \varphi_0) \le d_\eta(\varphi_n, \Phi) + d_\eta(\Phi, \varphi_0) \le d_\eta(\varphi, \Phi) + d_\eta(\Phi, \varphi_0)$ for all n, and hence $\hat{m}(\varphi_n, \Phi) \le d_\eta(\varphi, \Phi) + d_\eta(\Phi, \varphi_0)$ for all n. Thus, using the fact that $A^* \in \mathcal{C}_{\mathcal{L},\eta,1,\Phi}$, and hence $A^* \in \mathcal{C}_{\mathcal{L},\eta,2,\varphi_n}$ we get

$$\begin{aligned}
\|A_{n+1} - A^*\|_{\eta,b-1,\varphi_{n+1}} &\le e^{\lambda+\mu+(3+b)\eta}\|A_n - A^*\|_{\eta,b-1,\varphi_n} \\
&\quad + K_{b-1}\hat{m}(\varphi_n, \Phi)^\beta \|A^*\|_{\eta,2,\varphi_n} d_\eta(\varphi_n, \Phi)^\alpha \\
&\le e^{\lambda+\mu+(3+b)\eta}\|A_n - A^*\|_{\eta,b-1,\varphi_n} + c' e^{(\lambda+3\eta)n\alpha}
\end{aligned}$$

where $c' = K_{b-1}(d_\eta(\varphi, \Phi) + d_\eta(\Phi, \varphi_0))^\beta \|A^*\|_{\eta,2,\varphi_n} d_\eta(\varphi, \Phi)^\alpha$. We claim that this implies that for any ε such that $0 < \varepsilon < \max\{-(\lambda + \mu + (3 + b)\eta), -(\lambda + 3\eta)\alpha\}$ and c such that $c \ge \max\{\|A - A^*\|_{\eta,b-1,\varphi}, c' e^\varepsilon (1 - e^{\lambda+\mu+(3+b)\eta+\varepsilon})^{-1}\}$ we have

$$\|A_n - A^*\|_{\eta,b-1,\varphi_n} \le c e^{-\varepsilon n}.$$

This follows by induction. It trivially holds for $n = 0$. Suppose it holds for $n > 0$. Then

$$\begin{aligned}
\|A_{n+1} - A^*\|_{\eta,b-1,\varphi_{n+1}} &\le c e^{-\varepsilon n} e^{\lambda+\mu+(3+b)\eta} + c' e^{(\lambda+3\eta)n\alpha} \\
&\le c e^{-\varepsilon(n+1)} (e^{\lambda+\mu+(3+b)\eta+\varepsilon} + c' c^{-1} e^\varepsilon e^{((\lambda+3\eta)\alpha+\varepsilon)n})
\end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

and hence since $e^{((\lambda+3\eta)\alpha+\varepsilon)n} \leq 1$

$$\begin{aligned}\|A_{n+1} - A^*\|_{\eta,b-1,\varphi_{n+1}} &\leq ce^{-\varepsilon(n+1)}(e^{\lambda+\mu+(3+b)\eta+\varepsilon} + 1 - e^{\lambda+\mu+(4+b)\eta+\varepsilon}) \\ &\leq ce^{-\varepsilon(n+1)}\end{aligned}$$

as required. □

9.4. *Whitney smoothness.* Suppose $x, x' \in \Lambda_k$ and $x' \in N_x$, and recall that

$$\Delta((\varphi, A), x, x') = \varphi(x) - \varphi(x') - A(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x')).$$

Our aim now is to show that for some A^* we have a $c \geq 0$ and $0 < \gamma \leq 1$ such that

$$\|\Delta((\Phi, A^*), x, x')\| \leq ce^{b'\eta k} d(x, x')^{1+\gamma} \tag{9.3}$$

for all $x, x' \in \Lambda_k$ and $x' \in N_x$. As in §8, it is convenient to work in the adapted norm. We shall thus define $\mathcal{L}_{\eta,b}^{1+\gamma}$ to be the set of $(\varphi, A) \in \mathcal{L}_{\eta,b}$ such that

$$\|\Delta((\varphi, A), x, x')\|_{\eta,x,\varphi(x)} \leq ce^{b\eta k} d_{\eta,x}(x, x')^{1+\gamma}$$

for all $x, x' \in \Lambda_k'$ and $x' \in N_x$. For $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ we define

$$H_{1+\gamma,\eta,b}(\varphi, A) = \sup_{x,x' \in \Lambda_k, x' \in N_x} e^{-b\eta k} \frac{\|\Delta((\varphi, A), x, x')\|_{\eta,b,\varphi(x)}}{d_{\eta,x}(x, x')^{1+\gamma}}.$$

Note that if $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ then,

$$\|\Delta((\varphi, A), x, x')\| \leq ce^{(b+1+\gamma)\eta k} d(x, x')^{1+\gamma}$$

and hence φ satisfies (9.3) for $b' = b + 1 + \gamma$.

We now proceed very much as in §4.3, with the same modifications to account for the non-uniform contraction that we made in §8. First, we show that the limit of a convergent sequence in $\mathcal{L}_{\eta,b}$ for which $H_{1+\gamma,n,\beta}$ is uniformly bounded is in $\mathcal{L}_{\eta,b}^{1+\gamma}$; this corresponds to a combination of Lemma 4.6 and Lemma 8.4. We will then go on to prove that if $\lambda < -\mu(1+\gamma)$, then for a sufficiently large $C > 0$, if $H_{1+\gamma,\eta,b}(\varphi, A) \leq C$, then $\tilde{\mathbf{\Gamma}}^n(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ and $H_{1+\gamma,\eta,b}(\tilde{\mathbf{\Gamma}}^n(\varphi, A)) \leq C$, paralleling Propositions 4.7 and 8.5. The combination of these two results then immediately implies that $(\Phi, A^*) \in \mathcal{L}_{\eta,b}^{1+\gamma}$, as required.

LEMMA 9.6. *Let $(\varphi_n, A_n) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ be a sequence such that $H_{1+\gamma,\eta,b}(\varphi_n, A_n) \leq C$ for some $C \geq 0$ and such that $d_{\eta,b}((\varphi_n, A_n), (\varphi, A)) \to 0$ as $n \to \infty$ for some $(\varphi, A) \in \mathcal{L}_{\eta,b}$. Then $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ and $H_{1+\gamma,\eta,b}(\varphi, A) \leq C$.*

Proof. Suppose $x, x' \in \Lambda_k$ and $x' \in N_x$. Then given $\varepsilon > 0$ we have

$$\begin{aligned}\|\Delta((\varphi, A), x, x')\|_{\eta,x,\varphi(x)} \leq{}& \|\Delta((\varphi_n, A_n), x, x')\|_{\eta,x,\varphi(x)} \\ &+ \|\varphi(x) - \varphi_n(x)\|_{\eta,x,\varphi(x)} + \|\varphi(x') - \varphi_n(x')\|_{\eta,x,\varphi(x)} \\ &+ \|(A(x) - A_n(x))(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\|_{\eta,x,\varphi(x)}.\end{aligned}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Choose n sufficiently large such that $d_\eta(\varphi, \varphi') \leq \varepsilon e^{-2\eta k}$ and $\|A - A_n\|_{\eta,b,\varphi} \leq \varepsilon e^{-b\eta k} d_\eta(x, x')^{-1}$. Then since $x \in \Lambda_k$ we have $d_{\eta,x}(\varphi(x), \varphi_n(x)) \leq \varepsilon e^{-\eta k}$ and by (9.2) $\|\varphi(x) - \varphi_n(x)\|_{\eta,x,\varphi(x)} \leq e^{\eta k} d_{\eta,x}(\varphi(x), \varphi_n(x)) \leq \varepsilon$. Similarly, since $x' \in \Lambda_k$, we have

$$\begin{aligned}\|\varphi(x') - \varphi_n(x')\|_{\eta,x,\varphi'(x)} &\leq e^{\eta k}\|\varphi(x') - \varphi_n(x')\| \leq e^{\eta k} d(\varphi(x'), \varphi_n(x')) \\ &\leq e^{\eta k} d_{\eta,x'}(\varphi(x'), \varphi_n(x')) \leq \varepsilon.\end{aligned}$$

Since

$$\|\exp_x^{-1}(x) - \exp_x^{-1}(x')\|_{\eta,x} = \|Q_x.\exp_x^{-1}(x')\| = \|\Psi_x^{-1}(x')\| = d_{\eta,x}(x, x'),$$

we have

$$\begin{aligned}\|(A(x) - A_n(x))(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\|_{\eta,x,\varphi(x)} &\leq \|A(x) - A_n(x)\|_{\eta,x,\varphi(x)} d_{\eta,x}(x, x') \\ &\leq e^{b\eta k}\|A - A_n\|_{\eta,b,\varphi} d_{\eta,x}(x, x') \leq \varepsilon.\end{aligned}$$

Thus

$$\begin{aligned}\|\Delta((\varphi, A), x, x')\|_{\eta,x,\varphi(x)} &\leq 3\varepsilon + \|\Delta((\varphi_n, A_n), x, x')\|_{\eta,x,\varphi(x)} \\ &\leq 3\varepsilon + Ce^{b\eta k} d_{\eta,x}(x, x')^{1+\gamma}.\end{aligned}$$

But this holds for arbitrary $\varepsilon > 0$, and hence

$$\|\Delta((\varphi, A), x, x')\|_{\eta,x,\varphi(x)} \leq Ce^{b\eta k} d_{\eta,x}(x, x')^{1+\gamma}$$

as required. □

So finally we come to prove the crucial result that if $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ then there exists a $C \geq 0$ such that $H_{1+\gamma,n,b}(\tilde{\mathbf{\Gamma}}^n(\varphi, A)) \leq C$ for all $n \geq 0$. First note that if $\lambda < -\mu$, $0 < \gamma \leq \alpha$, $0 < \eta < -(\lambda + \mu)/7$, $b' = \max\{3, \beta + 1\}$, and $b \geq \alpha + \beta + 4$ then given $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$ such that $\varphi \in \mathcal{C}_{\eta,b}^{\mathrm{Lip}}$, by Corollary 7.4, Proposition 8.5 and Corollary 9.2, we can find constants C_0, C_1 and $C_{\mathcal{L}}$ such that for all $n \geq 0$ we have

$$\begin{aligned}d_\eta(\varphi_n, \varphi_0) &\leq C_0 \\ H_{\mathrm{Lip},\eta,b'}(\varphi_n) &\leq C_1 \\ \|A_n\|_{\eta,b-1,\varphi_n} &\leq C_{\mathcal{L}}\end{aligned}$$

where $(\varphi_n, A_n) = \tilde{\mathbf{\Gamma}}^n(\varphi, A)$ for $n \geq 1$, and as usual φ_0 is the constant function $\varphi_0(x) = y_0$. Without loss of generality we assume $C_0, C_1, C_{\mathcal{L}} \geq 1$. We then have the following.

PROPOSITION 9.7. *Suppose $\lambda < -\mu$, $\lambda + (1 + \gamma)\mu < 0$, with $0 < \gamma \leq \alpha$. Then for any η, b and b' such that $0 < \eta < -(\lambda + \mu(1 + \gamma))/(6 + b + 2\gamma)$, $b' = \max\{3, \beta + 1\}$ and $b \geq \max\{5, b'(\alpha + 1) + \beta + 2\}$ suppose that we are given $(\varphi, A) \in \mathcal{L}_{\eta,b}^{1+\gamma}$, such that $\varphi \in \mathcal{C}_{\eta,b'}^{\mathrm{Lip}}$. Let C_0, C_1 and $C_{\mathcal{L}}$ be as above, and set $C_{1+\gamma} = \max\{H_{1+\gamma,\eta,b}(\varphi, A), C_0, C_{\mathcal{L}}, (1 - e^{-\eta})^{-1}(c_0 + c_1 + c_2), e^{(\mu+4\eta)}(C_1 + C_{\mathcal{L}})\}$, where $c_0 = 2(K'_{g,0} + K'_g C_0^\beta C_1^{(1+\alpha)} e^{(\alpha+\beta+1)\eta} e^{(\mu+3\eta)(1+\alpha)})$, $c_1 = (K''_g + K'_{g,0} + K'_g C_0^\beta e^{\beta\eta}) C_0^\alpha e^{\alpha\eta} K_f$ and $c_2 = e^{\lambda+(2+b)\eta} C_{\mathcal{L}} K_f$. Then $H_{1+\gamma,\eta,b}(\varphi_n, A_n) \leq C_{1+\gamma}$, for all $n \geq 0$.*

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Proof. It suffices to prove that if $H_{1+\gamma,\eta,b}(\varphi, A) \le C_{1+\gamma}$, $d_\eta(\varphi_n, \varphi_0) \le C_0$, $H_{\mathrm{Lip},\eta,b-1}(\varphi) \le C_1$, and $\|A\|_{\eta,b-1,\varphi_n} \le C_{\mathcal{L}}$, then $\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\|_{\eta,x,\Gamma(\varphi)(x)} \le C_{1+\gamma}$ $e^{b\eta k} d_\eta(x, x')^{1+\gamma}$ for all $x, x' \in \Lambda_k$ and $x' \in N_x$. As in Proposition 8.5, first suppose that $d_\eta(x, x') \le e^{-\eta k} e^{-(\mu+4\eta)}$ so that $f^{-1}(x') \in N_{f^{-1}(x)}$. Then

$$\begin{aligned}\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x') &= g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) \\ &\quad - D_w g.(D_x f^{-1}, A(f^{-1}(x)).D_x f^{-1})(\exp_x^{-1}(x) - \exp_x^{-1}(x'))\end{aligned}$$

with $w = (f^{-1}(x), \varphi(f^{-1}(x)))$. Let $u_{x'} = -\exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))$ and $v_{x,x'} = \varphi(f^{-1}(x)) - \varphi(f^{-1}(x'))$. Then by (1.13)

$$\begin{aligned}&\|g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) - D_w g.(u_{x'}, v_{x,x'})\| \\ &\quad \le (K'_{g,0} + K'_g m(\varphi(f^{-1}(x)), \varphi(f^{-1}(x')))^\beta) \\ &\qquad \times (d(f^{-1}(x), f^{-1}(x'))^{1+\alpha} + d(\varphi(f^{-1}(x)), \varphi(f^{-1}(x')))^{1+\alpha}).\end{aligned}$$

Now, $d(\varphi(f^{-1}(x)), y_0) \le C_0 e^{\eta(k+1)}$, and similarly for $\varphi(f^{-1}(x'))$; thus $m(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))) \le C_0 e^{\eta(k+1)}$. Also $d(\varphi(f^{-1}(x)), \varphi(f^{-1}(x'))) \le C_1 e^{\eta(k+1)b'} d(f^{-1}(x), f^{-1}(x'))$, and hence

$$\begin{aligned}&\|g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) - D_w g.(u_{x'}, v_{x,x'})\| \\ &\quad \le (K'_{g,0} + K'_g C_0^\beta e^{\beta\eta(k+1)}) \\ &\qquad \times (1 + C_1^{(1+\alpha)} e^{\eta(k+1)b'(1+\alpha)}) d(f^{-1}(x), f^{-1}(x'))^{1+\alpha}).\end{aligned}$$

Since, using (6.5), $d(f^{-1}(x), f^{-1}(x')) \le d_{\eta, f^{-1}(x)}(f^{-1}(x), f^{-1}(x')) \le e^{\mu+3\eta} d_{\eta,x}(x, x')$ and $b \ge b'(\alpha+1) + \beta + 1$, we have

$$\begin{aligned}&\|g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) - D_w g.(u_{x'}, v_{x,x'})\| \\ &\quad \le c_0 e^{(b-1)\eta_k} d_{\eta,x}(x, x')^{1+\alpha}\end{aligned}$$

where $c_0 = 2(K'_{g,0} + K'_g) C_0^\beta C_1^{(1+\alpha)} e^{(b'(1+\alpha)+\beta+1)\eta} e^{(\mu+3\eta)(1+\alpha)}$. Hence

$$\begin{aligned}&\|g(f^{-1}(x), \varphi(f^{-1}(x))) - g(f^{-1}(x'), \varphi(f^{-1}(x'))) - D_w g.(u_{x'}, v_{x,x'})\|_{\eta,x,\Gamma(\varphi)(x)} \\ &\quad \le c_0 e^{(b-1)\eta k} d_{\eta,x}(x, x')^{1+\alpha}.\end{aligned}$$

On the other hand, since $\exp_x^{-1}(x) = \mathbf{0}$.

$$\begin{aligned}&D_w g.(u_{x'}, v_{x,x'}) - D_w g.(D_x f^{-1}, a(f^{-1}(x)).D_x f^{-1})(\exp_x^{-1}(x) - \exp_x^{-1}(x')) \\ &\quad = D_w g.(D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x')), \mathbf{0}) \\ &\qquad - D_w g.(\mathbf{0}, \varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) \\ &\qquad - A(f^{-1}(x)).D_x f^{-1}.\exp_x^{-1}(x')).\end{aligned}$$

Now

$$\|D_x f^{-1}.\exp_x^{-1}(x') - \exp^{-1}_{f^{-1}(x)}(f^{-1}(x'))\| \le K_f d(x, x')^{1+\alpha} \tag{9.4}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

where K_f is the Hölder norm of Df^{-1} and

$$\begin{aligned}\|D_w g\| &\le \|D_{(f^{-1}(x),y_0)}g\| + (K'_{g,0} + K'_g d(\varphi(f^{-1}(x)), y_0)^\beta) d(\varphi(f^{-1}(x)), y_0)^\alpha \\ &\le K''_g + (K'_{g,0} + K'_g C_0^\beta e^{\beta\eta(k+1)}) C_0^\alpha e^{\alpha\eta(k+1)},\end{aligned}$$

hence

$$\begin{aligned}&\|D_w g.(D_x f^{-1}.\exp_x^{-1}(x') - \exp_{f^{-1}(x)}^{-1}(f^{-1}(x')), \mathbf{0})\| \\ &\quad\le (K''_{g,0} + (K'_{g,0} + K'_g C_0^\beta e^{\beta\eta(k+1)}) C_0^\alpha e^{\alpha\eta(k+1)}) K_f d(x,x')^{1+\alpha}\end{aligned}$$

so that, since $b \ge \alpha + \beta + 1$

$$\begin{aligned}&\|D_w g.(D_x f^{-1}.\exp_x^{-1}(x') - \exp_{f^{-1}(x)}^{-1}(f^{-1}(x')), \mathbf{0})\|_{\eta,x,\Gamma(\varphi)(x)} \\ &\quad\le c_1 e^{b\eta k} d_{\eta,x}(x,x')^{1+\alpha}\end{aligned}$$

where $c_1 = (K''_g + (K'_{g,0} + K'_g C_0^\beta e^{\beta\eta}) C_0^\alpha e^{\alpha\eta}) K_f$. Also

$$\begin{aligned}&\|D_w g.(\mathbf{0}, A(f^{-1}(x)).(D_x f^{-1}.\exp_x^{-1}(x') - \exp_{f^{-1}(x)}^{-1}(f^{-1}(x'))))\|_{\eta,x,\Gamma(\varphi)(x)} \\ &\quad\le e^{\lambda+2\eta}\|A(f^{-1}(x))\|_{\eta,w} K_f d_{\eta,x}(x,x')^{1+\alpha}.\end{aligned}$$

Since $\|A(f^{-1}(x))\|_{\eta,w} \le C_{\mathcal{L}} e^{(b-1)\eta(k+1)}$ this gives

$$\begin{aligned}&\|D_w g.(\mathbf{0}, A(f^{-1}(x)).(D_x f^{-1}.\exp_x^{-1}(x') - \exp_{f^{-1}(x)}^{-1}(f^{-1}(x'))))\|_{\eta,x,\Gamma(\varphi)(x)} \\ &\quad\le c_2 e^{b\eta k} d_{\eta,x}(x,x')^{1+\alpha}\end{aligned}$$

where $c_2 = e^{\lambda+(2+b)\eta} C_{\mathcal{L}} K_f$. Finally

$$\begin{aligned}&\|\varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) - A(f^{-1}(x)).\exp_{f^{-1}(x)}^{-1}(f^{-1}(x'))\|_{\eta,w} \\ &\quad= \|\Delta((\varphi,A), f^{-1}(x), f^{-1}(x'))\|_{\eta,w} \\ &\quad\le C_{1+\gamma} e^{b\eta(k+1)} d_\eta(f^{-1}(x), f^{-1}(x'))^{1+\gamma} \\ &\quad\le C_{1+\gamma} e^{b\eta} e^{(\mu+2\eta)(1+\gamma)} e^{b\eta k} d_{\eta,x}(x,x')^{1+\gamma}\end{aligned}$$

and hence if $\lambda + (3+b)\eta + (\mu+2\eta)(1+\gamma) \le -\eta$

$$\begin{aligned}&\|D_w g.(\mathbf{0}, \varphi(f^{-1}(x)) - \varphi(f^{-1}(x')) - A(f^{-1}(x')).\exp_{f^{-1}(x)}^{-1}(f^{-1}(x')))\|_{\eta,x,\Gamma(\varphi)(x)} \\ &\quad\le C_{1+\gamma} e^{\lambda+3\eta} e^{b\eta} e^{(\mu+2\eta)(1+\gamma)} e^{b\eta k} d_{\eta,x}(x,x')^{1+\gamma} \\ &\quad\le C_{1+\gamma} e^{-\eta} e^{b\eta k} d_{\eta,x}(x,x')^{1+\alpha}.\end{aligned}$$

Combining all these estimates, we have

$$\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi,A), x, x')\|_{\eta,x,\Gamma(\varphi)(x)} \le (c_0 + c_1 + c_2 + C_{1+\gamma} e^{-\eta}) e^{b\eta k} d_{\eta,x}(x,x')^{1+\gamma}.$$

Hence if $C_{1+\gamma} \ge (1 - e^{-\eta})^{-1}(c_0 + c_1 + c_2)$ we have

$$\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi,A), x, x')\|_{\eta,x,\Gamma(\varphi)(x)} \le C_{1+\gamma} e^{b\eta k} d_{\eta,x}(x,x')^{1+\gamma}$$

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

as required. It remains for us to deal with the case $d_{\eta,x}(x, x') \geq e^{-\eta k} e^{-(\mu+4\eta)}$. By hypothesis, we have

$$d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) \leq C_1 e^{b'\eta k} d_{\eta,x}(x, x')$$
$$\|\Gamma_{\mathcal{L}}(\varphi, A)(x)\|_{\eta,x,\Gamma(\varphi)(x)} \leq C_{\mathcal{L}} e^{(b-1)\eta k}.$$

Note that $\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x') = \Gamma(\varphi)(x) - \Gamma(\varphi)(x') - \Gamma_{\mathcal{L}}(\varphi, A)(x)(\exp_x^{-1}(x) - \exp_x^{-1}(x'))$ and by (9.2)

$$\begin{aligned}\|\Gamma(\varphi)(x) - \Gamma(\varphi)(x')\|_{\eta,x,\Gamma(\varphi)(x)} &\leq e^{\eta k} d_{\eta,x}(\Gamma(\varphi)(x), \Gamma(\varphi)(x')) \\ &\leq C_1 e^{(b-1)\eta k} d_{\eta,x}(x, x')\end{aligned}$$

since $b' \leq b - 2$. Thus

$$\begin{aligned}\|\Delta(\tilde{\mathbf{\Gamma}}(\varphi, A), x, x')\|_{\eta,x,\Gamma(\varphi)(x)} &\leq C_1 e^{(b-1)\eta k} d_\eta(x, x') + C_{\mathcal{L}} e^{(b-1)\eta k} d(x, x') \\ &\leq (C_1 + C_{\mathcal{L}}) e^{(b-1)\eta k} d_{\eta,x}(x, x') \\ &\leq e^{(\mu+4\eta)} (C_1 + C_{\mathcal{L}}) e^{b\eta k} d_{\eta,x}(x, x') e^{-\eta k} e^{-(\mu+4\eta)} \\ &\leq e^{(\mu+4\eta)} (C_1 + C_{\mathcal{L}}) e^{b\eta k} d_{\eta,x}(x, x')^{1+\gamma} \\ &\leq C_{1+\gamma} e^{b\eta k} d_{\eta,x}(x, x')^{1+\gamma}\end{aligned}$$

as required. □

COROLLARY 9.8. $(\Phi, A^*) \in \mathcal{L}_{\eta,b}^{1+\gamma}$.

This completes the proof of Theorem 1.7.

Acknowledgements. Some of the preliminary ideas presented here were born at the International Conference on Dynamical Systems held at IMPA, Rio de Janeiro, Brazil, and I am extremely grateful to IMPA for their hospitality and financial support, and to the Royal Society for providing travel funds to attend this conference. I would like to thank R. de la Llave and Ya. Pesin for a number of helpful discussions during this conference. Subsequently, conversations with David Broomhead, Jerry Huke, Kevin Campbell and Mike Davies were absolutely crucial to the development of this work, and I am extremely grateful for all their help. I would also like to thank J. Harrison, J. Marsden and R. de la Llave for appraising me of Stein's version of the Whitney extension theorem. Part of this work was carried out at the Newton Institute programme 'From Finite to Infinite Dimensional Dynamical Systems' during the summer of 1995 and I would like to thank the Institute for their hospitality and financial support. Finally, I would like to gratefully acknowledge financial support from the UK Engineering and Physical Sciences Research Council (GR/J97243) and from the Nuffield Foundation (SCI/180/92/365/G).

REFERENCES

Ashwin, P., J. Buescu and I. Stewart. Bubbling of attractors and synchronisation of chaotic oscillators. *Phys. Lett.* **193A** (1994), 126–139.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press

Ashwin, P., J. Buescu and I. Stewart. From attractor to chaotic saddle: a tale of transverse instability. *Nonlinearity* **9** (1996), 706–737.

Arnold, L. *Random Dynamical Systems*. Springer, 1998 (to appear).

Badii, R., G. Broggi, B. Derighetti, M. Ravani, S. Cilberto, A. Politi and M. A. Rubio. Dimension increase in filtered chaotic signals. *Phys. Rev. Lett.* **60** (1998), 979–982.

Broomhead, D. S., J. P. Huke and M. R. Muldoon. Linear filters and non-linear systems. *J. Roy. Stat. Soc.* B **54** (1992), 373–382.

Campbell, K. M. and M. E. Davies. The existence of inertial functions in skew product systems. *Nonlinearity* **9** (1996), 801–817.

Davies, M. E. and K. M. Campbell. Linear recursive filters and nonlinear dynamics. *Nonlinearity* **9** (1996), 487–499.

Eells, J. A setting for global analysis. *Bull. Amer. Math. Soc.* **72** (1966), 751–807.

Eliasson, H. I. Geometry of manifolds of maps. *J. Diff. Geom.* **1** (1967), 169–194.

Fathi, A., M. R. Herman and J. C. Yoccoz. A proof of Pesin's stable manifold theorem. *Geometric Dynamics*. Ed. J. Palis. *Proc. Int. Conf. on Dynamical Systems (Rio de Janeiro, 1981) (Lecture Notes in Mathematics, 1007)*. Springer, 1983.

Hirsch, M. and C. Pugh. Stable manifolds and hyperbolic sets. *Global Analysis (Amer. Math. Soc. Proc. Symp. Pure Math. 14)*. American Mathematical Society, pp. 133–164.

Hirsch, M., C. Pugh and M. Shub. *Invariant Manifolds (Lecture Notes in Mathematics, 583)*. Springer, 1977.

Hunt, B. R., E. Ott and J. A. Yorke. Fractal dimensions of chaotic saddles of dynamical systems. *Phys. Rev.* E **54** (1996), 4819–4823.

Hunt, B. R., E. Ott and J. A. Yorke. Differentiable generalized synchronization of chaos. *Phys. Rev.* E **55** (1997), 4029–4034.

Katok, A. and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems (Encyclopaedia of Mathematics and its Applications, 54)*. Cambridge University Press, 1995.

Katok, A. and L. Mendoza. Dynamical systems with nonuniformly hyperbolic behaviour. In A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems (Encyclopaedia of Mathematics and Its Applications, 54)*. Cambridge University Press, 1995, appendix.

Kingman, J. F. C. The ergodic theory of subadditive stochastic processes. *J. R. Stat. Soc.* B **30** (1968), 499–510.

Ljung, S. and L. Ljung. Error propagation properties of recursive least-squares adaptation algorithms. *Automatica* **21** (1985), 157–167.

Mather, J. N. Characterization of Anosov diffeomorphisms. *Indagationes Mathematicae* **30** (1968), 479–483.

Nicol, M. and D. S. Broomhead. Random perturbation of skew product systems. *Dynamics and Stability of Systems* **14**(1), (1999) to appear.

Pecora, L. M. and T. L. Carroll. Synchronisation in chaotic system. *Phys. Rev. Lett.* **64** (1990), 821–824.

Pecora, L. M. and T. L. Carroll. Driving systems with chaotic signals. *Phys. Rev.* A **44** (1991), 2374–2383.

Pecora, L. M. and T. L. Carroll. Discontinuous and nondifferentiable functions and dimension increase induced by filtering chaotic data. *Chaos* **6** (1996), 432–439.

Pesin, Ya. B. Families of invariant manifolds corresponding to nonzero characteristic exponents. *Math. USSR Izvestija* **10** (1976), 1261–1305.

Pesin, Ya. B. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Math. Surveys* **32** (1977), 55–112.

Regalia, P. A. Numerical stability properties of a QR-based fast least squares algorithm. *IEEE Trans. Sig. Proc.* **41** (1993), 2096–2109.

Ruelle, D. *Elements of Differentiable Dynamics and Bifurcation Theory*. Academic Press, 1989.

Slock, D. T. Backward consistency concept and round-off error propagation dynamics in recursive least-squares algorithms. *Opt. Eng.* **31**, 1153–1169.

Stark, J. Recursive prediction of chaotic time series. *J. Nonlinear Sci.* **3** (1993), 197–223.

Stark, J. Invariant graphs for forced systems. *Physica* D **109** (1997), 163–179.

Stark, J. and M. E. Davies. Recursive filters driven by chaotic signals. *IEE Colloquium on Exploiting Chaos in Signal Processing, IEE Digest* **143** (1994), 5/1–5/16.

Stein, E. M. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.

Tanny, D. A zero-one law for stationary sequences. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **30** (1974), 139–148.

Temam, R. *Infinite Dimensional Dynamical Systems in Mechanics and Physics (Applied Mathematics Series, 68)*. Springer, 1988.

Whitney, H. Analytic extensions of differential functions defined in closed sets. *Amer. Math. Soc. Trans.* **36** (1934), 63–89.

https://doi.org/10.1017/S0143385799126555 Published online by Cambridge University Press