Journal of Functional Analysis **182**, 108–150 (2001)
doi:10.1006/jfan.2000.3723, available online at http://www.idealibrary.com on IDEAL®

Local Contractions of Banach Spaces and Spectral Gap Conditions

Mohamed S. ElBialy

Department of Mathematics, University of Toledo, Toledo, Ohio 43606
E-mail: melbialy@math.utoledo.edu

Communicated by L. Gross

Received March 9, 2000; revised November 20, 2000; accepted November 27, 2000

We study the linearization problem for a $C^{k,1}$, $k \geqslant 1$, contraction of a Banach space E near a fixed point which satisfies a spectral gap condition and a narrow band condition both of order k. We also assume that the part of the spectrum in each band satisfies a finite non-resonant condition of order k relative to itself together with the part that lies in the larger bands. We show that there is a $C^{k,\beta}$ linearization for sufficiently small $\beta > 0$. We give a precise estimate on β in terms of the gap and band conditions. © 2001 Academic Press

Key Words: linearization; infinite dimensional dynamical systems; contractions; invariant manifolds; weak-stable manifold.

1. INTRODUCTION

The linearization problem for a map of a Banach space near a hyperbolic fixed point is to find a C^k diffeomorphism in a neighbourhood of the fixed point that sends the map to its linear part. In this case we say that the map is C^k-linearizable. This problem plays an important rule in dynamical systems. It has been under investigation for many decades. In this work we study C^k partial linearization and linearization of local contractions Φ of a Banach space near a fixed point. Partial linearization is considered when the map is not invertible. We also consider local contractions of $\mathbb{R}^n$.

We will see that the smoothness of the linearizing transformation is sensitive to the *spectral gaps* in the spectrum of the linear part of the map at the fixed point.

We will see that many of the results that exist in the literature on strictly hyperbolic diffeomorphisms or vector fields in $\mathbb{R}^n$ lead to vacuous conditions when restricted to contractions. These results are stated in terms of *spectral width* conditions rather than *spectral gaps* conditions. We will elaborate on that in Subsection 3.2. Contractions of $\mathbb{R}^n$ were studied

0022-1236/01 $35.00

directly, and not by restricting statements on the strictly hyperbolic case, by Poincaré, Sternberg, and Hartman:

- Poincaré showed that if a contraction is analytic with analytic inverse and the fixed point is nonresonant, then there is an analytic linearization. See [1].
- In [25], Sternberg showed that if a contraction is C^∞ with C^∞ inverse and the fixed point is nonresonant, then there is a C^∞ linearization [25].
- In [12], Hartman showed that any $C^{1,1}$ contracting diffeomorphism is $C^{1,\beta}$ linearizable for sufficiently small $0<\beta<1$. The gap condition (A1) of Theorem 3 is trivially satisfied when $k=1$.

There are a few intricacies associated with this problem:

(1) The main intricacy of this problem is that one has to work with maps that vanish up to order k at the origin but defined on the whole space and their kth derivative has mixed differentiability which changes at each step of an induction. For these maps, the standard $C^{k,\delta}$-norm is not the appropriate norm to use. In Section 6.1 we introduce the Banach spaces of functions which are appropriate for this problem and prove Theorem 5 which is the centre piece in the proof of the main Theorem 1.

(2) When considering maps of an infinite dimensional Banach space one has to pay special attention to a few issues that do not arise in finite dimensional spaces, namely:

- An infinite dimensional Banach space does not always admit a smooth cut-off function [14, 16]. We will elaborate on this issue latter.
- In a finite dimensional space, the spectrum of the linearization $D\Phi(0)$ of the map at the fixed point $x=0$ lies on a finite number of circles and possibly $0\in\mathbb{C}$. In this case, it is enough to state our theorem in terms of gap conditions. In an infinite dimensional space we would like the spectrum to lie inside a finite number of bands and possibly in the interior of a small disc around $0\in\mathbb{C}$. In this case the gap conditions are not enough. We need two more conditions, one that relates the widths of the bands to the gap conditions, and the other requires that the part of the spectrum inside each band satisfies a finite nonresonant condition of order k relative to itself together with the part that lies in the larger bands.
- The fact that we can eliminate nonresonant terms from maps of infinite dimensional Banach spaces requires a more technical proof than in the case of finite dimensional spaces. A proof of this statement was given first in [15]. See also [10]. We include a proof in Section 8 for completeness.

In this work we prove two main theorems. Theorem 1 deals with partial linearization of a contraction when part of the spectrum lies inside a small disc around $0 \in \mathbb{C}$ and Theorem 2 deals with linearization when the map is a diffeomorphism. Theorem 2 is a special case of Theorem 1. We break down the proof of Theorem 1 into several parts each of which is proved under the minimum hypotheses needed. In this way we can see the structure of the proof. This paper is organized as follows:

- In Section 2 we state the main results.

- In Section 3 we restate our main results for maps of $\mathbb{R}^n$. In this case we do not need the band condition. In Subsection 3.2 we show that many of the results that exist in the literature on strictly hyperbolic diffeomorphisms or vector fields in $\mathbb{R}^n$ lead to vacuous conditions when restricted to contractions of $\mathbb{R}^n$. These results are stated in terms of *spectral width* conditions rather than *spectral gaps* conditions.

- In Section 4 we state two general theorems, Theorem 4 and Theorem 5. Theorem 4 gives certain estimates that arise from the existence of an invariant manifold under some additional conditions that can be obtained from finite nonresonance conditions. These estimates are independent of the type of the invariant manifold (stable, unstable, weak-stable, ...) and of the method used to show its existence. Theorem 5, which is really the heart of the proof of Theorem 1, is a general elimination theorem for any map that satisfies the estimates in the conclusion of Theorem 4. We prove Theorem 4 in Section 7 and Theorem 5 in Section 6 in order not to obstruct the flow of ideas.

- In Section 5 we prove the main Theorem 1 granting Theorems 4 and 5. We use a weak-stable manifold theorem under gap conditions. See for example Theorem 5.1 in [13, p. 53]. However, we need a weak-stable manifold for maps that satisfy the vanishing properties (5.14) and (5.15) below. These vanishing properties are essential for the linearization step in Theorem 5. We prove such a theorem as part of a more general theorem in [9] and we state it as Theorem 6 in the present notation.

- In Section 6 we prove Theorem 5. This is really the corner stone of the proof of Theorem 1 where we define the appropriate Banach space for this problem. This requires finding norms and semi-norms that are appropriate for this problem.

- In Section 7 we prove Theorem 4.

- In Section 8 we study the homological equation in a Banach space to eliminate nonresonant symmetric j-multilinear terms. We prove the first assertion of Lemma 5.4. This assertion was proved first in [15]. See also [9].

2. MAIN RESULTS

2.1. *The Map.* Let E be a Banach space and $\Phi: E \to E$ be a $C^{k,1}$, $k \geqslant 1$, map with a fixed point. Without loss of generality we may assume that the fixed point is $0 \in E$ and that the map takes the form $\Phi(x) = x^*$ where

$$\begin{aligned} x^* &= \Lambda x + F(x) \\ F(0) = 0, &\qquad DF(0) = 0 \\ D\Phi(0) &= \Lambda. \end{aligned} \tag{2.1}$$

We assume that Φ is a contraction near $0 \in E$. That is, $0 \leqslant |\lambda| < 1$ for all $\lambda \in \sigma(\Lambda)$, where $\sigma(\Lambda)$ is the spectrum of Λ.

The linearization problem is to find a $C^{k,\delta}$ diffeomorphism of a neighbourhood of $0 \in E$ which transforms (2.1) to the map

$$z^* = \Lambda z.$$

2.2. *Assumptions and Setup.* (1) Assume the spectrum of Λ takes the form

$$\begin{aligned} \sigma(\Lambda) &= \bigcup_{I=0}^{J} \sigma_I \\ \sigma_I &\subset \{\zeta \in \mathbb{C} \mid \chi_I < |\zeta| < b_I\}, \qquad I = 1, 2, ..., J \\ \sigma_o &\subset \{\zeta \in \mathbb{C} \mid 0 \leqslant |\zeta| < b_o\} \\ 0 < b_o < \chi_1 &< b_1 < \chi_2 < b_2 < \cdots < \chi_J < b_J < \gamma < 1. \end{aligned} \tag{2.2}$$

(2) By the Spectral Decomposition Theorem we may assume that

$$\begin{aligned} E &= E_o \times E_1 \times \cdots \times E_J \\ \Lambda &= \text{blockdiag}(\Lambda_o, \Lambda_1, ..., \Lambda_J) \\ x &= (x_o, x_1, ..., x_J) \in E_o \times E_1 \times \cdots \times E_J \\ \sigma(\Lambda_I) &= \sigma_I, \qquad I = 0, 1, ..., J \\ F &= (F_o, F_1, ..., F_J) \end{aligned}$$

and that the map takes the form

$$x_I^* = \Lambda_I x_I + F_I(x), \qquad I = 0, 1, ..., J. \tag{2.3}$$

(3) Let $\hat{\Lambda} = \text{blockdiag}(\Lambda_1, ..., \Lambda_J)$. We say that $\sigma(\hat{\Lambda})$ satisfies a gap condition of order k, which we denote by $G^{(k)}(\hat{\Lambda})$, iff

$$b_I < (\chi_{I+1})^k, \qquad I = 1, 2, ..., J. \tag{2.4}$$

Notice that the gap condition does not depend on σ_o.

(4) For $I = J, ..., 1$, let $0 < s_I < 1$. Let $\delta_J = 1$. For $I = J-1, ..., 1, 0$, let $\delta_I = \delta_{I+1}(1 - s_{I+1})$. Thus $\delta_o = (1 - s_1) \cdots (1 - s_J)$.

We say that $\sigma(\hat{\Lambda})$ satisfies a band condition of order k, which we denote by $B^{(k)}(\hat{\Lambda})$, iff the following holds for $I = J, J-1, ..., 2, 1$,

$$\theta_o(\chi_I, b_I, \delta_I) := \frac{b_I}{\chi_I} \gamma^{k + \delta_I - 1} < 1 \tag{2.5}$$

$$\theta_1(\chi_I, b_I, \delta_I, s_I) := \frac{b_I^{s_I}}{\chi_I} \gamma^{k + \delta_{I-1}} < 1 \tag{2.6}$$

$$\delta_{I-1} = \delta_I(1 - s_I).$$

We would like to point out that the exponent of γ in (2.5) is $(k + \delta_I - 1)$ while in (2.6) it is $(k + \delta_{I-1})$. There is no typographical error. Notice that the band condition does not depend on σ_o.

(5) Let A and B be two linear maps. The linear map B is said to satisfy be k-nonresonant with respect to A, which we denote by $NR_k(B; A)$, iff for all $2 \leqslant j \leqslant k$

$$\beta \neq \alpha_1 \alpha_2 \cdots \alpha_j, \qquad \beta \in \sigma(B), \alpha_i \in \sigma(A), \qquad 1 \leqslant i \leqslant j. \tag{2.7}$$

In other words,

$$\sigma(A)^j \cap \sigma(B) = \emptyset, \qquad 2 \leqslant j \leqslant k,$$

where for $\sigma \subset \mathbb{C}$

$$\sigma^j = \{\alpha_1 \alpha_2 \cdots \alpha_j \mid \alpha_i \in \sigma, 1 \leqslant i \leqslant j\}.$$

If B satisfies $NR_k(B; B)$, we say that it satisfies a self nonresonant condition of order k and write $NR_k(B)$. If B satisfies $NR_k(B)$, for all $k \geqslant 2$, it is said to be nonresonant.

(6) We use adapted norms on each E_I so that

$$\|\Lambda_I\| < b_I, \qquad I = 0, 1, ..., J$$

$$\|\Lambda_I^{-1}\| < \chi_I^{-1}, \qquad I = 1, 2, ..., J.$$

Before we can carry on with our discussion we need to define a certain mixed differentiability.

2.3. Definition. For a map $\psi(\eta, z)$ define

$$H_\delta(\psi) = \sup\left\{\frac{|\psi(\eta, z)-\psi(\eta', z')|}{|\eta-\eta'|^\delta+|z-z'|^\delta}\,\middle|\,(\eta, z)\neq(\eta', z')\right\}$$

$$H_\delta^{(z)}(\psi) = \sup\left\{\frac{|\psi(\eta, z)-\psi(\eta, z')|}{|z-z'|^\delta}\,\middle|\,z\neq z'\right\}$$

$$\mathrm{Lip}^{(\eta)}(\psi) = \sup\left\{\frac{|\psi(\eta, z)-\psi(\eta', z)|}{|\eta-\eta'|}\,\middle|\,\eta\neq\eta'\right\}.$$

We can define $\mathrm{Lip}(\psi)$, $\mathrm{Lip}^{(z)}(\psi)$ and $H_\delta^{(\eta)}(\psi)$ similarly.

Theorem 1. *Consider a $C^{k,1}(E, E)$ contraction $\Phi = \Lambda + F$ given by* (2.3) *in the setup of Section* 2.2. *Let $k \geqslant 1$ be an integer. Assume the following*:

(H1) *If $k\geqslant 2$, assume that $\sigma(\hat{\Lambda})$ satisfies a gap condition $G^{(k)}(\hat{\Lambda})$ given by* (2.4).

(H2) *If $k\geqslant 2$, assume that for all $1\leqslant I\leqslant J$, $(\Lambda_I, \Lambda_I\times\cdots\times\Lambda_J)$ satisfy a k-nonresonant condition $NR_k(\Lambda_I; \Lambda_I\times\cdots\times\Lambda_J)$ given by* (2.7).

(H3) *$\sigma(\hat{\Lambda})$ satisfies the band condition $B^{(k)}(\hat{\Lambda})$ given by* (2.5)–(2.6).

(H4) *$\|D^jF\|<\varepsilon$, $j=0, 1, ..., k$ and $\mathrm{Lip}(D^kF)<\varepsilon$.*

Let $\beta=\delta_o=(1-s_1)\cdots(1-s_J)$. Then the following are true for sufficiently small $\varepsilon>0$:

(1) *There is a $C^{k,\beta}$ diffeomorphism*

$$x=(x_o, x_1, ..., x_J)=Q(\eta, z_1, ..., z_J)$$

in a neighbourhood of $0\in E$ which sends the map (2.3) *to the map*

$$\begin{aligned}\eta^* &= \Lambda_o\eta + G(\eta, z)\\ z_I^* &= \Lambda_I z_I, \qquad I=1, 2, ..., J.\end{aligned}\tag{2.8}$$

(2) *If $b_o<(\chi_1)^k$, the $G(0, z)=0$. That is, $\{\eta=0\}$ is invariant.*

(3) *$H_\beta^{(z)}(D^kG)<\infty$ and $H_\beta^{(z)}(D^kQ)<\infty$.*

(4) *$\mathrm{Lip}(D_\eta^kG)<\infty$ and $\mathrm{Lip}(D_\eta^kQ)<\infty$.*

(5) *$\mathrm{Lip}^{(\eta)}(D^kG)<\infty$ and $\mathrm{Lip}^{(\eta)}(D^kQ)<\infty$.*

THEOREM 2. *Assume that Φ in Theorem 1 is a diffeomorphism and that $\sigma_o = \emptyset$. (Thus $E = E_1 \times \cdots \times E_J$.) Then the $C^{k,\beta}$ diffeomorphism of Theorem 1 sends (2.3) to*

$$z^* = \Lambda z.$$

2.4. *Remarks.* (1) Theorem 2 is a special case of Theorem 1. So, we prove Theorem 1.

(2) In the $C^{1,1}$ case the gap condition is automatically satisfied. Moreover, we do not need hypotheses (H2)–(H3). In [12], Hartman proves the existence of a $C^{1,\beta}$ linearization in finite dimensional spaces. In a finite dimensional space, we do not need the band condition (H3) since the spectrum lies on a finite number of circles and possibly inside a small disc around $0 \in \mathbb{C}$.

(3) The band conditions mean that each one of $\sigma_1, ..., \sigma_{J-1}$ and σ_J has to lie in a sufficiently narrow band.

(4) The smallness assumptions (H4) on $F(x)$ are essential for maps of infinite dimensional Banach spaces and need some discussion: At a certain point of the proof we will need to invoke Theorem 6 on the existence of a weak-stable manifold under a gap condition. The smallness assumptions on $F(x)$ are needed for such a theorem. This is the only place in the proof where they are needed.

First let us see how one can avoid the smallness assumptions in finite dimensional spaces: In a finite dimensional space we can replace "for sufficiently small $\varepsilon > 0$" in the conclusion by "in a sufficiently small closed ball $B_\alpha(0)$." Then we can obtain the smallness assumptions as follows:

(a) For $\alpha > 0$, we multiply $F(x)$ by a smooth cut-off function which vanishes outside a closed ball $B_{\alpha'}(0)$, $\alpha' < \alpha$. We will continue to denote the new nonlinear part of the map by $F(x)$.

(b) For $0 < \alpha < 1$ define

$$\Phi^{(\alpha)}(x) = \frac{1}{\alpha}\,\Phi(\alpha x) = D\Phi(0)\,x + F^{(\alpha)}(x)$$

$$= D\Phi(0)\,x + \frac{1}{\alpha}\,F(\alpha x), \qquad |x| \leqslant 1.$$

Straightforward calculations show that a diffeomorphism of the form

$$y = x + w(x), \qquad |x| \leqslant \alpha$$

partially linearizes the map Φ as in Theorem 1 iff the diffeomorphism

$$y = w^{(\alpha)}(x), \qquad |x| \leqslant 1$$

does the same for the map $\Phi^{(\alpha)}$ where

$$w^{(\alpha)}(x) = \frac{1}{\alpha} w(\alpha x), \qquad |x| \leqslant 1.$$

(c) Thus for all $\varepsilon > 0$ we can take $\alpha > 0$ sufficiently small so that our new map satisfies (H4).

We can perform step (b) in an infinite dimensional Banach space. However, we cannot perform step (a) because an infinite dimensional Banach space does not necessarily admit a smooth cutoff function. In [14], it is shown that the space of continuous functions on an interval does not admit a C^1 cut-off function. In [16], it is shown that a separable Banach space admits a Fréchet differentiable cut-off function iff its dual space is separable.

In general, we can perform step (a) in any space that admits smooth cut-off functions. Any space in which the map $x \mapsto |x|$ is differentiable away from the origin will admit cut-off functions which are as smooth as the map $|\cdot|$.

3. A SPECIAL CASE: CONTRACTIONS OF $\mathbb{R}^n$

When our map is a contraction of $\mathbb{R}^n$, we do not need many of the hypotheses in Theorem 1. The hypothesis that we need is the gap condition. For the sake of precision, we restate Theorems 1 and 2 for maps of $\mathbb{R}^n$, then we make a few comments on the hypotheses.

THEOREM 3. *Consider a $C^{k,1}$ map Φ: $\mathbb{R}^n \to \mathbb{R}^n$ which takes the following form on $\mathbb{R}^n = \mathbb{R}^{N_o} \times \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_J} = \mathbb{R}^{N_o} \times \mathbb{R}^N$, $N = N_1 + \cdots + N_J$:*

$$\begin{aligned}
x_o^* &= \Lambda_o \eta + F_o(x) \\
x_I^* &= \Lambda_I x_I + F_I(x), \qquad I = 1, 2, ..., J \\
F(0) &= 0, \qquad DF(0) = 0 \\
\sigma(\Lambda) &= \bigcup_{I=0}^{J} \sigma_I \qquad\qquad (3.1)\\
\sigma_o &= \sigma(\Lambda_o) \subset \{\zeta \in \mathbb{C} \mid 0 \leqslant |\zeta| < b_o\} \\
\sigma_I &= \sigma(\Lambda_I) \subset \{\zeta \in \mathbb{C} \mid |\zeta| = \mu_I\}, \qquad I = 1, 2, ..., J \\
0 &< b_o < \mu_1 < \mu_2 < \cdots < \mu_J < \gamma < 1.
\end{aligned}$$

[A1] *If* $k \geqslant 2$, assume *that* $\sigma(\hat{\Lambda})$ *satisfies a gap condition* $G^{(k)}(\hat{\Lambda})$:

$$\mu_I < (\mu_{I+1})^k, \qquad I = 1, 2, ..., J-1,$$

where $\hat{\Lambda} = \text{blockdiag}(\Lambda_1, ..., \Lambda_J)$.

Let $B_\alpha(0) \subset \mathbb{R}^n$ *be the closed ball with radius* α *centered at the origin.* *The following are true for sufficiently small* $\alpha > 0$ *and sufficiently small* $0 \lneqq \beta \lneqq 1$ *which depends on* $\mu_1, ..., \mu_{J-1}$ *and* μ_J:

(1) *There is a* $C^{k,\beta}$ *diffeomorphism* $x = Q(\eta, z)$, $(\eta, z) \in B_\alpha(0) \times B_\alpha(0) \subset \mathbb{R}^{N_o} \times \mathbb{R}^N$, *which sends* (3.1) *to the map* $(\eta, z) \mapsto (\eta^*, z^*)$ *where*

$$\begin{aligned} \eta^* &= \Lambda_o \eta + G(\eta, z) \\ z^* &= \hat{\Lambda} z. \end{aligned} \tag{3.2}$$

(2) *If in addition* $b_o < \mu_1^k$, *then* $G(0, z) = 0$.

(3) $H_\beta^{(z)}(D^k G) < \infty$ *and* $H_\beta^{(z)}(D^k Q) < \infty$.

(4) $\text{Lip}(D_\eta^k G) < \infty$ *and* $\text{Lip}(D_\eta^k Q) < \infty$.

(5) $\text{Lip}^{(\eta)}(D^k G) < \infty$ *and* $\text{Lip}^{(\eta)}(D^k Q) < \infty$.

(6) *Assume that* Φ *is a diffeomorphism near the fixed point, and hence* $\sigma_o = \varnothing$. *Then the* $C^{k,\beta}$ *diffeomorphism* Q *sends the map* (3.1) *to*

$$z^* = \Lambda z.$$

3.1. *Remarks.* (1) In the finite dimensional case, $\sigma(\Lambda)$ is a subset of a finite number of circles and, possibly, a small disc centered at $0 \in \mathbb{C}$.

(2) Thus we do not need to assume the nonresonant condition [H2] because it follows from the gap condition [A1].

(3) We do not need to assume the band condition [H3] because we can always choose $\chi_I < \mu_I < b_I$ close enough to μ_I so that (2.5) and (2.6) hold. Then we can construct the Hölder exponent β as in Theorem 1.

(4) Since smooth cutoff functions on $\mathbb{R}^n$ exist, we also do not need to assume the smallness condition [H4]. We can use the method of Remark 2.4.4(b) to obtain the smallness condition we need. Then we replace the "for sufficiently small $\varepsilon > 0$" in the conclusion of Theorem 1 by "for sufficiently small $\alpha > 0$" where α is the radius of a ball centered at $0 \in \mathbb{R}^n$.

3.2. *Classical Results on Contractions of* $\mathbb{R}^n$. In the following we summarize a few classical results on linearizing contractions of $\mathbb{R}^n$. Only Poincaré, Sternberg and Hartman addressed contractions directly. The rest studied the strictly hyperbolic case and then restricted their statements to contractions. We will see that those restricted statements are given in terms of *spectral width* conditions and, with the exception of Bryuno's theorem [6], lead to vacuous statements when the contracting rest point is resonant.

a. *The work of Poincaré, Sternberg, and Hartman.*

(1) Poincaré showed that if a contraction is analytic with analytic inverse and the fixed point is nonresonant, then there is an analytic linearization. Poincaré wrote a formal power series than showed that it converges. See [1].

(2) In [25], Sternberg showed that if a contraction is C^∞ with C^∞ inverse and the fixed point is nonresonant, then there is a C^∞ linearization. We will elaborate on the implications of his work on the resonant case later.

(3) In [11], Hartman showed that any $C^{1,1}$ contracting diffeomorphism is $C^{1,\beta}$ linearizable for sufficiently small $0<\beta<1$. The gap condition (A1) of order $k=1$ of Theorem 3 is trivially satisfied. Hartman's theorem is a special case of our Theorem 3.

(4) It is known that if a fixed point has a resonance of order 2, a C^2-linearization may not exist. In [25], Sternberg showed that the following contraction on $\mathbb{R}^2$ does not have a C^2 linearization for any $0<\alpha<1$,

$$\begin{aligned} x^* &= a^2x + y^2 \\ y^* &= ay. \end{aligned} \tag{3.3}$$

Notice that $\lambda_1 = a^2 = \lambda_2^2$. Thus the term y^2 is resonant and cannot be eliminated via a polynomial change of variables. On the other hand, we know from Hartman's work [11] that it is $C^{1,\beta}$ linearizable, for sufficiently small $\beta>0$.

b. *Example.* In order to make our discussion more concrete we consider the following simple diffeomorphism which is a generalization of the map (3.3). Let $p \geqslant 1$ be an integer, $0<a<1$, $\mu_1 = a^{p+1}$, $\mu_2 = a$. Consider the map

$$\begin{aligned} x^* &= a^{p+1}x + y^{p+1} \\ y^* &= ay. \end{aligned} \tag{3.4}$$

Notice that this map satisfies a *gap condition of order p*. However, it has a *resonance of order* $p+1$, namely, $\mu_1 = \mu_2^{p+1}$. Thus, the map satisfies only a *nonresonance condition* $NR(p)$. Moreover, the resonant monomial y^{p+1} is present. Hartman's theorem [11], which is a special case of our theorem, guarantees a $C^{1,\beta}$ linearization for sufficiently small $\beta > 0$. Our theorem guarantees a $C^{p,\beta}$ linearization for sufficiently small $\beta > 0$.

c. *Results that use* spectral width *conditions.* In the literature on linearization of contracting diffeomorphisms and vector fields near a hyperbolic fixed point in $\mathbb{R}^n$ one finds the following statements which we state in our notation for diffeomorphisms. Let $\sigma(\Lambda) = \{a_1, ..., a_n\} \subset \mathbb{C}$, and $0 < \mu_1 < \mu_2 \cdots < \mu_N < 1$ be the distinct ones among $\{|a_1|, ..., |a_n|\}$.

(1) The following statement appears in [25] where Sternberg is dealing with a nonresonant C^∞ contraction:
Assume that the map is C^k and the fixed point satisfies $NR(k)$. If

$$(\mu_N)^k < \mu_1, \tag{St}$$

then the contraction is C^k linearizable.

Sternberg's strategy is to first eliminate all monomials of degree $2 \leqslant d \leqslant k$ (all of which are nonresonant) via a polynomial change of variables, and then use (St) to eliminate the *tail* of the map, so to speak. He can afford to eliminate all monomials up to arbitrary high degree k and achieve the inequality (St) because he is dealing with a nonresonant fixed point of a C^∞ map.

The map (3.4) has a resonance of order $p+1$. Thus, we cannot use *Sternberg's strategy* for any $k \geqslant p+1$. For $1 \leqslant k \leqslant p$, since the map has no *tail*, we check inequality (St):

$$(\mu_2)^k = a^k \geqslant a^p > a^{p+1} = \mu_1.$$

Thus, condition (St) is not satisfied and a C^k linearization is not guaranteed even for $1 \leqslant k \leqslant p$ and hence for any $k \geqslant 1$.

All the work mentioned below follow *Sternberg's strategy* with the exception of Bryuno's which follows the opposite strategy of eliminating the *tail* of a C^∞ vector field first and than paying attention to the resonant terms.

(2) If we restrict the condition given by Nelson in [18] to vector fields with an attracting fixed point we obtain the vector field version of (St).

(3) In [2], Banyaga *et al.* study diffeomorphisms that preserve volume, contact or symplectic structures. In the case of a contraction that does not preserve any structure they obtain a sufficient condition which is slightly weaker than (St). Namely:

Assume that the map is C^k and the fixed point satisfies $NR(k)$. If

$$(\mu_N)^{k+1} < \mu_1, \tag{BLW}$$

then the contraction is C^k linearizable.

BL&W use Sternberg's strategy of eliminating all monomials of degree $2 \leqslant d \leqslant k$ first. As above, since the map (3.4) has a resonance of order $p+1$, it is enough to investigate $1 \leqslant k \leqslant p$:

$$(\mu_2)^{k+1} = a^{k+1} \geqslant a^{p+1} = \mu_1.$$

Thus, condition (BLW) is not satisfied and we cannot guarantee C^k linearization for $1 \leqslant k \leqslant p$ and hence for $k \geqslant 1$.

(4) In [24], Sell states the vector field version of the following statement:

Assume that the map is C^{3k} and that the fixed point satisfies $NR(k)$. If

$$(\mu_N)^k \leqslant (\mu_1)^m, \tag{Se}$$

then the contraction is C^m linearizable.

The map (3.4) satisfies only a nonresonance condition $NR(p)$. Now with $k = p$ and $m \geqslant 1$,

$$(\mu_2)^k = a^p > a^{m(p+1)} = (\mu_1)^m.$$

Hence, a C^m linearization is not guaranteed for $m \geqslant 1$.

(5) In [6], Bryuno follows a strategy which is opposite to *Sternberg's strategy*. He starts with a C^∞ vector field in $\mathbb{R}^n$ and invokes a theorem from [5] to conclude that one can find a C^k change of variables that eliminates all the nonlinearity except for possible resonant monomials of degrees $2 \leqslant d \leqslant nk$. That is, he eliminates the *tail* first then pays attention to the resonant terms. He has the following condition $T(k)$ which we write for a contracting diffeomorphism where $N = n$, i.e., $0 < \mu_1 \leqslant \cdots \leqslant \mu_n < 1$, $\mu_j = |a_j|$, $1 \leqslant j \leqslant n$:

$T(k)$. There exists a C^k change of variables that eliminates the remaining resonant monomials if the following inequality holds for every resonant monomial that is actually present

$$\begin{aligned} &\underline{\mu}^{\underline{m}} := \mu_1^{m_1} \cdots \mu_n^{m_n} < \mu_1^k, \qquad |\underline{m}| \leqslant nk \\ &|\underline{m}| := m_1 + \cdots + m_n \end{aligned} \tag{Br}$$

For the map (3.4) we have $n = 2$, and the resonant term y^{p+1} corresponds to the multi-index $\underline{m} = (0, p+1)$, $|m| = p+1$. Recall that $p \geqslant 1$

is an integer. In order to test (Br), let k be such that $|m| = p+1 \leqslant 2k = nk$. Then, $k \geqslant 1$ and

$$\mu_1^0 \mu_2^{p+1} = a^{p+1} \geqslant a^{k(p+1)} = \mu_1^k,$$

where equality holds iff $p = k = 1$. Thus, (Br) cannot be satisfied for $k \geqslant (p+1)/2$ and hence, a C^k linearization, $k \geqslant (p+1)/2$, is not guaranteed. Thus, Bryuno's theorem guarantees a C^k linearization only for $k < (p+1)/2$. For such a k, the resonant term y^{p+1} will be part of the tail that has been eliminated already via a C^k change of variables. Recall that our theorem guarantees a $C^{p,\beta}$ linearization for sufficiently small $\beta > 0$, which is an improvement over Bryuno's theorem for any $p \geqslant 1$. Also notice that when $p = 1$, i.e., the map (3.3), Bryuno's theorem does riot guarantee a C^1 linearization, although we know that the map (3.3) has a $C^{1,\beta}$ linearization for sufficiently small $\beta > 0$.

d. *Comments.* The following comments do not apply to the work of Poincaré. That is, *all* will mean *all except Poincaré*.

(1) All the above mentioned work, with the exception of Hartman's, rely on *spectral width conditions*. Hartman uses a *gap condition*, which is trivially satisfied for the $C^{1,1}$ case, and obtains a non-vacuous statement.

(2) In general, assume that we have a fixed point that satisfies $NR(p)$ but has a resonance of order $p+1 \geqslant 2$. That is, for some $1 \leqslant j \leqslant n$ we have

$$a_j = a_1^{m_1} a_2^{m_1} \cdots a_n^{m_n}, \qquad |\underline{m}| = p+1.$$

We consider first conditions (St), (Se), and (BLW). These three conditions are used after eliminating, via a polynomial change of variables, all monomials of degrees $2 \leqslant d \leqslant p$, all of which are nonresonant. Thus, we cannot hope for a C^k linearization with $k > p$ since we are assuming that a resonant term of order $p+1$ is present. Thus, it is enough to examine (St), (Se), and (BLW) for $1 \leqslant k \leqslant p$: For $m \geqslant 1$,

$$\mu_1 \leqslant (\mu_N)^{p+1} \leqslant (\mu_N)^{k+1}, \qquad 1 \leqslant k \leqslant p$$

$$(\mu_1)^m \leqslant \mu_1 \leqslant (\mu_N)^{p+1} < (\mu_N)^p \leqslant (\mu_N)^k, \qquad 1 \leqslant k \leqslant p.$$

It follows that (St), (Se), and (BLW) can never be attained for any $1 \leqslant k \leqslant p$ and hence, a C^k linearization with $k \geqslant 1$ cannot be guaranteed.

(3) As for condition (Br), if $|\underline{m}| \leqslant nk$, $n \geqslant 2$, we have

$$\mu_1^{m_1} \mu_2^{m_1} \cdots \mu_n^{m_n} = \mu_j \geqslant \mu_1 \geqslant \mu_1^p \geqslant \mu_1^k, \qquad k \geqslant p.$$

This shows that (Br) also can never be attained for a resonant fixed point satisfying $NR(p)$ but has a resonant of order $p+1$, and no C^k linearization, $k \geqslant p$, is guaranteed.

However, According to Bryuno's theorem, we can first eliminate a *tail* that includes the resonant terms of order $p+1$. That is, we choose an integer k such that $nk < p+1$. Since the map satisfies $NR(p)$, there is no resonant terms of order $i \leqslant nk < p+1$. Now, (Br) is trivially satisfied since it is required to hold only for resonant terms. Thus, the map is C^k linearizable for $k < (p+1)/n$. Notice that the k decreases when the dimension n increases.

Notice that a contracting map that satisfies a gap condition of order p automatically satisfies $NR(p)$. In view of our theorem, we can see that Bryuno's theorem suffers from a large loss of differentiability which gets worse if the dimension increases.

(4) It follows that all the above mentioned work, with the exception of Hartman's and Bryuno lead to vacuous statements for a resonant fixed point.

(5) All the above mentioned work, with the exception of that of Bryuno, follow *Sternberg's strategy* of assuming that the fixed point satisfies a nonresonant condition $NR(k)$, eliminating all monomials of degree $2 \leqslant d \leqslant k$ via a polynomial change of variables first, and than eliminating the *tail* with the aid of a *spectral width* condition (except for Hartman who uses *spectral gap* condition). Bryuno eliminates the *tail* first then eliminates the resonant monomials with the aid of a *spectral width* condition also. The nonresonant ones can always be eliminated via a polynomial change of variables.

(6) For the famous example (3.3) none of the above mentioned work, other than Hartman's guarantees a C^1 linearization.

(7) As for example (3.4), which is a generalization of example (3.3), the theorem of Bryuno guarantees a C^k linearization for $k < (p+1)/2$ when $p \geqslant 2$. If $p = 1$, even a C^1 linearization is not guaranteed. Our theorem guarantees a $C^{p,\beta}$ linearization, for sufficiently small $\beta > 0$ and any $p \geqslant 1$. Hartman's theorem guarantees a $C^{1,\beta}$ linearization, for sufficiently small $\beta > 0$ and any $p \geqslant 1$. The rest are vacuous.

(8) In view of the above we can see that there is no satisfactory theorem that provides a sufficient condition for C^k, $k \geqslant 2$, (partial) linearization near a resonant fixed point for contracting (maps) diffeomorphisms. Theorem 3 provides such a condition.

4. TWO GENERAL THEOREMS

In this section we state two theorems that are of general interest because they can be used in other situations. We give their proofs latter. The first, Theorem 4 gives some estimates that arise from the existence of an invariant manifold. These estimates are independent of the type of the invariant manifold and of the method used to show its existence. The second, Theorem 5 is really the heart of the proof of Theorem 1. It is a general elimination theorem for any map that satisfies the estimates in the conclusion of Theorem 4. We prove it in Section 6 where we define the the norms, semi-norms and Banach spaces that are appropriate for this problem.

4.1. *A General Estimate Resulting from the Existence of an Invariant Manifold.* (1) Let $E=X\times Y\times Z$ be a Banach space. Consider a $C^{k,\delta}$, $0\leqslant\delta\leqslant 1$, map

$$\begin{aligned} x^* &= Ax + R(x, y, z)\\ y^* &= By + S(x, y, z)\\ z^* &= Cz + T(x, y, z). \end{aligned} \tag{4.1}$$

(2) Let $\eta=(x, y)$, $F=(R, S)$ and $M=\text{blockdiag}(A, B)$.

(3) Assume that

$$\|D^jF\|<\varepsilon, \qquad \|D^jT\|<\varepsilon, \qquad 0\leqslant j\leqslant k \tag{4.2}$$

$$\begin{aligned} H_\delta(D^kF)&<\varepsilon, & H_\delta(D^kT)&<\varepsilon\\ \text{Lip}(D^k_\eta F)&<\varepsilon, & \text{Lip}(D^k_\eta T)&<\varepsilon\\ \text{Lip}^{(\eta)}(D^kF)&<\varepsilon, & \text{Lip}^{(\eta)}(D^kT)&<\varepsilon \end{aligned} \tag{4.3}$$

$$|S(\eta, z)|<\varepsilon(|\eta|^{k+\delta}+|z|^{k+\delta}). \tag{4.4}$$

(4) Assume that the map (4.1) has a $C^{k,\delta}$ invariant manifold of the form

$$\begin{aligned} &\eta = w(z)\\ &\|D^jw(z)\|\leqslant|z|^{k+\delta-j}, \qquad j=0, 1, ..., k. \end{aligned} \tag{4.5}$$

(5) We straighten up the invariant manifold by taking $\xi = \eta - w(z)$. Then the map (4.1) takes the form

$$\begin{aligned} x^* &= Ax + f(\xi, z) \\ y^* &= By + g(\xi, z) \\ z^* &= Cz + h(\xi, z) \\ f(\xi, z) &= R(\xi + w(z), z) - R(w(z), z) \\ g(\xi, z) &= S(\xi + w(z), z) - S(w(z), z) \\ h(\xi, z) &= T(\xi + w(z), z). \end{aligned} \tag{4.6}$$

(6) Let $\mathfrak{D}^j = D^j - D^j_z - D^j_\xi$. Let $p_j = r - j$, $q_j = r - j - 1$, $0 \leqslant j \leqslant k$.

Theorem 4. *Let* $0 \leqslant s \leqslant 1$, $r = k + \delta$ *and* $G = (f, g)$. *Then the map* (4.6) *satisfies the following for* $|\xi| \leqslant 1$ *and* $|z| \leqslant 1$:

$$|g(\xi, z)| \leqslant \varepsilon\, |\xi|\, [\,|\xi|^{r-1} + |z|^{r-1}] \tag{4.7}$$

$$\|D^j_z G(\xi, z)\| < \varepsilon\, |\xi|, \qquad 0 \leqslant j \leqslant k \tag{4.8}$$

$$\|D^j_z g(\xi, z)\| \leqslant \varepsilon\, |\xi|\, [\,|\xi|^{q_j} + |z|^{q_j}], \qquad 0 \leqslant j \leqslant k-1 \tag{4.9}$$

$$\|\mathfrak{D}^j g(\xi, z)\| \leqslant \varepsilon[\,|\xi|^{r-j} + |z|^{r-j}], \qquad 0 \leqslant j \leqslant k \tag{4.10}$$

$$\|D^j_\xi g(\xi, z)\| \leqslant \varepsilon[\,|\xi|^{r-j} + |z|^{r-j}], \qquad 0 \leqslant j \leqslant k \tag{4.11}$$

$$\mathrm{Lip}(D^k_\xi G) < \varepsilon \tag{4.12}$$

$$H_\delta(D^k G) < \varepsilon \tag{4.13}$$

$$\|D^k G(\xi, z) - D^k G(\xi', z)\| < \varepsilon\, |\xi - \xi'| \tag{4.14}$$

$$\|D^j_z G(\xi, z) - D^j_z G(\xi, z')\| < \varepsilon\, |\xi|, \qquad 0 \leqslant j \leqslant k \tag{4.15}$$

$$\|D^j_z G(\xi, z) - D^j_z G(\xi, z')\| < \varepsilon\, |\xi|^s\, |z - z'|^{\delta(1-s)}, \qquad 0 \leqslant j \leqslant k \tag{4.16}$$

$$\mathrm{Lip}(D^k_\xi h) < \varepsilon \tag{4.17}$$

$$\mathrm{Lip}^{(\xi)}(D^k h) < \varepsilon. \tag{4.18}$$

The proof of Theorem 4 will be given in Section 7 in order not to obstruct the flow of the discussion.

Theorem 5 (A General Elimination Theorem). *Consider a map* $\Phi(\xi, z)$ *which is* $C^{k,\delta}$, $0 < \delta \leqslant 1$, *and takes the form*

$$\begin{aligned} x^* &= Ax + f(\xi, z) \\ y^* &= By + g(\xi, z) \\ z^* &= Cz. \end{aligned} \tag{4.19}$$

Assume that the nonlinear part (f, g) *satisfies* (4.7)–(4.16) *of Theorem* 4. *Assume that for some* $0<\chi<b<\gamma<1$, $0<a\leqslant b$,

$$\begin{aligned} \|A\| &< a \\ \|B\| &< b \\ \|B^{-1}\| &< \chi^{-1} \end{aligned} \tag{4.20}$$

$$\begin{aligned} \theta_o &= \chi^{-1} b \gamma^{k+\delta-1} < 1 \\ \theta_1 &= \chi^{-1} b^s \gamma^{k+\beta} < 1 \\ \beta &= \delta(1-s), \qquad 0<s<1. \end{aligned} \tag{4.21}$$

Then the following are true for sufficiently small $\varepsilon>0$:

(1) *For any* $0<d\leqslant 1$ *there is a function* $u \in C^{k,\beta}(B_1(0), Y)$ *so that the change of variables*

$$\begin{aligned} (x, y, z) &\mapsto (X, Y, Z) \\ (X, Y, Z) &= Q(x, y, z) = (x, y+u(x, y, z), z) \end{aligned} \tag{4.22}$$

transforms the map (4.19) *to the map*

$$\begin{aligned} X^* &= AX + H(X, Y, Z) \\ Y^* &= BY \\ Z^* &= CZ. \end{aligned} \tag{4.23}$$

Moreover

$$\mathrm{Lip}(D_\xi^k u) \leqslant d \tag{4.24}$$

$$\mathrm{Lip}^{(\xi)}(D^k u) \leqslant d \tag{4.25}$$

$$H_\delta^{(z)}(\mathfrak{D}^k u) \leqslant d \tag{4.26}$$

$$\|D_z^k u(\xi, z) - D_z^k u(\xi, z')\| \leqslant d\, |\xi|^s\, |z-z'|^\beta. \tag{4.27}$$

(2) *For sufficiently small* $0<d<1$, $Q(x, y, z)$ *is a* $C^{k,\beta}$ *diffeomorphism of a neighbourhood of the origin.* (*This assertion follows immediately from the Inverse Function Theorem.*)

(3) *Let the inverse of* $Q(x, y, z)$ *be given by*

$$Q^{-1}(X, Y, Z) = (X, Y + v(X, Y, Z), Z).$$

Then $v(X, Y, Z) = -u(x, y, z)$. *Moreover,*

$$H(X, Y, Z) = f \circ Q^{-1}(X, Y, Z) = f(X, Y + v(X, Y, Z), Z)$$

satisfy the following with $\Xi = (X, Y)$:

$$\|D^j H\| < O(\varepsilon), \qquad 0 \leqslant j \leqslant k \tag{4.28}$$

$$H_\beta(D^k H) < O(\varepsilon) \tag{4.29}$$

$$\mathrm{Lip}(D^k_\Xi H) < O(\varepsilon) \tag{4.30}$$

$$\mathrm{Lip}^{(\Xi)}(D^k H) < O(\varepsilon). \tag{4.31}$$

We prove Theorem 5 in Section 6.

5. PROOF OF THEOREM 1 GRANTING THEOREM 4 AND THEOREM 5

We will proof Theorem 1 by reversed finite induction on $\sigma_1, ..., \sigma_{J-1}$ and σ_J. The induction is done as follows:

- After possibly adding a dummy variable with $z^* = \gamma_1 z$, we may assume that we linearized the part of the map which corresponds to $\sigma_{I+1}, ..., \sigma_J$.
- Then we eliminate certain nonresonant terms from the map. The resulting map will satisfy hypotheses (4.1)–(4.4) of Theorem 4 as stated in Subsection 4.1. The resulting map also satisfies the hypotheses of Theorem 6, Subsection 5.6 below. Thus, it has a weak-stable manifold of the form (4.5). Now we straighten up the weak-stable manifold which corresponds to $\sigma_{I+1}, ..., \sigma_J$, i.e., in the z-direction. The resulting map is of the form (4.6) (with $h \equiv 0$, which wouldn't hurt). By Theorem 4, the resulting map satisfies estimates (4.7)–(4.18).
- Estimates (4.7)–(4.18) are the main hypotheses of Theorem 5. Thus we eliminate F_I. The conclusion of Theorem 5 also states that after eliminating F_I, the resulting map satisfies the hypotheses of Theorem 4 given in Section 4.1 but with I replaced by $I-1$.
- We repeat this process until we eliminate F_1.
- If in addition $b_o < (\chi_1)^k$, we can straighten up the weak-stable manifold for the last time.

5.1. *Induction Hypothesis.* (1) We may assume that the map has been partially linearized for we can add a dummy variable, say z, and let

$z^* = \gamma_1 z$ with $0 < \gamma_1 < 1$ is close enough to 1 such that hypotheses (H1)–(H4) are satisfied.

(2) Let $1 \leqslant I < J$. We may assume that the map takes the form

$$\begin{aligned} \eta^* &= M\eta + F(\eta, z) \\ z^* &= Cz \\ \eta &= (x_o, x_1, ..., x_I) \\ z &= (x_{I+1}, ..., x_J) \\ M &= \text{blockdiag}(\Lambda_o, \Lambda_1, ..., \Lambda_I) \\ C &= \text{blockdiag}(\Lambda_{I+1}, ..., \Lambda_J) \\ F &= (F_o, F_1, ..., F_{I-1}, F_I), \end{aligned} \tag{5.1}$$

where for some $0 < \delta \leqslant 1$, $F(\eta, z)$ satisfies

$$\|D^i F\| < \varepsilon, \qquad 1 \leqslant i \leqslant k \tag{5.2}$$

$$\text{Lip}(D_\eta^k F) < \varepsilon \tag{5.3}$$

$$\text{Lip}^{(\eta)}(D^k F) < \varepsilon \tag{5.4}$$

$$H_\delta(D^k F) < \varepsilon. \tag{5.5}$$

5.2. *Preparing for the Induction Step.* In order to go one step further down the finite sequence $\langle \sigma_1, ..., \sigma_I \rangle$ (that is to eliminate $F_I(\eta, z)$) we observe that the Banach space E and the map have the following decomposition:

$$\begin{aligned} (\eta, z) &= (x, y, z) \in X \times Y \times Z = E \\ x &= (x_o, x_1, ..., x_{I-1}) \\ y &= x_I \\ z &= (x_{I+1}, ..., x_J) \\ F &= (R, S) \\ R &= (F_o, F_1, ..., F_{I-1}) \\ S &= F_I \\ \Lambda &= \text{blockdiag}(M, C) = \text{blockdiag}(A, B, C) \\ A &= \text{blockdiag}(\Lambda_o, \Lambda_1, ..., \lambda_{I-1}) \\ B &= \Lambda_I \\ C &= \text{blockdiag}(\Lambda_{I+1}, ..., \Lambda_J) \end{aligned} \tag{5.6}$$

$$
\begin{aligned}
\|A\| &< a := b_{I-1} < b_I \\
\|B\| &< b := b_I \\
\|B^{-1}\| &< \chi^{-1} := \chi_I^{-1} \\
\|C\| &< b_J < \gamma \\
\|C^{-1}\| &< c^{-1}, \qquad c = \chi_{I+1} \\
b\gamma^{k+\delta-1} &< \chi \\
b^s\gamma^{k+\beta} &< \chi \\
s &= s_I \\
\delta &= \delta_I \\
\beta &= \delta(1-s) = \delta_{I-1}.
\end{aligned}
\tag{5.7}
$$

Estimates (5.7) on $\|A\|$, $\|B\|$, $\|B^{-1}\|$, and $\|C\|$ follow from using adapted norms in Subsection 2.2.

The map takes the following form in the variables (x, y, z):

$$
\begin{aligned}
x^* &= Ax + R(x, y, z) \\
y^* &= By + S(x, y, z) \\
x^* &= Cz.
\end{aligned}
\tag{5.8}
$$

5.3. *Eliminating Nonresonant Monomials.* In a finite dimensional space $\mathbb{R}^n$, the next step in the proof would be to eliminate certain nonresonant monomials of degree 2 to k from the map. In an infinite dimensional Banach space we have to work with symmetric j-multilinear functions. As we mentioned above (in the introduction) eliminating such functions in the infinite dimensional case is more technical.

Let E_1 and E_2 be two Banach spaces. For $j \geqslant 2$, let $\mathscr{L}(E_1^j, E_2)$ be the Banach space of j-multilinear functions from E_1^j to E_2, and let $\mathscr{S}(E_1^j, E_2)$ be the Banach space of symmetric j-multilinear functions from E_1^j to E_2. Although E_1 and E_2 are infinite dimensional Banach spaces, we will continue to call elements in $\mathscr{S}(E_1^j, E_2)$ j-monomials or monomials of degree j. Consider a map

$$
\begin{aligned}
(\xi, \eta) \to (\xi^*, \eta^*)&: E_1 \times E_2 \to E_1 \times E_2 \\
\xi^* &= U\xi + \phi(\xi, \eta) \\
\eta^* &= V\eta + g_j(\xi) + \psi(\xi, \eta),
\end{aligned}
$$

where $U: E_1 \to E_1$ and $V: E_2 \to E_2$. Assuming that $j=2$ or we have already eliminated $g_i(\xi)$, $2 \leqslant i \leqslant j-1$, to eliminate $g_j(\xi)$ we need to find $p_j \in \mathcal{S}(E_1^j, E_2)$ which solves a homological equation of the form

$$L^j_{U,V} p_j := (I - K^j_{U,V})\, p_j = V^{-1} g_j \tag{5.9}$$

$$L^j_{U,V} p_j := V^{-1} p_j \circ U. \tag{5.10}$$

The crucial point is to find out whether $1 \notin \sigma(L^j_{U,V})$ or not. This requires an understanding of $\sigma(K^j_{U,V})$ and assuming that (U, V) satisfy $NR_j(V, U)$. We need the following lemma:

5.4. Lemma. *Let $U \in \mathcal{L}(E_1, E_1)$ and $V \in \mathcal{L}(E_2, E_2)$. For $j \geqslant 2$, define two linear operators*

$$K^j_{U,V}: \mathcal{S}(E_1^j, E_2) \to \mathcal{S}(E_1^j, E_2), \qquad K^j_{U,V} p_j = V^{-1} p_j \circ U$$

$$L^j_{U,V}: \mathcal{S}(E_1^j, E_2) \to \mathcal{S}(E_1^j, E_2), \qquad L^j_{U,V} = I - K^j_{U,V}.$$

Then for each fixed but arbitrary $j \geqslant 2$, the following are true:

(1) *The spectrum of $K^j_{U,V}$ satisfies*

$$\sigma(K^j_{U,V}) \subseteq \{ v^{-1} \mu_1 \cdots \mu_j \mid v \in \sigma(V), \mu_i \in \sigma(U), 1 \leqslant i \leqslant j \} \tag{5.11}$$

(2) *Assume also that (U, V) satisfy $NR_j(V, U)$. Then $1 \notin \sigma(K^j_{U,V})$.*

(3) *Thus, $L^j_{U,V}: \mathcal{S}(E_1^j, E_2) \to \mathcal{S}(E_1^j, E_2)$ is a bijection.*

(4) *$[L^j_{U,V}]^{-1}: \mathcal{S}(E_1^j, E_2) \to \mathcal{S}(E_1^j, E_2)$ is continuous.*

Proof. A proof of assertion (1) is given (first) in [9, 15]. In Section 8 we present a proof in our notation for completeness. Assertion (2) follows from (5.11) and the nonresonance condition. Assertion (3) follows from assertion (2) and the concept of a spectrum. Assertion (4) follows from the Inverse Mapping Theorem (which is an immediate corollary to the Open Mapping Theorem in the present case). ∎

5.5. Lemma. *Let A, B, and C be the linear operators in (5.8). Let $\Lambda = \text{blockdiag}(A, B, C)$ and let $E = X \times Y \times Z$. Let $2 \leqslant j \leqslant k$.*

(1) *The linear operator $K^j_{C,A}: S(Z^j, X) \to S(Z^j, X)$ satisfies*

$$\sigma(K^j_{C,A}) \subseteq \{ \alpha^{-1} \mu_1 \cdots \mu_j \mid \alpha \in \sigma(A), \mu_i \in \sigma(C), 1 \leqslant i \leqslant j \}. \tag{5.12}$$

Moreover, $1 \notin \sigma(L^j_{C,A})$ and $L^j_{C,A}: S(Z^j, X) \to S(Z^j, X)$ is a continuous bijection with continuous inverse.

(2) *The linear operator* $K^j_{\Lambda, B}: S(E^j, Y) \to S(E^j, Y)$ *satisfies*

$$\sigma(K^j_{\Lambda, B}) \subseteq \{\beta^{-1}\lambda_1 \cdots \lambda_j \mid \beta \in \sigma(B), \lambda_i \in \sigma(\Lambda), 1 \leqslant i \leqslant j\}. \tag{5.13}$$

Moreover, $1 \notin \sigma(L^j_{\Lambda, B})$ *and* $L^j_{\Lambda, B}: S(E^j, Y) \to S(E^j, Y)$ *is a continuous bijection with continuous inverse.*

(3) *There is a diffeomorphism of the form* $\psi = id + u : E \to E$, *where* u *is the sum of j-monomials,* $2 \leqslant j \leqslant k$, *which eliminates all j-monomials in* z, $2 \leqslant j \leqslant k$, *from* $R(x, y, z)$ *and all j-monomials in* (x, y, z), $2 \leqslant j \leqslant k$, $S(x, y, z)$. *In other words, we may assume that the map* (5.8) *satisfies*

$$|R(\eta, z)| < \varepsilon[\,|z|^{k+\delta} + |\eta|(|\eta| + |z|)\,] \tag{5.14}$$

$$|S(\eta, z)| < \varepsilon[\,|z|^{k+\delta} + |\eta|^{k+\delta}]. \tag{5.15}$$

Proof. (1) Notice that for $\alpha \in \sigma(A)$ and $\mu_i \in \sigma(C)$, $1 \leqslant i \leqslant j$, we have

$$|\alpha| < a = b_{I-1} < b_I < (\chi_{I+1})^k < |\mu_1| \cdots |\mu_j|.$$

Thus (A, C) satisfy the nonresonant condition $NR_k(A; C)$. It follows that $1 \notin \sigma(L^j_{C, A})$. The rest follows from Lemma 5.4.

(2) Notice that the spectra of A, B and C lie in the interior of the unit disc in $\mathbb{C}$. Thus, it suffices to consider possible resonances of the form

$$\beta = \beta_1 \cdots \beta_i \mu_1 \cdots \mu_l, \qquad i + l = j,\ \beta, \beta_s \in \sigma(B), \mu_r \in \sigma(C).$$

By the gap condition (H1), and since

$$b = b_I < (\chi_{I+1})^k < |\mu_1| \cdots |\mu_j|$$

such a resonance cannot exist if $i = 0$. So, assume that $i \geqslant 1$. By hypothesis (H2), $B = \Lambda_I$ satisfies a nonresonance condition $NR_k(\Lambda_I, \Lambda_I \times \cdots \times \Lambda_J)$. Thus such a resonance cannot exist. Thus, (B, Λ) satisfies a nonresonance condition $NR_k(B, \Lambda)$. It follows that $1 \notin \sigma(L^j_{B, \Lambda})$. The rest follows from Lemma 5.4.

(3) In order to eliminate a j-monomial $g_j(z)$ from $R(x, y, z)$ we need to find $p_j \in S(Z^j, X)$ which solves the homological equation

$$L^j_{C, A} p_j = A^{-1} g_j(z).$$

In view of assertion 1, $L^j_{C, A}: S(Z^j, X) \to S(Z^j, X)$ is invertible and the homological equation can be solved uniquely.

In order to eliminate a j-monomial $h_j(x, y, z)$ from $S(x, y, z)$ we need to find $q_j \in S(E^j, Y)$ which solves the homological equation

$$L^j_{\Lambda, B} q_j = B^{-1} h_j(x, y, z).$$

In view of assertion 1, $L^j_{\Lambda, B}: S(E^j, Y) \to S(E^j, Y)$ is invertible and the homological equation can be solved uniquely. ▮

5.6. *Weak-Stable Manifolds.* At this point we need to invoke a theorem that guarantees the existence of a weak-stable manifold under a gap condition. See, for example, Theorem 5.1 in [13, p. 53]. However, we need a weak-stable manifold for maps that satisfy the vanishing properties (5.14) and (5.15). We also need the resulting weak-stable manifold to satisfy certain vanishing properties. These vanishing properties are essential for the linearization step. Such a theorem appears as part of a more general theorem in [9] and we state it here in the present notation.

We would like to point out that this is the only point in the proof of Theorem 1 where we need the smallness assumptions (H4) on the nonlinear part. For the rest of the proof, since we work with a contraction, the smallness assumption can be obtained with the method of Subsection 2.4.4(b).

THEOREM 6. *Consider a $C^{k,\delta}$ map, $0 < \delta \leqslant 1$, $k = 1, 2, ...,$*

$$\begin{aligned} \eta^* &= M\eta + F(\eta, z) \\ z^* &= Cz + T(\eta, z). \end{aligned} \tag{5.16}$$

Assume that there is $b < c$ such that

(1) $b < c^{k+\delta}$.

(2) $|\mu| < b$, $\mu \in \sigma(M)$.

(3) $|\gamma| > c$, $\gamma \in \sigma(C)$.

(4) $|F(\eta, z)| < \varepsilon(|z|^{k+\delta} + |\eta|)$.

(5) $\|D^j F\| < \varepsilon$, $\|D^j T\| \leqslant \varepsilon$, $j = 0, 1, ..., k$.

(6) $H_\delta(D^k F) < \varepsilon$, $H_\delta(D^k T) < \varepsilon$.

Then, for sufficiently small $\varepsilon > 0$, the map has a $C^{k,\delta}$ weak stable manifold of the form $\eta = w(z)$ which is unique in the class of functions which satisfy:

$$\begin{aligned} w(z^*) &= Mw(z) + F(w(z), z) \\ z^* &= Cz + T(w(z), z) \\ \|D^j w(z)\| &\leqslant |z|^{k+\delta-j}, \qquad j = 0, 1, ..., k. \end{aligned} \tag{5.17}$$

Notice that the uniqueness of the weak-stable manifold is obtained assuming that a smallness condition holds. Recall the earlier remarks given in Subsection 2.45(b).

5.7. *Straightening Up the Weak-Stable Manifold.* The next step in the proof of Theorem 1 is to make a change of variables that straightens up the weak-stable manifold. Let $\xi = \eta - w(z)$. Thus (5.8) takes the form

$$\begin{aligned} \xi^* &= M\xi + G(\xi, z) \\ z^* &= Cz \\ G(\xi, z) &= F(\xi + w(z), z) - F(w(z), z) \\ G(0, z) &= 0. \end{aligned} \tag{5.18}$$

It follows that $\{(\xi, z) \mid \xi = 0\}$ is invariant under the map (5.18). The map (5.18) takes the following form where $G =: (f, g)$,

$$\begin{aligned} x^* &= Ax + f(\xi, z) \\ y^* &= By + g(\xi, z) \\ z^* &= Cz, \end{aligned} \tag{5.19}$$

where

$$\begin{aligned} f(\xi, z) &= R(\xi + w(z), z) - R(w(z), z) \\ g(\xi, z) &= S(\xi + w(z), z) - S(w(z), z) \\ \xi &= (x, y). \end{aligned} \tag{5.20}$$

5.8. *Proof of Theorem* 1 *Granting Theorems* 4 *and* 5. Notice that the map (5.19) and the weak stable manifold (5.17) satisfy the conditions of Theorem 4 of Subsection 4.1. Thus $G = (f, g)$ satisfies (4.7)–(4.16).

By (2.5) and (2.6), $(\chi, b, \gamma, s, \delta) = (\chi_I, b_I, \gamma, s_I, \delta_I)$ satisfy (4.20) and (4.21).

Now we can apply Theorem 5. But the new map (4.23) satisfies the Induction Hypothesis 5.1 with I replaced by $I-1$.

To prove assertion (2) notice that if $b_o < (\chi_1)^k$, we have a weak stable manifold $\eta = w(z)$. If we straighten it up as in Subsection 5.7 the new map will satisfy $G(0, z) = 0$. This finishes the proof of Theorem 1.

6. PROOF OF THEOREM 5

In this section we prove Theorem 5. A change of variables of the form (4.22) eliminates $g(x, y, z)$ from the map (4.19) iff $u(\xi, z)$ satisfies

$$u(\xi, z) = B^{-1}u(\xi^*, z^*) + B^{-1}g(\xi, z). \tag{6.1}$$

We should expect $u(\xi, z)$ to be a fixed point for a transformation on a closed subset of an appropriately constructed Banach space which we define presently.

6.1. *The Space.* Let $r=k+\delta$, $0<\delta\leqslant 1$. For $0\leqslant j\leqslant k-1$ let $p_j=r-j$ and $q_j=r-j-1=p_j-1$. Let $p_k=r-k=\delta$ and $q_k=0$. Thus, $q_j=p_{j+1}$, $0\leqslant j\leqslant k-1$. Let $0<d<1$. For $u\in C^{k,\delta}(B_1(0), Y)$ define

$$N_o(u)=\sup\left\{\frac{|u(\xi, z)|}{|\xi|\,[\,|\xi|^{q_o}+|z|^{q_o}]}\;\middle|\;(\xi, z)\neq(0, 0),\ \xi\neq 0\right\}.$$

For $j=1, ..., k$, let

$$K_j(u)=\sup\left\{\frac{\|D^j_\xi u(\xi, z)\|}{|\xi|^{p_j}+|z|^{p_j}}\;\middle|\;(\xi, z)\neq(0, 0)\right\}$$

$$N_j(u)=\sup\left\{\frac{\|D^j_z u(\xi, z)\|}{|\xi|\,[\,|\xi|^{q_j}+|z|^{q_j}]}\;\middle|\;\xi\neq 0\right\}$$

$$\mathfrak{D}^j=D^j-D^j_z-D^j_\xi=\{D^l_\xi D^i_z\mid i+l=j,\ 1\geqslant 1,\ l\geqslant 1\}$$

$$m_{l,i}(u)=\sup\left\{\frac{\|D^j_\xi D^i_z u(\xi, z)\|}{|\xi|^{p_j}+|z|^{p_j}}\;\middle|\;(\xi, z)\neq(0, 0)\right\},$$

$$M_j(u)=\max\{m_{l,i}(u)\mid i+l=j,\ i\geqslant 1,\ l\geqslant 1\}$$

$$\|u\|_*=N_o(u)+\sum_{j=1}^{k}\,[K_j(u)+N_j(u)+M_j(u)]$$

$$\beta=\delta(1-s),\qquad 0<s<1$$

$$\kappa(D^k_z u)=\sup\left\{\frac{\|D^k_z u(\xi, z)-D^k_z u(\xi, z')\|}{|\xi|^s\,|z-z'|^\beta}\;\middle|\;\xi\neq 0,\ z\neq z'\right\}$$

$$H^{(z)}_\delta(\mathfrak{D}^k u)=\max\{H^{(z)}_\delta(D^l_\xi D^i_z u)\mid l+i=k,\ l\geqslant 1,\ i\geqslant 1\}$$

$$\sigma(u)=\max\{\mathrm{Lip}^{(\xi)}(D^k u),\ \mathrm{Lip}(D^k_\xi u),\ \kappa(D^k_z u),\ H^{(z)}_\delta(\mathfrak{D}^k u)\}$$

$$\mathfrak{Z}=\{u\in C^k(B_1(0), Y)\mid \|u\|_*<\infty\}$$

$$\mathfrak{X}_d=\{u\in\mathfrak{Z}\mid \|u\|_*\leqslant d,\ \sigma(u)\leqslant d,\ D^j u(0)=0,\ 0\leqslant j\leqslant k\}.$$

Since convergence in the norm $\|\cdot\|_*$ implies uniform convergence on bounded sets, it follows that $(\mathfrak{Z}, \|\cdot\|_*)$ is a Banach space and that $\mathfrak{X}_d$ is a closed subset.

It is also obvious that $H_\beta(D^k u) \leqslant 2^{1-\beta} d$ for all $u \in \mathfrak{X}_d$. This is because for $0 < \beta < \alpha \leqslant 1$, $|X - X'|^\alpha \leqslant 2^{\alpha-\beta} |X - X'|^\beta$. Also we use the norm

$$\|D^k u\| = \max\{ \|D_\xi^l D_z^i u\| \mid l + i = k\}$$

It follows that $\|D^j u\| \leqslant 2d$, $j = 0, 1, ..., k$, for all $u \in \mathfrak{X}_d$.

6.2. *The Transformation.* Let Φ be the map (4.19). Define Ω: $\mathfrak{Z} \to \mathfrak{Z}$ and Γ: $\mathfrak{Z} \to \mathfrak{Z}$ by

$$\begin{aligned} \Omega u &= \Gamma u + B^{-1} g \\ \Gamma u &= B^{-1} u \circ \Phi. \end{aligned} \tag{6.2}$$

In other words,

$$\begin{aligned} \Omega u(\xi, z) &= B^{-1} u(\xi^*, z^*) + B^{-1} g(\xi, z) \\ \Gamma u(\xi, z) &= B^{-1} u(\xi^*, z^*). \end{aligned}$$

It is obvious that $B^{-1} g$ does not depend on u, Γ is linear and Ω is affine. These properties will simplify the proof drastically.

6.3. *The Steps.* We prove Theorem 5 in four steps:

Step 1. If $\varepsilon < d$, then $g \in \mathfrak{X}_d$.

Proof. This assertion follows from Theorem 4 which tells us that for $\varepsilon < d$, $\|g\|_* < \varepsilon$, $\sigma(g) < \varepsilon$) and $D^i g(0, 0) = 0$, $0 \leqslant i \leqslant k$. Thus $g \in \mathfrak{X}_d$.

Step 2. We will show that for sufficiently small $\varepsilon > 0$, there is $\theta \in (0, 1)$ such that for all $u \in \mathfrak{X}_d$

$$\|\Gamma u\|_* \leqslant \theta \|u\|_* \tag{6.3}$$

$$\sigma(\Gamma u) \leqslant \theta \sigma(u) + O(\varepsilon) < d. \tag{6.4}$$

Step 3. For sufficiently small $\varepsilon > 0$, Ω maps $\mathfrak{X}_d$ to itself and is a contraction and hence has a unique attractive fixed point in $\mathfrak{X}_d$.

Proof. Granting Step 2, it is obvious that if $D^j u(0) = 0$, $j = 0, ..., k$, then $D^j \Omega u(0) = 0$, $j = 0, ..., k$. By Step 1 and Step 2, for sufficiently small $\varepsilon > 0$ for all $u \in \mathfrak{X}_d$

$$\|\Omega u\|_* \leqslant \theta \|u\|_* + \|g\|_* < \theta d + O(\varepsilon) < d$$

$$\sigma(\Omega u) \leqslant \theta \sigma(u) + \sigma(g) < \theta d + O(\varepsilon) < d$$

Thus Ω maps $\mathfrak{X}_d$ to itself. Also by Step 2, since Γ is linear and $g(\xi, z)$ is independent of $u \in \mathfrak{X}_d$ we have for all $u, v \in \mathfrak{X}_d$

$$\|\Omega u - \Omega v\|_* = \|\Gamma u - \Gamma v\|_* \leqslant \theta \|u - v\|_*.$$

Thus Ω is a contraction and hence has a unique attractive fixed point in $\mathfrak{X}_d$.

Step 4. $H(X, Y, Z)$ in the new map (4.23) has the stated properties.

To prove Theorem 5, it remains to prove Steps 2 and 4.

6.4. *Notation.* (1) In what follows t will stand for a continuous function $t(\varepsilon) \to 0$ as $\varepsilon \to 0$ which is independent of $u \in \mathfrak{X}_d$ and depends only on the map.

(2) All norms are assumed to be box norms.

6.5. Lemma. *Let $|\xi| \leqslant 1$ and $|z| \leqslant 1$. The following are true for sufficiently small $\varepsilon > 0$:*

$$|z^*| \leqslant \gamma\, |z| < 1 \tag{6.5}$$

$$|\xi^*| \leqslant b(1+t)\, |\xi| < 1 \tag{6.6}$$

$$|\xi^*|^{p_j} + |z^*|^{p_j} \leqslant \gamma^{p_j}(1+t)[\,|\xi|^{p_j} + |z|^{p_j}] \tag{6.7}$$

$$|\xi^*|\,[\,|\xi^*|^{q_j} + |z^*|^{q_j}] \leqslant b\gamma^{q_j}(1+t)\,|\xi|\,[\,|\xi|^{q_j} + |z|^{q_j}] \tag{6.8}$$

$$\begin{aligned} |\xi|\,[\,|\xi|^{q_j} + |z|^{q_j}] &\leqslant O(|\xi|^{q_j+1} + |z|^{q_j+1}) \\ &\leqslant O(|\xi|^{p_j} + |z|^{p_j}). \end{aligned} \tag{6.9}$$

Since we are using the box norm, it follows from the first two estimates that $\Gamma u(\xi, z)$ is well defined on $|\xi| \leqslant 1$, $|z| \leqslant 1$.

Proof. The lemma follows immediately from Theorem 4. ∎

6.6. Proposition. *For sufficiently small $\varepsilon > 0$, for all $u \in \mathfrak{X}_d$*

$$N_o(\Gamma u) \leqslant \theta N_o(u)$$

$$\theta = b\chi^{-1}\gamma^{r-1}(1+t) = \theta_o(1+t) < 1.$$

Proof. Recall that $\theta_o < 1$ by (4.20). By (6.5) and (6.6) and Theorem 4

$$\begin{aligned} |\Gamma u(\xi, z)| &\leqslant \chi^{-1} N_o(u)\, |\xi^*|\,[\,|\xi^*|^{q_o} + |z^*|^{q_o}] \\ &\leqslant [b\chi^{-1}\gamma^{q_o}(1+t)]\, N_o(u)\, |\xi|\,[\,|\xi|^{q_o} + |z|^{q_o}] \end{aligned}$$

and $q_o = r - 1$ which proves the proposition. ∎

6.7. PROPOSITION. *For sufficiently small $\varepsilon > 0$, for all $u \in \mathfrak{X}_d$*

$$N_1(D_z \Gamma u) \leqslant \theta N_1(u) + t K_1(u)$$

$$K_1(D_\xi \Gamma u) \leqslant \theta K_1(u)$$

$$\theta = b\chi^{-1}\gamma^{r-1}(1+t) = \theta_o(1+t) < 1.$$

It is obvious that $\mathfrak{D}^1 = 0$.

Proof. Recall that $p_1 = q_1 + 1 = r - 1$, $|z^*| < 1$ and that $|\xi^*| < 1$. Thus, by (4.8) and Lemma 6.5 we have

$$D_z \Gamma u(\xi, z) = B^{-1} D_\xi u(\xi^*, z^*)\, D_z G(\xi, z) + B^{-1} D_z u(\xi^*, z^*)[C]$$

$$\begin{aligned} \|D_z \Gamma u(\xi, z)\| &\leqslant \chi^{-1} K_1(u)[\,|\xi^*|^{p_1} + |z^*|^{p_1}][\varepsilon\, |\xi|\,] \\ &\quad + \gamma\chi^{-1} N_1(u)\, |\xi^*|\, [\,|\xi^*|^{q_1} + |z^*|^{q_1}] \\ &\leqslant [\,t K_1(u) + b\chi^{-1}\gamma^{r-1}(1+t)\, N_1(u)]\, |\xi|\, [\,|\xi|^{q_1} + |z|^{q_1}]. \end{aligned}$$

Since $z^* = Cz$ is independent of ξ, we have

$$D_\xi \Gamma u(\xi, z) = B^{-1} D_\xi u(\xi^*, z^*)[M + D_\xi G(\xi, z)]$$

$$\begin{aligned} \|D_\xi \Gamma u(\xi, z)\| &\leqslant \chi^{-1} K_1(u)[\,|\xi^*|^{p_1} + |z^*|^{p_1}][b + t] \\ &\leqslant b\chi^{-1}\gamma^{r-1}(1+t)\, K_1(u)[\,|\xi|^{p_1} + |z|^{p_1}] \end{aligned}$$

which proves the lemma. ∎

6.8. PROPOSITION. *For sufficiently small $\varepsilon > 0$, for all $u \in \mathfrak{X}_d$ and $2 \leqslant j \leqslant k$*

$$N_j(\Gamma u) \leqslant \theta_o N_j(u) + t \sum_{i=1}^{j} [K_i(u) + N_i(u) + M_i(u)]$$

$$\theta_o = b\chi^{-1}\gamma^{r-1} < 1.$$

Proof. Notice that $D_\xi^n z^* = 0$, $n \geqslant 1$, and $D_z^n z^* = 0$, $n \geqslant 2$. Thus

$$\begin{aligned} D_z^j \Gamma u(\xi, z) &= B^{-1} D_z^j u(\xi^*, z^*)[C]^j + P_j(\xi, z) \\ P_j(\xi, z) &= \mathscr{P}_j(\tilde{\Delta}^{(j)} u(\xi^*, z^*), \Delta_z^{(j)} G(\xi, z)) \\ \mathscr{P}_j(0, \cdot) &\equiv 0 \\ \mathscr{P}_j(\cdot, 0) &\equiv 0 \\ \Delta^{(l)}\psi &= (D\psi, ..., D^l\psi) \\ \tilde{\Delta}^{(l)}\psi &= \Delta^{(l)}\psi - D_z^l\psi. \end{aligned} \tag{6.10}$$

Moreover, $\mathscr{P}_j(\Sigma, \Omega)$ is a polynomial in the components of Σ and Ω and linear in the components of Σ.

By (4.8) we have

$$\|\Delta_z^{(j)} G(\xi, z)\| \leqslant O(\varepsilon)\, |\xi|.$$

In view of this fact, the vanishing properties of $\mathscr{P}_j$ listed above and estimate (6.9) we obtain the estimate

$$\begin{aligned} \|P_j(\xi, z)\| &\leqslant t\, |\xi| \sum_{i=1}^{j} [K_i(u) + M_i(u) + N_i(u)][\,|\xi|^{p_i} + |z|^{p_i}] \\ &\leqslant t\, |\xi|\, [\,|\xi|^{q_j} + |z|^{q_j}] \sum_{i=1}^{j} K_i(u) + M_i(u) + N_i(u). \end{aligned} \tag{i}$$

Moreover, by estimate (6.8) and since $q_j = r - j - 1$, we have

$$\begin{aligned} \|B^{-1} D_z^j u(\xi^*, z^*)[C]^j\| &\leqslant \chi^{-1} \gamma^j N_j(u)\, |\xi^*|\, [\,|\xi^*|^{q_j} + |z^*|^{q_j}] \\ &\leqslant b\chi^{-1}\gamma^{r-1}(1+t)\, N_j(u)\, |\xi|\, [\,|\xi|^{q_j} + |z|^{q_j}]. \end{aligned} \tag{ii}$$

Adding up estimates (i) and (ii) and taking the sup, we obtain the desired estimate. ∎

6.9. PROPOSITION. *For sufficiently small $\varepsilon > 0$, for all $u \in \mathfrak{X}_d$ and $2 \leqslant j \leqslant k$*

$$K_j(\Gamma u) \leqslant \theta_o K_j(u) + t \sum_{i=1}^{j} K_i(u)$$

$$\theta_o = b\chi^{-1}\gamma^{r-1} < 1.$$

Proof. Notice that $D_\xi^n z^* = 0$, $n \geqslant 1$. Thus

$$\begin{aligned} D_\xi^j \Gamma u(\xi, z) &= B^{-1} D_\xi^j u(\xi^*, z^*)[M + D_\xi G(\xi, z)]^j + Q_j(\xi, z) \\ Q_j(\xi, z) &= \mathscr{Q}_j(\Delta_\xi^{(j-1)} u(\xi^*, z^*), \Delta_\xi^{(j)} G(\xi, z)) \\ \mathscr{Q}_j(0, \cdot) &\equiv 0 \\ \mathscr{Q}_j(\cdot, 0) &\equiv 0. \end{aligned} \tag{6.11}$$

Moreover, $\mathscr{Q}_j(\Sigma, \Omega)$ is a polynomial in the components of Σ and Ω and linear in the components of Σ.

By Theorem 4

$$\|\Delta_\xi^{(j)} G(\xi, z)\| \leqslant O(\varepsilon)[\,|\xi|^{p_j} + |z|^{p_j}].$$

Thus in view of the vanishing properties of $\mathscr{Q}_j$ we have

$$\|\mathscr{Q}_j(\xi, z)\| \leqslant t[|\xi|^{p_j} + |z|^{p_j}] \sum_{i=1}^{j-1} K_i(u). \tag{a}$$

Now we estimate $\mathscr{E} := \|B^{-1} D_\xi^j u(\xi^*, z^*)[M + D_z G(\xi, z)]^j\|$. Since $p_j = r - j$ and $b \leqslant \gamma$,

$$\begin{aligned} \mathscr{E} &\leqslant \chi^{-1} K_j(u)[|\xi^*|^{p_j} + |z^*|^{p_j}][b^j(1+t)] \\ &\leqslant b^j \chi^{-1} \gamma^{p_j}(1+t)\, K_j(u)[|\xi|^{p_j} + |z|^{p_j}] \\ &\leqslant b\chi^{-1}\gamma^{r-1}(1+t)\, K_j(u)[|\xi|^{p_j} + |z|^{p_j}]. \end{aligned} \tag{b}$$

Adding up estimates (a) and (b) and taking the sup, we obtain the desired estimate. ∎

6.10. PROPOSITION. *For sufficiently small $\varepsilon > 0$, for all $u \in \mathfrak{X}_d$ and $2 \leqslant j \leqslant k$.*

$$M_j(u) \leqslant \theta_o M_j(u) + t \sum_{i=1}^{j} [K_i(u) + N_i(u) + M_i(u)]$$
$$\theta_o = b\chi^{-1}\gamma^{r-1} < 1.$$

Proof. Let (i, l) be fixed but arbitrary such that $i + l = j$, $i \geqslant 1$ and $l \geqslant 1$. Recall that $D_\xi^n z^* = 0$, $n \geqslant 1$. Thus

$$\begin{aligned} D_\xi^l D_z^i \Gamma u(\xi, z) = B^{-1} D_\xi^l D_z^i u(\xi^*, z^*)[C]^i\,[M + D_\xi G(\xi, z)]^l \\ + \mathscr{W}_j(\xi, z) + R_{j-1}(\xi, z), \end{aligned} \tag{6.12}$$

where $\mathscr{W}_j(\xi, z)$ is a finite sum of the form

$$\mathscr{W}_j(\xi, z) = \sum_{n=1}^{i} c_n \theta_n$$
$$\theta_n = D_\xi^{l+n} D_z^{i-n} u(\xi^*, z^*)[C]^{i-n}\,[D_z G(\xi, z)]^n\,[M + D_\xi G(\xi, z)]^l$$
$$R_{j-1}(\xi, z) = \mathscr{R}_{j-1}(\Delta^{(j-1)} u(\xi^*, z^*), \Delta^{(j)} G(\xi, z))$$
$$\mathscr{R}_{j-1}(0, \cdot) \equiv 0$$
$$\mathscr{R}_{j-1}(\cdot, 0) \equiv 0,$$

where $\mathscr{R}_{j-1}(\Sigma, \Omega)$ is a polynomial in the components of Σ and Ω and linear in the components of Σ. In order to estimate $\|\mathscr{W}_j(\xi, z)\|$ we notice that $n \geqslant 1$ in the sum defining it. Thus

$$\|\mathscr{W}_j(\xi, z)\| \leqslant t[|\xi|^{p_j} + |z|^{p_j}] \sum_{v=1}^{j} [M_v(u) + K_v(u)]. \tag{a}$$

The vanishing properties of $\mathscr{R}_{j-1}$ and Theorem 4 imply that

$$\|\Delta^{(j-1)}G\| < O(\varepsilon)$$

and

$$\|R_{j-1}(\xi, z)\| \leqslant t[\,|\xi|^{p_j} + |z|^{p_j}] \sum_{n=1}^{j-1} [K_n(u) + M_n(u) + N_n(u)]. \qquad \text{(b)}$$

Now we estimate $\mathscr{A} := \|B^{-1} D_\xi^l D_z^i u(\xi^*, z^*)[C]^i [M + D_\xi G(\xi, z)]^l\|$. Since $i + l - 1 + p_j = p_j + j - 1 = r - 1$,

$$\begin{aligned} \mathscr{A} &\leqslant \chi^{-1} \gamma^i b^l (1+t)\, m_{l,i}(u)[\,|\xi^*|^{p_j} + |z^*|^{p_j}] \\ &\leqslant \chi^{-1} \gamma^{i+p_j} b^l (1+t)\, m_{l,i}(u)[\,|\xi|^{p_j} + |z|^{p_j}] \\ &\leqslant \chi^{-1} b \gamma^{r-1} (1+t)\, m_{l,i}(u)[\,|\xi|^{p_j} + |z|^{p_j}]. \end{aligned} \qquad \text{(c)}$$

Adding up estimates (a)–(c) we can see that

$$m_{l,i}(\Gamma u) \leqslant \chi^{-1} b \gamma^{r-1} m_{l,i}(u) + t \sum_{n=1}^{j} [K_n(u) + M_n(u) + N_n(u)].$$

Taking the maximum over all possible (l, i) such that $i + l = j$, $i \geqslant 1$ and $l \geqslant 1$ we obtain the desired estimate. ∎

6.11. Corollary. *For sufficiently small* $\varepsilon > 0$

$$\|\Gamma u\|_* \leqslant \theta\, \|u\|_*, \qquad u \in \mathfrak{X}_d$$
$$\theta = \chi^{-1} b \gamma^{r-1} + t = \theta_o + t < 1$$

which proves (6.3) *of Step* 2 *in Subsection* 6.3.

Proof. By Propositions 6.6, 6.7, 6.8, 6.9, and 6.10

$$\|\Gamma u\|_* \leqslant [\chi^{-1} b \gamma^{r-1} + t]\, \|u\|_*, \qquad u \in \mathfrak{X}_d.$$

By (4.20), $\theta_o < 1$. Thus $\theta < 1$ for sufficiently small $\varepsilon > 0$. ∎

Next we prove (6.4) of Step 2 in Subsection 6.3. This will be accomplished in four propositions. First we need the following lemma.

6.12. Lemma. *Let* $z \neq z'$, $\xi \neq \xi'$, *and define*

$$z^* = Cz, \qquad \hat{z} = Cz', \qquad \tilde{z} = Cz' = \hat{z}$$
$$\xi^* = M\xi + f(\xi, z), \qquad \hat{\xi} = M\xi + f(\xi, z'), \qquad \tilde{\xi} = M\xi' + f(\xi', z').$$

Then

$$|\xi^* - \hat{\xi}| \leqslant |f(\xi, z) - f(\xi, z')| \leqslant \varepsilon\, |z - z'| \tag{6.13}$$

$$|\xi^* - \hat{\xi}| \leqslant |f(\xi, z) - f(\xi, z')| \leqslant \varepsilon\, |\xi|^s\, |z - z'|^\beta \tag{6.14}$$

$$|\xi^* - \tilde{\xi}| \leqslant b\, |\xi - \xi'| + |f(\xi, z) - f(\xi', z')| \tag{6.15}$$

$$\leqslant b(1+t)\, |\xi - \xi'| + \varepsilon\, |z - z'|. \tag{6.16}$$

Proof. Estimate (6.14) follows from (4.16). The others follow from the fact that $G(x, y)$ is uniformly Lipschitz continuous. The smallness assumptions on $f(\xi, z)$ lead to the factor ε in all three estimates. ∎

6.13. Proposition. *For sufficiently small* $\varepsilon > 0$ *and for* $0 < s < 1$ *sufficiently close to* 1 *the following hold for all* $u \in \mathfrak{X}_d$:

$$\kappa(D_z^k \Gamma u) \leqslant \theta_1 \kappa(D_z^k u) + t < d$$
$$\theta_1 = \chi^{-1} b^s \gamma^{k+\beta} < 1.$$

Proof. Recall that by (4.21), $\theta_1 < 1$. Recall (6.10) with $j = k$.

$$D_z^k \Gamma u(\xi, z) = B^{-1} D_z^k u(\xi^*, z^*)[C]^k + P_k(\xi, z).$$

Let $z \neq z'$ and let $\hat{\xi}$ and $\hat{z} = Cz'$ be as in Lemma 6.12. Then

$$\mathscr{A} := \|[B^{-1}[D_z^k u(\xi^*, z^*) - D_z^k u(\hat{\xi}, \hat{z})]][C]^k\|$$
$$\mathscr{A} \leqslant \chi^{-1} \gamma^k [\, \|D_z^k u(\xi^*, z^*) - D_z^k u(\xi^*, \hat{z})\| + \|D_z^k u(\xi^*, \hat{z}) - D_z^k u(\hat{\xi}, \hat{z})\|\,].$$

Since $\mathrm{Lip}^{(\xi)}(D_z^k u) \leqslant d$, it follows from (6.14) that

$$\begin{aligned} \mathscr{A} &\leqslant \chi^{-1} \gamma^k (1+t)\, \kappa(D_z^k u)\, |\xi^*|^s\, |C(z - z')|^\beta + d\, |\xi^* - \hat{\xi}| \\ &\leqslant [\chi^{-1} b^s \gamma^{k+\beta} (1+t)\, \kappa(D_z^k u) + t]\, |\xi|^s\, |z - z'|^\beta \\ &\leqslant [\theta_1 \kappa(D_z^k u) + t]\, |\xi|^s\, |z - z'|^\beta. \end{aligned} \tag{1}$$

Let

$$\mathscr{B} := \|\mathscr{P}_k(\tilde{\Delta}^{(k)} u(\xi^*, z^*), \Delta_z^{(k)} G(\xi, z)) - \mathscr{P}_k(\tilde{\Delta}^{(k)} u(\hat{\xi}, \hat{z}), \Delta_z^{(k)} G(\xi, \hat{z}))\|.$$

By (4.8)

$$\|\Delta_z^{(k)} G(\xi, z)\| \leqslant t\, |\xi|.$$

By (4.16)

$$\|\Delta_z^{(k)} G(\xi, z) - \Delta_z^{(k)} G(\xi, \hat{z})\| \leqslant t\, |\xi|^s\, |z - z'|^\beta.$$

Also notice that $H_\beta(\tilde{\Delta}^{(k)}(u)) \leqslant 2d$ which is independent of $u \in \mathfrak{X}_d$. By the vanishing properties of $\mathscr{P}_k$, given in (6.10), and by (6.14) we have

$$\mathscr{B} \leqslant t\, |\xi|^s\, |z - z'|^\beta. \tag{2}$$

Adding (1) and (2) and taking the sup we obtain the desired estimate. ∎

6.14. PROPOSITION. *For sufficiently small $\varepsilon > 0$ for all $u \in \mathfrak{X}_d$,*

$$\begin{aligned} \mathrm{Lip}(D^k_\xi \Gamma u) &\leqslant \theta_2\, \mathrm{Lip}(D^k_\xi u) + t < d \\ \mathrm{Lip}^{(\xi)}(D^k_\xi \Gamma u) &\leqslant \theta_2\, \mathrm{Lip}^{(\xi)}(D^k_\xi u) + t < d \\ &\theta_2 := \chi^{-1} b^k \gamma < \theta_3 := \chi^{-1} b \gamma^k < \chi^{-1} b \gamma^{r-1} = \theta_o < 1. \end{aligned}$$

Proof. Recall that $\theta_o < 1$, $b < \gamma < 1$, and $r = k + \delta$. Thus $\theta_2 < \theta_3 < \theta_o < 1$. Consider (6.11) with $j = k$,

$$D^k_\xi \Gamma u(\xi, z) = B^{-1} D^k_\xi u(\xi^*, z^*)[M + D_\xi G(\xi, z)]^k + Q_k(\xi, z). \tag{6.17}$$

Recall that $\Delta^{(k-1)}_\xi u(\xi^*, z^*)$ is $C^{1,\delta}$ in (ξ, z). Moreover, $\mathrm{Lip}(\Delta^{(k)}_\xi G) < O(\varepsilon)$ by (4.12). It follows that

$$\mathrm{Lip}(Q_k) < t. \tag{a}$$

To estimate the first term in (6.17) let

$$\begin{aligned} \mathscr{A}(\xi, z) &= B^{-1} D^k_\xi u(\xi^*, z^*)[M + D_\xi G(\xi, z)]^k, \\ \mathscr{C} &= \|M + D_\xi G(\xi, z)\|^k\, \|D^k_\xi u(\xi^*, z^*) - D^k_\xi u(\tilde{\xi}, \tilde{z})\|. \end{aligned}$$

Let $\xi \neq \xi'$ and $z \neq z'$. Let $\tilde{\xi}$ and $\tilde{z} = Cz'$ be as in Lemma 6.12. Then

$$\begin{aligned} \mathscr{C} &\leqslant b^k(1+t)\, \mathrm{Lip}(D^k_\xi u)[\,|\xi^* - \tilde{\xi}| + \gamma\, |z - z'|\,] \\ &\leqslant b^k \gamma (1+t)\, \mathrm{Lip}(D^k_\xi u)[\,|\xi - \xi'| + |z - z'|\,]. \end{aligned}$$

Even if $k = 1$, $\mathrm{Lip}(D_\xi G(\xi, z)) < \varepsilon$. Thus

$$\begin{aligned} \mathscr{E} &:= \|D^k_\xi u(\tilde{\xi}, \tilde{z})\|\, \|[M + D_\xi G(\xi, z)]^k - [M + D_\xi G(\xi', z')]^k\| \\ \mathscr{E} &\leqslant t(|\xi - \xi'| + |z - z'|). \end{aligned}$$

Thus, for $u \in \mathfrak{X}_d$,

$$\|\mathscr{A}(\xi, z) - \mathscr{A}(\xi', z')\| \leqslant [\theta_2\, \mathrm{Lip}(D^k_\xi u) + t][\,|\xi - \xi'| + |z - z'|\,]. \tag{b}$$

The first estimate of the proposition follows from estimates (a) and (b).

It is obvious that the second estimate follows from the first since $z^* = Cz$ is independent of ξ. ∎

6.15. PROPOSITION. *The following hold for sufficiently small* $\varepsilon > 0$,

$$\mathrm{Lip}^{(\xi)}(D^k \Gamma u) \leqslant \theta_3 \,\mathrm{Lip}^{(\xi)}(D^k u) + t < d, \qquad u \in \mathfrak{X}_d$$
$$\theta_3 = \chi^{-1} b \gamma^k < \chi^{-1} b \gamma^{r-1} = \theta_o < 1.$$

Proof. Notice that $D^k u = (D^k_\xi u, D^k_z u, \mathfrak{D}^k u)$. So, we prove the estimate for the three parts separately.

We know from Proposition 6.14 that

$$\mathrm{Lip}^{(\xi)}(D^k_\xi \Gamma u) \leqslant \theta_2 \,\mathrm{Lip}^{(\xi)}(D^k_\xi u) + t < \theta_3 \,\mathrm{Lip}^{(\xi)}(D^k_\xi u) + t. \tag{a}$$

To consider $D^k_z u$ recall (6.10) with $j = k$,

$$D^k_z \Gamma u(\xi, z) = B^{-1} D^k_z u(\xi^*, z^*)[C]^k + P_k(\xi, z). \tag{6.18}$$

By Theorem 4 we conclude that

$$\mathrm{Lip}^{(\xi)}(\Delta^{(k)}_z G) < O(\varepsilon).$$

From the definition of $\mathfrak{X}_d$, it is obvious that $\mathrm{Lip}^{(\xi)}(\tilde{\Delta}^{(k)}_z u)$ is uniformly bounded with a bound which is independent of $u \in \mathfrak{X}_d$. Thus, from the vanishing properties of $\mathscr{P}_j$, it follows that

$$\mathrm{Lip}^{(\xi)}(P_k) = O(\varepsilon). \tag{1}$$

To consider the first term in (6.18) let $\bar{\xi} = M\xi' + f(\xi', z)$. Thus $|\xi^* - \bar{\xi}| \leqslant b(1+t)\,|\xi - \xi'|$. It follows that for

$$\mathscr{S} = \|B^{-1}[D^k_z u(\xi^*, z^*) - D^k_z u(\bar{\xi}, z^*)][C]^k\|$$
$$\mathscr{S} \leqslant \chi^{-1} b \gamma^k (1+t)\,\mathrm{Lip}^{(\xi)}(D^k_z u)\,|\xi - \xi'|. \tag{2}$$

It follows from (1) and (2) that

$$\mathrm{Lip}^{(\xi)}(D^k_z \Gamma u) \leqslant \chi^{-1} b \gamma^k \,\mathrm{Lip}^{(\xi)}(D^k_z u) + t = \theta_3 \,\mathrm{Lip}^{(\xi)}(D^k_z u) + t. \tag{b}$$

It remains to consider $\mathrm{Lip}^{(\xi)}(\mathfrak{D}^k \Gamma u)$. Recall (6.12) with $i + l = j = k$, $i \geqslant 1$, and $l \geqslant 1$,

$$\begin{aligned} D^l_\xi D^i_z \Gamma u(\xi, z) = {} & B^{-1} D^l_\xi D^i_z u(\xi^*, z^*)[C]^i\,[M + D_\xi G(\xi, z)]^l \\ & + \mathscr{W}_k(\xi, z) + R_{k-1}(\xi, z). \end{aligned} \tag{6.19}$$

It is obvious that for $0 \leqslant n \leqslant k$,

$$\text{Lip}^{(\xi)}(\mathscr{W}_k) = O(\varepsilon),$$
$$\text{Lip}^{(\xi)}([M + D_\xi G]^n) = O(\varepsilon),$$
$$\text{Lip}^{(\xi)}(R_{k-1}) = O(\varepsilon).$$

To consider the first term in (6.19) notice that for

$$\mathscr{A} = \chi^{-1} b^l \gamma^i (1+t) \, \| D_\xi^l D_z^i u(\xi^*, z^*) - D_\xi^l D_z^i u(\bar{\xi}, z^*) \|,$$
$$\mathscr{B} = \chi^{-1} b^{l+1} \gamma^i (1+t) \, \text{Lip}^{(\xi)}(D_\xi^l D_z^i u) \, |\xi - \xi'|,$$
$$\mathscr{A} \leqslant \mathscr{B}.$$

Since $\chi^{-1} b^{l+1} \gamma^i < \chi^{-1} b \gamma^k$, we have

$$\text{Lip}^{(\xi)}(D_\xi^l D_z^i \Gamma u) \leqslant \theta_3 \, \text{Lip}^{(\xi)}(D_\xi^l D_z^i u) + t.$$

Taking the maximum over all $l + i = k$, $i \geqslant 1$, $l \geqslant 1$, we obtain

$$\text{Lip}^{(\xi)}(\mathfrak{D}^k \Gamma u) \leqslant \theta_3 \, \text{Lip}^{(\xi)}(\mathfrak{D}^k u) + t. \tag{c}$$

Estimates (a), (b), and (c) prove the proposition. ∎

6.16. Proposition. *The following hold for sufficiently small* $\varepsilon > 0$:

$$H_\delta^{(z)}(\mathfrak{D}^k \Gamma u) \leqslant \theta_o H_\delta^{(z)}(\mathfrak{D}^k u) + t < d, \qquad u \in \mathfrak{X}_d$$
$$\theta_o = \chi^{-1} b \gamma^{r-1} < 1.$$

Proof. Recall (6.12) with $j = k = l + i$, $l \geqslant 1$, $i \geqslant 1$,

$$\begin{aligned} D_\xi^l D_z^i \Gamma u(\xi, z) = B^{-1} D_\xi^l D_z^i u(\xi^*, z^*)[C]^i \, [M + D_\xi G(\xi, z)]^l \\ + \mathscr{W}_k(\xi, z) + R_{k-1}(\xi, z). \end{aligned} \tag{6.20}$$

It is obvious that

$$\begin{aligned} H_\delta^{(z)}(\mathscr{W}_k) &= O(\varepsilon) \\ H_\delta^{(z)}(R_{k-1}) &= O(\varepsilon) \\ H_\delta^{(z)}([M + D_\xi G]^l) &= O(\varepsilon). \end{aligned} \tag{i}$$

Let $\hat{\xi} = M\xi + f(\xi, z')$. Thus, by (6.14), $|\hat{\xi} - \xi^*| \leqslant \varepsilon\, |z - z'|$. It follows that

$$\begin{aligned}\|D_\xi^l D_z^i u(\xi^*, Cz) - D_\xi^l D_z^i u(\hat{\xi}, Cz')\| &\leqslant H_\delta^{(z)}(D_\xi^l D_z^i u)\, |Cz - Cz'|^\delta \\ &\quad + \mathrm{Lip}^{(\xi)}(D_\xi^l D_z^i u)\, |\xi^* - \hat{\xi}| \\ &\leqslant \gamma^\delta (1 + t)\, H_\delta^{(z)}(D_\xi^l D_z^i u)\, |z - z'|.\end{aligned}$$

Since $i + l - 1 + \delta = k - 1 + \delta = r - 1$, it follows that

$$\begin{aligned}\mathcal{Q} &= \chi^{-1} \gamma^i b^l (1 + t)\, \|D_\xi^l D_z^i u(\xi^*, Cz) - D_\xi^l D_z^i u(\hat{\xi}, Cz')\| \\ \mathcal{Q} &\leqslant \chi^{-1} b^l \gamma^{i+\delta} (1 + t)\, H_\delta^{(z)}(D^l{}_\xi D_z^i u)\, |z - z'| \\ &\leqslant \chi^{-1} b \gamma^{r-1} (1 + t)\, H_\delta^{(z)}(D_\xi^l D_z^i u)\, |z - z'|.\end{aligned} \tag{ii}$$

Estimates (i) and (ii) imply that

$$H_\delta^{(z)}(D_\xi^l D_z^i \Gamma u) \leqslant \theta_o H_\delta^{(z)}(D_\xi^l D_z^i u) + t.$$

Taking the maximum over all $l + i = k$, $i \geqslant 1$, $l \geqslant 1$, we obtain the desired result. ∎

6.17. Corollary. *For sufficiently small* $\varepsilon > 0$, *there is* $\theta < 1$ *such that*

$$\sigma(\Gamma u) \leqslant \theta \sigma(u) + t < d, \qquad u \in \mathfrak{X}_d.$$

Proof. The corollary follows from Propositions 6.13, 6.14, 6.15, and 6.16. ∎

6.18. Proposition. *For sufficiently small* $\varepsilon > 0$, *there* $0 < \theta < 1$ *such that, if* $u \in \mathfrak{X}_d$, *then* $\Gamma u \in \mathfrak{X}_d$ *and* $\|\Gamma u\|_* < \theta\, \|u\|$.

Proof. This proposition follows from Corollaries 6.11 and 6.17. ∎

Proposition 6.18 finishes the proof of Step 2 in Section 6.3. Steps 1 and 3 were proved. So, to finish the proof of Theorem 5 it remains to prove Step 4, that is, to prove assertions (2) and (3) of the theorem. This will be done in Proposition 6.19.

6.19. Proposition. *For sufficiently small* $0 < d < 1$,

$$Q(x, y, z) = (x, u(x, y, z), z)$$

is a $C^{k,\beta}$ *diffeomorphism. The new map* (4.23) *satisfies estimates* (4.28)–(4.31).

Proof. Since $0<d<1$, $Q=id+(0, u, 0)$ has an inverse in a neighbourhood of $(0, 0, 0)$. By the Inverse Function Theorem, Q^{-1} is as smooth as Q and hence $v(X, Y, Z)$ is as smooth as $u(x, y, z)$.

Recall that $H=f\circ Q^{-1}$. The $O(\varepsilon)$ factors in estimates (4.28)–(4.31) follow from the presence of f. The smoothness part in these estimates follow from the fact that Q^{-1} is as smooth as Q. ∎

7. PROOF OF THEOREM 4

We have proved Theorem 5 granting Theorem 4. In this section we prove Theorem 4.

Proof of (4.7). Recall estimate (5.15) on S. Thus, for some $0\leqslant c\leqslant 1$, since $|\xi|\leqslant 1$ and $|z|\leqslant 1$

$$|g(\xi, z)|\leqslant \|D_\eta S(c\xi+w(z), z)\|\ |\xi|\leqslant O(\varepsilon)\ |\xi|\ [\,|\xi|^{r-1}+|z|^{r-1}]$$

which proves (4.7). ∎

Proof of (4.8). Let $\rho(\xi, z)=(\xi+w(z), z)$. It is obvious that $\rho(\xi, z)$ is $C^{k,\delta}$ and that $D^j\rho(\xi, z)$, $j=1, ..., k$, are independent of ξ. Thus, for $1\leqslant j\leqslant k$

$$D_z^jG(\xi, z)=[D_z^j(F\circ\rho)](\xi, z)-[D_z^j(F\circ\rho)](0, z) \tag{7.1}$$

$$\begin{aligned}
&[D_z^j(F\circ\rho)](\xi, z)=D^jF(\xi+w(z), z)[D_z\rho(\xi, z)]^j+C_{j-1}(\xi, z)\\
&C_{j-1}(\xi, z)=\mathscr{C}_{j-1}(\varDelta^{(j-1)}F(\xi+w(z), z), \varDelta_z^{(j)}\rho(\xi, z))\\
&\mathscr{C}_{j-1}(0, \cdot)\equiv 0\\
&\mathscr{C}_{j-1}(\cdot, 0)\equiv 0,
\end{aligned} \tag{7.2}$$

where $\mathscr{C}_{j-1}(\Sigma, \Omega)$ is a polynomial in the components of Σ and Ω and linear in the components of Σ. Recall that $\mathrm{Lip}_\eta(D_z^kG)<\varepsilon$. Thus, letting $j=k$ in (7.1) we obtain

$$\|D_z^kG(\xi, z)\|\leqslant\varepsilon\,|\xi|$$

which proves (4.8). ∎

Proof of (4.9). Consider (7.1) and (7.2) for (g, S). For $1\leqslant j\leqslant k-1$, let

$$C_{j-1}=(C_{j-1}^f, C_{j-1}^g).$$

Then $D_z^j(S\circ\rho)$ is $C^{1,\delta}$ and hence for some $0\leqslant c\leqslant 1$

$$\|D_z^jg(\xi, z)\|\leqslant\|[D_\xi D_z^j(S\circ\rho)](c\xi, z)\|\ |\xi|.$$

Notice that $\Delta^j \rho$ is independent of ξ. Thus by (5.5) and (5.15)

$$\|D_\xi C^{(g)}_{j-1}(c\xi, z)\| \leqslant O(\varepsilon)[\,|\xi|^{r-j} + |z|^{r-j}]$$
$$\|D_\xi D_z^j S(c\xi + w(z), z)[D_z \rho(\xi, z)]^j\| \leqslant O(\varepsilon)[\,|\xi|^{r-j-1} + |z|^{r-j-1}]$$

which proves (4.9). ∎

Proof of (4.10) *and* (4.11). Let

$$D_*^j = D^j - D_z^j = \{D_\xi^l D_z^i \mid i + l = j, l \geqslant 1\}.$$

Since $l \geqslant 1$ in each term of D_*^j and since $D^m \rho(\xi, z)$, $m = 1, ..., k$, are independent of ξ it follows from (7.2) that

$$\begin{aligned}\|D_\xi^l D_z^i g(\xi, z)\| &= \|[D_\xi^l D_z^i (S \circ \rho)](\xi, z)\| \\ &\leqslant \|D_\xi^l D_z^i S(\xi + w(z), z)[D_z \rho(\xi, z)]^i\| + \|D_\xi^l C^{(g)}_{i-1}(\xi, z)\| \\ &\leqslant O(\varepsilon)[(|\xi| + |w(z)|)^{r-j} + |z|^{r-j}] \\ &\quad + O(\varepsilon)[\,|\xi|^{r-(l+i-1)} + |z|^{r-(l+i-1)}] \\ &\leqslant O(\varepsilon)[\,|\xi|^{r-j} + |z|^{r-j}]\end{aligned}$$

which proves (4.10) and (4.11). ∎

Proof of (4.12). Notice that

$$D_\xi^k G(\xi, z) = [D_\xi^k (F \circ \rho)](\xi, z) = D_\eta^k F(\xi + w(z), z)\, I_1^k,$$

where I_1 is the identity map $I_1(\xi) = \xi$. Now, (4.12) follows from (5.3). ∎

Proof of (4.13). Estimate (4.13) follows from (5.5) and the facts that $G(\xi, z)$ is $C^{k,\delta}$ and that each term in $D^k G(\xi, z)$ has a factor of $\Delta^k F(\xi + w(z), z)$ or $\Delta^k F(w(z), z)$ which produces an $O(\varepsilon)$ factor. ∎

Proof of (4.14). In view of (4.12), $D_\xi^k G(\xi, z)$ satisfies (4.14). Thus to prove (4.14) we need to consider only $D_z^k G(\xi, z)$ and $\mathfrak{D}^k G(\xi, z)$. By (7.1), (7.2) and (5.3), it follows that $D_z^k G(\xi, z)$ satisfies (4.14).

To study $\mathfrak{D}^k G(\xi, z)$ let (i, l) be such that $i + l = k$, $i \geqslant 1$ and $l \geqslant 1$. Since $i \geqslant 1$ we have

$$\begin{aligned}D_z^l D_\xi^i G(\xi, z) &= D_z^l [D_\eta^i F(\xi + w(z), z)]\, I_1^i \\ &= D_z^l D_\eta^i F(\xi + w(z), z)\, I_1^i I_2^l + \Psi_k(\xi, z) + B_{l-1}(\xi, z)\end{aligned}$$

$$\Psi_k(\xi, z) = \sum_{s=1}^{l} c_s D_z^{l-s} D_\eta^{i+s} F(\xi + w(z), z)\, I_1^i [Dw(z)]^s\, I_2^{l-s}$$

$$B_{l-1}(\xi, z) = \mathscr{B}_{l-1}([\Delta^{(l-1)}(D_\eta^i F)](\rho(\xi, z)), \Delta^{(l-1)} \rho(\xi, z))$$

$$\mathscr{B}_{l-1}(0, \cdot) \equiv 0,$$

where $\mathscr{B}_{l-1}(\Sigma, \Omega)$ is a polynomial in the components of its arguments and linear in Σ. Thus, $B_{l-1}(\xi, z)$ is $C^{1,\delta}$. From the vanishing property of $\mathscr{B}_{l-1}$ it follows that $\mathrm{Lip}^{(\xi)}(B_{l-1}) = O(\varepsilon)$. Estimate (5.3) implies that $\mathrm{Lip}^{(\xi)}$ of the remaining terms is also $O(\varepsilon)$. This proves (4.14). ∎

Proof of (4.15). For $0 \leqslant j \leqslant k-1$, estimate (4.15) follows from (4.9), (4.8) and (4.8) because for some $0 \leqslant c \leqslant 1$

$$\|D_z^j G(\xi, z) - D_z^j G(\xi, z')\| \leqslant \|D_z^{j+1} G(\xi, z + c(z - z'))\| \, |z - z'|.$$

When $j = k$ estimate (4.15) follows from (5.3). ∎

Proof of (4.16). Let

$$\mathscr{J} = \|D_z^j G(\xi, z) - D_z^j G(\xi, z')\|.$$

Notice that for $0 \leqslant j \leqslant k-1$, $D_z^j G(\xi, z)$ is $C^{1,\delta}$ and $\mathrm{Lip}(D_z^j G) < O(\varepsilon)$. Thus by (4.13) and (4.15) we have

$$\begin{aligned} \mathscr{J} &= \mathscr{J}^s \mathscr{J}^{1-s} \\ &\leqslant O(\varepsilon) \, |\xi|^s \, |z - z'|^{\delta(1-s)}. \quad ∎ \end{aligned}$$

Proof of (4.17) *and* (4.18). The proof of theses two estimates are similar to those of (4.12) and (4.14), respectively. ∎

This finishes the proof of Theorem 4.

8. ON THE HOMOLOGICAL EQUATION IN BANACH SPACES

In order to eliminate a nonresonant symmetric j-multilinear term $g_j(x)$ from a map of Banach spaces, we need to solve the homological equation (5.9). In doing so, we use Lemma 5.4. The technical part of that lemma is the first assertion which is simple for maps of finite dimensional spaces. The proof for maps of Banach spaces can be found in [15]. We present the proof in our notation for completeness.

8.1. Definition. Let X and Y be Banach spaces.

- Let $\mathscr{L}(X^n, Y)$ be the Banach space of n-multilinear functions from X^j to Y.
- Let $\mathscr{S}(X^n, Y)$ be the Banach space of symmetric n-multilinear functions from X to Y.

- Let $A \in \mathscr{L}(X, Y)$ and $C \in \mathscr{L}(Y, Y)$. Define a linear transformation on $\mathscr{S}(X^n, Y)$ by

$$\begin{aligned} Q_{A,C} &: \mathscr{S}(X^n, Y) \to \mathscr{S}(X^n, Y), \\ [Q_{A,C}u](x_1, ..., x_n) &= Cu(Ax_1, ..., Ax_n). \end{aligned} \tag{8.1}$$

- For A and C as above, define

$$\begin{aligned} L_C &: \mathscr{L}(X^n, Y) \to \mathscr{L}(X^n, Y), \\ L_C u &= Cu, \\ R^i_A &: \mathscr{L}(X^n, Y) \to \mathscr{L}(X^n, Y), \\ [R^i_A u](x_1, ..., x_n) &= u(x_1, ..., x_{i-1}, Ax_i, x_{i+1}, ..., x_n). \end{aligned}$$

- For a linear map $T \in \mathscr{L}(E, E)$ let $\rho(T; E)$ and $\sigma(T; E)$ be its resolvent and spectrum respectively. When there is no ambiguity we write $\rho(T)$ and $\sigma(T)$.
- Let $\sigma_i \subset \mathbb{C}$, $i = 1, ..., n$. Let

$$\sigma_1 \sigma_1 \cdots \sigma_n = \{\alpha_1 \alpha_2 \cdots \alpha_n \mid \alpha_i \in \sigma_i, 1 \leqslant i \leqslant n\}.$$

- Let I_E denote the identity map on a space E.

8.2. Lemma. *Let $Q_{A,C}: \mathscr{S}(X^n, Y) \to \mathscr{S}(X^n, Y)$ be as in* (8.1). *Then*

$$\sigma(Q_{A,C}) \subset \sigma(C)\, \sigma(A)^n. \tag{8.2}$$

Proof. (1) Let $A_1, A_2 \in \mathscr{L}(X, X)$ and $C_1, C_2 \in \mathscr{L}(Y, Y)$. Let $1 \leqslant i < j \leqslant n$. Then

$$\begin{aligned} R^i_{A_2} R^i_{A_1} &= R^i_{A_1 A_2}, \\ R^i_{aA_1 + bA_2} &= aR^i_{A_1} + bR^i_{A_2}, \\ L_{C_2} L_{C_1} &= L_{C_2 C_1}, \\ L_{aC_1 + bC_2} &= aL_{C_1} + bL_{C_2}, \\ R^i_{I_X} &= I_{\mathscr{S}(X^n, Y)}, \\ L_{I_Y} &= I_{\mathscr{S}(X^n, Y)}, \\ R^i_{A_2} R^j_{A_1} &= R^j_{A_1} R^i_{A_2}, \\ L_C R^i_A &= R^i_A L_C. \end{aligned} \tag{8.3}$$

(2) If $[C-\gamma I_Y]$ is invertible, then by (8.3)

$$[L_C-\gamma I_{\mathscr{S}(X^n,\,Y)}]^{-1}=[L_C-\gamma L_{I_Y}]^{-1}=L_{[C-\gamma I_Y]^{-1}}.$$

If $[A-\alpha I_X]$ is invertible, then by (8.3)

$$[R^i_A-\alpha I_{\mathscr{S}(X^n,\,Y)}]^{-1}=[R^i_A-\alpha R^i_{I_X}]^{-1}=R^i_{[A-\alpha I_X]-1}.$$

It follows that $\rho(C)\subset\rho(L_C)$ and $\rho(A)\subset\rho(R^i_A)$. Thus

$$\begin{aligned}\sigma(L_C)&\subset\sigma(C)\\ \sigma(R^i_A)&\subset\sigma(A).\end{aligned}\tag{8.4}$$

(3) Theorem 11.23 of [20] states that if two elements T and S of a Banach algebra commute, then $\sigma(TS)=\sigma(T)\,\sigma(S)$. It follows that

$$\sigma(R^1_A\cdots R^n_A;\mathscr{L}(X^n,\,Y))\subset\sigma(A)^n.\tag{8.5}$$

(4) Now we would like to show that

$$\rho(R^1_A\cdots R^n_A;\mathscr{L}(X^n,\,Y))\subset\rho(R^1_A\cdots R^n_A;\mathscr{S}(X^n,\,Y)),\tag{8.6}$$

which implies that

$$\sigma(R^1_A\cdots R^n_A;\mathscr{S}(X^n,\,Y))\subset\sigma(R^1_A\cdots R^n_A;\mathscr{L}(X^n,\,Y))\subset\sigma(A)^n.\tag{8.7}$$

To simplify notation we let $R^{[n]}_A=R^1_A\cdots R^n_A$, $\mathscr{L}^n=\mathscr{L}(X^n,\,Y)$ and $\mathscr{S}^n=\mathscr{S}(X^n,\,Y)$. Let $\alpha\in\rho(R^{[n]}_A;\mathscr{L}^n)$. In order to show that $\alpha\in\rho(R^{[n]}_A;\mathscr{S}^n)$ we need to show the following: Given any $w\in\mathscr{S}^n$, there is $\phi\in\mathscr{S}^n$ such that

$$R^{[n]}_A\phi-\alpha\phi=w.\tag{8.8}$$

Since $\alpha\in\rho(R^{[n]}_A;\mathscr{L}^n)$, there is $\psi\in\mathscr{L}^n$ such that

$$\begin{aligned}R^{[n]}_A\psi-\alpha\psi&=w\\ \|\psi\|&\leqslant K\,\|w\|\\ K&=\|[R^{[n]}_A-\alpha I_{\mathscr{L}^n}]^{-1}\|.\end{aligned}\tag{8.9}$$

For any permutation τ of $\{1, 2, ..., n\}$, since $w\in\mathscr{S}^n$, we have

$$R^{[n]}_A\psi\circ\tau-\alpha\psi\circ\tau=w\circ\tau=w.\tag{8.10}$$

On averaging over all such permutations, we obtain the coveted $\phi \in \mathscr{S}^n$:

$$\phi := \frac{1}{n!} \sum_{\tau} \psi \circ \tau.$$

In view of (8.10), ϕ satisfies

$$R_A^{[n]} \phi \circ \tau - \alpha\phi = w.$$

This proves (8.6) which implies (8.7).

(5) Since $R_A^{[n]}$ and L_C commute, (8.4) and (8.7) show that

$$\sigma(K_{A,C}) = \sigma(L_C R_A^{[n]}) \subset \sigma(L_C)\,\sigma(R_A^{[n]}) \subset \sigma(C)\,\sigma(A)^n,$$

which finishes the proof of the lemma. ∎

REFERENCES

1. V. I. Arnold, "Geometrical Methods in the Theory of Differential Equations," Springer-Verlag, New York/Berlin, 1983.
2. A. Banyaga, R. de la Llave, and C. E. Wayne, Cohomology equations near hyperbolic points and geometric version of Sternberg linearization theorem, *Journal of Geometric Analysis* **6** (1997), 613–649.
3. G. R. Belitskii, Functional equations and conjugacy of local diffeomorphisms of a finite smoothness class, *Funct. Anal. Appl.* **7** (1973), 268–277.
4. G. R. Belitskii, Equivalence and normal forms of germs of smooth mappings, *Russian Math. Surveys* **11** (1978), 107–177.
5. A. Bryuno (Bruno), "Local Methods in Nonlinear Differential Equations," Springer-Verlag, New York/Berlin, 1989.
6. A. Bryuno (Bruno), Smooth linearization of differential equations, *Soviet Math. Dokl.* **45**, No. 1 (1992), 105–109.
7. A. Bryuno (Bruno), On a finitely smooth linearization of differential equations near a hyperbolic singular point, *Soviet Math. Dokl.* **43**, No. 3 (1992), 711–715.
8. S.-N. Chow and K. Lu, Invariant manifolds for flows in Banach spaces, *J. Differential Equations* **74** (1988), 285–317.
9. M. S. ElBialy, On pseudo-stable and pseudo-unstable manifolds for maps, *J. Math. Anal. Appl.* **232** (1999), 229–258.
10. M. S. ElBialy, Sub-stable and weak-stable manifolds associated with finitely non-resonant spectral subspaces, *Mathematisch Zeitschrift*, accepted.
11. P. Hartman, On local homeomorphisms of Euclidean spaces, *Bol. Soc. Mat. Mexicana* **5** (1960), 220–241.
12. P. Hartman, "Ordinary Differential Equations," 2nd ed., Birkhäuser, Basel, 1982.
13. M. W. Hirsh, C. C. Pugh, and M. Shub, "Invariant Manifolds," Lecture Notes in Mathematics, Vol. 583, Springer-Verlag, New York, 1977.
14. N. Kurzweil, On approximation in real Banach spaces, *Studia Math.* **14** (1954), 213–231.

15. R. de la Llave and C. E. Wayne, Invariant manifolds associated to non-resonant invariant subspaces for some non-invertible maps and flows, preprint.
16. E. B. Leach and J. H. M. Whitfiels, Differentiable functions and rough norms on Banach spaces, *Proc. Amer. Math. Soc.* **33** (1979), 120–126.
17. K. Lu, A Hartman–Grobman theorem for scalar reaction-diffusion equations, *J. Differential Equations* **93** (1991), 364–394.
18. E. Nelson, "Topics in Dynamics. I. Flows," Princeton Univ. Press and Univ. of Tokyo, Princeton, NJ, 1969.
19. C. Pugh, On a theorem of P. Hartman, *Amer. J. Math.* **91** (1969), 363–367.
20. W. Rudin, "Functional Analysis," 2nd ed., McGraw–Hill, New York, 1991.
21. V. S. Samovol, A criterion for C^1-smooth linearization of an autonomous system in a neighborhood of a nondegenerate singular point, *Math. Notes* **49**, No. 3–4 (1991), 288–292.
22. V. S. Samovol, On an obstruction to local smooth linearization of a smooth ordinary differential equation, *Math. Notes* **57**, No. 5–6 (1995), 662–664.
23. V. S. Samovol, A necessary condition for the local smooth linearization of a system of differential equations, *Dokl. Akad. Nauk* **351**, No. 2 (1996), 169–171.
24. G. R. Sell, Smooth linearization near a fixed point, *Amer. J. Math.* **107** (1985).
25. S. Sternberg, Local contractions and a theorem of Poincaré, *Amer. J. Math.* **79** (1957), 809–824.
26. D. Stowe, Linearization in two dimensions, *J. Differential Equations* **63** (1986), 183–226.
27. F. Takens, Partially hyperbolic fixed points, *Topology* **10** (1971), 133–147.
28. A. Vanderbauwhede and S. A. van Gils, Center manifolds and contractions on a scale of Banach spaces, *J. Funct. Anal.* **72** (1987), 209–224.