

Linearization via the Lie Derivative

Carmen Chicone* Richard Swanson†

November 16, 2000

*Department of Mathematics, University of Missouri, Columbia, MO 65211

†Department of Mathematical Sciences, Montana State University, Bozeman, MT
59717-0240

Contents

1	Introduction	2
2	Continuous Conjugacy	5
3	Smooth Conjugacy	9
3.1	Hyperbolic Sinks	12
3.1.1	Smooth Linearization on the Line	39
3.2	Hyperbolic Saddles	42
4	Linearization of Special Vector Fields	54
4.1	Special Vector Fields	55
4.2	Saddles	60
4.3	Infinitesimal Conjugacy and Fiber Contractions	61
4.4	Sources and Sinks	62

Abstract

The standard proof of the Grobman–Hartman linearization theorem for a flow at a hyperbolic rest point proceeds by first establishing the analogous result for hyperbolic fixed points of local diffeomorphisms. In this exposition we present a simple direct proof that avoids the discrete case altogether. We give new proofs for Hartman’s smoothness results: A \mathcal{C}^2 flow is \mathcal{C}^1 linearizable at a hyperbolic sink, and a \mathcal{C}^2 flow in the plane is \mathcal{C}^1 linearizable at a hyperbolic rest point. Also, we formulate and prove some new results on smooth linearization for special classes of quasi-linear vector fields where either the nonlinear part is restricted or additional conditions on the spectrum of the linear part (not related to resonance conditions) are imposed.

1 Introduction

This paper is divided into three parts. In the first part, a new proof is presented for the Grobman–Hartman linearization theorem: A \mathcal{C}^1 flow is \mathcal{C}^0 linearizable at a hyperbolic rest point. The second part is a discussion of Hartman’s results on smooth linearization where smoothness of the linearizing transformation is proved in those cases where resonance conditions are not required. For example, we will use the theory of ordinary differential

equations to prove two main theorems: A \mathcal{C}^2 vector field is \mathcal{C}^1 linearizable at a hyperbolic sink; and, a \mathcal{C}^2 vector field in the plane is \mathcal{C}^1 linearizable at a hyperbolic rest point. In the third part, we will study a special class of vector fields where the smoothness of the linearizing transformation can be improved.

The proof of the existence of a smooth linearizing transformation at a hyperbolic sink is delicate. It uses a version of the stable manifold theorem, consideration of the gaps in the spectrum of the linearized vector field at the rest point, carefully constructed Gronwall type estimates, and an induction argument. The main lemma is a result about partial linearization by near-identity transformations that are continuously differentiable with Hölder derivatives. The method of the proof requires the Hölder exponent of these derivatives to be less than a certain number, called the Hölder spectral exponent, that is defined for linear maps as follows. Suppose that $\{-b_1, -b_2, \dots, -b_N\}$ is the set of real parts of the eigenvalues of the linear transformation $A : \mathbf{R}^n \rightarrow \mathbf{R}^n$ and

$$-b_N < -b_{N-1} < \dots < -b_1 < 0. \quad (1.1)$$

The *Hölder spectral exponent* of A is the number

$$\frac{b_1(b_{j+1} - b_j)}{b_1(b_{j+1} - b_j) + b_{j+1}b_j}$$

where

$$\frac{b_{j+1} - b_j}{b_{j+1}b_j} = \min_{i \in \{1, 2, \dots, N-1\}} \frac{b_{i+1} - b_i}{b_{i+1}b_i}$$

in case $N > 1$; it is the number one in case $N = 1$. The Hölder spectral exponent of a linear transformation B whose eigenvalues all have positive real parts is the Hölder spectral exponent of $-B$.

Although a \mathcal{C}^2 flow in the plane is always \mathcal{C}^1 linearizable at a hyperbolic rest point, a \mathcal{C}^2 flow in \mathbf{R}^3 may not be \mathcal{C}^1 linearizable at a hyperbolic saddle point. For example, the flow of the system

$$\dot{x} = 2x, \quad \dot{y} = y + xz, \quad \dot{z} = -z \quad (1.2)$$

is not \mathcal{C}^1 linearizable at the origin (see Hartman's example (3.1)). We will prove that a flow in \mathbf{R}^n can be smoothly linearized at a hyperbolic saddle if the spectrum of the corresponding linearized system at the saddle point satisfies the following condition introduced by Hartman in [H60M]. Note first

that the real parts of the eigenvalues of the system matrix of the linearized system at a hyperbolic saddle lie in the union of two intervals, say $[-a_L, -a_R]$ and $[b_L, b_R]$ where $a_L, a_R, b_L,$ and b_R are all positive real numbers. Thus, the system matrix can be written as a direct sum $A \oplus B$ where the real parts of the eigenvalues of A are in $[-a_L, -a_R]$ and the real parts of the eigenvalues of B are in $[b_L, b_R]$. Let μ denote the Hölder spectral exponent of A and ν the Hölder spectral exponent of B . If Hartman's spectral condition

$$a_L - a_R < \mu b_L, \quad b_R - b_L < \nu a_R$$

is satisfied, then the C^2 nonlinear system is C^1 linearizable at the hyperbolic saddle point. It follows that, unlike system (1.2), the flow of

$$\dot{x} = 2x, \quad \dot{y} = y + xz, \quad \dot{z} = -4z$$

is C^1 linearizable at the origin.

In the case of hyperbolic saddles where one of the Hölder spectral exponents is small, Hartman's spectral condition is satisfied only if the corresponding real parts of the eigenvalues of the linear part of the field are contained in an accordingly small interval. Although the situation cannot be improved for general vector fields, stronger results (in the spirit of Hartman) are possible for a restricted class of vector fields. There are at least two ways to proceed: additional conditions can be imposed on the spectrum of the linearization, or restrictions can be imposed on the nonlinear part of the vector field. We will show that a C^3 vector field in "triangular form" with a hyperbolic saddle point at the origin can be C^1 linearized if Hartman's spectral condition is replaced by the inequalities $a_L - a_R < b_L$ and $b_R - b_L < a_R$ (see Theorem 4.6). Also, we will prove the following result: Suppose that $X = \mathcal{A} + \mathcal{F}$ is a quasi-linear C^3 vector field with a hyperbolic saddle at the origin, the set of negative real parts of eigenvalues of \mathcal{A} is given by $\{-\lambda_1, \dots, -\lambda_p\}$, the set of positive real parts is given by $\{\sigma_1, \dots, \sigma_q\}$, and

$$-\lambda_1 < -\lambda_2 < \dots < -\lambda_p < 0 < \sigma_q < \sigma_{q-1} < \dots < \sigma_1.$$

If $\lambda_{i-1}/\lambda_i > 3$, for $i \in \{2, 3, \dots, p\}$, and $\sigma_{i-1}/\sigma_i > 3$, for $i \in \{2, 3, \dots, q\}$, and if $\lambda_1 - \lambda_p < \sigma_q$ and $\sigma_1 - \sigma_q < \lambda_p$, then X is C^1 linearizable (see Theorem 4.7).

The important dynamical behavior of a nonlinear system associated with a hyperbolic sink is local: there is an open basin of attraction and every trajectory that enters this set is asymptotically attracted to the sink. This

behavior is adequately explained by using a linearizing homeomorphism, that is, by using the Grobman–Hartman theorem. On the other hand, the interesting dynamical behavior associated with saddles is global; for example, limit cycles are produced by homoclinic loop bifurcations and chaotic invariant sets are found near transversal intersections of homoclinic manifolds. Smooth linearizations at hyperbolic saddle points are used to analyze these global phenomena. It turns out that results on the smooth linearization at hyperbolic sinks are key lemmas required to prove the existence of smooth linearization for hyperbolic saddles. In fact, this is the main reason to study smooth linearization at hyperbolic sinks.

We treat only the case of rest points here, but we expect that our method can be applied to the problem of linearization near general invariant manifolds of differential equations.

Hartman’s article [H60M] is the main reference for our results on smoothness of linearizations. Other primary sources are the papers [G59], [H60], [H63], and [St57]. For historical remarks, additional references, and later work see [CL88], [CLL91], [KP90], [Se85], [St89], and [T99].

2 Continuous Conjugacy

A \mathcal{C}^1 vector field X on \mathbf{R}^n such that $X(0) = 0$ is called *locally topologically conjugate* to its linearization $A := DX(0)$ at the origin if there is a homeomorphism $h : U \rightarrow V$ of neighborhoods of the origin such that the flows of X and A are locally conjugated by h ; that is,

$$h(e^{tA}x) = X_t(h(x)) \tag{2.1}$$

whenever $x \in U$, $t \in \mathbf{R}^n$, and both sides of the conjugacy equation are defined. A matrix is *infinitesimally hyperbolic* if every one of its eigenvalues has a nonzero real part.

Theorem 2.1 (Grobman–Hartman). *Let X be a \mathcal{C}^1 vector field on \mathbf{R}^n such that $X(0) = 0$. If the linearization A of X at the origin is infinitesimally hyperbolic, then X is locally topologically conjugate to A at the origin.*

Proof. For each $r > 0$ there is a smooth bump function $\rho : \mathbf{R}^n \rightarrow [0, 1]$ with the following properties: $\rho(x) \equiv 1$ for $|x| < r/2$, $\rho(x) \equiv 0$ for $|x| > r$, and $|d\rho(x)| < 4/r$ for $x \in \mathbf{R}^n$. The vector field $Y = A + \xi$ where $\xi(x) :=$

$\rho(x)(X(x) - Ax)$ is equal to X on the open ball of radius $r/2$ at the origin. Thus, it suffices to prove that Y is locally conjugate to A at the origin.

Suppose that in equation (2.1) $h = \text{id} + \eta$ and $\eta : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is differentiable in the direction A . Rewrite equation (2.1) in the form

$$e^{-tA}h(e^{tA}x) = e^{-tA}X_t(h(x)) \quad (2.2)$$

and differentiate both sides with respect to t at $t = 0$ to obtain the *infinitesimal conjugacy equation*

$$L_A\eta = \xi \circ (\text{id} + \eta) \quad (2.3)$$

where

$$L_A\eta := \frac{d}{dt}(e^{-tA}\eta(e^{tA}))\Big|_{t=0} \quad (2.4)$$

is the Lie derivative of η along A . (We note that if h is a conjugacy, then the right-hand-side of equation (2.2) is differentiable; and therefore, the Lie derivative of h in the direction A is defined.)

We will show that if $r > 0$ is sufficiently small, then the infinitesimal conjugacy equation has a bounded continuous solution $\eta : \mathbf{R}^n \rightarrow \mathbf{R}^n$ (differentiable along A) such that $h := \text{id} + \eta$ is a homeomorphism of \mathbf{R}^n whose restriction to the ball of radius $r/2$ at the origin is a local conjugacy as in equation (2.1).

Since A is infinitesimally hyperbolic, $A = A^+ \oplus A^-$ having spectra, respectively, to the left and to the right of the imaginary axis. Put $\mathbf{E}^- = \text{Range}(A^-)$ and $\mathbf{E}^+ = \text{Range}(A^+)$. There are positive constants C and λ such that

$$|e^{tA}v^+| \leq Ce^{-\lambda t}|v^+|, \quad |e^{-tA}v^-| \leq Ce^{-\lambda t}|v^-| \quad (2.5)$$

for $t \geq 0$. The Banach space \mathcal{B} of bounded (in the supremum norm) continuous vector fields on \mathbf{R}^n splits into the complementary subspaces \mathcal{B}^+ and \mathcal{B}^- of vector fields with ranges, respectively, in \mathbf{E}^+ or \mathbf{E}^- . In particular, a vector field $\eta \in \mathcal{B}$ has a unique representation $\eta = \eta^+ + \eta^-$ where $\eta^+ \in \mathcal{B}^+$ and $\eta^- \in \mathcal{B}^-$.

The function G on \mathcal{B} defined by

$$G\eta(x) = \int_0^\infty e^{tA}\eta^+(e^{-tA}x) dt - \int_0^\infty e^{-tA}\eta^-(e^{tA}x) dt \quad (2.6)$$

is a bounded linear operator $G : \mathcal{B} \rightarrow \mathcal{B}$. The boundedness of G follows from the hyperbolic estimates (2.5). The continuity of the function $x \mapsto G\eta(x)$ is an immediate consequence of the following lemma from advanced calculus—essentially the Weierstrass M -test—and the hyperbolic estimates.

Lemma 2.2. *Suppose that $f : [0, \infty) \times \mathbf{R}^n \rightarrow \mathbf{R}^m$, given by $(t, x) \mapsto f(t, x)$, is continuous (respectively, the partial derivative f_x is continuous). If for each $y \in \mathbf{R}^n$ there is an open set $S \subset \mathbf{R}^n$ with compact closure \bar{S} and a function $M : [0, \infty) \rightarrow \mathbf{R}$ such that $y \in S$, the integral $\int_0^\infty M(t) dt$ converges, and $|f(t, x)| \leq M(t)$ (respectively, $|f_x(t, x)| \leq M(t)$) whenever $t \in [0, \infty)$ and x is in \bar{S} , then $F : \mathbf{R}^n \rightarrow \mathbf{R}^m$ given by $F(x) = \int_0^\infty f(t, x) dt$ is continuous (respectively, F is continuously differentiable and $DF(x) = \int_0^\infty f_x(t, x) dt$).*

Using the definition of L_A in display (2.4) and the fundamental theorem of calculus, we have the identity $L_A G = \text{id}_{\mathcal{B}}$. As a consequence, if

$$\eta = G(\xi \circ (\text{id} + \eta)) := F(\eta), \quad (2.7)$$

then η is a solution of the infinitesimal conjugacy equation (2.3).

Clearly, $F : \mathcal{B} \rightarrow \mathcal{B}$ and for η_1 and η_2 in \mathcal{B} we have that

$$\begin{aligned} \|F(\eta_1) - F(\eta_2)\| &\leq \|G\| \|\xi \circ (\text{id} + \eta_1) - \xi \circ (\text{id} + \eta_2)\| \\ &\leq \|G\| \|D\xi\| \|\eta_1 - \eta_2\|. \end{aligned}$$

Using the definitions of ξ and the properties of the bump function ρ , we have that

$$\|D\xi\| \leq \sup_{|x| \leq r} \|DX(x) - A\| + \frac{4}{r} \sup_{|x| \leq r} |X(x) - Ax|.$$

By the continuity of DX , there is some positive number r such that $\|DX(x) - A\| < 1/(10\|G\|)$ whenever $|x| \leq r$. By Taylor's theorem (applied to the \mathcal{C}^1 function X) and the obvious estimate of the integral form of the remainder, if $|x| \leq r$, then $|X(x) - Ax| < r/(10\|G\|)$. For the number $r > 0$ just chosen, we have the estimate $\|G\|\|D\xi\| < 1/2$; and therefore, F is a contraction on \mathcal{B} . By the contraction mapping theorem applied to the restriction of F on the closed subspace \mathcal{B}_0 of \mathcal{B} consisting of those elements that vanish at the origin, the equation (2.7) has a unique solution $\eta \in \mathcal{B}_0$, which also satisfies the infinitesimal conjugacy equation (2.3).

We will show that $h := \text{id} + \eta$ is a local conjugacy. To do this recall the following elementary fact about Lie differentiation: If U , V , and W are vector fields, ϕ_t is the flow of U , and $L_U V = W$, then

$$\frac{d}{dt} D\phi_{-t}(\phi_t(x))V(\phi_t(x)) = D\phi_{-t}(\phi_t(x))W(\phi_t(x)).$$

Apply this result to the infinitesimal conjugacy equation (2.3) to obtain the identity

$$\frac{d}{dt}(e^{-tA}\eta(e^{tA}x)) = e^{-tA}\xi(h(e^{tA}x)).$$

Using the definitions of h and Y , it follows immediately that

$$\frac{d}{dt}(e^{-tA}h(e^{tA}x)) = -e^{-tA}Ah(e^{tA}x) + e^{-tA}Y(h(e^{tA}x))$$

and (by the product rule)

$$e^{-tA}\frac{d}{dt}h(e^{tA}x) = e^{-tA}Y(h(e^{tA}x)).$$

Therefore, the function given by $t \mapsto h(e^{tA}x)$ is the integral curve of Y starting at the point $h(x)$. But, by the definition of the flow Y_t of Y , this integral curve is the function $t \mapsto Y_t(h(x))$. By uniqueness, $h(e^{tA}x) = Y_t(h(x))$. Because Y is linear on the complement of a compact set, Gronwall's inequality can be used to show that the flow of Y is complete. Hence, the conjugacy equation holds for all $t \in \mathbf{R}$.

It remains to show that the continuous function $h : \mathbf{R}^n \rightarrow \mathbf{R}^n$ given by $h(x) = x + \eta(x)$ is a homeomorphism. Since η is bounded on \mathbf{R}^n , the map $h = \text{id} + \eta$ is surjective. To see this, choose $y \in \mathbf{R}^n$, note that the equation $h(x) = y$ has a solution of the form $x = y + z$ if $z = -\eta(y + z)$, and apply Brouwer's fixed point theorem to the map $z \mapsto -\eta(y + z)$ on the ball of radius $\|\eta\|$ centered at the origin. (Using this idea, it is also easy to prove that h is proper; that is, the inverse image under h of every compact subset of \mathbf{R}^n is compact.) We will show that h is injective. If x and y are in \mathbf{R}^n and $h(x) = h(y)$, then $Y_t(h(x)) = Y_t(h(y))$ and, by the conjugacy relation, $A_t x + \eta(A_t x) = A_t y + \eta(A_t y)$. By the linearity of A_t , we have that

$$|A_t(x - y)| = |\eta(A_t y) - \eta(A_t x)|. \quad (2.8)$$

For each nonzero u in \mathbf{R}^n , the function $t \mapsto A_t u = e^{tA}u$ is unbounded on \mathbf{R} . Hence, either $x = y$ or the left side of equation (2.8) is unbounded for $t \in \mathbf{R}$. Since η is bounded, $x = y$; and therefore, the map h is injective. By Brouwer's theorem on invariance of domain, the bijective continuous map h is a homeomorphism. (Brouwer's theorem can be avoided by using instead the following elementary fact: A continuous, proper, bijective map from \mathbf{R}^n to \mathbf{R}^n is a homeomorphism.) \square

3 Smooth Conjugacy

In the classic paper [H60M], Hartman shows that if $a > b > 0$ and $c \neq 0$, then there is no \mathcal{C}^1 linearizing conjugacy at the origin for the analytic differential equation

$$\dot{x} = ax, \quad \dot{y} = (a - b)y + cxz, \quad \dot{z} = -bz. \quad (3.1)$$

On the other hand, he proved the following two important results. (1) If a \mathcal{C}^2 vector field has a rest point such that either all eigenvalues of its linearization have negative real parts or all eigenvalues have positive real parts, then the vector field is locally \mathcal{C}^1 conjugate to its linearization. (2) If a \mathcal{C}^2 planar vector field has a hyperbolic rest point, then the vector field is locally \mathcal{C}^1 conjugate to its linearization. Hartman proves the analogs of these theorems for maps and then derives the corresponding theorems for vector fields as corollaries. We will work directly with vector fields and thereby use standard methods from the theory of ordinary differential equations to obtain these results. We also note that S. Sternberg proved that the analytic planar system

$$\dot{x} = -x, \quad \dot{y} = -2y + x^2 \quad (3.2)$$

is not \mathcal{C}^2 linearizable. Hence, it should be clear that the proofs of Hartman's results on the existence of (maximally) smooth linearizations will require some delicate estimates. Nevertheless, as we will soon see, the strategy used in these proofs is easy to understand.

Although the starting point for the proof of Theorem 2.1, namely, the differentiation with respect to t of the desired conjugacy relation (2.1) and the inversion of the operator L_A as in display (2.6), leads to the simple proof of the existence of a conjugating homeomorphism given in Section 2, it turns out this strategy does not produce smooth conjugacies. This fact is illustrated by linearizing the scalar vector field given by $X(x) = -ax + f(x)$ where $a > 0$. Suppose that f vanishes outside a sufficiently small open subset of the origin with radius $r > 0$ so that $h(x) = x + \eta(x)$ is the continuous linearizing transformation where

$$\eta(x) = \int_0^\infty e^{-at} f(e^{at}x + \eta(e^{at}x)) dt$$

as in the proof of Theorem 2.1. With $F := f \circ (\text{id} + \eta)$, $u := e^{at}$, and $x \neq 0$, the function η is given by

$$\eta(x) = \frac{1}{a} \int_1^{e^{r/|x|}} \frac{F(ux)}{u^2} du.$$

Moreover, if $x > 0$, then (with $w = ux$)

$$\eta(x) = \frac{x}{a} \int_x^r \frac{F(w)}{w^2} dw,$$

and if $x < 0$, then

$$\eta(x) = -\frac{x}{a} \int_{-r}^x \frac{F(w)}{w^2} dw.$$

If η were continuously differentiable in a neighborhood of the origin, then we would have the identity

$$\eta'(x) = \frac{1}{a} \int_x^r \frac{F(w)}{w^2} dw - \frac{F(x)}{ax}$$

for $x > 0$ and the identity

$$\eta'(x) = -\frac{1}{a} \int_{-r}^x \frac{F(w)}{w^2} dw - \frac{F(x)}{ax}$$

for $x < 0$. Because the left-hand and right-hand derivatives agree at $x = 0$, it would follow that

$$\int_{-r}^r \frac{F(w)}{w^2} dw = 0.$$

But this equality is not true in general. For example, it is not true if $f(x) = \rho(x)x^2$ where ρ is a bump function as in the proof of Theorem 2.1. In this case, the integrand is nonnegative and not identically zero.

There are at least two ways to avoid the difficulty just described. First, note that the operator L_A , for the case $Ax = -ax$, is formally inverted by running time forward instead of backward. This leads to the formal inverse given by

$$(G\eta)(x) := - \int_0^\infty e^{at} \eta(e^{-at}x) dt$$

and the fixed point equation

$$\eta(x) = - \int_0^\infty e^{at} f(e^{-at}x + \eta(e^{-at}x)) dt.$$

In this case, no inconsistency arises from the assumption that $\eta'(0)$ exists. In fact, in the last chapter of this paper, we will show that this method does

produce a smooth conjugacy for certain “special vector fields”, for example, the scalar vector fields under consideration here (see Theorem 3.8).

Another idea that can be used to avoid the difficulty with smoothness is to differentiate both sides of the conjugacy relation

$$e^{tA}h(x) = h(X_t(x)) \tag{3.3}$$

with respect to t , or equivalently for the scalar differential equation, to use the change of coordinates $u = x + \eta(x)$. With this starting point, it is easy to see that η determines a linearizing transformation if it is a solution of the first order partial differential equation

$$D\eta(x)X(x) + a\eta(x) = -f(x).$$

To solve it, replace x by the integral curve $t \mapsto \phi_t(x)$ where ϕ_t denotes the flow of X , and note that (along this characteristic curve)

$$\frac{d}{dt}\eta(\phi_t(x)) + a\eta(\phi_t(x)) = -f(\phi_t(x)).$$

By variation of constants, we have the identity

$$\frac{d}{dt}e^{at}\eta(\phi_t(x)) = -e^{at}f(\phi_t(x)),$$

and (after integration on the interval $[0, t]$) it follows that the function η given by

$$\eta(x) = \int_0^\infty e^{at}f(\phi_t(x)) dt \tag{3.4}$$

determines a linearizing transformation $h = \text{id} + \eta$ if the improper integral converges on some open interval containing the origin. The required convergence is not obvious in general because the integrand of this integral contains the exponential growth factor e^{at} . In fact, to prove that η is continuous, a uniform estimate is required for the growth rate of the family of functions $t \mapsto |f(\phi_t(x))|$, and to show that η is continuously differentiable, a uniform growth rate estimate is required for their derivatives. The required estimates will be obtained in the next section where we will show that η is smooth for a hyperbolic sinks. For the scalar case as in equation (3.4), $f(x)$ is less than a constant times x^2 near the origin, and the solution $\phi_t(x)$ is approaching the origin like $e^{-at}x$. Because this quantity is squared by the function f , the integral converges.

To test the validity of this method, consider the example $\dot{x} = -ax + x^2$ where the flow can be computed explicitly and the integral (3.4) can be evaluated to obtain the smooth near-identity linearizing transformation $h : (-a, a) \rightarrow \mathbf{R}$ given by

$$h(x) = x + \frac{x^2}{a - x}.$$

3.1 Hyperbolic Sinks

The main result of this section is the following theorem.

Theorem 3.1 (Hartman). *Let X be a \mathcal{C}^2 vector field on \mathbf{R}^n such that $X(0) = 0$. If every eigenvalue of $DX(0)$ has negative real part, then X is locally \mathcal{C}^1 conjugate to its linearization at the origin.*

The full strength of the natural hypothesis that X is \mathcal{C}^2 is not used in the proof; rather, we will use only the weaker hypothesis that X is \mathcal{C}^1 and certain of its partial derivatives are Hölder on some fixed neighborhood of the origin. A function h is Hölder on a subset U of its domain if there is some (Hölder exponent) μ with $0 < \mu \leq 1$ and some constant $M > 0$ such that

$$|h(x) - h(y)| \leq M|x - y|^\mu$$

whenever x and y are in U . In the special case where $\mu = 1$, the function h is also called Lipschitz. As a convenient notation, let $\mathcal{C}^{1,\mu}$ denote the class of \mathcal{C}^1 functions whose first partial derivatives are all Hölder with Hölder exponent μ .

Recall the definition of Hölder spectral exponents given in Section 1. We will prove the following generalization of Theorem 3.1.

Theorem 3.2 (Hartman). *Let X be a $\mathcal{C}^{1,1}$ vector field on \mathbf{R}^n such that $X(0) = 0$. If every eigenvalue of $DX(0)$ has negative real part and $\mu > 0$ is smaller than the Hölder spectral exponent, then there is a near-identity $\mathcal{C}^{1,\mu}$ -diffeomorphism defined on some neighborhood of the origin that conjugates X to its linearization at the origin.*

The strategy for the proof of Theorem 3.2 is simple; in fact, the proof is by a finite induction. By a linear change of coordinates, the linear part of the vector field at the origin is transformed to a real Jordan canonical form where the diagonal blocks are ordered according to the real parts of

the corresponding eigenvalues, and the vector field is decomposed into (vector) components corresponding to these blocks. A theorem from invariant manifold theory is used to “flatten” the invariant manifold corresponding to the block whose eigenvalues have the largest real part onto the corresponding linear subspace. This transforms the original vector field into a special form which is then “partially linearized” by a near-identity diffeomorphism; that is, the flattened—but still nonlinear—component of the vector field is linearized by the transformation. This reduces the dimension of the linearization problem by the dimension of the flattened manifold. The process is continued until the system is completely linearized. Finally, the inverse of the linear transformation to Jordan form is applied to return to the original coordinates so that the composition of all the coordinate transformations is a near-identity map.

We will show that the nonlinear part of each near-identity partially linearizing transformation is given explicitly by an integral transform

$$\int_0^\infty e^{-tB} g(\varphi_t(x)) dt$$

where g is given by the nonlinear terms of the component function of the vector field corresponding to the linear block B and φ_t is the nonlinear flow. The technical part of the proof is to demonstrate that these transformations maintain the required smoothness. This is done by repeated applications of Lemma 2.2 to prove that “differentiation under the integral sign” is permitted. Because maximal smoothness is obtained, it is perhaps not surprising that some of the estimates required to majorize the integrand of the integral transform are rather delicate. In fact, the main difficulty is to prove that the exponential rate of decay toward zero of the functions $t \mapsto g(\varphi_t(x))$ and $t \mapsto g_x(\varphi_t(x))$, defined on some open neighborhood of $x = 0$, is faster than the exponential rate at which the linear flow e^{tB} moves points away from the origin in reverse time.

As in Section 2, the original vector field X can be expressed in the “almost linear” form $X(x) = Ax + (X(x) - Ax)$. There is a linear change of coordinates in \mathbf{R}^n such that X , in the new coordinates, is the almost linear vector field $Y(x) = Bx + (Y(x) - Bx)$ where the matrix B is in real Jordan canonical form with diagonal blocks B_1 and B_2 , every eigenvalue of B_2 has the same negative real part $-b_2$, and every eigenvalue of B_1 has its real part

strictly smaller than $-b_2$. The corresponding ODE has the form

$$\begin{aligned}\dot{x}_1 &= B_1 x_1 + P_1(x_1, x_2), \\ \dot{x}_2 &= B_2 x_1 + P_2(x_1, x_2)\end{aligned}\tag{3.5}$$

where $x = (x_1, x_2)$ and $(P_1(x_1, x_2), P_2(x_1, x_2)) = (Y(x) - Bx)$. Let c be a real number such that $-b < -c < 0$, and note that if the augmented system

$$\begin{aligned}\dot{x}_1 &= B_1 x_1 + P_1(x_1, x_2), \\ \dot{x}_2 &= B_2 x_1 + P_2(x_1, x_2), \\ \dot{x}_3 &= -cx_3\end{aligned}$$

is linearized by a near-identity transformation of the form

$$\begin{aligned}u_1 &= x_1 + \alpha_1(x_1, x_2, x_3), \\ u_2 &= x_2 + \alpha_2(x_1, x_2, x_3), \\ u_3 &= x_3,\end{aligned}$$

then the ODE (3.5) is linearized by the transformation

$$u_1 = x_1 + \alpha_1(x_1, x_2, 0), \quad u_2 = x_2 + \alpha_2(x_1, x_2, 0).$$

More generally, let $\mathcal{C}^{1,L,\mu}$ denote the class of all systems of the form

$$\begin{aligned}\dot{x} &= Ax + f(x, y, z), \\ \dot{y} &= By + g(x, y, z), \\ \dot{z} &= Cz\end{aligned}\tag{3.6}$$

where $x \in \mathbf{R}^k$, $y \in \mathbf{R}^\ell$, and $z \in \mathbf{R}^m$; where A , B , and C are square matrices of the corresponding dimensions; B is in real Jordan canonical form;

- every eigenvalue of B has real part $-b < 0$;
- every eigenvalue of A has real part less than $-b$;
- every eigenvalue of C has real part in an interval $[-c, -d]$ where $-b < -c$ and $-d < 0$;

and $F : (x, y, z) \mapsto (f(x, y, z), g(x, y, z))$ is a \mathcal{C}^1 function defined in a bounded product neighborhood

$$\Omega = \Omega_{xy} \times \Omega_z\tag{3.7}$$

of the origin in $(\mathbf{R}^k \times \mathbf{R}^\ell) \times \mathbf{R}^m$ such that

- $F(0, 0, 0) = 0$ and $DF(0, 0, 0) = 0$,
- the partial derivatives F_x and F_y are Lipschitz in Ω , and
- the partial derivative F_z is Lipschitz in Ω_{xy} uniformly with respect to $z \in \Omega_z$ and Hölder in Ω_z uniformly with respect to $(x, y) \in \Omega_{xy}$ with Hölder exponent μ .

System (3.6) satisfies the $(1, \mu)$ *spectral gap condition* if $(1 + \mu)c < b$.

We will show that system (3.6) can be linearized by a \mathcal{C}^1 near-identity transformation of the form

$$\begin{aligned} u &= x + \alpha(x, y, z), \\ v &= y + \beta(x, y, z), \\ w &= z. \end{aligned} \tag{3.8}$$

The proof of this result is given in three main steps: an invariant manifold theorem for a system with a spectral gap is used to find a preliminary near-identity \mathcal{C}^1 map, as in display (3.8), that transforms system (3.6) into a system of the same form but with the new function $F = (f, g)$ “flattened” along the coordinate subspace corresponding to the invariant manifold. Next, for the main part of the proof, a second near-identity transformation of the same form is constructed that transforms the flattened system to the partially linearized form

$$\begin{aligned} \dot{x} &= Ax + p(x, y, z), \\ \dot{y} &= By, \\ \dot{z} &= Cz \end{aligned} \tag{3.9}$$

where A , B , and C are the matrices in system (3.6) and the function p has the following properties:

- p is \mathcal{C}^1 on an open neighborhood $\Omega = \Omega_x \times \Omega_{yz}$ of the origin in $\mathbf{R}^k \times (\mathbf{R}^\ell \times \mathbf{R}^m)$;
- $p(0, 0, 0) = 0$ and $Dp(0, 0, 0) = 0$;
- The partial derivative p_x is Lipschitz in Ω ;
- The partial derivatives p_y and p_z are Lipschitz in Ω_x uniformly with respect to $(y, z) \in \Omega_{yz}$ and Hölder in Ω_{yz} uniformly with respect to $x \in \Omega_x$.

The final step of the proof consists of three observations: The composition of \mathcal{C}^1 near-identity transformations of the form considered here is again a \mathcal{C}^1 near-identity transformation; the dimension of the “unlinearized” part of the system is made strictly smaller after applying the partially linearizing transformation, and the argument can be repeated as long as the system is not linearized. In other words, the proof is completed by a finite induction.

The required version of the invariant manifold theorem is a special case of a more general theorem (see, for example, Yu. Latushkin and B. Layton [LL99]). For completeness, we will formulate and prove this special case. Our proof can be modified to obtain the general result.

For notational convenience, let us view system (3.6) in the compact form

$$\begin{aligned}\dot{\mathcal{X}} &= \mathcal{A}\mathcal{X} + F(\mathcal{X}, z), \\ \dot{z} &= Cz\end{aligned}\tag{3.10}$$

where $\mathcal{X} = (x, y)$, $\mathcal{A} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, and $F := (f, g)$. Hyperbolic estimates for the corresponding linearized equations are used repeatedly. In particular, in view of the hypotheses about the eigenvalues of A , B , and C , it follows that if $\epsilon > 0$ and

$$0 < \lambda < d,$$

then there is a constant $K > 1$ such that

$$\|e^{t\mathcal{A}}\| \leq Ke^{-(b-\epsilon)t}, \quad \|e^{tC}\| \leq Ke^{-\lambda t} \quad \|e^{-tC}\| \leq Ke^{(c+\epsilon)t}\tag{3.11}$$

for all $t \geq 0$.

Theorem 3.3. *If the $(1, \mu)$ spectral gap condition holds for system (3.10), then there is an open set $\Omega_z \subset \mathbf{R}^m$ containing $z = 0$ and a $\mathcal{C}^{1,\mu}$ function $\gamma : \Omega_z \rightarrow \mathbf{R}^{k+\ell}$ such that $\gamma(0) = D\gamma(0) = 0$ whose graph (the set $\{(\mathcal{X}, z) \in \mathbf{R}^{k+\ell} \times \mathbf{R}^m : \mathcal{X} = \gamma(z)\}$) is forward invariant.*

As a remark, we mention that the smoothness of γ cannot in general be improved by simply requiring additional smoothness of the vector field. Rather, the smoothness of the invariant manifold can be improved only if additional requirements are made on the smoothness of the vector field *and* on the length of the spectral gap (see [LL99] and the references therein). For these reasons, it seems that the technical burden imposed by working with

Hölder functions cannot be avoided by simply requiring additional smoothness of the vector field unless additional hypotheses are made on the eigenvalues of the linearization at the origin as well. Also, we mention that our proof illustrates the full power of the fiber contraction principle introduced by M. Hirsch and C. Pugh in [HP70] as a method for proving the smoothness of functions obtained as fixed points of contractions.

To describe the fiber contraction method in our setting, let us consider a metric subspace \mathcal{D} of a Banach space of continuous functions defined on $\Omega \subset \mathbf{R}^m$ with values in \mathbf{R}^p , and let us suppose that $\Gamma : \mathcal{D} \rightarrow \mathcal{D}$ is a contraction (on the complete metric space \mathcal{D}) with fixed point γ . (In the analysis to follow, Γ is given by an integral transform operator.) We wish to show that γ is differentiable. Naturally, we start by formally differentiating both sides of the identity $\eta(z) = \Gamma(\eta)(z)$ with respect to z to obtain the identity $D\gamma(z) = \Delta(\gamma, D\gamma)(z)$ where the map $\Phi \mapsto \Delta(\gamma, \Phi)$ is a linear operator on a metric—not necessarily complete—subspace \mathcal{J} of continuous functions from Ω to the bounded linear maps from \mathbf{R}^m to \mathbf{R}^p . We expect the derivative $D\eta$, if it exists, to satisfy the equation

$$\Phi = \Delta(\eta, \Phi).$$

Hence, \mathcal{J} is a space of “candidates for the derivative of γ ”.

The next step is to show that the bundle map $\Lambda : \mathcal{D} \times \mathcal{J} \rightarrow \mathcal{D} \times \mathcal{J}$ defined by

$$\Lambda(\gamma, \Phi) = (\Gamma(\gamma), \Delta(\gamma, \Phi))$$

is a fiber contraction; that is, for each $\gamma \in \mathcal{D}$, the map $\Phi \rightarrow \Delta(\gamma, \Phi)$ is a contraction on \mathcal{J} with respect to a contraction constant that does not depend on the choice of $\gamma \in \mathcal{D}$. The fiber contraction theorem (see [HP70] or, for more details, [C99]) states that *if γ is the globally attracting fixed point of Γ and if Φ is a fixed point of the map $\Phi \rightarrow \Delta(\gamma, \Phi)$, then (γ, Φ) is the globally attracting fixed point of Λ* . The fiber contraction theorem does not require \mathcal{J} to be a complete metric space. This leaves open the possibility to prove the existence of a fixed point in the fiber over γ by using, for example, Schauder’s theorem. But, for our applications, the space \mathcal{J} will be chosen to be complete so that the existence of the fixed point Φ follows from an application of the contraction mapping theorem.

After we show that Λ is a fiber contraction, the following argument can often be used to prove the desired equality $\Phi = D\gamma$. Find a point $(\gamma_0, \Phi_0) \in$

$\mathcal{D} \times \mathcal{J}$ such that $D\gamma_0 = \Phi_0$, define a sequence $\{(\gamma_j, \Phi_j)\}_{j=0}^\infty$ in $\mathcal{D} \times \mathcal{J}$ by

$$\gamma_j = \Gamma(\gamma_{j-1}), \quad \Phi_j = \Delta(\gamma_{j-1}, \Phi_{j-1}),$$

and prove by induction that $D\gamma_j = \Phi_j$ for every positive integer j . By the fiber contraction theorem, the sequence $\{\gamma_j\}_{j=0}^\infty$ converges to γ and the sequence $\{D\gamma_j\}_{j=0}^\infty$ converges to the fixed point Φ of the map $\Phi \rightarrow \Delta(\gamma, \Phi)$. If the convergence is *uniform*, then by a standard theorem from advanced calculus—the *uniform limit of a sequence of differentiable functions is differentiable and equal to the limit of the derivatives of the functions in the sequence provided that the sequence of derivatives is uniformly convergent*—the function γ is differentiable and its derivative is Φ .

Let us prove Theorem 3.3.

Proof. The graph of γ is forward invariant if and only if $\dot{\mathcal{X}} = D\gamma(z)\dot{z}$ whenever $\mathcal{X} = \gamma(z)$. Equivalently, the identity

$$D\gamma(z)Cz - \mathcal{A}\gamma(z) = F(\gamma(z), z) \tag{3.12}$$

holds for all z in the domain of γ .

The function γ will satisfy identity (3.12) if γ is \mathcal{C}^1 and

$$\frac{d}{d\tau} e^{-\tau\mathcal{A}}\gamma(e^{\tau C}z) = e^{-\tau\mathcal{A}}F(\gamma(e^{\tau C}z), e^{\tau C}z).$$

In this case, by integration on the interval $[-\tau, 0]$ followed by a change of variables in the integral on the right-hand side of the resulting equation, it follows that

$$\gamma(z) - e^{\tau\mathcal{A}}\gamma(e^{-\tau C}z) = \int_0^\tau e^{t\mathcal{A}}F(\gamma(e^{-tC}z), e^{-tC}z) dt.$$

If γ is a \mathcal{C}^1 function such that $\gamma(0) = D\gamma(0) = 0$ and $\lim_{\tau \rightarrow \infty} |e^{\tau\mathcal{A}}\gamma(e^{-\tau C}z)| = 0$, then the graph of γ will be (forward) invariant provided that

$$\gamma(z) = \Gamma(\gamma)(z) := \int_0^\infty e^{t\mathcal{A}}F(\gamma(e^{-tC}z), e^{-tC}z) dt. \tag{3.13}$$

For technical reasons, it is convenient to assume that γ is defined on all of \mathbf{R}^m and that F is “cut off” as in the proof of Theorem 2.1 to have support in an open ball at the origin in $\mathbf{R}^{k+\ell+m}$ of radius $r > 0$ so that the new function,

still denoted by the symbol F , is defined globally with $\|DF\|$ bounded by a small number $\rho > 0$ to be determined. Recall that both r and ρ can be chosen to be as small as we wish. This procedure maintains the smoothness of the original function and the modified function agrees with the original function on some open ball centered at the origin and with radius $r_0 < r$. Also, the graph of the restriction of a function γ that satisfies the equation (3.13) to an open subset of Ω_z , containing the origin and inside the ball of radius r_0 , is forward invariant for system (3.10) because the the modified differential equation agrees with the original differential equation in the ball of radius r_0 .

We will show that there is a solution γ of equation (3.13) such that $\gamma \in \mathcal{C}^{1,\mu}$ and $\gamma(0) = D\gamma(0) = 0$.

Let \mathcal{B} denote the Banach space of continuous functions from \mathbf{R}^m to $\mathbf{R}^{k+\ell}$ that are bounded with respect to the norm given by

$$\|\gamma\|_{\mathcal{B}} := \sup \left\{ \frac{|\gamma(z)|}{|z|} : z \in \mathbf{R}^m \setminus \{0\} \right\}.$$

Also, note that convergence of a sequence in the \mathcal{B} -norm implies uniform convergence of the sequence on compact subsets of \mathbf{R}^m .

Let $\mathcal{D} := \{\gamma \in \mathcal{B} : |\gamma(z_1) - \gamma(z_2)| \leq |z_1 - z_2|\}$. Note that if $\gamma \in \mathcal{D}$, then $\|\gamma\|_{\mathcal{B}} \leq 1$. We will also show that \mathcal{D} is a closed subset, and hence a complete metric subspace, of \mathcal{B} . In fact, if $\{\gamma^k\}_{k=1}^{\infty}$ is a sequence in \mathcal{D} that converges to γ in \mathcal{B} , then

$$\begin{aligned} |\gamma(z_1) - \gamma(z_2)| &\leq |\gamma(z_1) - \gamma^k(z_1)| + |\gamma^k(z_1) - \gamma^k(z_2)| + |\gamma^k(z_2) - \gamma(z_2)| \\ &\leq \|\gamma - \gamma^k\|_{\mathcal{B}}|z_1|^{\mu} + |z_1 - z_2| + \|\gamma - \gamma^k\|_{\mathcal{B}}|z_2|^{\mu}. \end{aligned} \quad (3.14)$$

To see that $\gamma \in \mathcal{D}$, pass to the limit as $k \rightarrow \infty$.

We will show that if ρ is sufficiently small, then Γ , the operator defined in display (3.13), is a contraction on \mathcal{D} .

Recall the hyperbolic estimates 3.11, choose ϵ so small that $c - b + 2\epsilon < 0$,

let $\gamma \in \mathcal{D}$, and note that

$$|\Gamma(\gamma)(z_1) - \Gamma(\gamma)(z_2)| \leq \int_0^\infty K e^{-(b-\epsilon)t} \|DF\| (|\gamma(e^{-tC} z_1) - \gamma(e^{-tC} z_2)| + |e^{-tC}(z_1 - z_2)|) dt \quad (3.15)$$

$$\begin{aligned} &\leq \int_0^\infty K e^{-(b-\epsilon)t} \rho (2|e^{-tC}(z_1 - z_2)|) dt \\ &\leq (2K\rho \int_0^\infty e^{(c-b+2\epsilon)t} dt) |z_1 - z_2|. \end{aligned} \quad (3.16)$$

By taking ρ sufficiently small, it follows that

$$|\Gamma(\gamma)(z_1) - \Gamma(\gamma)(z_2)| \leq |z_1 - z_2|.$$

Also, using similar estimates, it is easy to show that

$$|\Gamma(\gamma_1)(z) - \Gamma(\gamma_2)(z)| \leq \frac{K^2 \rho}{b - c - 2\epsilon} \|\gamma_1 - \gamma_2\|_{\mathcal{B}} |z|.$$

Hence, if $\rho = \|DF\|$ is sufficiently small, then Γ is a contraction on the complete metric space \mathcal{D} ; and therefore, it has a unique fixed point $\gamma_\infty \in \mathcal{D}$.

We will use the fiber contraction principle to show that $\gamma_\infty \in \mathcal{C}^{1,\mu}$.

Before modification by the bump function, there is some open neighborhood of the origin on which $F_{\mathcal{X}}$ is Lipschitz, F_z is Lipschitz in \mathcal{X} , and F_z is Hölder in z . Using this fact and the construction of the bump function, it is not difficult to show that there is a constant $\bar{M} > 0$ such that (for the modified function F)

$$|F_z(\mathcal{X}_1, z_1) - F_z(\mathcal{X}_2, z_2)| \leq \bar{M} (|\mathcal{X}_1 - \mathcal{X}_2|^\mu + |z_1 - z_2|^\mu) \quad (3.17)$$

and

$$|F_{\mathcal{X}}(\mathcal{X}_1, z_1) - F_{\mathcal{X}}(\mathcal{X}_2, z_2)| \leq \bar{M} (|\mathcal{X}_1 - \mathcal{X}_2|^\mu + |z_1 - z_2|^\mu). \quad (3.18)$$

Moreover, \bar{M} is independent of r as long as $r > 0$ is smaller than some preassigned positive number.

In view of the $(1, \mu)$ spectral gap condition, $-b + \epsilon + (1 + \mu)(c + \epsilon) < 0$ whenever $\epsilon > 0$ is sufficiently small. For ϵ in this class, let

$$\bar{K} := K^3 \int_0^\infty e^{(-b+\epsilon+(1+\mu)(c+\epsilon))t} dt.$$

Define the Banach space \mathcal{H} of continuous functions from \mathbf{R}^m to the bounded linear maps from \mathbf{R}^m to $\mathbf{R}^{k+\ell}$ that are bounded with respect to the norm

$$\|\Phi\|_{\mathcal{H}} := \sup_{z \in \mathbf{R}^m} \frac{|\Phi(z)|}{|z|^\mu},$$

note that each element Φ in this space is such that $\Phi(0) = 0$, and note that convergence of a sequence in the \mathcal{H} -norm implies uniform convergence of the sequence on compact subsets of \mathbf{R}^m .

For $r > 0$ the radius of the ball containing the support of F and for $\rho < 1/\bar{K}$, let \mathcal{J} denote the metric subspace consisting of those functions in \mathcal{H} that satisfy the following additional properties:

$$\|\Phi\|_{\mathcal{H}} \leq \frac{2\bar{K}\bar{M}}{1 - \bar{K}\rho}, \quad \sup\{|\Phi(z)| : z \in \mathbf{R}^m\} \leq \frac{2\bar{K}\bar{M}Kr}{1 - \bar{K}\rho}, \quad (3.19)$$

and

$$|\Phi(z_1) - \Phi(z_2)| \leq H|z_1 - z_2|^\mu \quad (3.20)$$

where $H := (1 - \bar{K}\rho + 2\bar{K}\bar{M}Kr)(2\bar{K}\bar{M})/(1 - \bar{K}\rho)^2$. Using estimates similar to estimate (3.14), it is easy to prove that \mathcal{J} is a complete metric space.

Moreover, for $\gamma \in \mathcal{D}$ and $\Phi \in \mathcal{J}$, define the bundle map

$$\Lambda(\gamma, \Phi) = (\Gamma(\gamma), \Delta(\gamma, \Phi))$$

where

$$\begin{aligned} \Delta(\gamma, \Phi)(z) := & \int_0^\infty e^{tA}(F_x(\gamma(e^{-tC}z), e^{-tC}z)\Phi(e^{-tC}z) \\ & + F_z(\gamma(e^{-tC}z), e^{-tC}z))e^{-tC} dt \end{aligned}$$

and note that the derivative of γ_∞ is a formal solution of the equation $\Phi = \Delta(\gamma_\infty, \Phi)$.

We will prove that Λ is a fiber contraction on $\mathcal{D} \times \mathcal{J}$ in two main steps: (1) $\Delta(\gamma, \Phi) \in \mathcal{J}$ whenever $\gamma \in \mathcal{D}$ and $\Phi \in \mathcal{J}$; (2) $\Phi \mapsto \Delta(\gamma, \Phi)$ is a contraction whose contraction constant is uniform for $\gamma \in \mathcal{D}$.

Because $|e^{tC}z| \leq Ke^{-\lambda t}|z|$ for $t \geq 0$, By substituting $e^{-tC}z$ for z in the hyperbolic estimate $|e^{tC}z| \leq Ke^{-\lambda t}|z|$, we have the inequality $|e^{tC}z| \geq \frac{1}{K}|z|$ for all $t \geq 0$. Hence, if $|z| \geq Kr$, then $|e^{tC}z| \geq r$.

Note that if $(\gamma, \Phi) \in \mathcal{D} \times \mathcal{J}$ and $|z| \geq Kr$, then $\Delta(\gamma, \Phi)(z) = 0$. On the other hand, for $|z| < Kr$, the Hölder estimate (3.17) can be used to obtain the inequalities

$$\begin{aligned}
|\Delta(\gamma, \Phi)(z)| &\leq \int_0^\infty |e^{tA}| |e^{-tC}| |(F_{\mathcal{X}}(\gamma(e^{-tC}z), e^{-tC}z)|\Phi(e^{-tC}z)| \\
&\quad + |F_z(\gamma(e^{-tC}z), e^{-tC}z)|) dt \\
&\leq \int_0^\infty K^2 e^{(c-b+2\epsilon)t} (\|DF\| \|\Phi\|_{\mathcal{H}} K^\mu e^{\mu(c+\epsilon)t} |z|^\mu \\
&\quad + \bar{M}(|\gamma(e^{-tC}z|^\mu + |e^{-tC}z|^\mu)) dt \\
&\leq (K^{2+\mu} \int_0^\infty e^{(-b+c+(1+\mu)(c+\epsilon))t} dt) (\|DF\| \|\Phi\|_{\mathcal{H}} + 2\bar{M}) |z|^\mu dt \\
&\leq \bar{K}(\rho \|\Phi\|_{\mathcal{H}} + 2\bar{M}) |z|^\mu \\
&\leq \frac{2\bar{K}\bar{M}}{1 - \bar{K}\rho} |z|^\mu.
\end{aligned}$$

It follows that $\Delta(\gamma, \Phi)$ satisfies both inequalities in display (3.19).

To show that $\Delta(\gamma, \Phi)$ satisfies the Hölder condition (3.20), estimate

$$Q := |\Delta(\gamma, \Phi)(z_1) - \Delta(\gamma, \Phi)(z_2)|$$

in the obvious manner, add and subtract $F_{\mathcal{X}}(\gamma(e^{-tC}z_1), e^{-tC}z_1)\Phi(e^{-tC}z_2)$, and then use the triangle inequality to obtain the inequality

$$\begin{aligned}
Q &\leq \int_0^\infty e^{(c-b+2\epsilon)t} (|F_{\mathcal{X}}(\gamma(e^{-tC}z_1), e^{-tC}z_1)| |\Phi(e^{-tC}z_1) - \Phi(e^{-tC}z_2)| \\
&\quad + |F_{\mathcal{X}}(\gamma(e^{-tC}z_1), e^{-tC}z_1) - F_{\mathcal{X}}(\gamma(e^{-tC}z_2), e^{-tC}z_2)| |\Phi(e^{-tC}z_2)| \\
&\quad + |F_z(\gamma(e^{-tC}z_1), e^{-tC}z_1) - F_z(\gamma(e^{-tC}z_2), e^{-tC}z_2)|) dt.
\end{aligned}$$

The first factor involving $F_{\mathcal{X}}$ is bounded above by $\rho = \|DF\|$; the second factor involving $F_{\mathcal{X}}$ is bounded above using the Hölder estimate (3.18) followed the Lipschitz estimate for γ in the definition of \mathcal{D} . Likewise, the second factor involving Φ is bounded using the supremum in display (3.19); the first factor involving Φ is bounded using the Hölder inequality (3.20). The term involving F_z is bounded above using the Hölder estimate (3.17). After some manipulation using the hyperbolic estimate for e^{-tC} and in view of the definition of \bar{K} , it follows that Q is bounded above by

$$\bar{K}(\rho H + 2\bar{M} \frac{2\bar{K}\bar{M}Kr}{1 - \bar{K}\rho} + 2\bar{M}) |z_1 - z_2|^\mu \leq H |z_1 - z_2|^\mu.$$

This completes the proof that $\Delta(\gamma, \Phi) \in \mathcal{J}$.

Finally, by similar estimation procedures, it is easy to show that

$$\begin{aligned} \|\Delta(\gamma_1, \Phi_1)(z) - \Delta(\gamma_2, \Phi_2)(z)\| \leq \\ \bar{K}(\rho\|\Phi_1 - \Phi_2\|_{\mathcal{H}} + \bar{M}(1 + \frac{2\bar{K}\bar{M}\bar{K}r}{1 - \bar{K}\rho}\|\gamma_1 - \gamma_2\|_{\mathcal{B}})|z|^\mu). \end{aligned}$$

Therefore, Δ and hence Λ is continuous. Also, because $\bar{K}\rho < 1$, the map $\Phi \mapsto \Delta(\gamma, \Phi)$ is a contraction, and the contraction constant is uniform over \mathcal{D} . This proves that Λ is a fiber contraction on $\mathcal{D} \times \mathcal{J}$.

Choose $(\gamma_0, \Phi_0) = (0, 0)$ and define a sequence in $\mathcal{D} \times \mathcal{J}$ inductively by

$$(\gamma_{j+1}, \Phi_{j+1}) := \Lambda(\gamma_j, \Phi_j).$$

In particular, by the contraction mapping theorem the sequence $\{\gamma_j\}_{j=0}^\infty$ converges to γ_∞ . Clearly, we have $D\gamma_0 = \Phi_0$. Proceeding by induction, let us assume that $\mathcal{D}\gamma_j = \Phi_j$. Since

$$\gamma_{j+1}(z) = \Gamma(\gamma_j)(z) = \int_0^\infty e^{tA} F(\gamma_j(e^{-tC}z), e^{-tC}z) dt$$

and since differentiation under the integral sign is permitted by Lemma 2.2 because we have the majorization (3.15), it follows that

$$D\gamma_{j+1}(z) = \int_0^\infty e^{tA} F_{\mathcal{X}}(\gamma_j(e^{-tC}z), e^{-tC}z) D\gamma_j(e^{-tC}z) e^{-tC} dt.$$

By the induction hypothesis $D\gamma_j(e^{-tC}z) = \Phi_j(e^{-tC}z)r$. Hence, by the definition of Δ , we have that $D\gamma_{j+1} = \Phi_{j+1}$. Finally, because convergence in the spaces \mathcal{D} and \mathcal{J} implies uniform convergence on compact sets, by using the fiber contraction theorem and the theorem from advanced calculus on uniform limits of differentiable functions, it follows that γ_∞ is \mathcal{C}^1 with its derivative in \mathcal{J} . Thus, in fact, $\gamma_\infty \in \mathcal{C}^{1,\mu}$. \square

We will now apply Theorem 3.3. After this is done, the remainder of this section will be devoted to the proof of the existence and smoothness of the partially linearizing transformation for the flattened system.

The mapping given by

$$U = \mathcal{X} - \gamma(z),$$

where the graph of γ is the invariant manifold in Theorem 3.3, transforms system (3.10) into the form

$$\begin{aligned}\dot{U} &= \mathcal{A}U + A\gamma(z) + F(U + \gamma(z), z) - D\gamma(z)Cz, \\ \dot{z} &= Cz.\end{aligned}\tag{3.21}$$

Because the graph of γ is invariant, $\dot{U} = 0$ whenever $U = 0$. In particular,

$$A\gamma(z) + F(\gamma(z), z) - D\gamma(z)Cz \equiv 0;$$

and therefore, the system (3.21) has the form

$$\begin{aligned}\dot{U} &= \mathcal{A}U + \mathcal{F}(U, z), \\ \dot{z} &= Cz\end{aligned}\tag{3.22}$$

where

$$\mathcal{F}(U, z) := F(U + \gamma(z), z) - F(\gamma(z), z).\tag{3.23}$$

Proposition 3.4. *The function \mathcal{F} in system (3.22) is $\mathcal{C}^{1,L,\mu}$ on a bounded neighborhood of the origin. In addition, if (U, z_i) and (U_i, z_i) for $i \in \{1, 2\}$ are in this neighborhood, then there are constants $M > 0$ and $0 \leq \vartheta < 1$ such that*

$$|\mathcal{F}(U_1, z_1) - \mathcal{F}(U_2, z_2)| \leq M(|U_1| + |z_1| + |U_2| + |z_2|)(|U_1 - U_2| + |z_1 - z_2|),\tag{3.24}$$

$$|\mathcal{F}_z(U_1, z_1) - \mathcal{F}_z(U_2, z_2)| \leq M(|z_1 - z_2|^\mu + |U_1 - U_2|^{1-\vartheta})(|U_1| + |U_2|)^\vartheta.\tag{3.25}$$

Proof. The function \mathcal{F} is \mathcal{C}^1 because it is the composition of \mathcal{C}^1 functions. Moreover, by definition (3.23) and because $F(0, 0) = DF(0, 0) = 0$, it is clear that $\mathcal{F}(0, 0) = D\mathcal{F}(0, 0) = 0$.

To show that the partial derivative \mathcal{F}_U is Lipschitz in a neighborhood of the origin, start with the equality $\mathcal{F}_U(U, z) = F_X(U + \gamma(z), z)$, note that F_X is Lipschitz, and conclude that there is a constant $K > 0$ such that

$$|\mathcal{F}_U(U_1, z_1) - \mathcal{F}_U(U_2, z_2)| \leq K(|U_1 - U_2| + |\gamma(z_1) - \gamma(z_2)| + |z_1 - z_2|).$$

By an application of the mean value theorem to the \mathcal{C}^1 function γ , it follows that

$$|\mathcal{F}_U(U_1, z_1) - \mathcal{F}_U(U_2, z_2)| \leq K(1 + \|D\gamma\|)(|U_1 - U_2| + |z_1 - z_2|).$$

Since $\|D\gamma\|$ is bounded in some neighborhood of the origin, we have the desired result.

Similarly, in view of the equality

$$\begin{aligned} \mathcal{F}_z(U, z) &= F_{\mathcal{X}}(U + \gamma(z), z)D\gamma(z) - F_{\mathcal{X}}(\gamma(z), z)D\gamma(z) \\ &\quad + F_z(U + \gamma(z), z) - F_z(\gamma(z), z), \end{aligned}$$

the properties of F , and the triangle law, there is a constant $K > 0$ such that

$$|\mathcal{F}_z(U_1, z) - \mathcal{F}_z(U_2, z)| \leq K|U_1 - U_2|$$

uniformly for z in a sufficiently small open neighborhood of $z = 0$.

Several easy estimates are required to show that \mathcal{F}_z is Hölder with respect to its second argument. For this, let $T := |\mathcal{F}_z(U, z_1) - \mathcal{F}_z(U, z_2)|$ and note that

$$\begin{aligned} T &\leq |F_{\mathcal{X}}(U + \gamma(z_1), z_1)D\gamma(z_1) - F_{\mathcal{X}}(U + \gamma(z_2), z_2)D\gamma(z_2)| \\ &\quad + |F_{\mathcal{X}}(\gamma(z_1), z_1)D\gamma(z_1) - F_{\mathcal{X}}(\gamma(z_2), z_2)D\gamma(z_2)| \\ &\quad + |F_z(U + \gamma(z_1), z_1) - F_z(U + \gamma(z_2), z_2)| \\ &\quad + |F_z(\gamma(z_1), z_1) - F_z(\gamma(z_2), z_2)|. \end{aligned}$$

Each term on the right-hand side of this inequality is estimated in turn. The desired result is obtained by combining these results. We will show how to obtain the Hölder estimate for the first term; the other estimates are similar.

To estimate the first term T_1 , add and subtract the quantity $F_{\mathcal{X}}(U + \gamma(z_1), z_1)D\gamma(z_2)$ and use the triangle inequality to obtain the upper bound

$$\begin{aligned} T_1 &\leq \|DF\| |D\gamma(z_1) - D\gamma(z_2)| \\ &\quad + \|D\gamma\| |F_{\mathcal{X}}(U + \gamma(z_1), z_1) - F_{\mathcal{X}}(U + \gamma(z_2), z_2)|. \end{aligned}$$

Because $D\gamma$ is Hölder and $F_{\mathcal{X}}$ is Lipschitz, there is a constant $K > 0$ such that

$$\begin{aligned} T_1 &\leq \|DF\|K|z_1 - z_2|^\mu + \|D\gamma\|K(|z_1 - z_2|^\mu + |z_1 - z_2|) \\ &\leq \|DF\|K|z_1 - z_2|^\mu + \|D\gamma\|K|z_1 - z_2|^\mu(1 + |z_1 - z_2|^{1-\mu}). \end{aligned}$$

Finally, by restricting to a sufficiently small neighborhood of the origin, it follows that there is a constant $M > 0$ such that

$$T_1 \leq M|z_1 - z_2|^\mu,$$

as required.

To prove the estimate (3.24), note that $\mathcal{F}(0, 0) = D\mathcal{F}(0, 0) = 0$ and, by Taylor's formula,

$$\mathcal{F}(U, z) = \int_0^1 (\mathcal{F}_U(tU, tz)U + \mathcal{F}_z(tU, tz)z) dt.$$

The desired estimate for $|\mathcal{F}(U_1, z_1) - \mathcal{F}(U_2, z_2)|$ is obtained by subtracting the integral expressions for $\mathcal{F}(U_1, z_1)$ and $\mathcal{F}(U_2, z_2)$, adding and subtracting $\mathcal{F}_U(U_1, z_1)U_2$ and $\mathcal{F}_z(U_1, z_1)z_2$, and then by using the triangle inequality, the Lipschitz estimates for F_U and F_z , and the observation that $|t| \leq 1$ in the integrand.

For the estimate (3.25), note that the function $U \mapsto \mathcal{F}_z(U, z)$ is (uniformly) Lipschitz and use the obvious triangle estimate to conclude that there is a constant $M > 0$ such that

$$|\mathcal{F}_z(U_1, z_1) - \mathcal{F}_z(U_2, z_2)| \leq M(|U_1| + |U_2|).$$

Also, a different upper bound for the same quantity is obtained as follows:

$$\begin{aligned} |\mathcal{F}_z(U_1, z_1) - \mathcal{F}_z(U_2, z_2)| &\leq |\mathcal{F}_z(U_1, z_1) - \mathcal{F}_z(U_1, z_2)| \\ &\quad + |\mathcal{F}_z(U_1, z_2) - \mathcal{F}_z(U_2, z_2)| \\ &\leq M(|z_1 - z_2|^\mu + |U_1 - U_2|) \end{aligned}$$

The desired inequality (3.25) is obtained from these two upper bounds and the following proposition: *Suppose that $a \geq 0$, $b > 0$, and $c > 0$. If $a \leq \max\{b, c\}$ and $0 \leq \vartheta < 1$, then $a \leq b^\vartheta c^{1-\vartheta}$.* The proposition is clearly valid in case $a = 0$. On the other hand, the case where $a \neq 0$ is an immediate consequence of the inequality

$$\ln a = \vartheta \ln a + (1 - \vartheta) \ln a \leq \vartheta \ln b + (1 - \vartheta) \ln c \leq \ln(b^\vartheta c^{1-\vartheta}).$$

□

As mentioned above, the main result of this section concerns the partial linearization of system (3.22). More precisely, let us fix the previously defined square matrices A , B , and C , and consider the system

$$\begin{aligned} \dot{x} &= Ax + f(x, y, z), \\ \dot{y} &= By + g(x, y, z), \\ \dot{z} &= Cz \end{aligned} \tag{3.26}$$

where $\mathcal{F} := (f, g)$ satisfies all of the properties mentioned in Proposition 3.4. Also, we will use the following hypothesis.

Hypothesis 3.5. *Let Ω be an open neighborhood of the origin given by $\Omega_{xy} \times \Omega_z$ as in display (3.7),*

$$U, U_1, U_2 \in \Omega_{xy}, \quad z, z_1, z_2 \in \Omega_z, \quad r := \sup_{(U,z) \in \Omega} |U|,$$

the numbers $K > 1$ and $\lambda > 0$ are the constants in display (3.11), the numbers $M > 0$ and ϑ are the constants in Proposition 3.4, and $-b$ is the real part of the eigenvalues of the matrix B in system (3.26). In addition, Ω is a sufficiently small open set, ϵ is a sufficiently small positive real number, and ϑ is a sufficiently large number in the open unit interval such that for $\delta := 2K^2Mr + \epsilon$ we have $\epsilon < 1$, $-b + \delta < -\lambda$, $-\lambda + 2\delta < 0$, $(1 - \vartheta)b - \lambda < 0$, and $\delta - \lambda + (1 - \vartheta)b < 0$.

Theorem 3.6. *There is a $\mathcal{C}^{1,L,\mu(1-\theta)}$ near-identity diffeomorphism defined on an open subset of the origin that transforms system (3.26) into system (3.9).*

Proof. We will show that there is a near-identity transformation of the form

$$\begin{aligned} u &= x, \\ v &= y + \alpha(x, y, z), \\ w &= z \end{aligned} \tag{3.27}$$

that transforms system (3.26) into

$$\begin{aligned} \dot{u} &= Au + p(u, v, w), \\ \dot{v} &= Bv, \\ \dot{w} &= Cw. \end{aligned} \tag{3.28}$$

The map (3.27) transforms system (3.26) into a system in the form of system (3.28) if and only if

$$\begin{aligned} Au + p(u, v, w) &= Ax + f(x, y, z), \\ Bv &= By + g(x, y, z) + D\alpha(x, y, z)V(x, y, z) \end{aligned}$$

where V denotes the vector field given by

$$(x, y, z) \mapsto (Ax + f(x, y, z), By + g(x, y, z), Cz).$$

Hence, to obtain the desired transformation, it suffices to show that the (first order partial differential) equation

$$D\alpha V + g = B\alpha \tag{3.29}$$

has a $\mathcal{C}^{1,L,\mu(1-\theta)}$ solution α with the additional property that $\alpha(0) = D\alpha(0) = 0$, and that p has the properties listed for system (3.9).

To solve equation (3.29), let us seek a solution along its characteristics; that is, let us seek a function α such that

$$\frac{d}{dt}\alpha(\varphi_t(x, y, z)) - B\alpha(\varphi_t(x, y, z)) = -g(\varphi_t(x, y, z)) \tag{3.30}$$

where φ_t is the flow of V . Of course, by simply evaluating at $t = 0$, it follows immediately that such a function α is also a solution of the equation (3.29).

By variation of parameters, equation (3.30) is equivalent to the differential equation

$$\frac{d}{dt}e^{-tB}\alpha(\varphi_t(x, y, z)) = -e^{-tB}g(\varphi_t(x, y, z)).$$

Hence, (after evaluation at $t = 0$) it suffices to solve the equation

$$J\alpha = -g$$

where J is the (Lie derivative) operator defined by

$$(J\alpha)(x, y, z) = \left. \frac{d}{dt}e^{-tB}\alpha(\varphi_t(x, y, z)) \right|_{t=0}.$$

In other words, it suffices to prove that the operator J is invertible in a space of functions containing g , that $\alpha := J^{-1}g$ is in $\mathcal{C}^{1,L,\mu(1-\theta)}$, and $\alpha(0) = D\alpha(0) = 0$.

Formally, J satisfies the ‘‘Lie derivative property’’; that is,

$$\frac{d}{dt}e^{-tB}\alpha(\varphi_t(x, y, z)) = e^{-tB}J\alpha(\varphi_t(x, y, z)).$$

Hence,

$$e^{-tB}\alpha(\varphi_t(x, y, z)) - \alpha(x, y, z) = \int_0^t e^{-sB}J\alpha(\varphi_s(x, y, z)) ds; \quad (3.31)$$

and therefore, if α is in a function space where $\lim_{t \rightarrow \infty} |e^{-tB}\alpha(\varphi_t(x, y, z))| = 0$, then the operator E defined by

$$(E\alpha)(x, y, z) = - \int_0^\infty e^{-tB}\alpha(\varphi_t(x, y, z)) dt \quad (3.32)$$

is the inverse of J . In fact, by passing to the limit as $t \rightarrow \infty$ on both sides of equation (3.31) it follows immediately that $EJ = I$. The identity $JE = I$ is proved using a direct computation and the fundamental theorem of calculus as follows:

$$\begin{aligned} JE\alpha(x, y, z) &= -\frac{d}{ds} \int_0^\infty e^{(t+s)B}\alpha(\varphi_{t+s}(x, y, z)) dt \Big|_{s=0} \\ &= -\frac{d}{ds} \int_s^\infty e^{uB}\alpha(\varphi_u(x, y, z)) du \Big|_{s=0} \\ &= \alpha(x, y, z). \end{aligned}$$

Let \mathcal{B} denote the Banach space consisting of all continuous functions $\alpha : \Omega \rightarrow \mathbf{R}^\ell$ with the norm

$$\|\alpha\|_{\mathcal{B}} := \sup_{(x,y) \neq 0} \frac{|\alpha(x, y, z)|}{(|x| + |y|)(|x| + |y| + |z|)}$$

where Ω is an open neighborhood of the origin with compact closure. We will show that E is a bounded operator on \mathcal{B} . If Ω is sufficiently small so that the function $\mathcal{F} = (f, g)$ satisfies property (3.24) in Proposition 3.4, then $g \in \mathcal{B}$. Thus, the near-identity transformation (3.27), with $\alpha := Eg$, is a candidate for the desired transformation that partially linearizes system (3.26). The proof is completed by showing that $Eg \in \mathcal{C}^{1,L,\mu(1-\theta)}$.

Because of the special decoupled form of system (3.26) and for the purpose of distinguishing solutions from initial conditions, it is convenient to recast system (3.26) in the form

$$\begin{aligned}\dot{\mathcal{U}} &= \mathcal{A}\mathcal{U} + \mathcal{F}(\mathcal{U}, \zeta), \\ \dot{\zeta} &= C\zeta\end{aligned}\tag{3.33}$$

so that we can write $t \mapsto (\mathcal{U}(t), \zeta(t))$ for the solution with the initial condition $(\mathcal{U}(0), \zeta(0)) = (U, z)$. The next proposition states the growth estimates for the components of the flow φ_t , its partial derivatives, and certain differences of its components and partial derivatives that will be used to prove that E is a bounded operator on \mathcal{B} and $Eg \in \mathcal{C}^{1,L,\mu(1-\theta)}$.

Proposition 3.7. *Suppose that for $i \in \{1, 2\}$ the function $t \mapsto (\mathcal{U}_i(t), \zeta_i(t))$ is the solution of system (3.33) such that $(\mathcal{U}_i(0), \zeta_i(0)) = (U_i, z_i)$ and $t \mapsto (\mathcal{U}(t), \zeta(t))$ is the solution such that $(\mathcal{U}(0), \zeta(0)) = (U, z)$. If Hypothesis 3.5 holds, then there are constants $K > 0$ and $\kappa > 0$ such that*

$$|\zeta(t)| \leq K e^{-\lambda t} |z|,\tag{3.34}$$

$$|\mathcal{U}(t)| \leq K e^{(\delta-b)t} |U|,\tag{3.35}$$

$$|\mathcal{U}(t)| + |\zeta(t)| \leq K e^{-\lambda t} (|U| + |z|),\tag{3.36}$$

$$|\zeta_1(t) - \zeta_2(t)| \leq K e^{-\lambda t} |z_1 - z_2|,\tag{3.37}$$

$$|\mathcal{U}_1(t) - \mathcal{U}_2(t)| \leq \mathcal{K} e^{(\delta-\lambda)t} (|U_1 - U_2| + |z_1 - z_2|),\tag{3.38}$$

$$|\mathcal{U}_U(t)| \leq K e^{(\delta-b)t},\tag{3.39}$$

$$|\mathcal{U}_z(t)| \leq \mathcal{K} e^{(\delta-b)t} |U|,\tag{3.40}$$

$$|\mathcal{U}_{1U}(t) - \mathcal{U}_{2U}(t)| \leq \kappa e^{(\delta-b)t} (|U_1 - U_2| + |z_1 - z_2|).\tag{3.41}$$

Moreover, if $z_1 = z_2$, then

$$|\mathcal{U}_1(t) - \mathcal{U}_2(t)| \leq K e^{(\delta-b)t} |U_1 - U_2|,\tag{3.42}$$

$$|\mathcal{U}_{1z}(t) - \mathcal{U}_{2z}(t)| \leq \kappa e^{(\delta-b)t} |U_1 - U_2|;\tag{3.43}$$

and, if $U_1 = U_2$, then

$$|\mathcal{U}_{1z}(t) - \mathcal{U}_{2z}(t)| \leq \kappa e^{(\delta-b)t} |z_1 - z_2|^{\mu(1-\theta)}.\tag{3.44}$$

Proof. By the definition of δ and the inequality $K > 1$, we have the inequalities $KMr + \epsilon < \delta$ and $2KMr + \epsilon < \delta$. (These inequalities are used so that

the single quantity $\delta - b$, rather than three different exponents, appears in the statement of the proposition.)

The estimate (3.34) follows immediately by solving the differential equation $\dot{\zeta} = C\zeta$ and using the hyperbolic estimate (3.11). To prove the inequality (3.35), start with the variation of parameters formula

$$\mathcal{U}(t) = e^{t\mathcal{A}}U + \int_0^t e^{(t-s)\mathcal{A}} \mathcal{F}(\mathcal{U}(s), \zeta(s)) ds,$$

use the hyperbolic estimates to obtain the inequality

$$|\mathcal{U}(t)| \leq Ke^{-(b-\epsilon)t} |U| + \int_0^t Ke^{-(b-\epsilon)(t-s)} |\mathcal{F}(\mathcal{U}(s), \zeta(s))| ds,$$

and then use the estimate (3.24) to obtain the inequality

$$|\mathcal{U}(t)| \leq Ke^{-(b-\epsilon)t} |U| + \int_0^t rMK e^{-(b-\epsilon)(t-s)} |\mathcal{U}(s)| ds.$$

Rearrange this last inequality to the equivalent form

$$e^{(b-\epsilon)t} |\mathcal{U}(t)| \leq K|U| + \int_0^t rMK e^{(b-\epsilon)s} |\mathcal{U}(s)| ds$$

and apply Gronwall's inequality to show

$$e^{(b-\epsilon)t} |\mathcal{U}(t)| \leq Ke^{rMKt} |U|,$$

an estimate that is equivalent to the desired result.

The inequality (3.37) is easy to prove and inequality (3.36) is a simple corollary of estimates (3.34) and (3.35).

To begin the proof for estimates (3.38) and (3.42), use variation of parameters to obtain the inequality

$$\begin{aligned} |\mathcal{U}_1(t) - \mathcal{U}_2(t)| &\leq Ke^{-(b-\epsilon)t} |U_1 - U_2| \\ &\quad + \int_0^t KMe^{-(b-\epsilon)(t-s)} |\mathcal{F}(\mathcal{U}_1(t), \zeta_1(t)) - \mathcal{F}(\mathcal{U}_2(t), \zeta_2(t))| ds. \end{aligned}$$

For estimate (3.38) use the inequalities (3.24) and $\delta - b < -\lambda$ to obtain the upper bound

$$|\mathcal{F}(\mathcal{U}_1(t), \zeta_1(t)) - \mathcal{F}(\mathcal{U}_2(t), \zeta_2(t))| \leq 2MKr(|\mathcal{U}_1(t) - \mathcal{U}_2(t)| + |\zeta_1(t) - \zeta_2(t)|).$$

Then, using this inequality, the estimates (3.37), and the inequality $-(b-\epsilon) \leq -(\lambda - \epsilon)$, it is easy to see that

$$e^{(\lambda-\epsilon)t} |\mathcal{U}_1(t) - \mathcal{U}_2(t)| \leq K|U_1 - U_2| + \frac{2K^3Mr}{\epsilon} |z_1 - z_2| \\ + \int_0^t e^{(\lambda-\epsilon)s} 2K^2Mr |\mathcal{U}_1(s) - \mathcal{U}_2(s)| ds.$$

The desired result follows by an application of Gronwall's inequality. The proof of estimate (3.42) is similar. The only difference is that the inequality

$$|\mathcal{F}(\mathcal{U}_1(t), \zeta(t)) - \mathcal{F}(\mathcal{U}_2(t), \zeta(t))| \leq 2MKr |\mathcal{U}_1(t) - \mathcal{U}_2(t)|$$

is used instead of inequality (3.37).

To obtain the bounds for the partial derivatives of solutions with respect to the space variables, note that the function $t \mapsto \mathcal{U}_U(t)$ is the solution of the variational initial value problem

$$\dot{\omega} = \mathcal{A}\omega + \mathcal{F}_U(\mathcal{U}(t), \zeta(t))\omega, \quad \omega(0) = I$$

whereas $t \mapsto \mathcal{U}_z(t)$ is the solution of the variational initial value problem

$$\dot{\omega} = \mathcal{A}\omega + \mathcal{F}_U(\mathcal{U}(t), \zeta(t))\omega + \mathcal{F}_z(\mathcal{U}(t), \zeta(t))e^{tC}, \quad \omega(0) = 0.$$

The proofs of the estimates (3.39) and (3.40) are similar to the proof of estimate 3.35. For (3.39), note that \mathcal{F}_U is Lipschitz and use the growth estimates for $|\mathcal{U}(t)|$ and $|\zeta(t)|$ to obtain the inequality $|\mathcal{F}_U(\mathcal{U}(t), \zeta(t))| \leq Mr$. For estimate (3.40) use variation of parameters, bound the term containing \mathcal{F}_z using the Lipschitz estimate, evaluate the resulting integral, and then apply Gronwall's inequality.

To prove estimate (3.41), subtract the two corresponding variational equations, add and subtract $\mathcal{F}_U(\mathcal{U}_1, \zeta_1)\mathcal{U}_{2U}$, and use variation of parameters to obtain the inequality

$$|\mathcal{U}_{1U} - \mathcal{U}_{2U}| \leq \int_0^t Ke^{-(b-\epsilon)(t-s)} |\mathcal{F}_U(\mathcal{U}_1, \zeta_1) - \mathcal{F}_U(\mathcal{U}_2, \zeta_2)| |\mathcal{U}_{2U}| ds \\ + \int_0^t Ke^{-(b-\epsilon)(t-s)} |\mathcal{F}_U(\mathcal{U}_1, \zeta_1)| |\mathcal{U}_{1U} - \mathcal{U}_{2U}| ds.$$

The second integral is bounded by using the Lipschitz estimate for \mathcal{F}_U , inequality (3.36), and the diameter of Ω . For the first integral, a suitable

bound is obtained by again using the Lipschitz estimate for \mathcal{F}_U followed by estimates (3.37) and (3.38), and by using estimate (3.39). After the replacement of the factor $e^{-\lambda s}$ (obtained from (3.37)) by $e^{(\delta-\lambda)s}$, multiplication of both sides of the inequality by $e^{(b-\epsilon)t}$, and an integration, it follows that

$$e^{(b-\epsilon)t} |\mathcal{U}_{1U} - \mathcal{U}_{2U}| \leq \frac{K\mathcal{K}(K + \mathcal{K})M}{\lambda - 2\delta - \epsilon} (|U_1 - U_2| + |z_1 - z_2|) + \int_0^t K^2 M r e^{(b-\epsilon)s} |\mathcal{U}_{1U} - \mathcal{U}_{2U}| ds.$$

The desired result is obtained by an application of Gronwall's inequality.

For the proof of estimate (3.43) subtract the two solutions of the appropriate variational equation, add and subtract $\mathcal{F}_U(\mathcal{U}_1, \zeta)\mathcal{U}_{2z}$ and use variation of parameters. After the Lipschitz estimates are employed, as usual, use inequality (3.36) to estimate $|\mathcal{U}_1| + |\zeta|$ and use inequality (3.42) to estimate $|\mathcal{U}_1 - \mathcal{U}_2|$.

The proof of estimate (3.44) again uses the same basic strategy, that is, variation of parameters and Gronwall's inequality; but several estimates are required before Gronwall's inequality can be applied. First, by the usual method, it is easy to see that

$$e^{(b-\epsilon)t} |\mathcal{U}_{1z}(t) - \mathcal{U}_{2z}(t)| \leq \int_0^t K e^{(b-\epsilon)s} |\mathcal{F}_U(\mathcal{U}_1, \zeta_1)| |\mathcal{U}_{1z} - \mathcal{U}_{2z}| ds + \int_0^t K e^{(b-\epsilon)s} |\mathcal{F}_U(\mathcal{U}_1, \zeta_1) - \mathcal{F}_U(\mathcal{U}_2, \zeta_2)| |\mathcal{U}_{2z}| ds + \int_0^t K^2 e^{(b-\epsilon-\lambda)s} |\mathcal{F}_z(\mathcal{U}_1, \zeta_1) - \mathcal{F}_z(\mathcal{U}_2, \zeta_2)| ds$$

To complete the proof, use Lipschitz estimates for the terms involving partial derivatives of \mathcal{F} in the first two integrals, and use the estimate (3.25) in the third integral. Next, use the obvious estimates for the terms involving \mathcal{U}_i and ζ_i ; but, for the application of inequality (3.38) in the second and third integrals, use the hypothesis that $U_1 = U_2$. Because $1 - \vartheta$ is such that $(1 - \vartheta)b - \lambda < 0$, the third integral converges as $t \rightarrow \infty$. By this observation together with some easy estimates and manipulations, the second integral is bounded above by a constant multiple of $|z_1 - z_2|^{\mu(1-\vartheta)}$. Because the second integral converges as $t \rightarrow \infty$, it is easy to show that the second integral is bounded above by a constant multiple of $|z_1 - z_2|$, a quantity that is itself

bounded above by $r^{1-\mu(1-\vartheta)} |z_1 - z_2|^{\mu(1-\vartheta)}$ where r is the radius of Ω . After the indicated estimates are made, the desired result follows in the usual manner by an application of Gronwall's inequality. \square

Let us return to the analysis of the operator E defined in display (3.32). We will show that if Hypothesis 3.5 is satisfied and $\alpha \in \mathcal{B}$, then $E\alpha \in \mathcal{B}$. The fundamental idea here is to apply the Weierstrass M -test in the form stated in Lemma 2.2. In particular, we will estimate the growth of the integrand in the integral representation of E .

Using the notation defined above, note first that $\varphi_i(x, y, z) = (\mathcal{U}(t), \zeta(t))$. Also, recall that the matrix $-B$ is in real Jordan canonical form; that is,

$$-B = bI + B^{\text{rot}} + B^{\text{nil}}$$

where the second summand has some diagonal or super-diagonal blocks of 2×2 infinitesimal rotation matrices of the form

$$\begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix}$$

where $\beta \neq 0$, and the third summand is a nilpotent matrix whose ℓ th power vanishes. The summands in this decomposition pairwise commute; and therefore,

$$e^{-tB} = e^{bt}Q(t)$$

where the components of the matrix $Q(t)$ are (real) linear combinations of functions given by

$$q_1(t), \quad q_2(t) \sin \beta t, \quad q_3(t) \cos \beta t$$

where q_i , for $i \in \{1, 2, 3\}$, is a polynomial of degree at most $\ell - 1$. It follows that there is a positive universal constant v such that

$$|e^{-tB}| \leq e^{bt}|Q(t)| \leq ve^{bt}(1 + |t|^{\ell-1})$$

for all $t \in \mathbf{R}$.

The integrand of E is bounded above as follows:

$$|e^{-tB} \alpha(\mathcal{U}(t), \zeta(t))| \leq e^{bt} |Q(t)| \|\alpha\|_{\mathcal{B}} (|\mathcal{U}(t)| + |\zeta(t)|) |\mathcal{U}(t)|. \quad (3.47)$$

In view of estimates (3.34) and (3.36), we have the inequality

$$|e^{-tB} \alpha(\mathcal{U}(t), \zeta(t))| \leq K^2 e^{(\delta-\lambda)t} |Q(t)| \|\alpha\|_{\mathcal{B}} (|x| + |y| + |z|) (|x| + |y|).$$

Because $|Q(t)|$ has polynomial growth and $\delta - b < 0$,

$$N := \sup_{t \geq 0} e^{(\delta-b)t/2} |Q(t)| < \infty. \quad (3.48)$$

Hence, we have that

$$|e^{-tB} \alpha(\mathcal{U}(t), \zeta(t))| \leq K^2 N e^{(\delta-\lambda)t/2} \|\alpha\|_{\mathcal{B}} (|x| + |y| + |z|)(|x| + |y|)$$

and

$$\int_0^\infty |e^{-tB} \alpha(\varphi_t(x, y, z))| dt \leq \frac{2K^2 N}{\lambda - \delta} \|\alpha\|_{\mathcal{B}} (|x| + |y| + |z|)(|x| + |y|);$$

and therefore, $E\alpha$ is continuous in Ω and

$$\|E\|_{\mathcal{B}} \leq \frac{2K^2 N}{\lambda - \delta}.$$

As a result, the equation $J\alpha = -g$ has a unique solution $\alpha \in \mathcal{B}$, namely,

$$\alpha(x, y, z) = \int_0^\infty e^{-tB} g(\varphi_t(x, y, z)) dt.$$

We will show that $\alpha \in \mathcal{C}^{1,L,\mu(1-\theta)}$.

In view of the form of system (3.33), to prove that $\alpha \in \mathcal{C}^1$ it suffices to demonstrate that the partial derivatives of α with respect to $U := (x, y)$ and z are both \mathcal{C}^1 , a fact that we will show by using Lemma 2.2.

The solution $t \mapsto (\mathcal{U}(t), \zeta(t))$ with initial condition (U, z) is more precisely written in the form $t \mapsto (\mathcal{U}(t, U, z), \zeta(t, z))$ where the dependence on the initial conditions is explicit. Although this dependence is suppressed in most of the formulas that follow, let us note here that ζ does not depend on U . At any rate, the partial derivatives of α are given formally by

$$\alpha_U(U, z) = \int_0^\infty e^{-tB} g_U(\mathcal{U}(t), \zeta(t)) \mathcal{U}_U(t) dt, \quad (3.49)$$

$$\alpha_z(U, z) = \int_0^\infty e^{-tB} (g_U(\mathcal{U}(t), \zeta(t)) \mathcal{U}_z(t) + g_z(\mathcal{U}(t), \zeta(t)) e^{tC}) dt. \quad (3.50)$$

To prove that α_U is \mathcal{C}^1 , use estimate 3.39, the definition (3.48) of N , and note that g_U is Lipschitz to show that the integrand of equation (3.49) is majorized by

$$\begin{aligned} |e^{-tB} g_U(\mathcal{U}(t), \zeta(t)) \mathcal{U}_U(t)| &\leq M e^{\delta t} |Q(t)| (|\mathcal{U}(t)| + |\zeta(t)|) \\ &\leq K M N e^{(\delta-\lambda)t/2} (|U| + |z|). \end{aligned}$$

By an application of Lemma 2.2, α_U is \mathcal{C}^1 . Moreover, because $|\alpha_U(U, z)|$ is bounded above by a constant multiple of $|U| + |z|$, we also have that $\alpha_U(0, 0) = 0$.

The proofs to show that α_z is \mathcal{C}^1 and $\alpha_z(0, 0) = 0$ are similar. For example, the integrand of equation (3.50) is majorized by

$$K^2 M N e^{(\delta-\lambda)t/2} \left(\frac{KM}{\lambda - \delta + \epsilon} (|U| + |z|) |U| + |U| \right).$$

To prove that α_U is Lipschitz, let us note first that by adding and subtracting $g_U(\mathcal{U}_1, \zeta_1) \mathcal{U}_{2U}$ and an application of the triangle law, we have the inequality

$$\begin{aligned} |\alpha_U(U_1, z_1) - \alpha_U(U_2, z_2)| &\leq \int_0^\infty |e^{-tB}| |g_U(\mathcal{U}_1, \zeta_1) \mathcal{U}_{1U} - g_U(\mathcal{U}_2, \zeta_2) \mathcal{U}_{2U}| dt \\ &\leq \int_0^\infty e^{bt} |Q(t)| (|g_U(\mathcal{U}_1, \zeta_1)| |\mathcal{U}_{1U} - \mathcal{U}_{2U}| \\ &\quad + |g_U(\mathcal{U}_1, \zeta_1) - g_U(\mathcal{U}_2, \zeta_2)| |\mathcal{U}_{2U}|) dt. \end{aligned}$$

By using the Lipschitz estimate for g_U inherited from \mathcal{F}_U , the obvious choices of the inequalities in Proposition 3.7, and an easy computation, it follows that the integrand in the above inequality is majorized up to a constant multiple by

$$e^{(2\delta-\lambda)t/2} (|U_1 - U_2| + |z_1 - z_2|).$$

There are two key points in the proof of this fact: the estimates (3.39) and (3.41) both contain the exponential decay factor e^{-bt} to compensate for the growth of e^{bt} ; and, after the cancelation of these factors, the majorizing integrand still contains the exponential factor $e^{(2\delta-\lambda)t}$, the presence of which ensures that the majorizing integral converges even though the factor $|Q(t)|$ has polynomial growth. After the majorization is established, the desired result follows from Lemma 2.2.

The proof that the function $U \mapsto \alpha_z(U, z)$ is (uniformly) Lipschitz is similar to the proof that α_U is Lipschitz. As before, by adding and subtracting $g_U(\mathcal{U}_1, \zeta) \mathcal{U}_{2z}$, it is easy to obtain the basic estimate

$$\begin{aligned} |\alpha_z(U_1, z) - \alpha_z(U_2, z)| &\leq \int_0^\infty e^{bt} |Q(t)| (|g_U(\mathcal{U}_1, \zeta)| |\mathcal{U}_{1z} - \mathcal{U}_{2z}| \\ &\quad + |g_U(\mathcal{U}_1, \zeta) - g_U(\mathcal{U}_2, \zeta)| |\mathcal{U}_{2z}| \\ &\quad + |g_z(\mathcal{U}_1, \zeta) - g_z(\mathcal{U}_2, \zeta)| K e^{-\lambda t}) dt. \end{aligned}$$

By first applying the Lipschitz estimates and then the inequalities (3.37), (3.42), and (3.43), the growth factor e^{bt} is again canceled; and, up to a constant multiple, the integrand is majorized by the integrable function

$$t \mapsto e^{(\delta-\lambda)t} |Q(t)| |U_1 - U_2|.$$

Finally, we will show that the function $z \mapsto \alpha_z(U, z)$ is (uniformly) Hölder. In this case $U_1 = U_2$. But this equality does not imply that $\mathcal{U}_1 = \mathcal{U}_2$. Thus, the basic estimate in this case is given by

$$\begin{aligned} |\alpha_z(U, z_1) - \alpha_z(U, z_2)| &\leq \int_0^\infty e^{bt} |Q(t)| (|g_U(\mathcal{U}_1, \zeta_1)| |\mathcal{U}_{1z} - \mathcal{U}_{2z}| \\ &\quad + |g_U(\mathcal{U}_1, \zeta_1) - g_U(\mathcal{U}_2, \zeta_2)| |\mathcal{U}_{2z}| \\ &\quad + |g_z(\mathcal{U}_1, \zeta_1) - g_z(\mathcal{U}_2, \zeta_2)| K e^{-\lambda t}) dt. \end{aligned}$$

Use Lipschitz estimates for the partial derivatives in the first two terms in the (expanded) integrand and the estimate (3.25) for the third term. Then, use the estimates (3.36), (3.37), and (3.38) to show that the first term is majorized, up to a constant multiple, by

$$e^{(\delta-\lambda)t} |Q(t)| |z_1 - z_2|^{\mu(1-\vartheta)}.$$

The second term is bounded above by a similar function that has the decay rate $2\delta - \lambda$. After some obvious manipulation, the third term is majorized by a similar term with the decay rate $\delta\vartheta - \lambda + b(1 - \vartheta)$. By Hypothesis 3.5, this number is negative; and therefore, the integrand is majorized by an integrable function multiplied by the required factor $|z_1 - z_2|^{\mu(1-\vartheta)}$.

The final step of the proof is to show that the function p in system (3.28) has the properties listed for system (3.9). There is an essential observation: p is obtained by composing f in system (3.26) with the inverse of the transformation (3.27). In particular, no derivatives of α are used to define p . More precisely, the inverse transformation has the form

$$x = u, \quad y = v + \beta(u, v, w), \quad w = z \tag{3.51}$$

and

$$p(u, v, w) = f(u, v + \beta(u, v, w), w).$$

Hence it is clear that $p(0, 0, 0) = 0$ and $Dp(0, 0, 0) = 0$.

Note that β is \mathcal{C}^1 and therefore Lipschitz in a bounded neighborhood of the origin. Because $\beta(u, v, w) = -\alpha(x, y, z)$, it follows that $\beta_u = -\alpha_x + \alpha_y \beta_u$.

By using a Neumann series representation, note that $I + \alpha_y$ is invertible (with Lipschitz inverse) if α_y is restricted to a sufficiently small neighborhood of the origin. Hence, we have that $\beta_u = -(I + \alpha_y)^{-1}\alpha_x$ and

$$p_u = f_x + f_y\beta_u = f_x - f_y(I + \alpha_y)^{-1}\alpha_x$$

where the right-hand side is to be viewed as a function composed with the inverse transformation (3.51). Moreover, since sums, products, and compositions of bounded Lipschitz (respectively, Hölder) maps are bounded Lipschitz (respectively, Hölder) maps, p_u is Lipschitz. Similarly,

$$p_v = f_y(I - (I + \alpha_y)^{-1}\alpha_y),$$

and it follows that p_v is (uniformly) Lipschitz. Hence, p_v is (uniformly) Lipschitz with respect to its first argument and (uniformly) Hölder with respect to its second and third arguments. Finally, we have that

$$p_w = -f_y(I + \alpha_y)^{-1}\alpha_w + f_z.$$

It follows that p_w is (uniformly) Lipschitz with respect to its first and second arguments and (uniformly) Hölder with respect to its third argument. Hence, p_w is (uniformly) Lipschitz with respect to its first argument and (uniformly) Hölder with respect to its second and third arguments. \square

To complete the proof of Theorem 3.2, we will show that it suffices to choose the Hölder exponent in the statement of the theorem less than the Hölder spectral exponent of $DX(0)$. Note that two conditions have been imposed on the Hölder exponent $\mu(1 - \vartheta)$ in Theorem 3.6: the $(1, \mu)$ *spectral gap condition* $(1 + \mu)c < b$ (or $\mu < (b - c)/c$) and the inequality $b(1 - \vartheta) - \lambda < 0$ (or $1 - \vartheta < \lambda/b$) in Hypothesis 3.5. Because the real parts of the eigenvalues of C lie in the interval $[-c, -d]$ and λ can be chosen anywhere in the interval $(0, d)$, the numbers μ and ϑ can be chosen so that the positive quantity

$$\frac{(b - c)d}{bc} - \mu(1 - \vartheta)$$

is as small as we wish. We will choose $\mu(1 - \vartheta)$, the Hölder exponent, as large as possible under the constraints imposed by the spectral gap condition and the inequality

$$\mu(1 - \vartheta) < \frac{(b - c)d}{bc}.$$

Suppose that the real parts of the eigenvalues of $A := DX(0)$ are as in display (1.1). At the first step of the finite induction on the dimension of the “unlinearized” part of the system, we artificially introduce a scalar equation $\dot{z} = -cz$ where $0 < c < b_1$. In this case $c = d$, and the exponent $\mu(1 - \vartheta)$ can be chosen to be as close as we like to the number $(b_1 - c)/b_1$. At the second step, the real parts of the eigenvalues of the new matrix C are in the interval $[-b_1, -c]$, the new exponent can be chosen as close as we like to the minimum of the numbers $(b_1 - c)/b_1$ and $(b_2 - b_1)c/(b_1b_2)$, and so on. Hence, the Hölder exponent in Theorem 3.2 can be chosen as close as we like to

$$\text{HSE} := \max_{0 < c < b_1} \min \left\{ \frac{b_1 - c}{b_1}, \frac{(b_2 - b_1)c}{b_2b_1}, \frac{(b_3 - b_2)c}{b_3b_2}, \dots, \frac{(b_N - b_{N-1})c}{b_Nb_{N-1}} \right\}.$$

By treating the rational expressions as linear functions of c defined on the interval $[0, b_1]$, it is easy to show that HSE is the Hölder spectral exponent for $DX(0)$. This completes the proof of Theorem 3.2.

3.1.1 Smooth Linearization on the Line

By Theorem 3.2, a $\mathcal{C}^{1,1}$ vector field on the line is $\mathcal{C}^{1,\mu}$ linearizable at a hyperbolic rest point. But in this case a stronger result is true (see [St57]).

Theorem 3.8. *If X is a $\mathcal{C}^{1,1}$ vector field on \mathbf{R}^1 with a hyperbolic rest point at the origin, then X is locally $\mathcal{C}^{1,1}$ linearizable at the origin by a near-identity transformation. If, in addition, X is \mathcal{C}^k with $k > 1$, then there is a \mathcal{C}^k near-identity linearizing transformation.*

Proof. Near the origin, the vector field has the form $X(x) = -ax + f(x)$ where $a \neq 0$ and f is a $\mathcal{C}^{1,1}$ -function with $f(0) = f'(0) = 0$. Let us assume that $a > 0$. The proof for the case $a < 0$ is similar.

We seek a linearizing transformation given by

$$u = x + \alpha(x)$$

where $\alpha(0) = \alpha'(0) = 0$. Clearly, it suffices to prove that

$$\alpha'(x)(-ax + f(x)) + a\alpha(x) = -f(x) \tag{3.52}$$

for all x in some open neighborhood of the origin.

Let ϕ_t denote the flow of X and (in the usual manner) note that if $\alpha \in \mathcal{C}^1$ is such that

$$\frac{d}{dt}\alpha(\phi_t(x)) + a\alpha(\phi_t(x)) = -f(\phi_t(x)),$$

then α satisfies the identity (3.52). Using variation of constants, it follows that α is given formally by

$$\alpha(x) = \int_0^\infty e^{at} f(\phi_t(x)) dt. \quad (3.53)$$

We will show that this formal expression defines a sufficiently smooth choice for α .

Using the assumption that f' is Lipschitz and Taylor's theorem, there is a constant $M > 0$ such that

$$|f(x)| \leq M|x|^2. \quad (3.54)$$

Also, the solution $t \mapsto \mathcal{X}(t) := \phi_t(x)$ is bounded if x is sufficiently small; in fact, for $0 < r < a/(2M)$, we have that

$$|\mathcal{X}(t)| \leq r \quad \text{whenever} \quad |x| < r.$$

To see this, write $X(x) = x(-a + (f(x)/x^2)x)$, use the inequality (3.54), and note the direction of $X(x)$ for x in the given interval.

By variation of constants and the inequality (3.54), we have that

$$e^{at}|\mathcal{X}(t)| \leq |x| + \int_0^t e^{as} Mr|\mathcal{X}(s)| ds.$$

Hence, by Gronwall's inequality, we have the estimate

$$|\mathcal{X}(t)| \leq |x|e^{(Mr-a)t}. \quad (3.55)$$

Note that the function given by $t \mapsto \mathcal{X}_x(t)$ is the solution of the variational initial value problem

$$\dot{w} = -aw + f'(\mathcal{X}(t))w, \quad w(0) = 1,$$

and in case $f \in \mathcal{C}^2$, the function given by $t \mapsto \mathcal{X}_{xx}(t)$ is the solution of

$$\dot{z} = -az + f'(\mathcal{X}(t))z + f''(\mathcal{X}(t))w^2(t), \quad z(0) = 0.$$

Their are similar variational equations for the higher order derivatives of \mathcal{X} with respect to x .

If $\rho > 0$ is given, there is a bounded open interval Ω containing the origin such that $|f'(x)| \leq \rho$ whenever $x \in \Omega$. Using this estimate, variation of constants, and Gronwall's inequality, it is easy to show that the solution \mathcal{W} of the first variational equation is bounded above as follows:

$$|\mathcal{W}(t)| \leq e^{(\rho-a)t}. \quad (3.56)$$

Likewise, if $|f''(x)| \leq \sigma$ whenever $x \in \Omega$, then, by a similar argument, we have that

$$|\mathcal{Z}(t)| \leq \frac{\sigma}{a-2\rho} e^{(\rho-a)t}, \quad (3.57)$$

and so on for higher order derivatives.

We will show the smoothness of α up to order two, the proof in the general case is similar.

Choose Ω with a sufficiently small radius so that $2\rho - a < 0$. (For the \mathcal{C}^k case, the inequality $k\rho - a < 0$ is required.) To prove that $\alpha \in \mathcal{C}^0$, bound the absolute value of the integrand in display (3.53) using the growth estimate (3.55) and the inequality (3.54), and then apply Lemma (2.2).

To show that $\alpha \in \mathcal{C}^1$, formally differentiate the integral representation and then bound the absolute value of the resulting integrand using the Lipschitz estimate for f' and the growth bound (3.56). This results in the upper bound

$$M|x|e^{(Mr+\rho-a)t}.$$

For $r > 0$ sufficiently small, the exponential growth rate $Mr + \rho - a$ is negative and the continuity of α_x follows from Lemma 2.2. Also, by using the same estimate, it is clear that $\alpha \in \mathcal{C}^{1,1}$.

In case $f \in \mathcal{C}^2$, the second derivative of the integrand in the integral representation of α is bounded above by

$$|e^{at}(|f''(\mathcal{X}(t))||\mathcal{X}_x(t)|^2 + |f'(\mathcal{X}(t))||\mathcal{X}_{xx}(t)|)|.$$

This term is majorized using the inequality $|f''(x)| \leq \sigma$, the Lipschitz estimate for f' , and the growth bounds (3.56) and (3.57). The exponential growth rates of the resulting upper bound are $2\rho - a$ and $Mr + \rho - a$. If Ω is chosen with a sufficiently small radius, then both rates are negative. \square

3.2 Hyperbolic Saddles

The main result of this section is the following theorem.

Theorem 3.9. *If X is a \mathcal{C}^2 vector field on \mathbf{R}^2 such that $X(0) = 0$ and $DX(0)$ is infinitesimally hyperbolic, then X is locally \mathcal{C}^1 conjugate to its linearization at the origin.*

We will formulate and prove a slightly more general result about the linearization of systems on \mathbf{R}^n with hyperbolic saddle points.

Consider a linear map $\mathcal{A} : \mathbf{R}^n \rightarrow \mathbf{R}^n$ and suppose that there are positive numbers $a_L, a_R, b_L,$ and b_R such that the real parts of the eigenvalues of \mathcal{A} are contained in the union of the intervals $[-a_L, -a_R]$ and $[b_L, b_R]$. By a linear change of coordinates, \mathcal{A} is transformed to a block diagonal matrix with two diagonal blocks: \mathcal{A}^s , a matrix whose eigenvalues have their real parts in the interval $[-a_L, -a_R]$, and \mathcal{A}^u , a matrix whose eigenvalues have their real parts in the interval $[b_L, b_R]$. Suppose that $0 < \mu < 1$ and every quasi-linear $\mathcal{C}^{1,1}$ vector field of the form $\mathcal{A}^s + F$ (that is, $F(0) = DF(0) = 0$) is $\mathcal{C}^{1,\mu}$ linearizable at the origin. Likewise, suppose that $0 < \nu < 1$ and every quasi-linear $\mathcal{C}^{1,1}$ vector field of the form $\mathcal{A}^u + G$ is $\mathcal{C}^{1,\nu}$ linearizable at the origin. In particular, this is true if μ is the Hölder spectral exponent (1.1) associated with the real parts of eigenvalues in the interval $[-a_L, -a_R]$ and ν is the Hölder spectral exponent associated with the real parts of eigenvalues in the interval $[b_L, b_R]$. We say that \mathcal{A} satisfies *Hartman's (μ, ν) -spectral condition* if

$$a_L - a_R < \mu b_L, \quad b_R - b_L < \nu a_R.$$

A linear transformation of \mathbf{R}^2 with one negative and one positive eigenvalue satisfies Hartman's (μ, ν) -spectral condition for every pair of Hölder exponents (μ, ν) . Hence, Theorem 3.9 is a corollary of the following more general result.

Theorem 3.10. *If X is a $\mathcal{C}^{1,1}$ vector field on \mathbf{R}^n such that $X(0) = 0$ and $DX(0)$ satisfies Hartman's (μ, ν) -spectral condition, then X is locally \mathcal{C}^1 conjugate to its linearization at the origin.*

The proof of Theorem 3.10 has two main ingredients: a change of coordinates into a normal form where the stable and unstable manifolds of the saddle point at the origin are flattened onto the corresponding linear subspaces of \mathbf{R}^n in such a way that the system is linear on each of these invariant

subspaces, and the application of a linearization procedure for systems in this normal form.

A vector field on \mathbf{R}^n with a hyperbolic saddle point at the origin is in (μ, ν) -flattened normal form if it is given by

$$(x, y) \mapsto (Ax + f(x, y), y = By + g(x, y)) \quad (3.58)$$

where $(x, y) \in \mathbf{R}^k \times \mathbf{R}^\ell$ for $k + \ell = n$, all eigenvalues of A have negative real parts, all eigenvalues of B have positive real parts, $F := (f, g)$ is a \mathcal{C}^1 function defined on an open subset Ω of the origin in $\mathbf{R}^k \times \mathbf{R}^\ell$ with $F(0, 0) = DF(0, 0) = 0$, and there are real numbers M , μ , and ν with $M > 0$, $0 < \mu \leq 1$, and $0 < \nu \leq 1$ such that for $(x, y) \in \Omega$,

$$|f_y(x, y)| \leq M|x|, \quad |g_x(x, y)| \leq M|y|, \quad (3.59)$$

$$|f_x(x, y)| \leq M|y|^\mu, \quad |g_y(x, y)| \leq M|x|^\nu, \quad (3.60)$$

$$|f(x, y)| \leq M|x||y|, \quad |g(x, y)| \leq M|x||y|. \quad (3.61)$$

Theorem 3.10 is a corollary of the following two results.

Theorem 3.11. *If X is a $\mathcal{C}^{1,1}$ vector field on \mathbf{R}^n such that $X(0) = 0$, the linear transformation $DX(0)$ satisfies Hartman's (μ, ν) -spectral condition, and $0 < \nu < \min\{\mu, \nu\}$, then there is an open neighborhood of the origin on which X is $\mathcal{C}^{1,\nu}$ conjugate to a vector field in (μ, ν) -flattened normal form.*

Theorem 3.12. *If X is a vector field on \mathbf{R}^n such that $X(0) = 0$, the linear transformation $DX(0)$ satisfies Hartman's (μ, ν) -spectral condition, and X is in (μ, ν) -flattened normal form, then there is an open neighborhood of the origin on which X is \mathcal{C}^1 conjugate to its linearization at the origin.*

The proof of Theorem 3.11 uses three results: the stable manifold theorem, Dorroh smoothing, and Theorem 3.2. The required version of the stable manifold theorem is a standard result, which is a straightforward generalization of the statement and the proof of Theorem 3.3. On the other hand, since Dorroh smoothing is perhaps less familiar, we will formulate and prove the required result.

Suppose that X is a \mathcal{C}^k vector field on \mathbf{R}^n and $h : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is a \mathcal{C}^k diffeomorphism. The flow of X is automatically \mathcal{C}^k . If we view h as a change of coordinates $y = h(x)$, then the vector field in the new coordinate, given by the push forward of X by h , is

$$y \mapsto Dh(h^{-1}(y))X(h^{-1}(y));$$

and, because the derivative of h appears in the transformation formula, the maximal smoothness of the representation of X in the new coordinate is (generally) at most \mathcal{C}^{k-1} . But the transformed flow, given by $h(\phi_t(h^{-1}(y)))$, is \mathcal{C}^k . Thus, in the transformation theory for differential equations, it is natural to encounter vector fields that are less smooth than their flows. Fortunately, this problem is often proved to be inessential by applying the following version of a result of J. R. Dorroh [D71].

Theorem 3.13. *Suppose that X is a \mathcal{C}^0 vector field on \mathbf{R}^n and k is a positive integer. If X has a $\mathcal{C}^{k,\mu}$ (local) flow ϕ_t and $\phi_t(0) \equiv 0$, then there is a number $T > 0$ and an open set $\Omega \subset \mathbf{R}^n$ with $0 \in \Omega$ such that*

$$h(x) := \frac{1}{T} \int_0^T \phi_t(x) dt$$

defines a $\mathcal{C}^{k,\mu}$ -diffeomorphism in Ω that conjugates X to a $\mathcal{C}^{k,\mu}$ vector field Y on $h(\Omega)$. In particular, $Y(0) = 0$ and Y has a $\mathcal{C}^{k,\mu}$ -flow that is $\mathcal{C}^{k,\mu}$ conjugate to ϕ_t .

Proof. Because the flow is $\mathcal{C}^{k,\mu}$, the function h defined in the statement of the theorem is $\mathcal{C}^{k,\mu}$ for each fixed $T > 0$. Also, we have that

$$Dh(0) = \frac{1}{T} \int_0^T D\phi_t(0) dt.$$

Recall that the space derivative of a flow satisfies the cocycle property, that is, the identity

$$D\phi_t(\phi_s(x))D\phi_s(x) = D\phi_{t+s}(x).$$

In particular, because $x = 0$ is a rest point, $D\phi_t(0)D\phi_s(0) = D\phi_{t+s}(0)$, and because $\phi_0(x) \equiv x$, we also have that $D\phi_0(0) = I$. Hence, $\{D\phi_t(0)\}_{t \in \mathbf{R}}$ is a one-parameter group of linear transformations on \mathbf{R}^n . Moreover, the function $t \mapsto D\phi_t(0)$ is continuous.

If the function $t \mapsto D\phi_t(0)$ were differentiable and $C := \frac{d}{dt}D\phi_t(0)|_{t=0}$, then

$$\frac{d}{dt}D\phi_t(0) = \frac{d}{ds}D\phi_{s+t}(0)|_{s=0} = \frac{d}{ds}D\phi_s(0)D\phi_t(0)|_{s=0} = CD\phi_t(0);$$

and therefore, $D\phi_t(0) = e^{tC}$. Using elementary semigroup theory, this result follows without the a priori assumption that $t \mapsto D\phi_t(0)$ is differentiable (see, for example, [DS58, p. 614]).

There is a constant $M > 0$ such that

$$\|e^{tC} - I\| \leq M|t|$$

whenever $|t| \leq 1$. Hence, we have

$$\begin{aligned} Dh(0) &= \frac{1}{T} \int_0^T e^{tC} dt \\ &= \frac{1}{T} \left(\int_0^T I dt + \int_0^T (e^{tC} - I) dt \right) \\ &= I + \frac{1}{T} \int_0^T (e^{tC} - I) dt. \end{aligned}$$

For $0 < T < 1$, the norm of the operator

$$B := \frac{1}{T} \int_0^T (e^{tC} - I) dt$$

is bounded by MT . If $T < 1/M$, then $Dh(0) = I+B$ with $\|B\| < 1$. It follows that $Dh(0)$ is invertible, and by the inverse function theorem, h is a $\mathcal{C}^{k,\mu}$ -diffeomorphism defined on some neighborhood of the origin. (The “usual” inverse function theorem does not mention Hölder derivatives. But the stated result can be proved with an easy modification of the standard proof that uses the contraction principle. For example, use the fiber contraction method to prove smoothness and note that the fiber contraction preserves a space of candidate derivatives that also satisfy the Hölder condition.)

Let us note that

$$\frac{d}{ds} \phi_t(\phi_s(x)) = D\phi_t(\phi_s(x))F(\phi_s(x))$$

and

$$\frac{d}{ds} \phi_t(\phi_s(x)) = \frac{d}{ds} \phi_{t+s}(x) = F(\phi_{s+t}(x)).$$

Hence, we have the identity

$$D\phi_t(\phi_s(x))F(\phi_s(x)) = F(\phi_{s+t}(x))$$

and, at $s = 0$,

$$D\phi_t(x)F(x) = F(\phi_t(x)).$$

It follows that

$$\begin{aligned}
Dh(x)F(x) &= \frac{1}{T} \int_0^T D\phi_t(x)F(x) dt \\
&= \frac{1}{T} \int_0^T F(\phi_t(x)) dt \\
&= \frac{1}{T} \int_0^T \frac{d}{dt} \phi_t(x) dt \\
&= \frac{1}{T} (\phi_T(x) - x);
\end{aligned}$$

and therefore, $x \mapsto Dh(x)F(x)$ is a $\mathcal{C}^{k,\mu}$ -function. The push forward of X by h , namely,

$$y \mapsto Dh(h^{-1}(y))F(h^{-1}(y))$$

is the composition of two $\mathcal{C}^{k,\mu}$ functions, hence it is $\mathcal{C}^{k,\mu}$. \square

As a remark, we mention another variant of Dorroh's theorem: A \mathcal{C}^1 vector field with a \mathcal{C}^k flow is locally conjugate to a \mathcal{C}^k vector field. In particular, this result is valid at an arbitrary point p in the phase space, which we may as well assume is $p = 0$. The essential part of the proof is to show that $Dh(0)$ is invertible where h is the function defined in the statement of Theorem 3.13. In fact, by choosing a bounded neighborhood Ω of the origin so that $M := \sup\{\|DF(x)\| : x \in \Omega\}$ is sufficiently small, by using Gronwall's lemma to obtain the estimate $|D\phi_t(0)| \leq e^{Mt}$, and by also choosing $T > 0$ sufficiently small, it is easy to show that $\|I - Dh(0)\| < 1$. Hence, because $Dh(0) = I - (I - Dh(0))$, the inverse of $Dh(0)$ is given by the Neumann series $\sum_{i=0}^{\infty} (I - Dh(0))^i$.

One nice feature of Dorroh's theorem is the explicit formula for the smoothing diffeomorphism h . In particular, since h is an average over the original flow, most dynamical properties of this flow are automatically inherited by the smoothed vector field. For example, invariant sets of the flow are also h -invariant. This fact will be used in the following proof of Theorem 3.11.

Proof. Suppose that the vector field (3.58) is such that $F := (f, g)$ is a \mathcal{C}^1 function. If the inequalities (3.59)-(3.60) are satisfied, then so are the inequalities (3.61). In fact, by using (3.60) we have the identity $f(x, 0) \equiv 0$ and by Taylor's theorem the estimate

$$|f(x, y)| \leq |f_y(x, 0)||y| + \int_0^1 (|f_y(x, ty)||y| - |f_y(x, 0)||y|) dt.$$

The first inequality in display (3.61) is an immediate consequence of this estimate and the first inequality in display (3.59). The second inequality in display (3.61) is proved similarly.

By an affine change of coordinates, the differential equation associated with X has the representation

$$\dot{p} = \tilde{A}p + f_4(p, q), \quad \dot{q} = \tilde{B}q + g_4(p, q) \quad (3.62)$$

where $(p, q) \in \mathbf{R}^k \times \mathbf{R}^\ell$ with $k + \ell = n$, all eigenvalues of \tilde{A} have negative real parts, all eigenvalues of \tilde{B} have positive real parts, and $F_4 := (f_4, g_4)$ is $\mathcal{C}^{1,1}$ with $F_4(0, 0) = DF_4(0, 0) = 0$.

By the (local) stable manifold theorem, there are open sets $U_4 \subset \mathbf{R}^k$ and $V_4 \subset \mathbf{R}^\ell$ such that $(0, 0) \in U_4 \times V_4$ and $\mathcal{C}^{1,1}$ functions $\eta : V_4 \rightarrow \mathbf{R}^k$ and $\gamma : U_4 \rightarrow \mathbf{R}^\ell$ such that $\eta(0) = D\eta(0) = 0$, $\gamma(0) = D\gamma(0) = 0$, the set $\{(p, q) : p = \eta(q)\}$ is overflowing invariant, and $\{(p, q) : q = \gamma(p)\}$ is inflowing invariant.

By the inverse function theorem, the restriction of the near-identity transformation given by

$$u = p - \eta(q), \quad v = q - \gamma(p)$$

to a sufficiently small open set containing the origin in $\mathbf{R}^k \times \mathbf{R}^\ell$ is a $\mathcal{C}^{1,1}$ diffeomorphism. Moreover, the differential equation (3.62) is transformed to

$$\dot{u} = \tilde{A}u + f_3(u, v), \quad \dot{v} = \tilde{B}v + g_3(u, v) \quad (3.63)$$

where $F_3 := (f_3, g_3)$ is $\mathcal{C}^{0,1}$ with $F_3(0, 0) = DF_3(0, 0) = 0$. In view of the fact that the stable and unstable manifolds are invariant, we also have the identities

$$f_3(0, v) \equiv 0, \quad g_3(u, 0) \equiv 0. \quad (3.64)$$

Hence, the transformed invariant manifolds lie on the respective coordinate planes.

Because system (3.63) has a $\mathcal{C}^{1,1}$ flow, Dorroh's smoothing transformation h (defined in Theorem 3.13) conjugates system (3.63) to a $\mathcal{C}^{1,1}$ system. Moreover, by the definition of h , it is clear that it preserves the coordinate planes in the open neighborhood of the origin where it is defined. In fact, h is given by $h(u, v) = (\bar{h}_1(u, v), \bar{h}_2(u, v))$ where

$$\bar{h}_1(0, v) \equiv 0, \quad \bar{h}_2(u, 0) \equiv 0.$$

The invertible derivative of h at the origin has the block diagonal form

$$Dh(0,0) = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}.$$

Hence, the diffeomorphism h is given by

$$(\xi, \zeta) = h(u, v) = (C_1 u + h_1(u, v), C_2 v + h_2(u, v))$$

where C_1 and C_2 are invertible, $\tilde{H} := (h_1, h_2)$ is $\mathcal{C}^{1,1}$ with $\tilde{H}(0,0) = 0$, $D\tilde{H}(0,0) = 0$, and

$$h_1(0, v) \equiv 0, \quad h_2(u, 0) \equiv 0. \quad (3.65)$$

The system (3.63) is transformed by h to

$$\dot{\xi} = \bar{A}\xi + f_2(\xi, \zeta), \quad \dot{\zeta} = \bar{B}\zeta + g_2(\xi, \zeta) \quad (3.66)$$

where $\bar{A} = C_1 \tilde{A} C_1^{-1}$, $\bar{B} = C_2 \tilde{B} C_2^{-1}$, $F_2 := (f_2, g_2)$ is $\mathcal{C}^{1,1}$ with $F_2(0,0) = DF_2(0,0) = 0$, and

$$f_2(0, \zeta) \equiv 0, \quad g_2(\xi, 0) \equiv 0. \quad (3.67)$$

In view of the identities (3.67), the dynamical system (3.66) restricted to a neighborhood of the origin in $\mathbf{R}^k \times \{0\}$ is given by the $\mathcal{C}^{1,1}$ system

$$\dot{\xi} = \bar{A}\xi + f_2(\xi, 0).$$

Moreover, since this system satisfies the hypotheses of Theorem 3.2, it is linearized by a near-identity $\mathcal{C}^{1,\mu}$ diffeomorphism $H_1 : \xi \mapsto \xi + h_3(\xi)$. Likewise, there is a near-identity $\mathcal{C}^{1,\nu}$ diffeomorphism H_2 , given by $\zeta \mapsto \zeta + h_4(\zeta)$, that linearizes

$$\dot{\zeta} = \bar{B}\zeta + g_2(0, \zeta).$$

These maps define a diffeomorphism $H := (H_1, H_2)$ that transforms system (3.66) to a system of the form

$$\dot{\psi} = \bar{A}\psi + f_1(\psi, \omega), \quad \dot{\omega} = \bar{B}\omega + g_1(\psi, \omega) \quad (3.68)$$

where $F_1 := (f_1, g_1)$ is \mathcal{C}^0 with $F_1(0,0) = DF_1(0,0) = 0$,

$$f_1(0, \omega) \equiv 0, \quad g_1(\psi, 0) \equiv 0, \quad (3.69)$$

and

$$f_1(\psi, 0) \equiv 0, \quad g_1(0, \omega) \equiv 0. \quad (3.70)$$

Let $\phi_t = (\phi_t^1, \phi_t^2)$ denote the flow of system (3.66) (even though the same notation has been used to denote other flows). The first component of the flow of system (3.68) is given by

$$H_1(\phi_t^1(H_1^{-1}(\psi), H_2^{-1}(\omega))).$$

Its partial derivative with respect to ψ is clearly in \mathcal{C}^μ , where for notational convenience we use μ and ν to denote Hölder exponents that are strictly smaller than the corresponding Hölder spectral exponents. On the other hand, its partial derivative with respect to ω is bounded by a constant times

$$|(\frac{\partial}{\partial \omega} \phi_t^1)(H_1^{-1}(\psi), H_2^{-1}(\omega))|.$$

Because $f_2(0, \zeta) \equiv 0$, it follows that $\phi_t^1(0, \zeta) \equiv 0$, and therefore,

$$(\frac{\partial}{\partial \zeta} \phi_t^1)(0, \zeta) \equiv 0.$$

Because system (3.66) is in $\mathcal{C}^{1,1}$, there is a constant $M > 0$ such that

$$|(\frac{\partial}{\partial \zeta} \phi_t^1)(\xi, \zeta)| \leq M|\xi|,$$

and consequently,

$$|(\frac{\partial}{\partial \omega} \phi_t^1)(H_1^{-1}(\psi), H_2^{-1}(\omega))| \leq M|H_1^{-1}(\psi)| \leq M\|DH_1^{-1}\|\|\psi\|.$$

Similarly, the partial derivative with respect to ψ of the second component of the flow is bounded above by a constant times $|\omega|$.

By a second application of Dorroh's theorem (3.13), there is a \mathcal{C}^1 diffeomorphism, whose partial derivatives satisfy Hölder and Lipschitz conditions corresponding to those specified for the flow of system (3.68), that transforms system (3.68) to a system of the form

$$\dot{x} = Ax + f(x, y), \quad \dot{y} = By + g(x, y) \quad (3.71)$$

where A is similar to \bar{A} and B is similar to \bar{B} , where $F := (f, g)$ is \mathcal{C}^1 with $F_1(0, 0) = DF_1(0, 0) = 0$ and with corresponding Hölder partial derivatives, and where

$$f(0, y) \equiv 0, \quad g(x, 0) \equiv 0, \quad (3.72)$$

and

$$f(x, 0) \equiv 0, \quad g(0, y) \equiv 0. \quad (3.73)$$

The identities (3.72) are equivalent to the invariance of the coordinate planes, whereas the identities (3.73) are equivalent to the linearity of the system on each coordinate plane. The preservation of linearity on the coordinate planes by the Dorroh transformation is clear from its definition; to wit, a linear flow produces a linear Dorroh transformation.

Because $f(x, 0) \equiv 0$, it follows that $f_x(x, 0) \equiv 0$. Also, $f_x \in \mathcal{C}^\mu$. Hence, there is some $M > 0$ such that

$$|f_x(x, y)| = |f_x(x, y) - f_x(x, 0)| \leq M|y|^\mu.$$

Likewise, we have the estimate

$$|g_y(x, y)| \leq M|x|^\nu.$$

Because $f(0, y) \equiv 0$, it follows that $f_y(0, y) \equiv 0$, and because f_y is Lipschitz, there is a constant $M > 0$ such that

$$|f_y(x, y)| = |f_y(x, y) - f_y(0, y)| \leq M|x|.$$

Similarly, we have that $|g_x(x, y)| \leq M|y|$. □

For a \mathcal{C}^2 vector field on the plane with a hyperbolic saddle point at the origin, there is a stronger version of Theorem 3.11. In fact, in this case, the vector field is conjugate to a \mathcal{C}^2 vector field in flattened normal form. To prove this, use the \mathcal{C}^2 stable manifold theorem to flatten the stable and the unstable manifolds onto the corresponding coordinate axes. Dorroh smoothing conjugates the resulting vector field to a \mathcal{C}^2 vector field that still has invariant coordinate axes. Apply Theorem 3.8 to \mathcal{C}^2 linearize on each coordinate axis, and then use Dorroh smoothing to obtain a \mathcal{C}^2 vector field that is also linearized on each coordinate axis.

The following proof of Theorem 3.12 is similar to a portion of the proof of Theorem 3.2. In particular, an explicit integral formula for the nonlinear part of the linearizing transformation is obtained and its smoothness is proved using Lemma 2.2.

Proof. Let X denote the vector field (3.58) in flattened normal form. We will construct a smooth near-identity linearizing transformation given by

$$u = x + \alpha(x, y), \quad v = y + \beta(x, y). \quad (3.74)$$

The smooth transformation (3.74) linearizes the vector field X if and only if the pair of functions α, β satisfies the system of partial differential equations

$$D\alpha X - A\alpha = -f, \quad D\beta X - B\beta = -g.$$

The first equation is equivalent to the differential equation

$$\frac{d}{dt} e^{tA} \alpha(\phi_{-t}(x, y)) = e^{tA} f(\phi_{-t}(x, y));$$

and therefore, it has the solution

$$\alpha(x, y) = - \int_0^\infty e^{tA} f(\phi_{-t}(x, y)) dt \quad (3.75)$$

provided that the improper integral converges. Similarly, the second equation is equivalent to the differential equation

$$\frac{d}{dt} e^{-tB} \beta(\phi_t(x, y)) = -e^{-tB} g(\phi_t(x, y))$$

and has the solution

$$\beta(x, y) = \int_0^\infty e^{-tB} g(\phi_t(x, y)) dt.$$

We will prove that α , as defined in display (3.75), is a \mathcal{C}^1 function. The proof for β is similar.

By using a smooth bump function as in Section 2, there is no loss of generality if we assume that X is bounded on \mathbf{R}^n . Under this assumption, f is bounded; and because A is a stable matrix, it follows immediately that α is a continuous function defined on an open ball Ω at the origin with radius r .

Let $t \mapsto (\mathcal{X}(t), \mathcal{Y}(t))$ denote the solution of the system

$$\dot{x} = -Ax - f(x, y), \quad (3.76)$$

$$\dot{y} = -By - g(x, y) \quad (3.77)$$

with initial condition $(\mathcal{X}(0), \mathcal{Y}(0)) = (x, y)$ and note that (formally)

$$\alpha_x(x, y) = - \int_0^\infty e^{tA} (f_x(\mathcal{X}(t), \mathcal{Y}(t)) \mathcal{X}_x(t) + f_y(\mathcal{X}(t), \mathcal{Y}(t)) \mathcal{Y}_x(t)) dt \quad (3.78)$$

where $t \mapsto (\mathcal{X}_x(t), \mathcal{Y}_x(t))$ is the solution of the variational initial value problem

$$\begin{aligned} \dot{w} &= -Aw - f_x(\mathcal{X}(t), \mathcal{Y}(t))w - f_y(\mathcal{X}(t), \mathcal{Y}(t))z, \\ \dot{z} &= -Bz - g_x(\mathcal{X}(t), \mathcal{Y}(t))w - g_y(\mathcal{X}(t), \mathcal{Y}(t))z, \\ w(0) &= I, \\ z(0) &= 0. \end{aligned} \tag{3.79}$$

We will show that α_x is a continuous function defined on an open neighborhood of the origin. The proof for α_y is similar.

Several (Gronwall) estimates are required. To set notation, let $F := (f, g)$ and $\rho := \sup\{\|DF(x, y)\| : (x, y) \in \Omega\}$; and assume that every eigenvalue of A has its real part in the interval $[-a_L, -a_R]$ where $0 < a_R < a_L$ and every eigenvalue of B has its real part in the interval $[b_L, b_R]$ where $0 < b_L < b_R$. As before, if a, b, λ , and σ are numbers with $0 < \lambda < a_R < a_L < a$ and $0 < \sigma < b_L < b_R < b$, then there is a number $K > 0$ such that the following hyperbolic estimates hold whenever $t \geq 0$:

$$\begin{aligned} |e^{tA}x| &\leq Ke^{-\lambda t}|x|, & |e^{-tA}x| &\leq Ke^{at}|x|, \\ |e^{tB}y| &\leq Ke^{bt}|y|, & |e^{-tB}y| &\leq Ke^{-\sigma t}|y|. \end{aligned}$$

We will show that there is a constant $\mathcal{K} > 0$ such that the following Gronwall estimates hold:

$$|\mathcal{X}(t)| \leq \mathcal{K}e^{(K\rho+a)t}, \tag{3.80}$$

$$|\mathcal{Y}(t)| \leq \mathcal{K}|y|e^{(KM\rho-\sigma)t}, \tag{3.81}$$

$$|\mathcal{X}_x(t)| \leq \mathcal{K}e^{(K\rho+a)t}, \tag{3.82}$$

$$|\mathcal{Y}_x(t)| \leq \mathcal{K}|y|e^{(KM\rho+2K\rho+a-\sigma)t} \tag{3.83}$$

where M is the constant that appears in the definition of the flattened normal form.

The inequality (3.80) is proved in the usual manner: apply the variation of constants formula to equation (3.76) to derive the estimate

$$\begin{aligned} |\mathcal{X}(t)| &\leq |e^{-tA}x| + \int_0^t |e^{(s-t)A}| |f(\mathcal{X}(s), \mathcal{Y}(s))| ds \\ &\leq Ke^{at}|x| + \int_0^t Ke^{a(t-s)} \rho(|\mathcal{X}(s)| + r) ds, \end{aligned}$$

rearrange and integrate to obtain the estimate

$$e^{-at}|\mathcal{X}(t)| \leq Kr + \frac{K\rho r}{a} + \int_0^t K\rho e^{-as}|\mathcal{X}(s)| ds, \quad (3.84)$$

and then apply Gronwall's inequality.

The proof of inequality (3.81) is similar to the proof of inequality (3.80) except that the estimate in display (3.61) is used for $|g(\mathcal{X}(t), \mathcal{Y}(t))|$ instead of the mean value estimate used for $|f(\mathcal{X}(t), \mathcal{Y}(t))|$.

The estimates (3.82) and (3.83) are proved in two main steps. First, define \mathcal{A} to be the block diagonal matrix with blocks A and B , $U := (x, y)$, and $F := (f, g)$ so that the system (3.76)–(3.77) is expressed in the compact form

$$\dot{U} = -\mathcal{A}U - F(U), \quad (3.85)$$

and the corresponding variational equation (also corresponding to equation (3.79)) is

$$\dot{V} = -\mathcal{A}V - DF(\mathcal{U}(t))V$$

where $t \mapsto \mathcal{U}(t)$ is the solution of system (3.85) with initial condition $\mathcal{U}(0) = U$. An easy Gronwall estimate shows that

$$|\mathcal{V}(t)| \leq Ke^{(K\rho+a)t}$$

where $t \mapsto \mathcal{V}(t)$ is the corresponding solution of the variational equation. Because $|V|$ can be defined to be $|w| + |z|$, it follows that

$$|\mathcal{W}(t)| \leq Ke^{(K\rho+a)t}, \quad |\mathcal{Z}(t)| \leq Ke^{(K\rho+a)t}. \quad (3.86)$$

Next, the estimate for \mathcal{Z} is improved. In fact, using equation (3.79), the corresponding initial condition for $\mathcal{Z}(t)$, and variation of constants, we have that

$$\begin{aligned} |\mathcal{Z}(t)| &\leq \int_0^t |e^{(s-t)B}| |g_x(\mathcal{X}(s), \mathcal{Y}(s))| |\mathcal{W}(s)| ds \\ &\quad + \int_0^t |e^{(s-t)B}| |g_y(\mathcal{X}(s), \mathcal{Y}(s))| |\mathcal{Z}(s)| ds \\ &\leq \int_0^t Ke^{-\sigma(t-s)} M |\mathcal{Y}(s)| Ke^{(K\rho+a)s} ds \\ &\quad + \int_0^t Ke^{-\sigma(t-s)} \rho |\mathcal{Z}(s)| ds. \end{aligned} \quad (3.87)$$

The inequality

$$\begin{aligned} e^{\sigma t} |\mathcal{Z}(t)| &\leq \int_0^t K \mathcal{K}^2 M r e^{\sigma s} e^{(KMr-\sigma)s} e^{(K\rho+a)s} ds \\ &\quad + \int_0^t K \rho e^{\sigma s} |\mathcal{Z}(s)| ds \end{aligned}$$

is obtained by rearrangement of inequality (3.87) and by using the hyperbolic estimate (3.81). After the first integral is bounded above by its value on the interval $[0, \infty)$, the desired result is obtained by an application of Gronwall's inequality.

To show that α_x is continuous, it suffices to show that the absolute value of the integrand J of its formal representation (3.78) is majorized by an integrable function. In fact,

$$\begin{aligned} J &\leq K e^{-\lambda t} (|f_x(\mathcal{X}(t), \mathcal{Y}(t))| |\mathcal{W}(t)| + |f_y(\mathcal{X}(t), \mathcal{Y}(t))| |\mathcal{Z}(t)|) \\ &\leq K e^{-\lambda t} (M |\mathcal{Y}(t)|^\mu |\mathcal{W}(t)| + \rho |\mathcal{Z}(t)|) \\ &\leq K e^{-\lambda t} (M K^\mu r^\mu e^{(KMr\mu-\sigma\mu)t} \mathcal{K} e^{(K\rho+a)t} + \mathcal{K} \rho e^{(KMr+2K\rho+a-\sigma)t}). \end{aligned}$$

Thus, we have proved that J is bounded by a function with two exponential growth rates:

$$KMr\mu + K\rho + a - \lambda - \sigma\mu, \quad KMr\mu + 2K\rho + a - \lambda - \sigma.$$

Note that $a - \lambda - \sigma < a - \lambda - \sigma\mu$, and recall Hartman's spectral condition $a_L - a_R - \mu b_L < 0$. By choosing admissible values of a , λ , and σ such that the three quantities $|a - a_L|$, $|\lambda - a_R|$, and $|\sigma - b_L|$ are sufficiently small, it follows that $a - \lambda - \sigma\mu < 0$. Moreover, once this inequality is satisfied, if $r > 0$ and $\rho > 0$ are sufficiently small, then the two rate factors are both negative. This proves that $\alpha \in \mathcal{C}^1$. \square

4 Linearization of Special Vector Fields

As we have seen, a $\mathcal{C}^{1,1}$ vector field is $\mathcal{C}^{1,\mu}$ linearizable at a hyperbolic sink if the Hölder exponent μ is less than the Hölder spectral exponent of its linearization. Also, a $\mathcal{C}^{1,1}$ vector field is \mathcal{C}^1 linearizable at a hyperbolic saddle point if Hartman's (μ, ν) -spectral condition is satisfied. Can these results be improved?

In view of Sternberg's example (3.2), there is no hope of improving the smoothness of the linearization at a hyperbolic sink from class \mathcal{C}^1 to class \mathcal{C}^2 even for polynomial vector fields.

For hyperbolic saddle points, on the other hand, the existence of a \mathcal{C}^1 linearization is in doubt unless Hartman's (μ, ν) -spectral condition is satisfied. In view of Hartman's example (3.1), it is not possible to remove this condition. Note, however, that this spectral condition is imposed, in the course of the proof of the linearization theorem, under the assumption that the nonlinear part of the vector field at the rest point is arbitrary. Clearly, this result can be improved by restricting the type of nonlinearities that appear. As a trivial example, note that no restriction is necessary if the vector field is linear. It can also be improved by placing further restrictions on the spectrum of the linear part of the vector field.

We will define a class of nonlinear vector fields with hyperbolic sinks at the origin where there is a linearizing transformation of class $\mathcal{C}^{1,\mu}$ for every $\mu \in (0, 1)$. In particular, the size of the Hölder exponent μ of the derivative of the linearizing transformation is not restricted by the Hölder spectral exponent of the linear part of the vector field at the origin. This result will be used to enlarge the class of nonlinear vector fields with hyperbolic saddles at the origin that can be proved to be \mathcal{C}^1 linearizable.

Vector fields corresponding to systems of differential equations of the form

$$\begin{aligned}\dot{u}_1 &= -a_1 u_1 + f_{11}(u_1, \dots, u_n) u_1, \\ \dot{u}_2 &= -a_2 u_2 + f_{21}(u_1, \dots, u_n) u_1 + f_{22}(u_1, \dots, u_n) u_2, \\ &\vdots \\ \dot{u}_n &= -a_n u_n + f_{n1}(u_1, \dots, u_n) u_1 + f_{n2}(u_1, \dots, u_n) u_2 \\ &\quad + \dots + f_{nn}(u_1, \dots, u_n) u_n\end{aligned}$$

where $a_1 > a_2 > \dots > a_n > 0$ and the functions f_{ij} are all of class \mathcal{C}^2 with $f_{ij}(0) = 0$ are in the special class. They are $\mathcal{C}^{1,\mu}$ linearizable at the origin for every $\mu \in (0, 1)$.

4.1 Special Vector Fields

The next definition lists the properties of the special vector fields that will be used in the proofs of the results in this section. The following propositions give simple and explicit criteria that can be easily checked to determine if

a \mathcal{C}^3 vector field and some vector field in this special class are equal when restricted to some open neighborhood of the origin.

We will use the notation $D_j H$ to denote the partial derivative of the function H with respect to its j th variable. Also, for $r > 0$, let

$$\Omega_r := \{x \in \mathbf{R}^n : |x| < r\}.$$

Sometimes we will view Ω_r as a subset of $\mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_p}$ where $n_1 + \cdots + n_p = n$. In this case, a point $x \in \Omega_r$ is expressed in components as $x = (x_1, \dots, x_p)$.

Let \mathcal{P}_r denote the set of all vector fields on Ω_r of the form

$$(x_1, \dots, x_p) \mapsto (A_1 x_1 + F_1(x_1, \dots, x_p), \dots, A_p x_p + F_p(x_1, \dots, x_p)), \quad (4.1)$$

with the following additional properties:

- (1) There are real numbers $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$ such that, for each $i \in \{1, 2, \dots, p\}$, every eigenvalue of the matrix A_i has real part $-\lambda_i$.
- (2) For each $i \in \{1, 2, \dots, p\}$, the function $F_i : \Omega_r \rightarrow \mathbf{R}^{n_i}$ is in class $\mathcal{C}^1(\Omega_r)$. (In particular, F_i and DF_i are bounded functions).
- (3) There is a constant $M > 0$ such that

$$|F_i(x) - F_i(y)| \leq M \left((|x| + |y|) \sum_{k=1}^i |x_k - y_k| + |x - y| \sum_{k=1}^i (|y_k| + |x_k|) \right)$$

whenever $x, y \in \Omega_r$.

- (4) There is a constant $M > 0$ such that

$$|D_j F_i(x) - D_j F_i(y)| \leq M |x - y|$$

whenever $i, j \in \{1, 2, \dots, p\}$ and $x, y \in \Omega_r$. (In particular,

$$|D_j F_i(x_1, \dots, x_p)| \leq M |x|$$

whenever $i \in \{1, 2, \dots, p\}$, $j \in \{1, 2, \dots, i\}$, and $x \in \Omega_r$.)

- (5) There is a constant $M > 0$ such that

$$|D_j F_i(x) - D_j F_i(y)| \leq M \left(\sum_{k=1}^i |x_k - y_k| + |x - y| \sum_{k=1}^i (|x_k| + |y_k|) \right)$$

whenever $i \in \{1, 2, \dots, p\}$, $j \in \{i+1, i+2, \dots, p\}$, and $x \in \Omega_r$. (In particular, $|D_j F_i(x_1, \dots, x_p)| \leq M(|x_1| + \cdots + |x_i|)$ whenever $i \in \{1, 2, \dots, p\}$, $j \in \{i+1, i+2, \dots, p\}$, and $x \in \Omega_r$.)

Definition 4.1. A vector field Y , given by

$$(x_1, \dots, x_p) \mapsto (A_1 x_1 + G_1(x_1, \dots, x_p), \dots, A_p x_p + G_p(x_1, \dots, x_p)) \quad (4.2)$$

where the matrices A_1, \dots, A_p satisfy the property (1) listed in the definition of \mathcal{P}_r , the function $G := (G_1, \dots, G_p)$ is defined on an open neighborhood U of the origin in \mathbf{R}^n , and $G(0) = DG(0) = 0$, is called *lower triangular* if for each $i \in \{1, 2, \dots, p-1\}$

$$G_i(0, 0, \dots, 0, x_{i+1}, x_{i+2}, \dots, x_p) \equiv 0.$$

For a quasi-linear vector field in the form of Y , as given in display (4.2), let A denote the block diagonal matrix with diagonal blocks A_1, A_2, \dots, A_p so that Y is expressed in the compact form $Y = A + G$.

Proposition 4.2. *If the \mathcal{C}^3 vector field $Y = A + G$ is lower triangular on the open set U containing the origin and the closure of Ω_r is in U , then there is a vector field of the form $X = A + F$ in \mathcal{P}_r such that the restrictions of the vector fields X and Y to Ω_r are equal.*

Proof. Fix $r > 0$ such that the closure of Ω_r is contained in U . Because Y is \mathcal{C}^3 , there is a constant $K > 0$ such that G and its first three derivatives are bounded by K on Ω_r .

Because DG is \mathcal{C}^1 , the mean value theorem implies that

$$|D_j G_i(x) - D_j G_i(y)| \leq K|x - y|$$

whenever $x, y \in \Omega_r$. This proves property (4).

Note that (as in the proof of Taylor's theorem)

$$G_i(x_1, x_2, \dots, x_p) = \int_0^1 \sum_{k=1}^i D_k G_i(tx_1, tx_2, \dots, tx_i, x_{i+1}, x_{i+2}, \dots, x_p) x_k dt.$$

Hence, with

$$\begin{aligned} u &:= (x_1, x_2, \dots, x_i), & v &:= (x_{i+1}, x_{i+2}, \dots, x_p), \\ w &:= (y_1, y_2, \dots, y_i), & z &:= (y_{i+1}, y_{i+2}, \dots, y_p), \end{aligned}$$

we have

$$|G_i(x) - G_i(y)| \leq \int_0^1 \sum_{k=1}^i |D_k G_i(tu, v) x_k - D_k G_i(tw, z) y_k| dt.$$

Using the mean value theorem applied to the \mathcal{C}^1 function $f := D_k G_i$, we have the inequalities

$$\begin{aligned}
|f(tu, v)x_k - f(tw, z)y_k| &\leq |f(tu, v)x_k - f(tu, v)y_k| \\
&\quad + |f(tu, v)y_k - f(tw, z)y_k| \\
&\leq |f(tu, v)||x_k - y_k| + |f(tu, v) - f(tw, z)||y_k| \\
&\leq K((|t||u| + |v|)|x_k - y_k| \\
&\quad + (|t||u - w| + |v - z|)|y_k|) \\
&\leq K(|x||x_k - y_k| + |x - y||y_k|); \tag{4.3}
\end{aligned}$$

and as a consequence,

$$\begin{aligned}
|G_i(x) - G_i(y)| &\leq K\left(|x| \sum_{k=1}^i |x_k - y_k| + |x - y| \sum_{k=1}^i |y_k|\right) \\
|G_i(x) - G_i(y)| &\leq K\left((|x| + |y|) \sum_{k=1}^i |x_k - y_k| \tag{4.4}
\end{aligned}$$

$$\begin{aligned}
&\quad + |x - y| \sum_{k=1}^i (|x_k| + |y_k|)) \tag{4.5}
\end{aligned}$$

whenever $x, y \in \Omega_r$. This proves property (3).

Using the integral representation of G , note that

$$D_j G_i(x_1, x_2, \dots, x_p) = \int_0^1 \sum_{k=1}^i (t(D_j D_k G_i)(tu, v)x_k + G_i(tu, v)D_j x_k) dt.$$

If $j > i$, then $D_j x_k = 0$; and therefore, the estimate for $|D_j G_i(x) - D_j G_i(y)|$ required to prove property (5) is similar to the proof of estimate (4.4). The only difference in the proof occurs because the corresponding function f is not required to vanish at the origin. For this reason, the estimate $|f(tu, v)| < K$ is used in place of the Lipschitz estimate for $|f(tu, v)|$ in the chain of inequalities (4.3). \square

Definition 4.3. Suppose that A is an $n \times n$ real matrix, $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$, and the real part of each eigenvalue of A is one of the real numbers $-\lambda_i$ for $i \in \{1, 2, \dots, p\}$. The matrix A has the k -spectral gap condition if $\lambda_{i-1}/\lambda_i > k$ for $i \in \{2, 3, \dots, p\}$.

Proposition 4.4. *Suppose that $Y = A + G$ is a quasi-linear \mathcal{C}^3 vector field defined on the open set U containing the origin. If the matrix A has the 3-spectral gap condition and the closure of Ω_r is contained in U , then there is a vector field of the form X in \mathcal{P}_r such that the restrictions of the vector fields X and Y to Ω_r are equal.*

Proof. We will outline the proof, the details are left to the reader.

There is a linear change of coordinates such that the linear part of the transformed vector field is block diagonal with diagonal blocks A_i , for $i \in \{1, 2, \dots, p\}$, such that $-\lambda_i$ is the real part of each eigenvalue of the matrix A_i and $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$. Thus, without loss of generality, we may as well assume that A has this block diagonal form.

Consider the vector field Y in the form

$$(Bx + G_1(x, y), A_p y + G_2(x, y))$$

where B is block diagonal with blocks A_1, A_2, \dots, A_{p-1} . Because the 3-gap condition is satisfied for the \mathcal{C}^3 vector field Y viewed in this form, an application of the smooth spectral gap theorem (see [LL99]) can be used to obtain a \mathcal{C}^3 function ϕ , defined for $|y|$ sufficiently small, such that $\phi(0) = 0$ and $\{(x, y) : x = \phi(y)\}$ is an invariant manifold.

In the new coordinates $u = x - \phi(y)$ and $v = y$, the vector field Y is given by

$$(Bu + G_2(u, v), A_p v + G_3(u, v))$$

where G_2 and G_3 are \mathcal{C}^2 and $G_2(0, v) \equiv 0$. By an application of Dorroh's theorem (as in the proof of Theorem 3.11), this system is \mathcal{C}^3 conjugate to a \mathcal{C}^3 vector field Y_1 of the same form.

Next, consider the vector field Y_1 in the form

$$(Cu + G_4(u, v, w), A_{p-1}v + G_5(u, v, w), A_p w + G_6(u, v, w))$$

where the variables are renamed. We have already proved that $G_5(0, 0, w) \equiv 0$. By an application of the smooth spectral gap theorem, there is a \mathcal{C}^3 function ψ such that $\{(u, v, w) : u = \psi(v, w)\}$ is an invariant manifold. In the new coordinates $a = u - \psi(v, w)$, $b = v$, and $c = w$, the vector field has the form

$$(Ca + G_7(a, b, c), A_{p-1}b + G_8(a, b, c), A_p w + G_9(a, b, c))$$

where $G_8(0, 0, c) \equiv 0$ and $G_7(0, b, c) \equiv 0$. By Dorroh smoothing we can assume that the functions G_7 , G_8 , and G_9 are class \mathcal{C}^3 .

To complete the proof, repeat the argument to obtain a lower triangular vector field and then apply Proposition 4.4. \square

We will prove the following theorem.

Theorem 4.5. *For each $\mu \in (0, 1)$, a \mathcal{C}^3 lower triangular vector field (or a \mathcal{C}^3 quasi-linear vector field whose linear part satisfies the 3-spectral gap condition) is linearizable at the origin by a $\mathcal{C}^{1,\mu}$ near-identity diffeomorphism.*

In particular, for the restricted class of vector fields mentioned in Theorem 4.5, the Hölder exponent of the linearizing transformation is not required to be less than the Hölder spectral exponent of A ; rather, the Hölder exponent can be chosen as close to the number one as we wish.

4.2 Saddles

Theorem 4.5 together with Theorem 3.10 can be used, in the obvious manner, to obtain improved results on the smooth linearization of special systems with hyperbolic saddles. For example, suppose that $X = \mathcal{A} + \mathcal{F}$ is quasi-linear such that \mathcal{A} is in block diagonal form $\mathcal{A} = (\mathcal{A}^s, \mathcal{A}^u)$ where all eigenvalues of \mathcal{A}^s have negative real parts and all eigenvalues of \mathcal{A}^u have positive real parts. In this case, the vector field has the form $X(x, y) = (\mathcal{A}^s x + G(x, y), \mathcal{A}^u y + H(x, y))$ where $\mathcal{F} = (G, H)$. The vector field X is called *triangular* if $G(0, y) \equiv 0$ and $H(x, 0) \equiv 0$, and the vector fields $x \mapsto \mathcal{A}^s x + G(x, 0)$ and $y \mapsto -(\mathcal{A}^u y + H(0, y))$ are both lower triangular.

Theorem 4.6. *Suppose that $X = \mathcal{A} + \mathcal{F}$ is a quasi-linear \mathcal{C}^3 triangular vector field and there are positive numbers a_L , a_R , b_L , and b_R such that the real parts of the eigenvalues of \mathcal{A} are contained in the union of the intervals $[-a_L, -a_R]$ and $[b_L, b_R]$. If $a_L - a_R < b_L$ and $b_R - b_L < a_R$, then X is \mathcal{C}^1 linearizable.*

The next theorem replaces the requirement that the vector field be triangular with a spectral gap condition.

Theorem 4.7. *Suppose that $X = \mathcal{A} + \mathcal{F}$ is a quasi-linear \mathcal{C}^3 vector field with a hyperbolic saddle at the origin, the set of negative real parts of eigenvalues*

of \mathcal{A} is given by $\{-\lambda_1, \dots, -\lambda_p\}$, the set of positive real parts is given by $\{\sigma_1, \dots, \sigma_q\}$, and

$$-\lambda_1 < -\lambda_2 < \dots < -\lambda_p < 0 < \sigma_q < \sigma_{q-1} < \dots < \sigma_1.$$

If $\lambda_{i-1}/\lambda_i > 3$, for $i \in \{2, 3, \dots, p\}$, and $\sigma_{i-1}/\sigma_i > 3$, for $i \in \{2, 3, \dots, q\}$, and if $\lambda_1 - \lambda_p > \sigma_q$ and $\sigma_1 - \sigma_q < \lambda_p$, then X is \mathcal{C}^1 linearizable.

4.3 Infinitesimal Conjugacy and Fiber Contractions

Recall from Section 2 that the near-identity map $h = \text{id} + \eta$ on \mathbf{R}^n conjugates the quasi-linear vector field $X = A + F$ to the linear vector field given by A if η satisfies the infinitesimal conjugacy equation

$$L_A \eta = F \circ (\text{id} + \eta).$$

In case the nonlinear vector field X is in \mathcal{P}_r , we will invert the Lie derivative operator L_A on a Banach space \mathcal{B} of continuous functions, defined on an open neighborhood Ω of the origin, that also satisfy a Hölder condition at the origin. The inverse G of L_A is used to obtain a fixed point equation,

$$\alpha = G(F \circ (\text{id} + \alpha)),$$

that can be solved by the contraction principle. Its unique fixed point η is a solution of the infinitesimal conjugacy equation and $h = \text{id} + \eta$ is the desired near-identity continuous linearizing transformation. To show that η is smooth, we will use fiber contraction (see the discussion following Theorem 3.3).

The candidates for the (continuous) derivative of η belong to the space \mathcal{H} of continuous functions from Ω to the bounded linear operators on \mathcal{B} . Moreover, the derivative of η , if it exists, satisfies the fixed point equation

$$\Psi = \mathcal{G}(DF \circ (\text{id} + \eta)(I + \Psi))$$

on \mathcal{H} , where \mathcal{G} is an integral operator that inverts the differential operator \mathcal{L}_A given by

$$\mathcal{L}_A \Psi(x) = \frac{d}{dt} e^{-tA} \Psi(e^{tA} x) e^{tA} \Big|_{t=0}. \quad (4.6)$$

For appropriately defined subsets $\mathcal{D} \subset \mathcal{B}$ and $\mathcal{J} \subset \mathcal{H}$, we will show that the bundle map $\Lambda : \mathcal{D} \times \mathcal{J} \rightarrow \mathcal{D} \times \mathcal{J}$ given by

$$\Lambda(\alpha, \Psi) = (G(F \circ (\text{id} + \alpha)), \mathcal{G}(DF \circ (\text{id} + \alpha)(I + \Psi)))$$

is the desired fiber contraction.

4.4 Sources and Sinks

Theorem 4.5 is an immediate consequence of Proposition 4.2 and the following result.

Theorem 4.8. *If $\mu \in (0, 1)$ and X is in \mathcal{P}_r , then X is linearizable at the origin by a $C^{1,\mu}$ near-identity diffeomorphism.*

The remainder of this section is devoted to the proof of Theorem 4.8

By performing a linear change of coordinates (if necessary), there is no loss of generality if we assume that the block matrices A_1, \dots, A_p on the diagonal of A are each in real Jordan canonical form. In this case, it is easy to see that there is a real valued function $t \mapsto Q(t)$, given by $Q(t) = C_Q(1 + |t|^{n-1})$ where C_Q is a constant, such that

$$|e^{tA_i}| \leq e^{-\lambda_i t} Q(t) \quad (4.7)$$

for each $i \in \{1, \dots, p\}$. Also, for each λ with

$$-\lambda_1 < -\lambda_2 < \dots < -\lambda_p < -\lambda < 0, \quad (4.8)$$

there is an adapted norm on \mathbf{R}^n such that

$$|e^{tA}x| \leq e^{-\lambda t}|x| \quad (4.9)$$

whenever $x \in \mathbf{R}^n$ and $t \geq 0$ (see, for example, [C99] for the standard construction of the adapted norm).

Unfortunately, the adapted norm is not necessarily natural with respect to the decomposition $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_p}$ of \mathbf{R}^n . The natural norm is the ℓ_1 -norm. For $i \in \{1, 2, \dots, p\}$, we will use the notation

$$|x|_i := \sum_{k=1}^i |x_k|. \quad (4.10)$$

In particular, $|\cdot|_p$ is a norm on \mathbf{R}^n that does respect the decomposition. It is also equivalent to the adapted norm; that is, there is a constant $K > 1$ such that

$$\frac{1}{K}|x|_p \leq |x| \leq K|x|_p. \quad (4.11)$$

Because A is block diagonal and in view of the ordering of the real parts of the eigenvalues in display (4.8), we have the useful estimate

$$\begin{aligned} |e^{tA}x|_i &= \sum_{k=1}^i |e^{tA_k}x_k| \\ &= \sum_{k=1}^i e^{-\lambda_k t} Q(t) |x_k| \\ &\leq e^{-\lambda_i t} Q(t) |x|_i. \end{aligned} \tag{4.12}$$

Recall that for $r > 0$, $\Omega_r := \{x \in \mathbf{R}^n : |x| < r\}$. Also, note that a function $\alpha : \Omega_r \rightarrow \mathbf{R}^n$ is given by $\alpha = (\alpha_1, \dots, \alpha_p)$ corresponding to the decomposition $x = (x_1, \dots, x_p)$.

For $r > 0$ and $0 < \mu < 1$, let $\mathcal{B}_{r,\mu}$ denote the space of all continuous functions from Ω_r to \mathbf{R}^n such that the norm

$$\|\alpha\|_{r,\mu} := \max_{i \in \{1, \dots, p\}} \sup_{0 < |x| < r} \frac{|\alpha_i(x)|}{|x|_i |x|^\mu}$$

is finite.

Proposition 4.9. *The set $\mathcal{B}_{r,\mu}$ endowed with the norm $\|\cdot\|_{r,\mu}$ is a Banach space.*

For $\alpha \in \mathcal{B}_{r,\mu}$, define

$$(G\alpha)(x) := - \int_0^\infty e^{-tA} \alpha(e^{tA}x) dt. \tag{4.13}$$

Note that the “natural” definition of G —in view of the definitions in Section 2—would be $\int_0^\infty e^{tA} \alpha(e^{-tA}x) dt$. Although this definition does lead to the existence of a linearizing homeomorphism, the homeomorphism thus obtained may not be smooth.

Proposition 4.10. *The function G is a bounded linear operator on $\mathcal{B}_{r,\mu}$; and, for fixed μ , the operator norms are uniformly bounded for $r > 0$. Moreover, $L_A G = I$ on $\mathcal{B}_{r,\mu}$, and $G L_A$, restricted to the domain of L_A on $\mathcal{B}_{r,\mu}$, is the identity.*

Proof. The k th component of $G\alpha$ is given by

$$(G\alpha)_i(x) = - \int_0^\infty e^{-tA_i} \alpha_i(e^{tA}x) dt.$$

Since we are using an adapted norm on \mathbf{R}^n , if $x \in \Omega_r$, then so is $e^{tA}x$. Using this fact, the definition of the space $\mathcal{B}_{r,\mu}$, and the inequalities (4.7) and (4.12), we have the estimate

$$\begin{aligned} |e^{-tA_i} \alpha_i(e^{tA}x)| &\leq e^{\lambda_i t} Q(t) \frac{|\alpha_i(e^{tA}x)|}{|e^{tA}x|_i |e^{tA}x|^\mu} |e^{tA}x|_i |e^{tA}x|^\mu \\ &\leq e^{-\lambda\mu t} Q^2(t) \|\alpha\|_{r,\mu} |x|_i |x|^\mu. \end{aligned}$$

Because $Q(t)$ has polynomial growth, there is a universal constant $c > 0$ such that

$$e^{-\lambda\mu t/2} Q^2(t) \leq c$$

whenever $t \geq 0$. Hence, it follows that

$$|e^{-tA_i} \alpha_i(e^{tA}x)| \leq c e^{-\lambda\mu t/2} \|\alpha\|_{r,\mu} |x|_i |x|^\mu.$$

By Lemma 2.2, the function $x \mapsto (G\alpha)_i(x)$ is continuous in Ω_r and clearly

$$\sup_{0 < |x| < r} \frac{|(G\alpha)_i(x)|}{|x|_i |x|^\mu} < \frac{2c}{\lambda\mu} \|\alpha\|_{r,\mu}.$$

The fundamental theorem of calculus and the properties of the Lie derivative are used to show that G is a right inverse of L_A and that GL_A is the identity operator on the domain of L_A . \square

Proposition 4.11. *If $\alpha \in \mathcal{B}_{r,\mu}$, then $F \circ (\text{id} + \alpha) \in \mathcal{B}_{r,\mu}$. Moreover, if $\epsilon > 0$ is given and $r > 0$ is sufficiently small, then the map $\alpha \mapsto F \circ (\text{id} + \alpha)$ restricted to the closed unit ball in $\mathcal{B}_{r,\mu}$ has range in the ball with radius ϵ centered at the origin.*

Proof. Clearly the function $F \circ (\text{id} + \alpha)$ is continuous in Ω_r . We will show first that this function is in $\mathcal{B}_{r,\mu}$.

The i th component of $F \circ (\text{id} + \alpha)$ is

$$[F \circ (\text{id} + \alpha)]_i = F_i \circ (\text{id} + \alpha)$$

Using property (3) in the definition of \mathcal{P}_r , the equivalence of norms, and the triangle law, we have the estimate

$$|F_i(x + \alpha(x))| \leq KM(|x|_p + |\alpha(x)|_p)(|x|_i + |\alpha(x)|_i),$$

and for $k \in \{1, 2, \dots, p\}$ we have the inequality

$$\begin{aligned} |\alpha(x)|_k &\leq \|\alpha\|_{r,\mu} |x|^\mu \sum_{\ell=1}^k |x|_\ell \\ &\leq p \|\alpha\|_{r,\mu} |x|^\mu |x|_k. \end{aligned} \tag{4.14}$$

By combining these estimates and restricting x to lie in Ω_r where $|x| < r$, it follows that

$$|F_i(x + \alpha(x))| \leq MK^2(1 + p\|\alpha\|_{r,\mu}r^\mu)^2|x||x|_i;$$

and therefore,

$$\|F \circ (\text{id} + \alpha)\|_{r,\mu} \leq MK^2(1 + p\|\alpha\|_{r,\mu}r^\mu)^2r^{1-\mu}.$$

This proves the first statement of the proposition. The second statement of the proposition follows from this norm estimate because $0 < \mu < 1$. \square

The *special Lipschitz number* for $\alpha \in \mathcal{B}_{r,\mu}$ is defined as follows:

$$\text{sLip}(\alpha) := \max_{i \in \{1, 2, \dots, p\}} \sup \left\{ \frac{|\alpha(x) - \alpha(y)|_i}{|x - y|_i} : x, y \in \Omega_r; x \neq y \right\}.$$

Also, let \mathcal{D} denote the set of all α in $\mathcal{B}_{r,\mu}$ such that $\|\alpha\|_{r,\mu} \leq 1$ and $\text{sLip}(\alpha) \leq 1$.

Proposition 4.12. *The set \mathcal{D} is a complete metric subspace of $\mathcal{B}_{r,\mu}$.*

Proof. Suppose that $\{\alpha^m\}_{m=1}^\infty$ is a sequence in \mathcal{D} that converges to α in $\mathcal{B}_{r,\mu}$. To show $\text{sLip}(\alpha) \leq 1$, use the inequality (4.14) to obtain the estimate

$$\begin{aligned} |\alpha(x) - \alpha(y)|_i &\leq |\alpha(x) - \alpha^m(x)|_i + |\alpha^m(x) - \alpha^m(y)|_i + |\alpha^m(y) - \alpha(y)|_i \\ &\leq 2p\|\alpha - \alpha^m\|_{r,\mu}r^{1+\mu} + \text{sLip}(\alpha^m)|x - y|_i \end{aligned}$$

and pass to the limit as $m \rightarrow \infty$. \square

Proposition 4.13. *If $r > 0$ is sufficiently small, then the function*

$$\alpha \mapsto G(F \circ (\text{id} + \alpha)) \tag{4.15}$$

is a contraction on \mathcal{D} .

Proof. Let $R := 1/\|G\|_{r,\mu}$ and suppose that $\|\alpha\|_{r,\mu} \leq 1$. By Proposition 4.11, if $r > 0$ sufficiently small, then

$$\|G(F \circ (\text{id} + \alpha))\|_{r,\mu} \leq \|G\|_{r,\mu} \|F \circ (\text{id} + \alpha)\|_{r,\mu} \leq 1.$$

Hence, the closed unit ball in $\mathcal{B}_{r,\mu}$ is an invariant set for the map $\alpha \mapsto G(F \circ (\text{id} + \alpha))$.

To prove that \mathcal{D} is an invariant set, we will show the following proposition: If $\alpha \in \mathcal{D}$ and $r > 0$ is sufficiently small, then $\text{sLip}(G(F \circ (\text{id} + \alpha))) < 1$.

Start with the basic inequality

$$\begin{aligned} |(G(F \circ (\text{id} + \alpha)))_i(x) - (G(F \circ (\text{id} + \alpha)))_i(y)| \leq \\ \int_0^\infty e^{\lambda_i t} Q(t) |F_i(e^{tA}x + \alpha(e^{tA}x)) - F_i(e^{tA}y + \alpha(e^{tA}y))| dt, \end{aligned}$$

and then use property (3) in the definition of \mathcal{P}_r to estimate the third factor of the integrand. Note that the resulting estimate has two terms, each of the form $|\cdot|_i$. After making the obvious triangle law estimates using the linearity of e^{tA} , the inequality $|\alpha(x) - \alpha(y)|_i \leq |x - y|_i$, and the inequality (4.12); it is easy to see that the first factor is majorized by a bounded multiple of $e^{-\lambda t}$ and the second factor is majorized by a bounded multiple of $e^{-\lambda_i t} Q(t)$. One of the multipliers is bounded above by a constant multiple of r ; the other is bounded above by a constant multiple of $|x - y|$. The integral converges because its integrand is thus majorized by a constant (in t) multiple of $e^{-\lambda t} Q(t)$. In fact, there is a constant $c > 0$ such that

$$|(G(F \circ (\text{id} + \alpha)))(x) - (G(F \circ (\text{id} + \alpha)))(y)| \leq cr|x - y|;$$

and therefore, if $r > 0$ is sufficiently small, then $\text{sLip}(G(F \circ (\text{id} + \alpha))) < 1$, as required.

We have just established that the complete metric space \mathcal{D} is an invariant set for the map $\alpha \mapsto G(F \circ (\text{id} + \alpha))$. To complete the proof, we will show that this map is a contraction on \mathcal{D} .

Fix α and β such that $\|\alpha\|_{r,\mu} \leq 1$ and $\|\beta\|_{r,\mu} \leq 1$, and note that

$$\begin{aligned} & \|G(F \circ (\text{id} + \alpha)) - G(F \circ (\text{id} + \beta))\|_{r,\mu} \leq \\ & \|G\|_{r,\mu} \|F \circ (\text{id} + \alpha) - F \circ (\text{id} + \beta)\|_{r,\mu}. \end{aligned} \quad (4.16)$$

The i th component function of the function

$$x \mapsto F \circ (\text{id} + \alpha) - F \circ (\text{id} + \beta)$$

is given by

$$C_i := F_i(x + \alpha(x)) - F_i(x + \beta(x)).$$

Using the inequality (4.14) and property (3) of the definition of \mathcal{P}_r , we have the estimate

$$\begin{aligned} |C_i| & \leq MK(|x + \alpha(x)|_p |\alpha(x) - \beta(x)|_i + |\alpha(x) - \beta(x)|_p |x + \beta(x)|_i) \\ & \leq KM((|x|_p + p\|\alpha\|_{r,\mu} r^\mu |x|_p) p \|\alpha - \beta\|_{r,\mu} |x|^\mu |x|_i \\ & \quad + p\|\alpha - \beta\|_{r,\mu} |x|^\mu |x|_p (|x|_i + p\|\alpha\|_{r,\mu} r^\mu |x|_i)) \\ & \leq 2K^2 Mp(1 + pr^\mu) r \|\alpha - \beta\|_{r,\mu} |x|^\mu |x|_i. \end{aligned}$$

Hence, there is a constant $\bar{M} > 0$ such that

$$\|F(x + \alpha(x)) - F(x + \beta(x))\|_{r,\mu} \leq \bar{M} r \|\alpha - \beta\|_{r,\mu}.$$

Using the inequality (4.16) and Proposition 4.10, it follows that if $r > 0$ is sufficiently small, then the map $\alpha \mapsto G(F \circ (\text{id} + \alpha))$ is a contraction on the closed unit ball in $\mathcal{B}_{r,\mu}$. \square

We will prove that the unique fixed point η of the contraction (4.15) is $\mathcal{C}^{1,\mu}$ for all $\mu \in (0, 1)$.

Let $\mathcal{H}_{r,\mu}$ denote the space of all continuous maps $\Psi : \Omega_r \rightarrow \mathbf{L}(\mathbf{R}^n, \mathbf{R}^n)$ such that, for $j \leq i$,

$$\sup_{0 < |x| < r} \frac{|\Psi_{ij}(x)|}{|x|^\mu} < \infty$$

and, for $j > i$,

$$\sup_{0 < |x| < r} \frac{|\Psi_{ij}(x)|}{|x|_i^\mu} < \infty$$

where the subscripts refer to the components of the matrix valued function Ψ with respect to the decomposition $\mathbf{R}^n = \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_p}$. Also, the norm $\|\Psi\|_\mu$ of $\Psi \in \mathcal{H}_{r,\mu}$ is defined to be the maximum of these suprema.

Proposition 4.14. *The space $\mathcal{H}_{r,\mu}$ endowed with the norm $\|\cdot\|_\mu$ is a Banach space.*

For $\Psi \in \mathcal{H}_{r,\mu}$, define

$$(\mathcal{G}\Psi)(x) := - \int_0^\infty e^{-tA} \Psi(e^{tA}x) e^{tA} dt. \quad (4.17)$$

Proposition 4.15. *If $\mu < 1$ is sufficiently large, then the operator \mathcal{G} is a bounded linear operator on $\mathcal{H}_{r,\mu}$. Also, $\|\mathcal{G}\|_\mu$ is uniformly bounded with respect to r .*

Proof. In view of the inequality (4.7) and for $x \in \Omega_r$, the ij -component of the integrand in the definition of \mathcal{G} is bounded above as follows:

$$|e^{-tA_i} \Psi_{ij}(e^{tA}x) e^{tA_j}| \leq e^{\lambda_i t} Q(t) |\Psi_{ij}(e^{tA}x)| e^{-\lambda_j t} Q(t).$$

For $j \leq i$, we have the inequality

$$\begin{aligned} |e^{-tA_i} \Psi_{ij}(e^{tA}x) e^{tA_j}| &\leq e^{(\lambda_i - \lambda_j)t} Q^2(t) \|\Psi\|_\mu |e^{tA}x|^\mu \\ &\leq e^{(\lambda_i - \lambda_j - \mu\lambda)t} Q^3(t) \|\Psi\|_\mu |x|^\mu \\ &\leq e^{-\mu\lambda t} Q^3(t) \|\Psi\|_\mu |x|^\mu; \end{aligned}$$

and for $j > i$, by using the estimate (4.12), we have

$$\begin{aligned} |e^{-tA_i} \Psi_{ij}(e^{tA}x) e^{tA_j}| &\leq e^{(\lambda_i - \lambda_j)t} Q^2(t) \|\Psi\|_\mu |e^{tA}x|_i^\mu \\ &\leq e^{((1-\mu)\lambda_i - \lambda_j)t} Q^3(t) \|\Psi\|_\mu |x|_i^\mu. \end{aligned}$$

Because Q has polynomial growth, if $\mu < 1$ is sufficiently large, then the integrals

$$\int_0^\infty e^{((1-\mu)\lambda_i - \lambda_j)t} Q^3(t) dt, \quad \int_0^\infty e^{-\lambda_j t} Q^3(t) dt$$

both converge. By the definition of the norm, $\|\mathcal{G}\|_\mu$ is bounded by a constant that does not depend on r . \square

The hypothesis of Proposition 4.15 is the first instance where $\mu < 1$ is required to be sufficiently large. This restriction is compatible with the conclusion of Theorem 4.8. Indeed, if a function is Hölder on a bounded set with Hölder exponent μ , then it is Hölder with exponent ν whenever $0 < \nu \leq \mu$.

A map $\Psi \in \mathcal{H}_{r,\mu}$ is called *special μ -Hölder* if, for all $i, j \in \{1, 2, \dots, p\}$,

$$|\Psi_{ij}(x) - \Psi_{ij}(y)| \leq |x - y|^\mu \quad (4.18)$$

and, for all $j > i$,

$$|\Psi_{ij}(x) - \Psi_{ij}(y)| \leq |x - y|_i^\mu + |x - y|^\mu(|x|_i^\mu + |y|_i^\mu) \quad (4.19)$$

whenever $x, y \in \Omega_r$.

Let \mathcal{J} denote the subset of $\mathcal{H}_{r,\mu}$ consisting of those functions in $\mathcal{H}_{r,\mu}$ such that Ψ is special μ -Hölder and $\|\Psi\|_\mu \leq 1$. Also, for $\alpha \in \mathcal{D}$ and $\Psi \in \mathcal{J}$, let

$$\begin{aligned} \Gamma(\alpha) &:= G(F \circ (\text{id} + \alpha)), \\ \Upsilon(\alpha, \Psi) &:= DF \circ (\text{id} + \alpha)(I + \Psi), \\ \Delta(\alpha, \Psi) &:= \mathcal{G}\Upsilon(\alpha, \Psi), \\ \Lambda(\alpha, \Psi) &:= (\Gamma(\alpha), \Delta(\alpha, \Psi)). \end{aligned} \quad (4.20)$$

Proposition 4.16. *The set \mathcal{J} is a complete metric subspace of $\mathcal{H}_{r,\mu}$.*

Proof. It suffices to show that \mathcal{J} is closed in $\mathcal{H}_{r,\mu}$. The proof of this fact is similar to the proof of Proposition 4.12. \square

Proposition 4.17. *If $r > 0$ is sufficiently small and $\mu < 1$ is sufficiently large, then the bundle map Λ defined in display (4.20) is a fiber contraction on $\mathcal{D} \times \mathcal{J}$.*

Proof. We will show first that there is a constant C such that

$$\|DF \circ (\text{id} + \alpha)(I + \Psi)\|_\mu \leq Cr^{1-\mu}$$

for all $\alpha \in \mathcal{D}$ and $\Psi \in \mathcal{J}$.

Using the properties listed in the definition of \mathcal{P}_r , note that if $j \leq i$, then

$$\begin{aligned} |D_j F_i(x + \alpha(x))| &\leq M|x + \alpha(x)| \\ &\leq MK(|x|_p + |\alpha(x)|_p) \\ &\leq MK^2(1 + p\|\alpha\|_{r,\mu}r^\mu)|x| \\ &\leq MK^2(1 + pr^\mu)r^{1-\mu}|x|^\mu, \end{aligned} \quad (4.21)$$

and if $j > i$, then

$$\begin{aligned} |D_j F_i(x + \alpha(x))| &\leq MK(|x|_i + |\alpha(x)|_i) \\ &\leq MK^{1+\mu}(1 + pr^\mu)r^{1-\mu}|x|_i^\mu. \end{aligned} \quad (4.22)$$

It follows that there is a constant $c_1 > 0$ such that

$$\|DF \circ (\text{id} + \alpha)\|_\mu \leq c_1 r^{1-\mu}. \quad (4.23)$$

Suppose that Φ and Ψ are in $\mathcal{H}_{r,\mu}$. We will show that there is a constant $c_2 > 0$ such that

$$\|\Phi\Psi\|_\mu \leq c_2 r^\mu \|\Phi\|_\mu \|\Psi\|_\mu. \quad (4.24)$$

First, note that

$$|(\Phi\Psi)_{ij}| \leq \sum_{k=1}^p |\Phi_{ik}| |\Psi_{kj}|.$$

There is a constant \bar{c} such that

$$|\Phi_{ij}| \leq \bar{c} \|\Phi\|_\mu |x|^\mu \quad (4.25)$$

for all $i, j \in \{1, 2, \dots, p\}$. In fact, for $j \leq i$, this estimate is immediate from the definition of the norm; for $j > i$, it is a consequence of the inequality

$$|\Phi_{ij}| \leq \|\Phi\|_\mu |x|_i^\mu \leq K^\mu \|\Phi\|_\mu |x|_i^\mu.$$

Using estimate (4.25), it follows that

$$\begin{aligned} |(\Phi\Psi)_{ij}| &\leq \sum_{k=1}^p \bar{c}^2 \|\Phi\|_\mu \|\Psi\|_\mu |x|^{2\mu} \\ &\leq (p\bar{c}^2 |r|^\mu) \|\Phi\|_\mu \|\Psi\|_\mu |x|^\mu \end{aligned}$$

whenever $j \leq i$, and

$$\begin{aligned} |(\Phi\Psi)_{ij}| &\leq \sum_{k=1}^i \bar{c}^2 \|\Phi\|_\mu |x|^\mu \|\Psi\|_\mu |x|_i^\mu + \sum_{k=i+1}^p \|\Phi\|_\mu |x|_i^\mu \bar{c} \|\Psi\|_\mu |x|^\mu \\ &\leq (p\bar{c} |r|^\mu) \|\Phi\|_\mu \|\Psi\|_\mu |x|_i^\mu \end{aligned}$$

whenever $j > i$. This completes the proof of estimate (4.24).

It is now clear that there is a constant $C > 0$ such that

$$\|DF \circ (\text{id} + \alpha)(I + \Psi)\|_\mu \leq \|DF \circ (\text{id} + \alpha)\|_\mu + \|DF \circ (\text{id} + \alpha)\Psi\|_\mu \leq Cr^{1-\mu} \quad (4.26)$$

for all $\alpha \in \mathcal{D}$ and $\Psi \in \mathcal{J}$. Hence, if r is sufficiently small, then

$$\|\Delta(\alpha, \Psi)\|_\mu \leq \|\mathcal{G}\|_\mu \|DF \circ (\text{id} + \alpha)(I + \Psi)\|_\mu \leq 1$$

for all $\alpha \in \mathcal{D}$ and $\Psi \in \mathcal{J}$.

To complete the proof that $\mathcal{J} \times \mathcal{D}$ is an invariant set for Λ , we will prove the following proposition: If $r > 0$ is sufficiently small, then $\Delta(\alpha, \Psi)$ is special μ -Hölder whenever $\alpha \in \mathcal{D}$ and $\Psi \in \mathcal{J}$.

Recall from display (4.20) that

$$\Upsilon(\alpha, \Psi) := DF \circ (\text{id} + \alpha)(I + \Psi).$$

We will use the following uniform estimates: There is a constant $c > 0$ such that

$$|\Upsilon(\alpha, \Psi)_{ij}(x) - \Upsilon(\alpha, \Psi)_{ij}(y)| \leq cr^\mu |x - y|^\mu \quad (4.27)$$

and, for all $j > i$,

$$|\Upsilon(\alpha, \Psi)_{ij}(x) - \Upsilon(\alpha, \Psi)_{ij}(y)| \leq c(r^\mu + r^{1-\mu})|x - y|_i^\mu + |x - y|^\mu(|x|_i^\mu + |y|_i^\mu) \quad (4.28)$$

whenever $0 < r < 1$, $x, y \in \Omega_r$, $\alpha \in \mathcal{D}$, and $\Psi \in \mathcal{J}$.

To prove the inequalities (4.27) and (4.28) we will show (the key observation) that for $\alpha \in \mathcal{D}$ there are Lipschitz (hence Hölder) estimates for $DF \circ (\text{id} + \alpha)$. In fact, using property (4) in the definition of \mathcal{P}_r , there is a constant \bar{M} such that

$$\begin{aligned} |D_j F_i(x + \alpha(x)) - D_j F_i(y + \alpha(y))| &\leq M(|x - y| + |\alpha(x) - \alpha(y)|) \\ &\leq M(|x - y| + K|\alpha(x) - \alpha(y)|_p) \\ &\leq M(|x - y| + K^2 \text{sLip}(\alpha)|x - y|) \\ &\leq M(1 + K^2)|x - y| \\ &\leq M(1 + K^2)|x - y|^{1-\mu}|x - y|^\mu \\ &\leq M(1 + K^2)2^{1-\mu}r^{1-\mu}|x - y|^\mu \\ &\leq \bar{M}r^{1-\mu}|x - y|^\mu \end{aligned} \quad (4.29)$$

for all $i, j \in \{1, 2, \dots, p\}$ and $x, y \in \Omega_r$. Using property (5) in the definition of \mathcal{P}_r and the special Lipschitz estimates for α , we have the following similar result for $j > i$:

$$|D_j F_i(x + \alpha(x)) - D_j F_i(y + \alpha(y))| \leq \bar{M}r^{1-\mu}(|x - y|_i^\mu + |x - y|^\mu(|x|_i^\mu + |y|_i^\mu)). \quad (4.30)$$

We have just obtained “special Hölder” estimates the first summand in the representation

$$\Upsilon(\alpha, \Psi) = DF \circ (\text{id} + \alpha) + DF \circ (\text{id} + \alpha)\Psi;$$

to obtain estimates for the second summand, and hence for $\Upsilon(\alpha, \Psi)$, let $\Phi := DF \circ (\text{id} + \alpha)$ and note that

$$\begin{aligned} |(\Phi\Psi)_{ij}(x) - (\Phi\Psi)_{ij}(y)| &\leq \sum_{k=1}^p |\Phi_{ik}(x)\Psi_{kj}(x) - \Phi_{ik}(y)\Psi_{kj}(y)| \\ &\leq \sum_{k=1}^p \Xi_k \end{aligned}$$

where

$$\Xi_k := |\Phi_{ik}(x)| |\Psi_{kj}(x) - \Psi_{kj}(y)| + |\Phi_{ik}(x) - \Phi_{ik}(y)| |\Psi_{kj}(y)|.$$

Because Ψ and Φ are in \mathcal{J} and $0 < r < 1$, there is a constant $\bar{c} > 0$ such that

$$\begin{aligned} \Xi_k &\leq \|\Phi\|_\mu |x|^\mu |x - y|^\mu + \bar{M} r^{1-\mu} |x - y|^\mu \|\Psi\|_\mu |y|^\mu \\ &\leq (1 + \bar{M}) |r|^\mu |x - y|^\mu \\ &\leq \bar{c} |r|^\mu |x - y|^\mu. \end{aligned} \tag{4.31}$$

The desired inequality (4.27) is obtained by summing over k and adding the result to the estimate (4.29).

Suppose that $j > i$. If $k \leq i$, then $j > k$ and

$$\begin{aligned} \Xi_k &\leq \|\Phi\|_\mu |x|^\mu (|x - y|_k^\mu + |x - y|^\mu (|x|_k^\mu + |y|_k^\mu)) + \bar{M} r^{1-\mu} |x - y|^\mu \|\Psi\|_\mu |y|_k^\mu \\ &\leq |r|^\mu |x - y|_k^\mu + |r|^\mu |x - y|^\mu (|x|_k^\mu + |y|_k^\mu) + \bar{M} r^{1-\mu} |x - y|^\mu (|x|_k^\mu + |y|_k^\mu) \\ &\leq (1 + \bar{M}) (r^\mu + r^{1-\mu}) (|x - y|_k^\mu + |x - y|^\mu (|x|_k^\mu + |y|_k^\mu)) \\ &\leq (1 + \bar{M}) (r^\mu + r^{1-\mu}) (|x - y|_i^\mu + |x - y|^\mu (|x|_i^\mu + |y|_i^\mu)); \end{aligned}$$

and if $k > i$, then

$$\begin{aligned} \Xi_k &\leq \bar{M} r^{1-\mu} (|x|_i^\mu + |x|^\mu |x|_i^\mu) |x - y|^\mu \\ &\quad + \bar{M} r^{1-\mu} (|x - y|_i^\mu + |x - y|^\mu (|x|_i^\mu + |y|_i^\mu)) |y|^\mu \\ &\leq \bar{M} |r|^{1-\mu} (|x - y|_i^\mu + (1 + r^\mu) |x - y|^\mu (|x|_i^\mu + |y|_i^\mu)) \\ &\quad + r^\mu r^{1-\mu} |x - y|^\mu (|x|_i^\mu + |y|_i^\mu) \\ &\leq 2\bar{M} r^{1-\mu} (|x - y|_i^\mu + |x - y|^\mu (|x|_i^\mu + |y|_i^\mu)). \end{aligned}$$

The desired inequality (4.28) is obtained by summing over k and adding the result to the estimate (4.30).

Note that

$$\begin{aligned}
& |(\mathcal{G}\Upsilon(\alpha, \Psi))_{ij}(x) - (\mathcal{G}\Upsilon(\alpha, \Psi))_{ij}(y)| \\
& \leq \int_0^\infty |e^{-tA_i}| |\Upsilon(\alpha, \Psi)_{ij}(e^{tA}x) - \Upsilon(\alpha, \Psi)_{ij}(e^{tA}y)| |e^{tA_j}| dt \\
& \leq \int_0^\infty e^{(\lambda_i - \lambda_j)t} Q^2(t) |\Upsilon(\alpha, \Psi)_{ij}(e^{tA}x) - \Upsilon(\alpha, \Psi)_{ij}(e^{tA}y)| dt.
\end{aligned}$$

Using the estimates (4.27) and (4.9), for $j \leq i$ we have

$$|(\mathcal{G}\Upsilon(\alpha, \Psi)_{ij})(x) - (\mathcal{G}\Upsilon(\alpha, \Psi)_{ij})(y)| \leq (cr^{1-\mu} \int_0^\infty e^{(\lambda_i - \lambda_j - \mu\lambda)t} Q^2(t) dt) |x - y|^\mu$$

with $\lambda_i - \lambda_j - \mu\lambda < 0$. On the other hand, using inequality (4.28) and (4.12) for the case $j > i$, it follows that

$$|(\mathcal{G}\Upsilon(\alpha, \Psi)_{ij})(x) - (\mathcal{G}\Upsilon(\alpha, \Psi)_{ij})(y)| \leq \bar{M}(|x - y|_i^\mu + |x - y|^\mu(|x|_i^\mu + |y|_i^\mu))$$

where

$$\bar{M} := c(r^{1-\mu} + r^\mu) \int_0^\infty e^{((1-\mu)\lambda_i - \lambda_j)t} Q^{2+\mu}(t) dt$$

Hence, if $\mu < 1$ is sufficiently large, then $\mathcal{G}\Psi_{ij}$ satisfies the inequality (4.19); and because

$$|x - y|_i^\mu + |x - y|^\mu(|x|_i^\mu + |y|_i^\mu) \leq (1 + 2r^\mu)|x - y|^\mu,$$

the previous estimate also shows that $\mathcal{G}\Psi_{ij}$ is μ -Hölder. This completes the proof that $\mathcal{D} \times \mathcal{J}$ is an invariant set for the fiber contraction.

To show that the function $\Psi \mapsto \Delta(\alpha, \Psi)$ is a uniform contraction, use the linearity of \mathcal{G} together with inequalities (4.23) and (4.24) to obtain the estimate

$$\begin{aligned}
\|\Delta(\alpha, \Psi_1) - \Delta(\alpha, \Psi_2)\|_\mu & \leq \|\mathcal{G}\|_\mu \|DF \circ (\text{id} + \alpha)(\Psi_1 - \Psi_2)\|_\nu \\
& \leq \|\mathcal{G}\|_\mu cr^\mu \|DF \circ (\text{id} + \alpha)\|_\nu \|(\Psi_1 - \Psi_2)\|_\mu \\
& \leq \|\mathcal{G}\|_\mu c^2 r \|(\Psi_1 - \Psi_2)\|_\mu.
\end{aligned}$$

Hence, if $r > 0$ is sufficiently small, then Δ is a uniform contraction; and therefore, Λ is a fiber contraction on $\mathcal{D} \times \mathcal{H}_{r,\mu}$. \square

Proposition 4.18. *If $\alpha \in \mathcal{D}$, $D\alpha \in \mathcal{J}$, then $D(G(F \circ (\text{id} + \alpha))) \in \mathcal{J}$ and*

$$D(G(F \circ (\text{id} + \alpha))) = \mathcal{G}(DF \circ (\text{id} + \alpha)(I + D\alpha)).$$

Proof. Let

$$(G(F \circ (\text{id} + \alpha)))(x) := - \int_0^\infty e^{-tA} F(e^{tA}x + \alpha(e^{tA}x)) dt.$$

Since $D\alpha$ exists, the integrand is differentiable. Moreover, the derivative of the integrand has the form $e^{-tA}\Psi(e^{tA}x)e^{tA}$ where, by the estimate (4.26),

$$\Psi := DF \circ (\text{id} + \alpha)(I + D\alpha)$$

is in \mathcal{J} . Using the same estimates as in Proposition 4.15, it follows that the derivative of the original integrand is majorized by an integrable function. The result now follows from an application of Lemma 2.2 and the definition of \mathcal{G} . \square

We are now ready to prove Theorem 4.8

Proof. By Proposition 4.17, Γ is a fiber contraction.

Choose a function $\alpha_0 \in \mathcal{D}$ such that $\Psi_0 := D\alpha_0 \in \mathcal{J}$ —for example, take $\alpha_0 = 0$, and consider the sequence in $\mathcal{D} \times \mathcal{J}$ given by the forward Λ -orbit of (α_0, Ψ_0) , namely, the sequence $\{(\alpha_k, \Psi_k)\}_{k=0}^\infty$ where

$$\alpha_k := \Gamma(\alpha_{k-1}), \quad \Psi_k := \Delta(\alpha_{k-1}, \Psi_{k-1}).$$

We will prove, by induction, that $\Psi_k = D\alpha_k$.

By definition,

$$\alpha_k = G(F \circ (\text{id} + \alpha_{k-1}))$$

Also, by the induction hypothesis, $\Psi_{k-1} = D\alpha_{k-1}$. Because $\alpha_{k-1} \in \mathcal{D}$ and $D\alpha_{k-1} \in \mathcal{J}$, by an application of Proposition 4.18 we have that

$$\begin{aligned} D\alpha_k &= \mathcal{G}DF \circ (\text{id} + \alpha_{k-1})(I + \Psi_{k-1}) \\ &= \Psi_k, \end{aligned}$$

as required.

By an application of the fiber contraction theorem, if η is the fixed point of Γ and Φ is the fixed point of the map $\Psi \rightarrow \Delta(\eta, \Psi)$, then

$$\lim_{k \rightarrow \infty} \alpha_k = \eta, \quad \lim_{k \rightarrow \infty} D\alpha_k = \Phi$$

where the limits exist in the respective spaces \mathcal{D} and \mathcal{J} .

The following lemma will be used to finish the proof.

Lemma 4.19. *If a sequence converges in either of the spaces $\mathcal{B}_{r,\mu}$ or $\mathcal{H}_{r,\mu}$, then the sequence converges uniformly.*

To prove the lemma, recall that the functions in the spaces $\mathcal{B}_{r,\mu}$ and $\mathcal{H}_{r,\mu}$ are continuous functions defined on Ω_r , the ball of radius r at the origin in \mathbf{R}^n with respect to the adapted norm. Also, by the equivalence of the norms (see display (4.11)) there is a positive constant K such that $|x|_i < K|x|$.

If $\lim_{k \rightarrow \infty} \alpha_k = \alpha$ in $\mathcal{B}_{r,\mu}$, then for each $\epsilon > 0$ there is an integer $\kappa > 0$ such that

$$\frac{|(\alpha_k)_i(x) - \alpha_i(x)|}{|x|_i|x|^\mu} < \frac{\epsilon}{Kr^{1+\mu}}$$

whenever $0 < |x| < r$, $k \geq \kappa$, and $i \in \{1, \dots, p\}$. Using the inequality $|x|_i < K|x|$ and the norm equivalence (4.11), it follows that

$$\|\alpha_k - \alpha\| < \epsilon$$

whenever $0 < |x| < 1$ and $k \geq \kappa$; that is, the sequence of continuous functions $\{\alpha_k\}_{k=0}^\infty$ converges uniformly to α . The proof of the uniform convergence of a convergent sequence in $\mathcal{H}_{r,\mu}$ is similar.

As mentioned previously, the equality $\Psi = D\eta$ follows from the uniform convergence and a standard result in advanced calculus on the differentiability of the limit of a uniformly convergent sequence of functions. Thus, the conjugating homeomorphism $h = \text{id} + \eta$ is continuously differentiable. Moreover, using the equality $Dh(0) = I$ and the inverse function theorem, the conjugacy h is a diffeomorphism when restricted to a sufficiently small open ball at the origin. \square

References

- [C99] C. Chicone, *Ordinary Differential Equations with Applications*, New York: Springer-Verlag, 1999
- [CL88] S-N. Chow and K. Lu, \mathcal{C}^k centre unstable manifolds, *Proc. Roy. Soc. Edinburgh*, **108A** (1988) 303–320.
- [CLL91] S-N. Chow, X-B. Lin, and K. Lu, Smooth invariant foliations in infinite-dimensional spaces, *J. Diff. Eqs.*, **94** (2) (1988) 266–291.
- [D71] J. R. Dorroh, Local groups of differentiable transformations, *Math. Ann.*, **192** (1971) 243–249.

- [DS58] N. Dunford and J. Schwartz, *Linear Operators*, Vol. 1, New York: Interscience Pub. 1958.
- [G59] D. Grobman, Homeomorphisms of systems of differential equations, *Dokl. Akad. Nauk., SSSR* **128** (1959) 880–881.
- [H60] P. Hartman, A lemma in the theory of structural stability of differential equations, *Proc. Amer. Math. Soc.* **11** (1960) 610–620.
- [H60M] P. Hartman, On local homeomorphisms of Euclidean space, *Bol. Soc. Mat. Mexicana* **5** (2) (1960) 220–241.
- [H63] P. Hartman, On the local linearization of differential equations, *Proc. Amer. Math. Soc.* **14** (1963) 568–573.
- [HP70] M. Hirsch and C. Pugh, Stable manifolds and hyperbolic sets, In *Global Analysis XIV*, Amer. Math. Soc. (1970) 133–164.
- [KP90] U. Kirchgraber and K. Palmer, *Geometry in the Neighborhood of Invariant Manifolds of Maps and Flows and Linearization*, New York: Wiley 1990.
- [LL99] Yu. Latushkin and B. Layton, The optimal gap condition for invariant manifolds, *Disc. Cont. Dy. Sys.* **5** (2) (1999) 233–268.
- [Se85] G. Sell, Smooth linearization near a fixed point, *Amer. J. Math.*, **107**(1985) 1035–1091.
- [St57] S. Sternberg, Local contractions and a theorem of Poincaré, *Amer. J. Math.*, **79**(1957), 809–824.
- [St89] D. Stowe, Linearization in two dimensions, *J. Diff. Eqs.*, **63** (1989) 183–226.
- [T99] B. Tan, σ -Hölder continuous linearizations near hyperbolic fixed points in \mathbf{R}^n , Preprint, Georgia Tech., 1999.