Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 245 (2008) 411–441
Journal of Differential Equations
www.elsevier.com/locate/jde

Exchange lemmas 2: General Exchange Lemma ☆

Stephen Schecter

Mathematics Department, North Carolina State University, Box 8205, Raleigh, NC 27695, USA

Received 9 July 2007; revised 4 October 2007
Available online 26 November 2007

Abstract

Exchange lemmas are used in geometric singular perturbation theory to track flows near normally hyperbolic invariant manifolds. We prove a General Exchange Lemma, and show that it implies versions of existing exchange lemmas for rectifiable slow flows and loss-of-stability turning points.

MSC: 34E15; 34C30

Keywords: Geometric singular perturbation theory; Normally hyperbolic invariant manifold

1. Introduction

This paper is the second in a series of three; the others are [11] and [12]. An introduction to the series is in [11]. In this paper, we state and prove a General Exchange Lemma, and show that it implies versions of existing exchange lemmas for rectifiable slow flows and loss-of-stability turning points.

We begin in Section 2 by reviewing exchange lemmas for rectifiable slow flows and loss-of-stability turning points. In Section 3 we state the General Exchange Lemma. In Section 4 we show that it implies versions of existing exchange lemmas. We then state in Section 5 a version of the Implicit Function Theorem that is useful in proving the General Exchange Lemma. The proof of the General Exchange Lemma is given in Section 6. In addition to the Implicit Function Theorem, the proof uses the generalization of Deng's lemma that was proved in [11].

DOI of related article: 10.1016/j.jde.2007.08.011.
☆ This work was supported in part by the National Science Foundation under grant DMS-0406016.
E-mail address: schecter@math.ncsu.edu.

0022-0396/$ – see front matter
doi:10.1016/j.jde.2007.10.021

In the third paper in this series [12], we shall use the General Exchange Lemma to prove an exchange lemma for gain-of-stability turning points.

2. Exchange lemmas

2.1. Slow–fast systems [4,5]

A slow–fast system has the form

$$\dot{a} = f(a, b, \epsilon), \tag{2.1}$$

$$\dot{b} = \epsilon g(a, b, \epsilon), \tag{2.2}$$

with $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $\epsilon \geqslant 0$ a small parameter. The variable a is fast; the variable b is slow. The dot represents derivative with respect to t, the fast time.

Let $\tau = \epsilon t$, the slow time. Using prime to denote derivative with respect to τ, the system (2.1)–(2.2) becomes

$$\epsilon a' = f(a, b, \epsilon), \tag{2.3}$$

$$b' = g(a, b, \epsilon). \tag{2.4}$$

System (2.1)–(2.2) is the fast form of the slow–fast system; system (2.3)–(2.4) is the slow form.

The fast subsystem is obtained by setting $\epsilon = 0$ in (2.1)–(2.2):

$$\dot{a} = f(a, b, 0), \tag{2.5}$$

$$\dot{b} = 0. \tag{2.6}$$

The slow subsystem is obtained by setting $\epsilon = 0$ in (2.3)–(2.4):

$$0 = f(a, b, 0), \tag{2.7}$$

$$b' = g(a, b, 0). \tag{2.8}$$

The fast subsystem (2.5)–(2.6) can be viewed as a parameterized family of differential equations on $\mathbb{R}^n$. Its equilibria are pairs (a, b) such that $f(a, b, 0) = 0$. Suppose there is a manifold E_0 of such equilibria parameterized by b, each a hyperbolic equilibrium for the differential equation (2.5) with b fixed. More precisely, suppose:

(SF1) There is an open set V in $\mathbb{R}^m$, and a smooth function $\check{a}(b)$ defined on V, such that
- (a) for all $b \in V$, $f(\check{a}(b), b, 0) = 0$, and
- (b) there are numbers $\lambda_0 < 0 < \mu_0$ such that for all $b \in V$, $D_a f(\check{a}(b), b, 0)$ has k eigenvalues with real part in $(-\infty, \lambda_0)$ and $l = n - k$ eigenvalues with real part in (μ_0, ∞).

(SF2) The differential equation $b' = g(\check{a}(b), b, 0) \neq 0$ is rectifiable on V.

According to Fenichel theory [4,5], after an ϵ-dependent change of coordinates, (2.1)–(2.2) takes the form

$$\dot{x} = A(x, y, c, \epsilon)x, \tag{2.9}$$

$$\dot{y} = B(x, y, c, \epsilon)y, \tag{2.10}$$

$$\dot{c} = \epsilon\big((1, 0, \dots, 0) + L(x, y, c, \epsilon)xy\big), \tag{2.11}$$

with $x \in \mathbb{R}^k$, $y \in \mathbb{R}^l$, $c \in \mathbb{R}^m$; the matrices $A(0, 0, c, 0)$ and $B(0, 0, c, 0)$ have the real parts of their eigenvalues in $(-\infty, \lambda_0)$ and (μ_0, ∞), respectively.

2.2. *Exchange Lemma of Jones and Kopell in the case $m = 1$*

Consider the system (2.9)–(2.11) with m, the dimension of c-space, equal to 1. The third equation is then just $\dot{c} = \epsilon(1 + l(x, y, c, \epsilon)xy)$.

For each $\epsilon \geqslant 0$, let M_ϵ be a submanifold of xyc-space of dimension l. Assume

(JK1) $M = \{(x, y, c, \epsilon)\colon (x, y, c) \in M_\epsilon\}$ is itself a manifold.
(JK2) M_0 meets the space $y = 0$ transversally at a point $(x_*, 0, 0)$.

In applications one usually has $x_* \neq 0$, but this is not necessary to the statement of the result.

Under the forward flow of (2.9)–(2.11), M_ϵ becomes a manifold M_ϵ^* of dimension $l + 1$.

Theorem 2.1. *(See [7,8].) Assume* (2.9)–(2.11) *is a C^{r+1} system and M is a C^{r+1} manifold, $r \geqslant 1$. Let $0 < c^*$, and let $y^* \neq 0$ be small. Let Δ be a small neighborhood of (y^*, c^*) in yc-space. Then for $\epsilon_0 > 0$ sufficiently small there is C^r function $\tilde{x} : \Delta \times [0, \epsilon_0) \to \mathbb{R}^k$ such that*:

(1) $\tilde{x}(y, c, 0) = 0$.
(2) *As $\epsilon \to 0$, $\tilde{x} \to 0$ exponentially, along with its derivatives through order r with respect to all variables.*
(3) *For $0 < \epsilon < \epsilon_0$, $\{(x, y, c)\colon (y, c) \in \Delta$ and $x = \tilde{x}(y, c, \epsilon)\}$ is contained in M_ϵ^*.*

See Fig. 1(a) and (b). When we say that a function $h(\epsilon) \to 0$ exponentially as $\epsilon \to 0$, we mean that there are numbers $K > 0$ and $L > 0$ such that for small $\epsilon > 0$, $\|h(\epsilon)\| \leqslant Ke^{-\frac{L}{\epsilon}}$.

We remark that using [3], one can show that if (2.1)–(2.2) is C^{r+3}, then the coordinate change can be chosen so that (2.9)–(2.11) is C^{r+1}.

2.3. *Reformulation of Jones and Kopell's Exchange Lemma as an Inclination Lemma*

In the formulation of Jones and Kopell, the subject of the Exchange Lemma is the entrance of a manifold of orbits into a neighborhood of a normally hyperbolic invariant manifold and the subsequent exit of these orbits from that neighborhood. In Brunovsky's reformulation, the subject is the entrance of a manifold of orbits into a neighborhood of a normally hyperbolic invariant manifold and their subsequent behavior, whether or not they exit the neighborhood.

Theorem 2.2. *(See [1].) Assume* (2.9)–(2.11) *is a C^{r+1} system and M is a C^{r+1} manifold, $r \geqslant 1$. Let $0 < c^*$. Let Δ be a small neighborhood of $(0, c^*)$ in yc-space. Then for $\epsilon_0 > 0$ sufficiently small there is C^r function $\tilde{x} : \Delta \times [0, \epsilon_0) \to \mathbb{R}^k$ such that*:

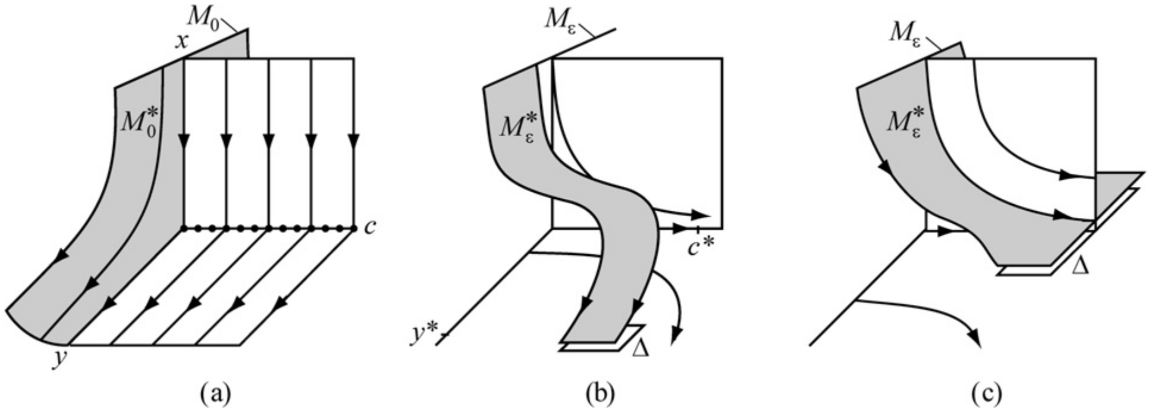

Fig. 1. Jones and Kopell's Exchange Lemma in the case $m = 1$. (a) $\epsilon = 0$. (b) $\epsilon > 0$, Jones and Kopell's formulation. (c) $\epsilon > 0$, Brunovsky's formulation.

(1) $\tilde{x}(y, c, 0) = 0$.
(2) *As* $\epsilon \to 0$, $\tilde{x} \to 0$ *exponentially, along with its derivatives through order r with respect to all variables.*
(3) *For* $0 < \epsilon < \epsilon_0$, $\{(x, y, c)\colon (y, c) \in \Delta$ *and* $x = \tilde{x}(y, c, \epsilon)\}$ *is contained in* M_ϵ^*.

See Fig. 1(c). We shall give exchange lemmas in Brunovsky's formulation rather than the original formulation of Jones and Kopell.

2.4. *Exchange Lemma of Jones and Tin*

Consider the system (2.9)–(2.11) with m, the dimension of c-space, greater than or equal to 1.

For each $\epsilon \geqslant 0$, let M_ϵ be a submanifold of xyc-space of dimension $l + p$, $0 \leqslant p \leqslant m - 1$. Assume

(JT1) $M = \{(x, y, c, \epsilon)\colon (x, y, c) \in M_\epsilon\}$ is itself a manifold.
(JT2) M_0 meets the space $y = 0$ transversally at a point $(x_*, 0, 0)$.
(JT3) $T_{(x^*,0,0)}M_0$ contains no nonzero vectors with $y = 0$ and $c_2 = \cdots = c_m = 0$.

From (JT2), each M_ϵ meets the space $y = 0$ transversally in a manifold N_ϵ of dimension p. From (JT3), N_ϵ projects to a submanifold P_ϵ of c-space of dimension p, and the vector $(1, 0, \ldots, 0)$ is not tangent to P_0 at the origin.

After an ϵ-dependent change of coordinates $c(u, v, w, \epsilon)$, $(u, v, w) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m-1-p}$, that takes each P_ϵ to v-space, (2.9)–(2.11) can be put in the form

$$\dot{x} = A(x, y, u, v, w, \epsilon)x, \tag{2.12}$$

$$\dot{y} = B(x, y, u, v, w, \epsilon)y, \tag{2.13}$$

$$\dot{u} = \epsilon\bigl(1 + e(x, y, u, v, w, \epsilon)xy\bigr), \tag{2.14}$$

$$\dot{v} = \epsilon F(x, y, u, v, w, \epsilon)xy, \tag{2.15}$$

$$\dot{w} = \epsilon G(x, y, u, v, w, \epsilon)xy. \tag{2.16}$$

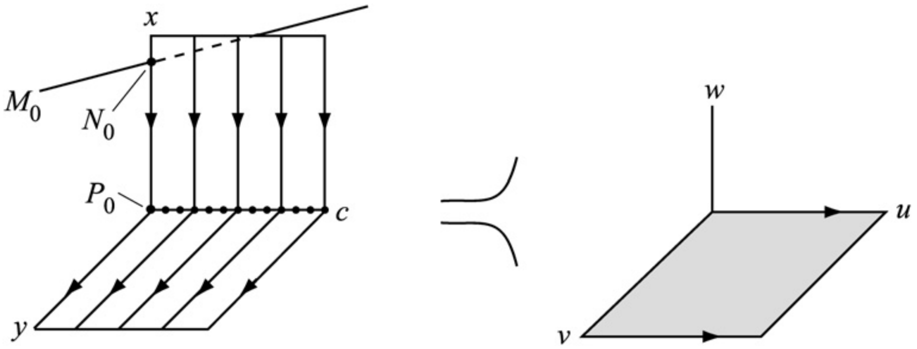

Fig. 2. The Exchange Lemma of Jones and Tin. The first picture shows xyc-space with $\epsilon = 0$. The second shows a more detailed view of c-space for $\epsilon > 0$. In uvw-coordinates on c-space, P_ϵ corresponds to v-space and P_ϵ^* to uv-space.

Under the forward flow of (2.9)–(2.11), M_ϵ and P_ϵ become manifolds M_ϵ^* and P_ϵ^* of dimension $l + p + 1$ and $p + 1$, respectively. P_ϵ^* corresponds to uv-space. See Fig. 2.

Theorem 2.3. *(See [6,13].) Assume* (2.12)–(2.16) *is a C^{r+1} system and M is a C^{r+1} manifold, $r \geqslant 1$. Let $0 < u^*$. Let Δ be a small neighborhood of $(0, u^*, 0)$ in yuv-space. Then for $\epsilon_0 > 0$ sufficiently small there are C^r functions $\tilde{x} : \Delta \times [0, \epsilon_0) \to \mathbb{R}^k$ and $\tilde{w} : \Delta \times [0, \epsilon_0) \to \mathbb{R}^{m-1-p}$ such that*:

(1) $\tilde{x}(y, u, v, 0) = 0$.
(2) $\tilde{w}(y, u, v, 0) = \tilde{w}(0, u, v, \epsilon) = 0$.
(3) *As $\epsilon \to 0$, $(\tilde{x}, \tilde{w}) \to 0$ exponentially, along with its derivatives through order r with respect to all variables.*
(4) *For $0 < \epsilon < \epsilon_0$, $\{(x, y, u, v, w)\colon (y, u, v) \in \Delta$ and $(x, w) = (\tilde{x}, \tilde{w})(y, u, v, \epsilon)\}$ is contained in M_ϵ^*.*

The original Jones–Kopell Exchange Lemma, which we stated only in the case $m = 1$, is actually the Jones–Tin Exchange Lemma in the case $p = 0$, in which case assumption (JT3) is automatic.

We remark that using [3], one can show that if (2.1)–(2.2) is C^{r+3}, then coordinate change can be chosen so that (2.12)–(2.16) is C^{r+1}.

2.5. Normally hyperbolic manifolds of equilibria

We consider a differential equation $\dot{\xi} = F(\xi, \epsilon)$ on $\mathbb{R}^n$ such that $\dot{\xi} = F(\xi, 0)$ has an m-dimensional normally hyperbolic manifold E_0 of equilibria. We assume there are numbers $\lambda_0 < 0 < \mu_0$ such that for all $\xi \in E_0$, $D_\xi F(\xi, 0)$ has k eigenvalues with real part in $(-\infty, \lambda_0)$ and l eigenvalues with real part in (μ_0, ∞), with $k + l + m = n$.

Such a system can be put in the form

$$\dot{x} = A(x, y, c, \epsilon)x, \tag{2.17}$$

$$\dot{y} = B(x, y, c, \epsilon)y, \tag{2.18}$$

$$\dot{b} = \epsilon \check{K}(b, \epsilon) + \check{L}(x, y, b, \epsilon)xy, \tag{2.19}$$

with $x \in \mathbb{R}^k$, $y \in \mathbb{R}^l$, $c \in \mathbb{R}^m$; the matrices $A(0,0,c,0)$ and $B(0,0,c,0)$ have the real parts of their eigenvalues in $(-\infty, \lambda_0)$ and (μ_0, ∞), respectively.

Near a compact nontrivial orbit segment of $b' = \check{K}(b,0)$, the system can be put in the form

$$\dot{x} = A(x,y,c,\epsilon)x, \tag{2.20}$$

$$\dot{y} = B(x,y,c,\epsilon)y, \tag{2.21}$$

$$\dot{c} = \epsilon(1,0,\dots,0) + L(x,y,c,\epsilon)xy. \tag{2.22}$$

For each $\epsilon \geqslant 0$, let M_ϵ be a submanifold of xyc-space of dimension $l+p$, $0 \leqslant p \leqslant m-1$, that satisfies (JT1)–(JT3) of Section 2.4. Define N_ϵ and P_ϵ as in that section. After an ϵ-dependent change of coordinates $c(u,v,w,\epsilon)$, $(u,v,w) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m-1-p}$, that takes each P_ϵ to v-space, (2.20)–(2.22) can be put in the form

$$\dot{x} = A(x,y,u,v,w,\epsilon)x, \tag{2.23}$$

$$\dot{y} = B(x,y,u,v,w,\epsilon)y, \tag{2.24}$$

$$\dot{u} = \epsilon + e(x,y,u,v,w,\epsilon)xy, \tag{2.25}$$

$$\dot{v} = F(x,y,u,v,w,\epsilon)xy, \tag{2.26}$$

$$\dot{w} = G(x,y,u,v,w,\epsilon)xy. \tag{2.27}$$

Under the forward flow of (2.23)–(2.27), M_ϵ becomes a manifold M_ϵ^* of dimension $l+p+1$. Theorem 2.3 holds exactly as stated. (This fact is remarked in [7] and [1].)

2.6. *Loss-of-stability turning points*

Liu [9] considers a slow–fast system

$$\dot{a} = f(a,b,\epsilon) = f_\epsilon(a,b), \tag{2.28}$$

$$\dot{b} = \epsilon g(a,b,\epsilon) = \epsilon g_\epsilon(a,b), \tag{2.29}$$

with $a \in \mathbb{R}^{k+l+1}$ and $b \in \check{V} \subset \mathbb{R}^{m-1}$, $\check{V}$ open, and $m \geqslant 2$. The system has the following properties.

(L1) $f(0,b,\epsilon) = 0$. Hence $\{0\} \times \check{V}$ is locally invariant for each ϵ, and consists of equilibria for $\epsilon = 0$.

(L2) There are numbers $\lambda_0 < 0 < \mu_0$, and a codimension-one submanifold $\check{V}_0$ of $\check{V}$, such that for all $b \in \check{V}$, $D_a f_0(0,b)$ has
- k eigenvalues with real part less than λ_0;
- l eigenvalues with real part greater than μ_0;
- one eigenvalue $\nu(b)$ with $\lambda_0 < \nu(b) < \mu_0$;
- $\nu(b) = 0$ if $b \in \check{V}_0$.

(L3) For $b \in \check{V}_0$, $D\nu(b)g_0(0,b) > 0$.

(L1)–(L3) imply that for $\epsilon = 0$, the $(m-1)$-dimensional manifold of equilibria $\{0\} \times \check{V}$ loses normal hyperbolicity along $\{0\} \times \check{V}_0$. More precisely, as one crosses $\{0\} \times \check{V}_0$ along solutions of $\dot{b} = g_0(0, b)$, an eigenvalue of $D_a f_0(0, b)$ changes from negative to positive (loss of stability).

From [9], for each small ϵ, the $(m-1)$-dimensional invariant set $\{0\} \times \check{V}$ is contained in an m-dimensional normally hyperbolic invariant manifold K_ϵ. Let $K = \{(a, b, \epsilon)\colon (a, b) \in K_\epsilon\}$, an $(m+1)$-dimensional normally hyperbolic invariant manifold of $ab\epsilon$-space.

Let $b_0 \in \check{V}_0$. According to [9], near $(0, b_0, 0)$, we can choose coordinates $(x, y, z, \omega, \epsilon)$ on a neighborhood of $\{0\} \times \check{V} \times \{0\}$ in $ab\epsilon$-space, with $x \in \mathbb{R}^k$, $y \in \mathbb{R}^l$, $z \in \mathbb{R}$, $\omega \in \mathbb{R}^{m-1}$, and (z, ω, ϵ) coordinates on K, such that $(0, b_0, 0)$ corresponds to the origin, and the system (2.28)–(2.29) becomes

$$\dot{x} = \check{A}(x, y, z, \omega, \epsilon)x, \tag{2.30}$$

$$\dot{y} = \check{B}(x, y, z, \omega, \epsilon)y, \tag{2.31}$$

$$\dot{z} = \check{h}(z, \omega, \epsilon)z + \check{k}(x, y, z, \omega, \epsilon)xy, \tag{2.32}$$

$$\dot{\omega} = \epsilon\big((1, 0, \dots, 0) + \check{L}(z, \omega, \epsilon)z + \check{M}(x, y, z, \omega, \epsilon)xy\big), \tag{2.33}$$

with $\operatorname{sgn} \check{h}(0, (\omega_1, \omega_2, \dots, \omega_{m-1}), 0) = \operatorname{sgn} \omega_1$ and $\frac{\partial \check{h}}{\partial \omega_1}(0, (0, \omega_2, \dots, \omega_{m-1}), 0) > 0$. The matrices $\check{A}(0, 0, z, \omega, 0)$ and $\check{B}(0, 0, z, \omega, 0)$ have the real parts of their eigenvalues in $(-\infty, \lambda_0)$ and (μ_0, ∞), respectively. After dividing by the positive function $1 + \check{L}_1(z, \omega, \epsilon)z$, we have $\dot{\omega}_1 = \epsilon(1 + \text{terms of order } xy)$, i.e., $\check{L}_1(z, \omega, \epsilon)z = 0$; we shall assume that this has been done.

Let

$$I = \left\{\omega \in \mathbb{R}^{m-1}\colon\ \omega_1 < 0 \text{ and there exists } \omega_1^* > 0 \text{ such that} \right.$$
$$(1)\ (t, \omega_2, \dots, \omega_{m-1}) \in \check{V} \text{ for } \omega_1 \leqslant t \leqslant \omega_1^*, \text{ and}$$
$$\left. (2) \int_{\omega_1}^{\omega_1^*} \check{h}\big(0, (t, \omega_2, \dots, \omega_{m-1}), 0\big)\, dt = 0 \right\}.$$

Define a map $\pi : I \to \mathbb{R}$ as follows: $\pi(\omega)$ is the smallest number ω_1^* such that

$$\int_{\omega_1}^{\omega_1^*} \check{h}\big(0, (t, \omega_2, \dots, \omega_{m-1}), 0\big)\, dt = 0.$$

For later use we define $\Pi_0 : I \to \mathbb{R}^{m-1}$ by $\Pi_0(\omega_1, \omega_2, \dots, \omega_{m-1}) = (\pi(\omega), \omega_2, \dots, \omega_{m-1})$.

Choose numbers $\beta_0 > 0$ and $\eta > 0$, and a neighborhood $\check{V}$ of the origin in ω-space, such that

- for all (z, ω, ϵ) with $|z| < \eta$, $\omega \in \check{V}$, and $|\epsilon| < \eta$, $\frac{\partial}{\partial z}(z\hat{h}) < \beta_0$;
- $\lambda_0 + \mu_0 + r\beta_0 < 0 < \mu_0 - \max(6, 2r + 1)\beta_0$;
- if $\omega \in \check{V}$ with $\omega_1 < 0$, then $\omega \in I$, and the points $(t, \omega_2, \dots, \omega_{m-1})$ with $\omega_1 \leqslant t \leqslant \pi(\omega)$ are in $\check{V}$.

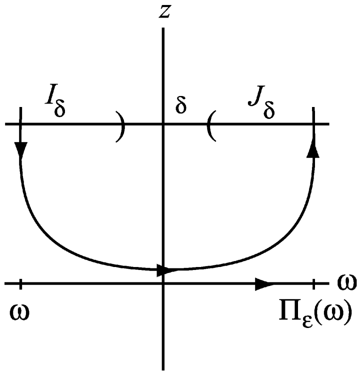

Fig. 3. The mapping Π_ϵ with $\epsilon > 0$ and $m = 2$.

(The numbers λ_0 and μ_0 used in assumption (L2) may need to be adjusted to allow this.) Let $V = \{(z, \omega, z)\colon |z| < \eta \text{ and } \omega \in \check{V}\}$.

Choose a small number $\delta_1 > 0$. Let $W_* = \{\omega \in \check{V}\colon \omega_1 < -\delta_1\}$. Then for the system (2.30)–(2.33) with $\epsilon = 0$, the subset $\{0\} \times \{0\} \times \{0\} \times W_*$ of $xyz\omega$-space is a normally hyperbolic invariant manifold of equilibria, with stable manifold equal to an open subset of $xz\omega$-space, and unstable manifold equal to an open subset of $y\omega$-space. Similarly, let $W^* = \{\omega \in \check{V}\colon \omega_1 > \pi(-\delta_1, \omega_2, \dots, \omega_{m-1})\}$. Then for the system (2.30)–(2.33) with $\epsilon = 0$, the subset $\{0\} \times \{0\} \times \{0\} \times W^*$ of $xyz\omega$-space is a normally hyperbolic invariant manifold of equilibria, with stable manifold equal to an open subset of $x\omega$-space, and unstable manifold equal to an open subset of $yz\omega$-space. For $\epsilon > 0$, W_* and W^* remain normally hyperbolic invariant manifolds with the same properties.

For a given δ, $0 < \delta < \eta$, let I_δ denote the set of points in $z\omega$-space with $z = \delta$ and $\omega \in W_*$, and let J_δ denote the set of points in $z\omega$-space with $z = \delta$ and $\omega \in W^*$.

For $\epsilon > 0$ there is a Poincaré map from a large open subset of I_δ to J_δ given by $(\delta, \omega) \to (\delta, \Pi_\epsilon(\omega))$. See Fig. 3.

For each $\epsilon \geqslant 0$, let M_ϵ be a submanifold of $xyz\omega$-space of dimension $l + p$, $0 \leqslant p \leqslant m - 2$.

Assume:

(L4) $M = \{(x, y, z, \omega, \epsilon)\colon (x, y, z, \omega) \in M_\epsilon\}$ is itself a manifold.
(L5) M_0 meets the space $y = 0$ transversally at a point $(x_*, 0, \delta, \omega_*)$ in the stable manifold of $\{0\} \times \{0\} \times \{0\} \times W_*$.
(L6) $T_{(x_*,0,\delta,\omega_*)} M_0$ contains no nonzero vectors with $y = 0$ and $\omega_2 = \cdots = \omega_{m-1} = 0$.

We may assume that $M \subset \{(x, y, z, \omega, \epsilon)\colon z = \delta\}$. See Fig. 4.

Each M_ϵ meets the space $y = 0$ transversally in a manifold N_ϵ of dimension p. Each N_ϵ projects along stable fibers to a p-dimensional submanifold P_ϵ of $z\omega$-space, which in turn projects along stable fibers to a p-dimensional submanifold Q_ϵ of ω-space. The vector $(1, 0, \dots, 0)$ is not tangent to Q_ϵ.

Under the flow, M_ϵ and P_ϵ become manifolds M^*_ϵ and P^*_ϵ of dimensions $l + p + 1$ and $p + 1$, respectively.

For each $\epsilon \geqslant 0$ the mapping $(\delta, \omega) \to (\delta, \Pi_\epsilon(\omega))$ takes P_ϵ to a p-dimensional submanifold $P^\dagger_\epsilon$ of J_δ.

Assume:

(L7) The functions $\check{h}(z, \omega, \epsilon)$ and $\check{L}(z, \omega, \epsilon)z$ are C^{r+1}.

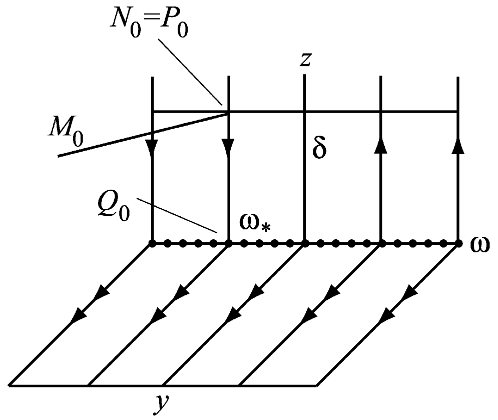

Fig. 4. A loss-of-stability turning point with $k = 0$, $l = 1$, $m = 2$, $p = 0$. The flow for $\epsilon = 0$ is shown. Since $k = 0$, $N_0 = P_0$.

Then by [2], the mappings Π_ϵ, $\epsilon \geqslant 0$, fit together to form a C^{r+1} mapping $\Pi(\omega, \epsilon) = \Pi_\epsilon(\omega)$, $\epsilon \geqslant 0$. Therefore the sets $P_\epsilon^\dagger$, $\epsilon \geqslant 0$, fit together to form a C^{r+1} manifold $P^\dagger$ of $J_\delta \times \mathbb{R}$.

On a neighborhood V^* of $(\delta, \Pi_0(\omega_*))$ in $z\omega$-space, there is an ϵ-dependent change of coordinates $(z, \omega)(u^1, v^1, w^1, \epsilon)$, $(u^1, v^1, w^1) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m-1-p}$, such that:

- $z = \delta$ if and only if $u^1 = 0$.
- $(z, \omega)(0, 0, 0, 0) = (\delta, \Pi_0(\omega_*))$.
- $(z, \omega)(0, v^1, w^1, \epsilon) \in P_\epsilon^\dagger$ if and only if $w^1 = 0$.
- In the new coordinates, (2.32)–(2.33) becomes

$$\dot{u}^1 = 1, \quad \dot{v}^1 = 0, \quad \dot{w}^1 = 0.$$

See Fig. 5. In these coordinates, $P^\dagger$ is $v\epsilon$-space.

Theorem 2.4. *(See [9].) Assume* (2.30)–(2.33) *is a* C^{r+1} *system and* M *is a* C^{r+1} *manifold,* $r \geqslant 1$. *On a neighborhood of* $(0, 0, \delta, \Pi_0(\omega_*), 0)$ *in* $xyz\omega\epsilon$*-space, with* $\epsilon \geqslant 0$, *use the coordinates* $(x, y, u^1, v^1, w^1, \epsilon)$.

Let Δ *be a small neighborhood of* $(0, 0, 0)$ *in* yu^1v^1*-space. Then for* $\epsilon_0 > 0$ *sufficiently small there is a* C^r *function* $(\tilde{x}, \tilde{w}) : \Delta \times [0, \epsilon_0) \to \mathbb{R}^k \times \mathbb{R}^{m-1-p}$ *such that*:

(1) $\tilde{x}(y, u^1, v^1, 0) = 0$.
(2) $\tilde{w}(y, u^1, v^1, 0) = \tilde{w}(0, u^1, v^1, \epsilon) = 0$.
(3) *As* $\epsilon \to 0$, $(\tilde{x}, \tilde{w}) \to 0$ *exponentially, along with its derivatives through order r with respect to all variables.*
(4) *For* $0 < \epsilon < \epsilon_0$, $\{(x, y, u^1, v^1, w^1)$: $(y, u^1, v^1) \in \Delta$ *and* $(x, w^1) = (\tilde{x}, \tilde{w})(y, u^1, v^1, \epsilon)\}$ *is contained in* M_ϵ^*.

See Fig. 5.

We remark that using [3], one can show that if (2.28)–(2.29) is C^{r+4}, $r \geqslant 1$, then the coordinate change can be chosen so that (2.30)–(2.33) is C^{r+2}, in which case (L7) and the differentiability assumption of the theorem are both satisfied.

I would like to emphasize the importance of the fact that the mappings Π_ϵ, $\epsilon \geqslant 0$, fit together to form a C^{r+1} mapping defined for $\epsilon \geqslant 0$, which is essential to the proof of this theorem that

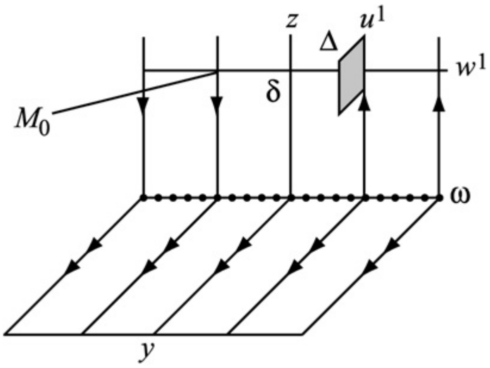

Fig. 5. Theorem 2.4 with $k=0$, $l=1$, $m=2$, $p=0$. The flow for $\epsilon=0$ is shown. Since $p=0$, there is no v^1-coordinate.

we will give in Section 4. This fact was recently proved by Peter De Maesschalck; to the best of my knowledge, it was not previously in the literature. The asymptotic expansion of $\Pi_\epsilon(\omega)$ is calculated in [10], but the existence of an asymptotic expansion does not imply the result. Liu [9] makes a weaker assertion, namely that Π_ϵ approaches Π_0 in the C^{r+1} topology if (2.30)–(2.33) is sufficiently differentiable, but he does not give a reference even for the weaker result.

3. General Exchange Lemma

On $\mathbb{R}^n$ we use coordinates $\xi=(x,y,c)$, with $x\in\mathbb{R}^k$, $y\in\mathbb{R}^l$, $c\in\mathbb{R}^m$, $k+l+m=n$. Let V be an open subset of $\mathbb{R}^m$. We consider a C^{r+1} differential equation $\dot\xi=F(\xi,\epsilon)$, $r\geqslant 1$, on a neighborhood of $\{0\}\times\{0\}\times V\times\{0\}$ in $\mathbb{R}^n\times\mathbb{R}$ of the following form:

$$\dot x = A(x,y,c,\epsilon)x, \tag{3.1}$$

$$\dot y = B(x,y,c,\epsilon)y, \tag{3.2}$$

$$\dot c = C(c,\epsilon)+E(x,y,c,\epsilon)xy. \tag{3.3}$$

Let $\phi_\epsilon(t,c)$ be the flow of $\dot c=C(c,\epsilon)$. For each $c\in V$ there is a maximal interval I_c containing 0 such that $\phi_0(t,c)\in V$ for all $t\in I_c$. Let the linearized solution operator of (3.1)–(3.3), with $\epsilon=0$, along the solution $(0,0,\phi_0(t,c^0))$ be

$$\begin{pmatrix}\bar x(t)\\ \bar y(t)\\ \bar c(t)\end{pmatrix}=\begin{pmatrix}\Phi^s(t,s,c^0) & 0 & 0\\ 0 & \Phi^u(t,s,c^0) & 0\\ 0 & 0 & \Phi^c(t,s,c^0)\end{pmatrix}\begin{pmatrix}\bar x(s)\\ \bar y(s)\\ \bar c(s)\end{pmatrix}. \tag{3.4}$$

We shall make five types of assumptions: (1) assumptions on the *linearized solution operator* (3.4); (2) assumptions on the *incoming manifolds* M_ϵ; (3) assumptions on the *flow on V near the starting points*; (4) assumptions on the *flow on V near the ending points*; and (5) an assumption on the *time spent flowing*.

Assumptions on the linearized solution operator (3.4).

(E1) There are numbers $\lambda_0<0<\mu_0$, $\beta_0>0$, and $M>0$ such that for all $c^0\in V$ and $s,t\in I_{c^0}$,

$$\left\|\Phi^s\left(t,s,c^0\right)\right\|\leqslant Me^{\lambda_0(t-s)}\quad \text{if } t\geqslant s, \tag{3.5}$$

$$\left\| \Phi^u\left(t, s, c^0\right) \right\| \leqslant M e^{\mu_0(t-s)} \quad \text{if } t \leqslant s, \tag{3.6}$$

$$\left\| \Phi^c\left(t, s, c^0\right) \right\| \leqslant M e^{\beta_0|t-s|} \quad \text{for all } t,\ s. \tag{3.7}$$

(E2) $\lambda_0 + \mu_0 + r\beta_0 < 0 < \mu_0 - \max(6, 2r+1)\beta_0$.

After slightly changing λ_0, μ_0, β_0, and M, we may assume that these estimates also hold for linearization around solutions of $\dot{c} = C(c, \epsilon)$ on V for ϵ small.

Assumptions on the incoming manifolds M_ϵ. For each $\epsilon \geqslant 0$, let M_ϵ be a C^{r+1} submanifold of xyc-space of dimension $l + p$, $0 \leqslant p \leqslant m-1$.

(E3) $M = \{(x, y, c, \epsilon)\colon (x, y, c) \in M_\epsilon\}$ is itself a C^{r+1} manifold.
(E4) M_0 meets the space $y = 0$ transversally at a point $(x_*, 0, c_*)$.
(E5) $T_{(x_*,0,c_*)} M_0$ contains no nonzero vectors $(\bar{x}, \bar{y}, \bar{c})$ with $\bar{y} = 0$ and $\bar{c} = 0$.

From (E3) and (E4), each M_ϵ meets the space $y = 0$ transversally in a manifold N_ϵ of dimension p. From (E5), each N_ϵ projects to a submanifold P_ϵ of c-space of dimension p. Let $P = \{(c, \epsilon)\colon c \in P_\epsilon\}$.

Assumptions on the flow on V near the starting points. There is an open neighborhood V_* of c_* in V, and coordinates $c(u^0, v^0, w^0, \epsilon)$ on V_*, with $(u^0, v^0, w^0) \in U_* \subset \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m-1-p}$ and $\epsilon \geqslant 0$ small, such that $c(0, 0, 0, 0) = c_*$ and P is contained in $v^0\epsilon$-space. After shrinking M if necessary, we may assume that U_* contains the closed ball of radius 3γ about the origin, and each $P_\epsilon \subset \{(u^0, v^0, w^0)\colon (u^0, w^0) = (0, 0) \text{ and } \|v^0\| < \gamma\}$. We assume the coordinates can be chosen so that:

(E6) $\dot{w}^0 = 0$.
(E7) $\dot{u}^0$ is a function of ϵ only. Moreover, there are numbers $a > 0$ and $K_3 > 0$ such that for small $\epsilon > 0$, $\dot{u}_0 > K_3 \epsilon^a$.
(E8) The coordinate change $c(u^0, v^0, w^0, \epsilon)$ is C^{r+1}, and the differential equation in these coordinates is C^{r+1}.

Assumptions on the flow on V near the ending points. For $\epsilon > 0$ we have the following: From (E7), under the flow of $\dot{c} = C(c, \epsilon)$, each P_ϵ becomes a manifold P_ϵ^* of dimension $p+1$. Similarly, under the flow of $\dot{\xi} = F(\xi, \epsilon)$, each M_ϵ becomes a manifold M_ϵ^* of dimension $l + p + 1$.

We assume there is an open subset V^* of V with the following properties:

(E9) There are coordinates $c(v^1, w^1, \epsilon)$ on V^*, with $(v^1, w^1) \in U^* \subset \mathbb{R}^{p+1} \times \mathbb{R}^{m-1-p}$ and $\epsilon \geqslant 0$ small, such that for each $\epsilon > 0$, the set $P_\epsilon^* \cap V^*$ is given by $w^1 = 0$.
(E10) The coordinate change $c(v^1, w^1, \epsilon)$ is C^{r+1}, and the differential equation in these coordinates is C^{r+1}.

See Fig. 6.

Assumption on the time spent flowing. For each $c^1 \in P_\epsilon^* \cap V^*$ there is a positive number $\tau(c^1, \epsilon)$ such that $\phi_\epsilon(-\tau(c^1, \epsilon), c^1) \in P_\epsilon$.

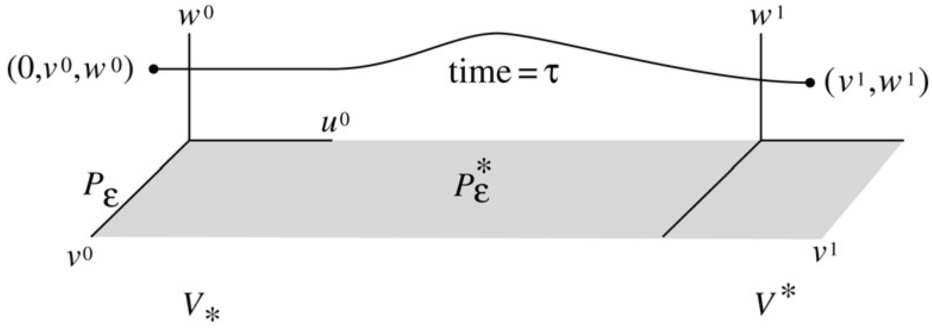

Fig. 6. Coordinates on V_* and V^*. Note that v^0-space is p-dimensional (it is P_ϵ) but v^1-space is $(p+1)$-dimensional.

(E11) There are positive numbers β, β_1, K_1, and K_2 such that (1) $\beta_0 < \beta < \beta_1$, (2) $0 < \mu_0 - \max(6, 2r+1)\beta_1$, (3) $K_1 < K_2 < \frac{\beta_1}{\beta} K_1$, and (4) for small $\epsilon > 0$ and all $c^1 \in P^*_\epsilon \cap V^*$, $\frac{K_1}{\epsilon^a} \leqslant \tau(c^1, \epsilon) \leqslant \frac{K_2}{\epsilon^a}$.

We are now ready to state the General Exchange Lemma.

Theorem 3.1. *Let $(v^{1*}, 0) \in U^*$. Let A be a small neighborhood of $(0, v^{1*})$ in yv^1-space. Then for $\epsilon_0 > 0$ sufficiently small there are C^r functions $\tilde{x} : A \times [0, \epsilon_0) \to \mathbb{R}^k$ and $\tilde{w} : A \times [0, \epsilon_0) \to \mathbb{R}^{m-p-1}$ such that:*

(1) $\tilde{x}(y, v^1, 0) = 0$.
(2) $\tilde{w}(y, v^1, 0) = \tilde{w}(0, v^1, \epsilon) = 0$.
(3) *As $\epsilon \to 0$, $(\tilde{x}, \tilde{w}) \to 0$ exponentially, along with its derivatives through order r with respect to all variables. More precisely, any first partial derivative of $(\tilde{x}, \tilde{w})$ is of order $e^{-\frac{K_1}{\epsilon^a}(\mu_0 - 4\beta_1)}$, and, for $2 \leqslant i \leqslant r$, any partial derivative of $(\tilde{x}, \tilde{w})$ of order i is of order $e^{-\frac{K_1}{\epsilon^a}(\mu_0 - (2i+1)\beta_1)}$.*
(4) *For $0 < \epsilon < \epsilon_0$, $\{(x, y, v^1, w^1)$: $(y, v^1) \in A$ and $(x, w) = (\tilde{x}, \tilde{w})(y, v^1, \epsilon)\}$ is contained in M^*_ϵ.*

Remark 3.2. Two of the assumptions of the General Exchange Lemma deserve comment.

Assumption (E7) requires that one be able to choose coordinates on V_* in which $\dot{u}_0$ depends only on ϵ. It would be desirable to remove this assumption.

Assumption (E11) only requires looking at the time of transit from P_ϵ to points of P^*_ϵ in V^*. It does not require looking in general at the time of transit from V_* to V^*. This is important because it may be impossible to define an open set V^* in which, for small $\epsilon > 0$, every point comes from a point in V_*. The proof will show, however, that points in V^* with w^1 exponentially small as $\epsilon \to 0$ (i.e., points in V^* that are very close to P^*_ϵ) do come from points in V_*, and this is essential to the proof.

4. The General Exchange Lemma and existing exchange lemmas

4.1. Normally hyperbolic manifolds of equilibria

As in Section 2.5, we consider a differential equation $\dot{\xi} = F(\xi, \epsilon)$ such that $\dot{\xi} = F(\xi, 0)$ has an m-dimensional normally hyperbolic manifold of equilibria, and the assumptions of Section 2.5 on the eigenvalues and the manifold M_ϵ are satisfied. (The situations described in Sections 2.3

and 2.4 are special cases of this one.) We write the system in the form (2.23)–(2.27), and let $c = (u, v, w)$.

By decreasing μ_0 if necessary, we can make $\lambda_0 + \mu_0 < 0$. Then for any sufficiently small $\beta_0 > 0$, (E1) and (E2) are satisfied.

Let c_* be the origin of uvw-space. Then (E3)–(E5) are satisfied. Let V_* be a small neighborhood of c_*, with coordinates $(u, v, w) = (u^0, v^0, w^0)$. Since $\dot{w} = 0$ and $\dot{u} = \epsilon$ on V_*, (E6)–(E8) are satisfied; in (E7) we use $a = 1$.

Let $u^* > 0$ and let V^* be a small neighborhood of $(u^*, 0, 0)$ in uvw-space. For $\epsilon > 0$, $P_\epsilon \cap V_* = \{(u, v, w) \in V_*\colon\ (u, w) = (0, 0)\}$, so $P^*_\epsilon \cap V^* = \{(u, v, w) \in V^*\colon\ w = 0\}$. Hence (E9) and (E10) are satisfied with the coordinates

$$u = v^1_0, \quad v_1 = v^1_1, \quad \ldots, \quad v_p = v^1_p, \quad w = w^1.$$

We have $\tau(u, v, 0, \epsilon) = \frac{u}{\epsilon}$, so (E11) is satisfied if V^* is small enough.

4.2. Liu's Exchange Lemma

As in Section 2.6, we consider a differential equation (2.28)–(2.29) that satisfies (L1)–(L3). We write the system in the form (2.30)–(2.33), let $c = (z, \omega)$, and choose β_0 and V as described in Section 2.6. Then (E1) and (E2) are satisfied.

Given a family of manifolds M_ϵ that satisfy (L4)–(L6), let $c_* = (\delta, \omega_*)$. Then (E3)–(E5) are satisfied.

On a neighborhood V_* of (δ, ω_*) in $z\omega$-space, there is an ϵ-dependent change of coordinates $(z, \omega)(u^0, v^0, w^0, \epsilon)$, $(u^0, v^0, w^0) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m-1-p}$, such that:

- $z = \delta$ if and only if $u^0 = 0$.
- $(z, \omega)(0, 0, 0, 0) = (\delta, \omega_*)$.
- $(z, \omega)(0, v^0, w^0, \epsilon) \in P_\epsilon$ if and only if $w^0 = 0$.
- In the new coordinates, (2.32)–(2.33) becomes

$$\dot{u}^0 = 1, \quad \dot{v}^0 = 0, \quad \dot{w}^0 = 0.$$

Thus (E6)–(E8) are satisfied. In (E7) we could use any $a > 0$; we use $a = 1$.

Let V^* be as defined in Section 2.6. The ϵ-dependent coordinates (u^1, v^1, w^1) defined on V^* there show that (E9) and (E10) are satisfied; (u^1, w^1) plays the role of w^1 in the statement of those conditions. Note that assumption (L7) was used to construct this coordinate system.

(E11) is satisfied if V^* is small enough because in the system (2.30)–(2.33), $\dot{\omega}_1 = \epsilon$.

5. Implicit Function Theorem

Let Y be an open set in $\mathbb{R}^l$, let Z be an open neighborhood of 0 in $\mathbb{R}^m$, and let $s : Y \times Z \to \mathbb{R}$ be a positive continuous function. Let

$$\Omega = \big\{(x, y, z) \in \mathbb{R}^k \times \mathbb{R}^l \times \mathbb{R}^m\colon\ (y, z) \in Y \times Z \text{ and } \|x\| \leqslant s(y, z)\big\}.$$

Let $G : \Omega \to \mathbb{R}^k$ be a C^r function.

Theorem 5.1. *Assume there are real-valued functions m on $Y \times Z$, n on Y, and p on Ω, and a number λ, $0 < \lambda < 1$, such that*:

(I1) $\|G(0, y, z)\| \leqslant m(y, z)$ *for all* $(y, z) \in Y \times Z$.
(I2) $D_x G(0, y, 0)$ *is invertible and* $\|D_x G(0, y, 0)^{-1}\| \leqslant n(y)$ *for all* $y \in Y$.
(I3) $\|D_x G(x, y, z) - D_x G(0, y, 0)\| \leqslant p(y, z)$ *for all* $(x, y, z) \in \Omega$.
(I4) $n(y)p(y, z) \leqslant \lambda$ *for all* $(y, z) \in Y \times Z$.

Let

$$\delta(y, z) = \frac{n(y)m(y, z)}{1 - n(y)p(y, z)}.$$

Assume:

(I5) *For each* $(y, z) \in Y \times Z$, $\delta(y, z) \leqslant s(y, z)$.

Then for each $(y, z) \in Y \times Z$ *there is a unique* x *with* $\|x\| \leqslant \delta(y, z)$ *such that* $G(x, y, z) = 0$. *If we let* $x = g(y, z)$, *then* g *is* C^r.

Proof. Write

$$G(x, y, z) = G(0, y, z) + D_x G(0, y, 0)x + R(x, y, z).$$

$G(x, y, z) = 0$ if and only if

$$x = T(x, y, z) = -D_x G(0, y, 0)^{-1}\bigl(G(0, y, z) + R(x, y, z)\bigr).$$

Note that for $\|x\| \leqslant \delta(y, z)$,

$$\begin{aligned}\bigl\|R(x, y, z)\bigr\| &= \bigl\|R(x, y, z) - R(0, y, z) + R(0, y, z)\bigr\| = \bigl\|R(x, y, z) - R(0, y, z)\bigr\| \\ &\leqslant \sup_{\|x'\| \leqslant \delta(y,z)} \bigl\|D_x R(x', y, z)\bigr\| \|x\| \leqslant p(y, z)\delta(y, z).\end{aligned}$$

Hence, if $\|x\| \leqslant \delta(y, z)$, then

$$\bigl\|T(x, y, z)\bigr\| \leqslant n(y)\bigl(m(y, z) + p(y, z)\delta(y, z)\bigr) = \delta(y, z).$$

In addition, if $\|x_1\| \leqslant \delta(y, z)$ and $\|x_2\| \leqslant \delta(y, z)$, then

$$\begin{aligned}\bigl\|T(x_1, y, z) - T(x_2, y, z)\bigr\| &\leqslant \bigl\|DG(0, y, 0)^{-1}\bigr\| \bigl\|R(x_1, y, z) - R(x_2, y, z)\bigr\| \\ &\leqslant n(y)p(y, z)\|x_1 - x_2\| \leqslant \lambda\|x_1 - x_2\|.\end{aligned}$$

Hence for each (y, z), T is a contraction of $\{x\colon \|x\| \leqslant \delta(y, z)\}$. The result follows from the C^r Contraction Mapping Theorem. $\square$

6. Proof of the General Exchange Lemma

We consider the system (3.1)–(3.3) with assumptions (E1)–(E11). On V_* and V^* we use the coordinates (u^0, v^0, w^0) and (v^1, w^1) defined in Section 3. On $\mathbb{R}^k \times \mathbb{R}^l \times V_* \times \mathbb{R}$, we set $x = x^0$ and $y = y^0$, obtaining coordinates $(x^0, y^0, u^0, v^0, w^0, \epsilon)$. On $\mathbb{R}^k \times \mathbb{R}^l \times V^* \times \mathbb{R}$, we set $x = x^1$ and $y = y^1$, obtaining coordinates $(x^1, y^1, v^1, w^1, \epsilon)$.

In our coordinates on $\mathbb{R}^k \times \mathbb{R}^l \times V_* \times \mathbb{R}$, M takes the form

$$\left(x^0, u^0, w^0\right) = (\hat{x}, \hat{u}, \hat{w})\left(y^0, v^0, \epsilon\right), \quad \hat{x}(0,0,0) = x_*, \quad (\hat{u}, \hat{w})\left(0, v^0, \epsilon\right) = (0,0). \tag{6.1}$$

We have used the fact that P is contained in $v^0\epsilon$-space. The mapping $(\hat{x}, \hat{u}, \hat{w})$ is C^{r+1}.

We wish to consider Silnikov's second boundary value problem, i.e., (3.1)–(3.3) together with the boundary conditions

$$x(0) = x^0, \quad y(\tau) = y^1, \quad c(\tau) = c^1.$$

The solution is denoted $(x, y, c)(t, \tau, x^0, y^1, c^1, \epsilon)$ and is C^{r+1}. We shall always take $c^1 \in V^*$ and values of τ such that $c(0, \tau, x^0, y^1, c^1, \epsilon) \in V_*$. Hence we will write $c(0, \tau, x^0, y^1, c^1, \epsilon) = (u^0, v^0, w^0)$ and $c^1 = (v^1, w^1)$. Thus for t near 0 we will denote the solution of the boundary value problem by $(x^0, y^0, u^0, v^0, w^0)(t, \tau, x^0, y^1, v^1, w^1, \epsilon)$. Deng's lemma (Theorem 2.2 of [11]) provides estimates on the solution, which remain valid despite the coordinate changes.

To prove the General Exchange Lemma, given $(y^1, v^1) \in A$ and a small $\epsilon > 0$, we want to find (τ, x^0, w^1) such that

$$\left(x^0, \left(u^0, w^0\right)\left(0, \tau, x^0, y^1, v^1, w^1, \epsilon\right)\right) = (\hat{x}, \hat{u}, \hat{w})\left(\left(y^0, v^0\right)\left(0, \tau, x^0, y^1, v^1, w^1, \epsilon\right), \epsilon\right). \tag{6.2}$$

Once (τ, x^0, w^1) is found, the desired functions $\tilde{x}$ and $\tilde{w}$ are

$$\tilde{x}\left(y^1, v^1, \epsilon\right) = x\left(\tau, \tau, x^0, y^1, v^1, w^1, \epsilon\right), \qquad \tilde{w}\left(y^1, v^1, \epsilon\right) = w^1. \tag{6.3}$$

The estimates on $\tilde{x}$ and $\tilde{w}$ required by conclusion (3) of the General Exchange Lemma are obtained with the help of Deng's lemma (Theorem 2.2 of [11]).

Recall the function $\tau(c^1, \epsilon)$ defined in Section 3. In our coordinates on V^* it becomes a function $\tau(v^1, 0, \epsilon)$.

We gather some simple facts in the following lemma.

Lemma 6.1.

(1) $(u^0, v^0, w^0)(0, \tau, 0, 0, v^1, w^1, \epsilon) = (u^0, v^0, w^0)(0, \tau, x^0, 0, v^1, w^1, \epsilon) = (u^0, v^0, w^0)(0, \tau, 0, y^1, v^1, w^1, \epsilon)$.
(2) $y^0(0, \tau, x^0, 0, v^1, w^1, \epsilon) = 0$.
(3) $u^0(0, \tau(v^1, 0, \epsilon), 0, 0, v^1, 0, \epsilon) = 0$.
(4) $w^0(0, \tau(v^1, 0, \epsilon), 0, 0, v^1, 0, \epsilon) = 0$.

Proof. To show (1), note that $(u^0, v^0, w^0)(0, \tau, 0, 0, v^1, w^1, \epsilon)$ is given by the backward flow in c-space; it is just $\phi_\epsilon(-\tau, c(v^1, w^1, \epsilon))$. From (3.3), the evolution of the c variables is unchanged as long as $y = 0$ or $x = 0$. (2) follows from (3.2): if $y = 0$, then $\dot{y} = 0$, so y remains 0. (3) is just the definition of $\tau(v^1,0,\epsilon)$. (4) says that for the flow in c-space, if $w^1=0$, then $w^0=0$; this is a consequence of our choice of coordinates on V^*. $\square$

We wish to extend the function $\tau(v^1,0,\epsilon)$ to a function $\tau(v^1,w^1,\epsilon)$ with the property

$$u^0\big(0,\tau\big(v^1,w^1,\epsilon\big),0,0,v^1,w^1,\epsilon\big)=0. \tag{6.4}$$

Lemma 6.2. *$\tau(v^1,w^1,\epsilon)$ satisfying* (6.4) *is defined and C^{r+1} for $(v^1,w^1)\in U^*$ with $\|w^1\|\leqslant e^{-\frac{2K_2}{\epsilon^a}\beta_0}$ and $\epsilon>0$ sufficiently small. Moreover, for slightly smaller K_1 and slightly larger K_2, still satisfying* (3) *of* (E11),

$$\frac{K_1}{\epsilon^a}\leqslant\tau\big(v^1,w^1,\epsilon\big)\leqslant\frac{K_2}{\epsilon^a}.$$

Proof. For $\|w^1\|\leqslant e^{-\frac{2K_2}{\epsilon^a}\beta_0}$ and small $\epsilon>0$, (3.7) and assumption (E11) imply

$$\begin{aligned}
&\big\|\big(u^0,v^0,w^0\big)\big(0,\tau\big(v^1,0,\epsilon\big),0,0,v^1,w^1,\epsilon\big)-\big(u^0,v^0,w^0\big)\big(0,\tau\big(v^1,0,\epsilon\big),0,0,v^1,0,\epsilon\big)\big\|\\
&\quad\leqslant\sup_{\sigma\in[0,1]}\left\|\frac{\partial(u^0,v^0,w^0)}{\partial w^1}\big(0,\tau\big(v^1,0,\epsilon\big),0,0,v^1,\sigma w^1,\epsilon\big)\right\|\big\|w^1\big\|\\
&\quad\leqslant Me^{\beta_0\tau(v^1,0,\epsilon)}\big\|w^1\big\|\leqslant Me^{\frac{K_2}{\epsilon^a}\beta_0}e^{-\frac{2K_2}{\epsilon^a}\beta_0}=Me^{-\frac{K_2}{\epsilon^a}\beta_0}.
\end{aligned} \tag{6.5}$$

From the description of the (u^0,v^0,w^0)-coordinate system in Section 3, $(u^0,v^0,w^0)(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon)=(0,v^0,0)$ with $\|v^0\|<\gamma$. Therefore for $(u^0,v^0,w^0)(0,\tau(v^1,0,\epsilon),0,0,v^1,w^1,\epsilon)$ with ϵ sufficiently small, (6.5) implies

$$\max\big(\big|u^0\big|,\big\|v^0\big\|-\gamma,\big\|w^0\big\|\big)\leqslant Me^{-\frac{K_2}{\epsilon^a}\beta_0}<\gamma. \tag{6.6}$$

Choose L such that for (u^0,v^0,w^0) in the closed ball of radius 3γ about the origin and ϵ small, $\|(\dot u^0,\dot v^0,\dot w^0)\|\leqslant L$. Then for (u^0,v^0,w^0) in the closed ball of radius 2γ about the origin, $\phi_\epsilon(t,(u^0,v^0,w^0))$ is defined for $|t|\leqslant\frac{\gamma}{L}$. For (u^0,v^0,w^0) satisfying (6.6), on the other hand, assumption (E7) implies that the value of t for which $\phi_\epsilon(t,(u^0,v^0,w^0))$ has u^0-coordinate equal to 0 satisfies

$$|t|\leqslant\frac{1}{K_3\epsilon^a}Me^{-\frac{K_2}{\epsilon^a}\beta_0}\leqslant Ke^{-\frac{K_1}{\epsilon^a}\beta_0},$$

which is smaller than $\frac{\gamma}{L}$ for ϵ small. Therefore, for (u^0,v^0,w^0) satisfying (6.6), we can define $t(u^0,v^0,w^0,\epsilon)$, with $|t(u^0,v^0,w^0,\epsilon)|\leqslant Ke^{-\frac{K_1}{\epsilon^a}\beta_0}$, such that $\phi_\epsilon(t(u^0,v^0,w^0),(u^0,v^0,w^0))$ has u^0-coordinate equal to 0.

Let

$$\tau\big(v^1,w^1,\epsilon\big)=\tau\big(v^1,0,\epsilon\big)+t\big(\big(u^0,v^0,w^0\big)\big(0,\tau\big(v^1,0,\epsilon\big),0,0,v^1,w^1,\epsilon\big),\epsilon\big).$$

Then (6.4) holds, and $\tau(v^1,w^1,\epsilon)$ satisfies the required estimate. $\square$

Because of Lemma 6.2, we shall always assume $\|w^1\| \leqslant e^{-\frac{2K_2}{\epsilon^d}\beta_0}$.

From Lemma 6.1, we see that a family of solutions of (6.2), with (v^1, ϵ) arbitrary, is

$$\tau = \tau\big(v^1, 0, \epsilon\big), \tag{6.7}$$

$$x^0 = \hat{x}\big(0, v^0\big(0, \tau\big(v^1, 0, \epsilon\big), 0, 0, v^1, 0, \epsilon\big), \epsilon\big), \tag{6.8}$$

$$y^1 = 0, \tag{6.9}$$

$$w^1 = 0. \tag{6.10}$$

Let

$$x^0 = \hat{x}\big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big) + \bar{x}^0, \tag{6.11}$$

$$\tau = \tau\big(v^1, w^1, \epsilon\big) + \bar{\tau}. \tag{6.12}$$

Let $\check{x}(\bar{\tau}, \bar{x}^0, v^1, w^1, \epsilon)$ be the composite of (6.11) and (6.12):

$$x^0 = \check{x}\big(\bar{\tau}, \bar{x}^0, v^1, w^1, \epsilon\big) = \hat{x}\big(0, v^0\big(0, \tau\big(v^1, w^1, \epsilon\big) + \bar{\tau}, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big) + \bar{x}^0. \tag{6.13}$$

Let Y be the product of a neighborhood of 0 in $\mathbb{R}^{p+1}$ and an interval $(0, \epsilon_0)$, and let Z be a neighborhood of 0 in $\mathbb{R}^l$. For $(\bar{\tau}, \bar{x}^0, w^1)$ near $(0, 0, 0)$ and $((v^1, \epsilon), y^1) \in Y \times Z$, define

$$G\big(\big(\bar{\tau}, \bar{x}^0, w^1\big), \big(v^1, \epsilon\big), y^1\big) = \begin{pmatrix} x^0 - \hat{x}((y^0, v^0)(0, \tau, x^0, y^1, v^1, w^1, \epsilon), \epsilon) \\ u^0(0, \tau, x^0, y^1, v^1, w^1, \epsilon) - \hat{u}((y^0, v^0)(0, \tau, x^0, y^1, v^1, w^1, \epsilon), \epsilon) \\ w^0(0, \tau, x^0, y^1, v^1, w^1, \epsilon) - \hat{w}((y^0, v^0)(0, \tau, x^0, y^1, v^1, w^1, \epsilon), \epsilon) \end{pmatrix}, \tag{6.14}$$

with x^0 and τ given in terms of $(\bar{\tau}, \bar{x}^0, v^1, w^1, \epsilon)$ by (6.13) and (6.12). G is C^{r+1}. According to (6.2), we need to find solutions of $G = 0$.

We will prove the General Exchange Lemma in the following steps, in which the notation of Section 3 is used. We use the letter K to denote a variety of different constants. Let $\tau_1(\epsilon) = \inf_{(v^1, w^1)} \tau(v^1, w^1, \epsilon)$, $\tau_2(\epsilon) = \sup_{(v^1, w^1)} \tau(v^1, w^1, \epsilon)$.

(1) $\|G((0,0,0), (v^1, \epsilon), y^1)\| \leqslant K e^{-\mu_0 \tau(v^1, 0, \epsilon)}$.

(2) $\|D_{(\bar{\tau}, \bar{x}^0, w^1)} G((0,0,0), (v^1, \epsilon), 0)^{-1}\| \leqslant K e^{\beta \tau(v^1, 0, \epsilon)}$.

(3) $$\big\|D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big(\big(\bar{\tau}, \bar{x}^0, w^1\big), (v^1, \epsilon), y^1\big) - D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((0,0,0), \big(v^1, \epsilon\big), 0\big)\big\| \leqslant K e^{-(\mu_0 - 2\beta)(\tau_1(\epsilon) - |\bar{\tau}|)} + K e^{2\beta(\tau_2(\epsilon) + |\bar{\tau}|)} \big\|\big(\bar{\tau}, \bar{x}^0, w^1\big)\big\|.$$

(4) Using the Implicit Function Theorem (Theorem 5.1), we show that for each $((v^1, \epsilon), y^1) \in Y \times Z$, the equation $G((\bar{\tau}, \bar{x}^0, w^1), (v^1, \epsilon), y^1) = 0$ has a unique solution with $\|(\bar{\tau}, \bar{x}^0, w^1)\| \leqslant K e^{-\frac{K_1}{\epsilon^d}(\mu_0 - \beta_1)}$. Moreover, $(\bar{\tau}, \bar{x}^0, w^1)$ is a C^{r+1} function of $((v^1, \epsilon), y^1)$.

(5) For $\|(\bar{\tau}, \bar{x}^0, w^1)\| \leqslant K e^{-\frac{K_1}{\epsilon^d}(\mu_0 - \beta_1)}$ (consistent with step (4)), any first partial derivative of G with respect to $((v^1, \epsilon), y^1)$ is of order $e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 3\beta_1)}$ (i.e., is bounded in norm by a constant times this function), and, for $2 \leqslant i \leqslant r$, any partial derivative of order i of G with respect to $((v^1, \epsilon), y^1)$ is of order $e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 2i\beta_1)}$.

(6) For $\|(\bar{\tau}, \bar{x}^0, w^1)\| \leqslant K e^{-\frac{K_1}{\epsilon^a}(\mu_0 - \beta_1)}$ (consistent with step (4)), $\|D_{(\bar{\tau}, \bar{x}^0, w^1)} G((\bar{\tau}, \bar{x}^0, w^1), (v^1, \epsilon), y^1)^{-1}\|$ is of order $e^{\frac{K_2}{\epsilon^a}\beta}$.

(7) Any first partial derivative of $(\bar{\tau}, \bar{x}^0, w^1)$ with respect to $((v^1, \epsilon), y^1)$ is of order $e^{-\frac{K_1}{\epsilon^a}(\mu_0 - 4\beta_1)}$, and, for $2 \leqslant i \leqslant r$, any partial derivative of order i of $(\bar{\tau}, \bar{x}^0, w^1)$ with respect to $((v^1, \epsilon), y^1)$ is of order $e^{-\frac{K_1}{\epsilon^a}(\mu_0 - (2i+1)\beta_1)}$.

The last step implies the result: using (6.3), the desired estimates on $\tilde{w}$ for $0 < \epsilon < \epsilon_0$ are immediate, and those on $\tilde{x}$ for $0 < \epsilon < \epsilon_0$ follow from Deng's lemma. If we extend $\tilde{w}$ and $\tilde{x}$ to be 0 for $\epsilon = 0$, then these estimates, together with l'Hopital's rule, imply that the extended $\tilde{w}$ and $\tilde{x}$ are C^r.

We gather some more useful facts in the following lemma. Here and throughout this section, we shall use, for example, $\frac{\partial}{\partial w^1}$ to denote the matrix of partial derivatives more properly denoted by D_{w^1}.

Lemma 6.3.

(1) $\frac{\partial u^0}{\partial \tau}(0, \tau, 0, 0, v^1, w^1, \epsilon)$ *depends only on* ϵ *and is* $\geqslant K_3 \epsilon^a$.
(2) $\frac{\partial w^0}{\partial \tau}(0, \tau, 0, 0, v^1, w^1, \epsilon) = 0$.
(3) *For* $1 \leqslant i \leqslant r+1$, *any partial derivative of* $(u^0, v^0, w^0)(0, \tau, 0, 0, v^1, w^1, \epsilon)$ *of order* i *that includes* j *derivatives with respect to* τ *is of order* $e^{(i-j)\beta_0 \tau}$.
(4) *For* $1 \leqslant i \leqslant r+1$, *any partial derivative of* $\tau(v^1, w^1, \epsilon)$ *of order* i *is of order* $e^{i\beta\tau}$.
(5) *For* $1 \leqslant i \leqslant r$, *any partial derivative of* x^0 *given by* (6.11) *of order* i *that includes* j *derivatives with respect to* τ *is of order* $e^{(i-j)\beta_0 \tau}$.
(6) *For* $1 \leqslant i \leqslant r$, *any partial derivative of* $\check{x}$ *of order* i *is of order* $e^{i\beta\tau}$.

Proof. The functions $u^0(0, \tau, 0, 0, v^1, w^1, \epsilon)$, etc., are just components of $\phi_\epsilon(-\tau, c(u^1, v^1, \epsilon))$. (1) and (2) follow from (E7) and (E6). (3) is based on (3.7) and is a general fact about the derivatives of the flow of a C^{r+1} differential equation; compare [11, Proposition 3.2]. To prove (4), note that from (6.4),

$$\begin{aligned} &\frac{\partial u^0}{\partial \tau}\bigl(0, \tau\bigl(v^1, w^1, \epsilon\bigr), 0, 0, v^1, w^1, \epsilon\bigr) \frac{\partial \tau}{\partial w^1}\bigl(v^1, w^1, \epsilon\bigr) \\ &\quad + \frac{\partial u^0}{\partial w^1}\bigl(0, \tau\bigl(v^1, w^1, \epsilon\bigr), 0, 0, v^1, w^1, \epsilon\bigr) = 0. \end{aligned} \tag{6.15}$$

Therefore

$$\begin{aligned} \frac{\partial \tau}{\partial w^1}\bigl(v^1, w^1, \epsilon\bigr) = &-\left(\frac{\partial u^0}{\partial \tau}\bigl(0, \tau\bigl(v^1, w^1, \epsilon\bigr), 0, 0, v^1, w^1, \epsilon\bigr)\right)^{-1} \\ &\times \frac{\partial u^0}{\partial w^1}\bigl(0, \tau\bigl(v^1, w^1, \epsilon\bigr), 0, 0, v^1, w^1, \epsilon\bigr). \end{aligned}$$

Then from (1), (3), and Lemma 6.2,

$$\left\|\frac{\partial\tau}{\partial w^1}\right\| \leqslant \frac{1}{K_3\epsilon^a}Ke^{\beta_0\tau} = \frac{K}{K_3K_1}\frac{K_1}{\epsilon^a}e^{\beta_0\tau} \leqslant \frac{K}{K_3K_1}\tau e^{\beta_0\tau} = \frac{K}{K_3K_1}e^{\ln\tau+\beta_0\tau} \leqslant \tilde{K}e^{\beta\tau}.$$

The same estimates hold for $\frac{\partial\tau}{\partial v^1}$ and $\frac{\partial\tau}{\partial\epsilon}$. The general result follows by induction using (3).

To prove (5) for $i=1$, note that partial derivatives of $\hat{x}$ are bounded, and partial derivatives of $v_0(0,\tau,0,0,v^1,w^1,\epsilon)$ can be estimated by (3). The general result follows by induction. (6) follows from (4) and (5). □

It is convenient to write

$$\begin{aligned} G\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1\big) = {} & G\big((0,0,0),(v^1,\epsilon),0\big) + I\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big) \\ & + J\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1\big) \end{aligned}$$

with

$$\begin{aligned} I\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big) &= \big(G\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),0\big) - G\big((0,0,0),(v^1,\epsilon),0\big)\big), \\ J\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1\big) &= \big(G\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1\big) - G\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),0\big)\big). \end{aligned}$$

Using (6.1) and Lemma 6.1, we see that

$$\begin{aligned} & G\big((0,0,0),(v^1,\epsilon),0\big) \\ & = \begin{pmatrix} \hat x(0,v^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon),\epsilon) - \hat x(0,v^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon),\epsilon) \\ u^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon) - \hat u(0,v^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon),\epsilon) \\ w^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon) - \hat w(0,v^0(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon),\epsilon) \end{pmatrix} \\ & = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{6.16}$$

Equation (6.16) expresses the fact that (6.7)–(6.10) is a family of solutions of (6.2). It follows that $G = I + J$.

Again using (6.1) and Lemma 6.1, we obtain

$$\begin{aligned} & I\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big) \\ & = \begin{pmatrix} \hat x(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon) + \bar x^0 - \hat x(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon) \\ u^0(0,\tau,x^0,0,v^1,w^1,\epsilon) - \hat u(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon) \\ w^0(0,\tau,x^0,0,v^1,w^1,\epsilon) - \hat w(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon) \end{pmatrix} \\ & = \begin{pmatrix} \bar x^0 \\ u^0(0,\tau,0,0,v^1,w^1,\epsilon) \\ w^0(0,\tau,0,0,v^1,w^1,\epsilon) \end{pmatrix}, \end{aligned} \tag{6.17}$$

with τ given by (6.12). From Lemma 6.1(1) we have

$$\begin{aligned}
&J\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1\big)\\
&=\begin{pmatrix}0\\ u^0(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-u^0(0,\tau,x^0,0,v^1,w^1,\epsilon)\\ w^0(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-w^0(0,\tau,x^0,0,v^1,w^1,\epsilon)\end{pmatrix}\\
&\quad-\begin{pmatrix}\hat x((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat x(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon)\\ \hat u((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat u(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon)\\ \hat w((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat w(0,v^0(0,\tau,x^0,0,v^1,w^1,\epsilon),\epsilon)\end{pmatrix}\\
&=\begin{pmatrix}0\\ u^0(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-u^0(0,\tau,0,0,v^1,w^1,\epsilon)\\ w^0(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-w^0(0,\tau,0,0,v^1,w^1,\epsilon)\end{pmatrix}\\
&\quad-\begin{pmatrix}\hat x((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat x(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon)\\ \hat u((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat u(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon)\\ \hat w((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon)-\hat w(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon)\end{pmatrix} \qquad (6.18)
\end{aligned}$$

with τ given by (6.12) and x^0 given by (6.13).

Step 1. Using $G = I + J$, (6.17), and (6.18), we have

$$\begin{aligned}
&G\big((0,0,0),(v^1,\epsilon),y^1\big)\\
&=\begin{pmatrix}0\\ u^0(0,\tau,0,0,v^1,0,\epsilon)\\ w^0(0,\tau,0,0,v^1,0,\epsilon)\end{pmatrix}+\begin{pmatrix}0\\ u^0(0,\tau,x^0,y^1,v^1,0,\epsilon)-u^0(0,\tau,0,0,v^1,0,\epsilon)\\ w^0(0,\tau,x^0,y^1,v^1,0,\epsilon)-w^0(0,\tau,0,0,v^1,0,\epsilon)\end{pmatrix}\\
&\quad-\begin{pmatrix}\hat x((y^0,v^0)(0,\tau,x^0,y^1,v^1,0,\epsilon),\epsilon)-\hat x(0,v^0(0,\tau,0,0,v^1,0,\epsilon),\epsilon)\\ \hat u((y^0,v^0)(0,\tau,x^0,y^1,v^1,0,\epsilon),\epsilon)-\hat u(0,v^0(0,\tau,0,0,v^1,0,\epsilon),\epsilon)\\ \hat w((y^0,v^0)(0,\tau,x^0,y^1,v^1,0,\epsilon),\epsilon)-\hat w(0,v^0(0,\tau,0,0,v^1,0,\epsilon),\epsilon)\end{pmatrix} \qquad (6.19)
\end{aligned}$$

with $x^0 = \hat x(0, v^0(0,\tau,0,0,v^1,0,\epsilon),\epsilon)$ and $\tau = \tau(v^1,0,\epsilon)$. The first matrix is 0 by Lemma 6.1(3) and (4). The second is of order $e^{-\mu_0\tau(v^1,0,\epsilon)}$ by Deng's lemma (Theorem 2.2 of [11]). As for the third, let us consider its first line. In norm it is at most a bound on the first partial derivatives of $\hat x$ times

$$\big\|(y^0,v^0)\big(0,\tau,x^0,y^1,v^1,w^1,\epsilon\big)-\big(0,v^0\big(0,\tau,0,0,v^1,w^1,\epsilon\big)\big)\big\|,$$

which by Deng's lemma is of order $e^{-\mu_0\tau(v^1,0,\epsilon)}$. The other lines of the matrix are treated analogously.

Step 2. Since $J = 0$ when $y^1 = 0$, we have $G((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),0) = I((\bar\tau,\bar x^0,w^1),(v^1,\epsilon))$. $D_{(\bar\tau,\bar x^0,w^1)}G((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),y^1)$ can be regarded as 3×3 block-partitioned matrix. Using (6.17) and Lemma 6.3(2) we have

$$
\begin{aligned}
&D_{(\bar\tau,\bar x^0,w^1)}G\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon),0\big)\\
&\quad = D_{(\bar\tau,\bar x^0,w^1)}I\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big)\\
&\quad = \begin{pmatrix} 0 & I & 0\\ \frac{\partial u^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon) & 0 & \frac{\partial u^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon)\frac{\partial\tau}{\partial w^1}(v^1,w^1,\epsilon)+\frac{\partial u^0}{\partial w^1}(0,\tau,0,0,v^1,w^1,\epsilon)\\ 0 & 0 & \frac{\partial w^0}{\partial w^1}(0,\tau,0,0,v^1,w^1,\epsilon)\end{pmatrix},
\end{aligned}
$$

with $\tau=\tau(v^1,w^1,\epsilon)+\bar\tau$. From (6.15), it follows that

$$
\begin{aligned}
&D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0),(v^1,\epsilon),0\big)\\
&\quad = D_{(\bar\tau,\bar x^0,w^1)}I\big((0,0,0),(v^1,\epsilon)\big)\\
&\quad = \begin{pmatrix} 0 & I & 0\\ \frac{\partial u^0}{\partial\tau}(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon) & 0 & 0\\ 0 & 0 & \frac{\partial w^0}{\partial w^1}(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon)\end{pmatrix}.
\end{aligned}
\tag{6.20}
$$

If $\frac{\partial w^0}{\partial w^1}$ is invertible, we have

$$
\begin{aligned}
&\big(D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0),(v^1,\epsilon),0\big)\big)^{-1}\\
&\quad = \begin{pmatrix} 0 & (\frac{\partial u^0}{\partial\tau}(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon))^{-1} & 0\\ I & 0 & 0\\ 0 & 0 & (\frac{\partial w^0}{\partial w^1}(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon))^{-1}\end{pmatrix}.
\end{aligned}
\tag{6.21}
$$

By Lemma 6.3(1), $(\frac{\partial u^0}{\partial\tau}(0,\tau(v^1,0,\epsilon),0,0,v^1,0,\epsilon))^{-1}$ is of order ϵ^{-a}, hence of order $e^{\frac{K_1}{\epsilon^a}\beta}$, hence of order $e^{\beta\tau(v_1,0,\epsilon)}$.

For $c\in V_*$ and $\phi_\epsilon(t,c)\in V^*$, let us write $\phi_\epsilon(t,c)$ as $(v^1,w^1)(t,u^0,v^0,w^0,\epsilon)$. Then

$$
w^1\big(\tau\big(v^1,w^1,\epsilon\big),0,\big(v^0,w^0\big)\big(0,\tau\big(v^1,w^1,\epsilon\big),0,0,v^1,w^1,\epsilon\big),\epsilon\big)=w^1.
$$

Therefore

$$
\frac{\partial w^1}{\partial\tau}\frac{\partial\tau}{\partial w^1}+\frac{\partial w^1}{\partial v^0}\left(\frac{\partial v^0}{\partial\tau}\frac{\partial\tau}{\partial w^1}+\frac{\partial v^0}{\partial w^1}\right)+\frac{\partial w^1}{\partial w^0}\left(\frac{\partial w^0}{\partial\tau}\frac{\partial\tau}{\partial w^1}+\frac{\partial w^0}{\partial w^1}\right)=I. \tag{6.22}
$$

From Lemma 6.3(2), $\frac{\partial w^0}{\partial\tau}=0$. From the definition of w^1 in Section 3, if $w^1=0$, the terms $\frac{\partial w^1}{\partial\tau}$ and $\frac{\partial w^1}{\partial v^0}$ also vanish. Hence, for $w^1=0$, (6.22) reduces to

$$
\frac{\partial w^1}{\partial w^0}\frac{\partial w^0}{\partial w^1}=I.
$$

We therefore see that the inverse of $\frac{\partial w^0}{\partial w^1}(0, \tau(v^1, 0, \epsilon), 0, 0, v^1, 0, \epsilon)$ is

$$\frac{\partial w^1}{\partial w^0}\big(\tau\big(v^1, 0, \epsilon\big), 0, v^0\big(0, \tau\big(v^1, 0, \epsilon\big), 0, 0, v^1, 0, \epsilon\big), 0, \epsilon\big).$$

By the analogue of Lemma 6.3(3) for the forward flow $\phi_\epsilon(t, c^0)$, this expression is of order $e^{\beta\tau(v^1,0,\epsilon)}$.

Step 3. We have

$$\begin{aligned}
&D_{(\bar\tau,\bar x^0,w^1)}G\big(\big(\bar\tau, \bar x^0, w^1\big), \big(v^1, \epsilon\big), y^1\big) - D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0), \big(v^1, \epsilon\big), 0\big)\\
&\quad = D_{(\bar\tau,\bar x^0,w^1)}I\big(\big(\bar\tau, \bar x^0, w^1\big), \big(v^1, \epsilon\big)\big) - D_{(\bar\tau,\bar x^0,w^1)}I\big((0,0,0), \big(v^1, \epsilon\big)\big)\\
&\qquad + D_{(\bar\tau,\bar x^0,w^1)}J\big(\big(\bar\tau, \bar x^0, w^1\big), \big(v^1, \epsilon\big), y_1\big).
\end{aligned}$$

We will show in Step 5 that all first partial derivatives of J are of order $e^{-(\mu_0-2\beta)\tau}$ (see Propositions 6.6 and 6.7). We therefore consider

$$D_{(\bar\tau,\bar x^0,w^1)}I\big(\big(\bar\tau, \bar x^0, w^1\big), \big(v^1, \epsilon\big)\big) - D_{(\bar\tau,\bar x^0,w^1)}I\big((0,0,0), \big(v^1, \epsilon\big)\big).$$

Both matrices were calculated in Step 2. The difference has three nonzero terms:

(1) $\frac{\partial u^0}{\partial \tau}(0, \tau, 0, 0, v^1, w^1, \epsilon) - \frac{\partial u^0}{\partial \tau}(0, \tau(v^1, 0, \epsilon), 0, 0, v^1, 0, \epsilon)$,
(2) $\frac{\partial u^0}{\partial \tau}(0, \tau, 0, 0, v^1, w^1, \epsilon)\frac{\partial \tau}{\partial w^1}(v^1, w^1, \epsilon) + \frac{\partial u^0}{\partial w^1}(0, \tau, 0, 0, v^1, w^1, \epsilon)$,
(3) $\frac{\partial w^0}{\partial w^1}(0, \tau, 0, 0, v^1, w^1, \epsilon) - \frac{\partial w^0}{\partial w^1}(0, \tau(v^1, 0, \epsilon), 0, 0, v^1, 0, \epsilon)$,

with $\tau = \tau(v^1, w^1, \epsilon) + \bar\tau$. We consider each term.

(1) By Lemma 6.3(1), $\frac{\partial u^0}{\partial \tau}$ depends only on ϵ, so this term is 0.

(2) $\frac{\partial u^0}{\partial w^1}(0, \tau, 0, 0, v^1, w^1, \epsilon)$ is independent of τ, because

$$\frac{\partial^2 u^0}{\partial \tau \partial w^1} = \frac{\partial^2 u^0}{\partial w^1 \partial \tau} = 0.$$

Since $\frac{\partial u^0}{\partial \tau}$ is also independent of τ, we rewrite this term as (6.15) and see that it is 0.

(3) We rewrite this term as

$$\begin{aligned}
&\frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big) + \bar\tau, 0, 0, v^1, w^1, \epsilon\big) - \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, 0, \epsilon\big), 0, 0, v^1, 0, \epsilon\big)\\
&\quad = \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big) + \bar\tau, 0, 0, v^1, w^1, \epsilon\big) - \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big), 0, 0, v^1, w^1, \epsilon\big)\\
&\qquad + \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big), 0, 0, v^1, w^1, \epsilon\big) - \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big), 0, 0, v^1, 0, \epsilon\big)\\
&\qquad + \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, w^1, \epsilon\big), 0, 0, v^1, 0, \epsilon\big) - \frac{\partial w^0}{\partial w^1}\big(0, \tau\big(v^1, 0, \epsilon\big), 0, 0, v^1, 0, \epsilon\big).
\end{aligned}$$

In norm this is at most

$$\sup_{\sigma\in[0,\bar\tau]}\left\|\frac{\partial^2 w^0}{\partial\tau\partial w^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\sigma,0,0,v^1,w^1,\epsilon\big)\right\||\bar\tau|$$
$$+\sup_{w\in[0,w^1]}\left\|\frac{\partial^2 w^0}{\partial(w^1)^2}\big(0,\tau\big(v^1,w^1,\epsilon\big),0,0,v^1,w,\epsilon\big)\right\|\|w^1\|$$
$$+\sup_{\sigma\in[\tau(v^1,0,\epsilon),\tau(v^1,w^1,\epsilon)]}\left\|\frac{\partial^2 w^0}{\partial\tau\partial w^1}\big(0,\sigma,0,0,v^1,0,\epsilon\big)\right\|\big|\tau\big(v^1,w^1,\epsilon\big)-\tau\big(v^1,0,\epsilon\big)\big|.$$

By Lemma 6.3(3), the first two summands are at most

$$Ke^{2\beta(\tau(v^1,w^1,\epsilon)+|\bar\tau|)}\big\|\big(\bar\tau,\bar x^0,w^1\big)\big\|.$$

By Lemma 6.3(3) and (4), the third summand is at most

$$\sup_{\sigma\in[\tau(v^1,0,\epsilon),\tau(v^1,w^1,\epsilon)]}\left\|\frac{\partial^2 w^0}{\partial\tau\partial w^1}\big(0,\sigma,0,0,v^1,0,\epsilon\big)\right\|\sup_{w\in[0,w^1]}\left\|\frac{\partial\tau}{\partial w^1}\big(v^1,w,\epsilon\big)\right\|\|w^1\|$$
$$\leqslant Ke^{\beta\tau_2(\epsilon)}\cdot Ke^{\beta\tau(v^1,w^1,\epsilon)}\|w^1\|.$$

Putting everything together, we have the result.

Step 4. From Step 1 and (E11),

$$\big\|G\big((0,0,0),\big(v^1,\epsilon\big),y^1\big)\big\|\leqslant Ke^{-\frac{K_1}{\epsilon^a}\mu_0}. \tag{6.23}$$

From Step 2 and (E11),

$$\big\|D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0),\big(v^1,\epsilon\big),0\big)^{-1}\big\|\leqslant Ke^{\frac{K_2}{\epsilon^a}\beta}. \tag{6.24}$$

From Step 3 and Lemma 6.2, if $\|(\bar\tau,\bar x^0,w^1)\|\leqslant\delta$, then

$$\big\|D_{(\bar\tau,\bar x^0,w^1)}G\big(\big(\bar\tau,\bar x^0,w^1\big),\big(v^1,\epsilon\big),y^1\big)-D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0),\big(v^1,\epsilon\big),0\big)\big\|$$
$$\leqslant Ke^{-(\mu_0-2\beta)(\frac{K_1}{\epsilon^a}-\delta)}+Ke^{2\beta(\frac{K_2}{\epsilon^a}+\delta)}\delta. \tag{6.25}$$

For $\delta<1$ and ϵ small, using $\beta<\beta_1$ and $K_2\beta<K_1\beta_1$, (6.25) implies

$$\big\|D_{(\bar\tau,\bar x^0,w^1)}G\big(\big(\bar\tau,\bar x^0,w^1\big),\big(v^1,\epsilon\big),y^1\big)-D_{(\bar\tau,\bar x^0,w^1)}G\big((0,0,0),\big(v^1,\epsilon\big),0\big)\big\|$$
$$\leqslant Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-2\beta_1)}+Ke^{\frac{2K_1}{\epsilon^a}\beta_1}\delta. \tag{6.26}$$

We define

$$m=Ke^{-\frac{K_1}{\epsilon^a}\mu_0},\quad n=Ke^{\frac{K_2}{\epsilon^a}\beta},\quad r=Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-2\beta_1)},\quad q=Ke^{\frac{2K_1}{\epsilon^a}\beta_1},\quad p=r+q\delta.$$

With these definitions, (6.23), (6.24), and (6.26) show that hypotheses (I1)–(I3) of Theorem 5.1 are satisfied.

Motivated by the proof of Theorem 5.1, we wish to choose δ to be the smaller of the two solutions of the equation

$$\delta = \frac{nm}{1-np} = \frac{nm}{1-n(r+q\delta)}.$$

Therefore

$$\delta = \frac{1}{2nq}\big(1-nr-\big((1-nr)^2-4n^2qm\big)^{\frac{1}{2}}\big). \tag{6.27}$$

Using $K_2\beta < K_1\beta_1$, we have

$$nr = K^2 e^{\frac{1}{\epsilon^a}(K_2\beta - K_1\mu_0 + 2K_1\beta_1)} \leqslant K^2 e^{-\frac{K_1}{\epsilon^a}(\mu_0 - 3\beta_1)},$$
$$4n^2qm = 4K^4 e^{\frac{1}{\epsilon^a}(2K_2\beta + 2K_1\beta_1 - K_1\mu_0)} \leqslant 4K^4 e^{-\frac{1}{\epsilon^a}(\mu_0 - 4\beta_1)K_1}.$$

Therefore nr and $4n^2qm$ approach 0 as $\epsilon \to 0$, so for small $\epsilon > 0$ the definition (6.27) yields a positive number for δ.

We now easily see that $\delta < \frac{1}{2nq}$, so

$$np = nr + nq\delta < nr + nq\frac{1}{2nq} = nr + \frac{1}{2}.$$

Therefore $np < \frac{3}{4}$ for small $\epsilon > 0$, so hypothesis (I4) of Theorem 5.1 is satisfied. Then

$$\delta = \frac{nm}{1-np} < 4nm = 4K^2 e^{\frac{1}{\epsilon^a}(K_2\beta - K_1\mu_0)} \leqslant 4K^2 e^{\frac{1}{\epsilon^a}(K_1\beta_1 - K_1\mu_0)} = 2K^2 e^{-\frac{K_1}{\epsilon^a}(\mu_0 - \beta_1)}. \tag{6.28}$$

By (E11), $\mu_0 - 6\beta_1 > 0$, so in particular $\mu_0 - \beta_1 > 3\beta_1$. Therefore

$$\delta < 2K^2 e^{-\frac{3K_1}{\epsilon^a}\beta_1} < 2K^2 e^{-\frac{3K_2}{\epsilon^a}\beta} < e^{-\frac{2K_2}{\epsilon^a}\beta}$$

for ϵ small. Recall from Lemma 6.2 that $\tau(v^1, w^1, \epsilon)$, and hence G, is defined for $\|w^1\| < e^{-\frac{2K_2}{\epsilon^a}\beta}$. Therefore, for small ϵ, hypothesis (I5) of Theorem 5.1 holds. Therefore Theorem 5.1 applies, and the desired estimate on δ is given by (6.28).

Step 5. We use $G = I + J$, and consider separately I and J. From (6.17),

$$\begin{pmatrix} I_1((\bar\tau, \bar x^0, w^1), (v^1, \epsilon)) \\ I_2((\bar\tau, \bar x^0, w^1), (v^1, \epsilon)) \\ I_3((\bar\tau, \bar x^0, w^1), (v^1, \epsilon)) \end{pmatrix} = \begin{pmatrix} \bar x^0 \\ u^0(0, \tau(v^1, w^1, \epsilon) + \bar\tau, 0, 0, v^1, w^1, \epsilon) \\ w^0(0, \tau(v^1, w^1, \epsilon) + \bar\tau, 0, 0, v^1, w^1, \epsilon) \end{pmatrix}.$$

Proposition 6.4. *For* $1 \leqslant i \leqslant r$:

(1) $\|\frac{\partial^i I_2}{\partial \epsilon^i}\|$ *is of order* $e^{-\frac{K_1}{\epsilon^a}(\mu_0 - (i+1)\beta_1)}$.

(2) *Any* i*th partial derivative of* I_2 *with respect to* (v^1,ϵ) *in which* I_2 *is differentiated at least once with respect to* v^1 *is* 0.

(3) *Any* i*th partial derivative of* I_2 *with respect to* (v^1,ϵ) *is of order* $e^{-\frac{K_1}{\epsilon^d}(\mu_0-(i+2)\beta_1)}$.

Proof. (1) We first consider $\frac{\partial I_2}{\partial\epsilon}$. From Lemma 6.3(1), $\frac{\partial u^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon)$ is independent of τ. Using this fact and Lemma 6.1(3), we have

$$\begin{aligned}
&\frac{\partial I_2}{\partial\epsilon}\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big)\\
&\quad=\frac{\partial u^0}{\partial\tau}\big(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\frac{\partial\tau}{\partial\epsilon}+\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\\
&\quad=\frac{\partial u^0}{\partial\tau}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big)\frac{\partial\tau}{\partial\epsilon}+\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big)\\
&\qquad+\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon\big)-\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big)\\
&\quad=\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon\big)-\frac{\partial u^0}{\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big).
\end{aligned}\tag{6.29}$$

In norm this is at most

$$\sup_{\sigma\in[0,\bar\tau]}\left\|\frac{\partial^2u^0}{\partial\tau\,\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon)+\sigma,0,0,v^1,w^1,\epsilon\big)\right\||\bar\tau|.\tag{6.30}$$

Lemma 6.3(1) implies that $\frac{\partial^2u^0}{\partial\tau\partial\epsilon}(0,\tau,0,0,v^1,w^1,\epsilon)$ depends only on ϵ, because

$$\frac{\partial^2u^0}{\partial\tau\,\partial\epsilon}\big(0,\tau,0,0,v^1,w^1,\epsilon\big)=\frac{\partial^2u^0}{\partial\epsilon\,\partial\tau}\big(0,\tau,0,0,v^1,w^1,\epsilon\big).$$

Therefore, using Lemma 6.3(3) and $K_2\beta<K_1\beta_1$, (6.30) is at most

$$\left\|\frac{\partial^2u^0}{\partial\tau\,\partial\epsilon}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big)\right\||\bar\tau|\leqslant Ke^{\frac{K_2}{\epsilon^d}\beta}\cdot Ke^{-\frac{K_1}{\epsilon^d}(\mu-\beta_1)}\leqslant K^2e^{-\frac{K_1}{\epsilon^d}(\mu-2\beta_1)}.$$

Next we consider $\frac{\partial^2I_2}{\partial\epsilon^2}$. Since $\frac{\partial^2u^0}{\partial\tau\partial\epsilon}(0,\tau,0,0,v^1,w^1,\epsilon)$ depends only on ϵ, (6.29) yields

$$\begin{aligned}
&\frac{\partial^2I_2}{\partial\epsilon^2}\big((\bar\tau,\bar x^0,w^1),(v^1,\epsilon)\big)\\
&\quad=\frac{\partial^2u^0}{\partial\epsilon^2}\big(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon\big)-\frac{\partial^2u^0}{\partial\epsilon^2}\big(0,\tau(v^1,w^1,\epsilon),0,0,v^1,w^1,\epsilon\big),
\end{aligned}$$

which by a similar argument has norm of order $e^{-\frac{K_1}{\epsilon^d}(\mu_0-3\beta_1)}$.

The general result follows by induction.

(2) $\frac{\partial u^0}{\partial v^1}(0,\tau,0,0,v^1,w^1,\epsilon)$ is independent of τ, because, from Lemma 6.3(1),

$$\frac{\partial^2 u^0}{\partial \tau \partial v^1}\big(0,\tau,0,0,v^1,w^1,\epsilon\big)=\frac{\partial^2 u^0}{\partial v^1 \partial \tau}\big(0,\tau,0,0,v^1,w^1,\epsilon\big)=0.$$

Therefore, from Lemma 6.1(3),

$$\begin{aligned}
&\frac{\partial I_2}{\partial v^1}\big(\big(\bar\tau,\bar x^0,w^1\big),\big(v^1,\epsilon\big)\big)\\
&\quad=\frac{\partial u^0}{\partial \tau}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\frac{\partial \tau}{\partial v^1}+\frac{\partial u^0}{\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\\
&\quad=\frac{\partial u^0}{\partial \tau}\big(0,\tau\big(v^1,w^1,\epsilon\big),0,0,v^1,w^1,\epsilon\big)\frac{\partial \tau}{\partial v^1}+\frac{\partial u^0}{\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big),0,0,v^1,w^1,\epsilon\big)=0.
\end{aligned}$$

The result follows.

(3) From Lemma 6.3(2),

$$\begin{aligned}
&\frac{\partial I_3}{\partial v^1}\big(\big(\bar\tau,\bar x^0,w^1\big),\big(v^1,\epsilon\big)\big)\\
&\quad=\frac{\partial w^0}{\partial \tau}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\frac{\partial \tau}{\partial v^1}+\frac{\partial w^0}{\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\\
&\quad=\frac{\partial w^0}{\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big).
\end{aligned}$$

Since $\frac{\partial w^0}{\partial v^1}(0,\tau,0,0,v^1,0,\epsilon)=0$ (from the choice of the coordinate w^1 in Section 3),

$$\begin{aligned}
&\left\|\frac{\partial w^0}{\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big)\right\|\\
&\quad\leqslant \sup_{w\in[0,w^1]}\left\|\frac{\partial^2 w^0}{\partial w^1\partial v^1}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w,\epsilon\big)\right\|\big\|w^1\big\|\\
&\quad\leqslant Ke^{2\beta(\tau(v^1,w^1,\epsilon)+1)}\cdot Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-\beta_1)}\leqslant Ke^{2\beta(\frac{K_2}{\epsilon^a}+1)}\cdot Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-\beta_1)}\\
&\quad=K^2e^{\frac{1}{\epsilon^a}(2\beta(K_2+\epsilon^a)-K_1\mu_0+K_1\beta_1)}\leqslant K^2e^{\frac{K_1}{\epsilon^a}(3\beta_1-\mu_0)}=K^2e^{-\frac{K_1}{\epsilon^a}(\mu_0-3\beta_1)}.
\end{aligned}$$

We have used Lemma 6.3(3), $\bar\tau<1$ for ϵ small, and $K_2\beta<K_1\beta_1$.

The same estimate holds for $\frac{\partial I_3}{\partial \epsilon}((\bar\tau,\bar x^0,w^1),(v^1,\epsilon))$. Proceeding inductively, we find that for $1\leqslant j+k\leqslant r$,

$$\frac{\partial^{j+k} I_3}{\partial (v^1)^j\epsilon^k}\big(\big(\bar\tau,\bar x^0,w^1\big),\big(v^1,\epsilon\big)\big)=\frac{\partial^{j+k} w^0}{\partial (v^1)^j\epsilon^k}\big(0,\tau\big(v^1,w^1,\epsilon\big)+\bar\tau,0,0,v^1,w^1,\epsilon\big),$$

and $\|\frac{\partial^j w^0}{\partial (v^1)^k\epsilon^j}(0,\tau(v^1,w^1,\epsilon)+\bar\tau,0,0,v^1,w^1,\epsilon)\|\leqslant Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-(j+k+2)\beta_1)}$. □

Before continuing, we record the following immediate consequence of Deng's lemma (Theorem 2.2 of [11]).

Lemma 6.5. *For $1 \leqslant k \leqslant r$, any partial derivative of a component of $(u^0, v^0, w^0)(0, \tau, x^0, y^1, v^1, w^1, \epsilon)$ of order k, provided we differentiate at least once with respect to x^0 or y^0, is of order $e^{-(\mu_0 - k\beta_0)\tau}$.*

Next we turn to J, which is given in (6.18) as the sum of two matrices.

Proposition 6.6. *For the first of the two matrices that sum to J in* (6.18)*, any ith partial derivative, $1 \leqslant i \leqslant r$, is of order $e^{-(\mu_0 - 2i\beta)\tau}$.*

Proof. We first consider the expression

$$u^0\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - u^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big) \tag{6.31}$$

in (6.18), with τ and x^0 as given by (6.12) and (6.13). The partial derivative with respect to v^1 is

$$\begin{aligned}&\left(\frac{\partial u^0}{\partial \tau}\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - \frac{\partial u^0}{\partial \tau}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\right)\frac{\partial \tau}{\partial v^1} \\ &\quad + \frac{\partial u^0}{\partial x^0}\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big)\frac{\partial \check{x}}{\partial v^1} \\ &\quad + \left(\frac{\partial u^0}{\partial v^1}\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - \frac{\partial u^0}{\partial v^1}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\right).\end{aligned}$$

By Deng's lemma (Theorem 2.2 of [11]), the two differences are of order $e^{-(\mu_0 - \beta_0)\tau}$, and $\frac{\partial u^0}{\partial x^0}(0, \tau, x^0, y^1, v^1, w^1, \epsilon)$ is also of order $e^{-(\mu_0 - \beta_0)\tau}$. From Lemma 6.3, $\frac{\partial \tau}{\partial v^1}$ and $\frac{\partial \check{x}}{\partial v^1}$ are of order $e^{\beta\tau}$. We conclude that the entire expression is of order $e^{-(\mu_0 - 2\beta)\tau}$. A similar argument applies to the partial derivative of (6.31) with respect to any variable.

The result follows by induction. It is important to note that according to Lemma 6.5, any partial derivative of $v^0(0, \tau, x^0, y^1, v^1, w^1, \epsilon)$ of order k, provided we differentiate at least once with respect to x^0 or y^0, is of order $e^{-(\mu_0 - k\beta_0)\tau}$.

The same argument applies to $w^0(0, \tau, x^0, y^1, v^1, w^1, \epsilon) - w^0(0, \tau, 0, 0, v^1, w^1, \epsilon)$ in (6.18). □

Proposition 6.7. *For the second of the two matrices that sum to J in* (6.18)*, any ith partial derivative, $1 \leqslant i \leqslant r$, is of order $e^{-(\mu_0 - 2i\beta)\tau}$.*

Proof. We consider the expression

$$\hat{x}\big(\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big) - \hat{x}\big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big) \tag{6.32}$$

in (6.18), with τ given by (6.12) and x^0 given by (6.13).

The partial derivative with respect to v^1 of this composite function is

$$\frac{\partial\hat{x}}{\partial y^0}\left(\frac{\partial y^0}{\partial\tau}\frac{\partial\tau}{\partial v^1}+\frac{\partial y^0}{\partial x^0}\frac{\partial\check{x}}{\partial v^1}+\frac{\partial y^0}{\partial v^1}\right)+\frac{\partial\hat{x}}{\partial v^0}\left(\frac{\partial v^0}{\partial\tau}\frac{\partial\tau}{\partial v^1}+\frac{\partial v^0}{\partial x^0}\frac{\partial\check{x}}{\partial v^1}+\frac{\partial v^0}{\partial v^1}\right)$$
$$-\frac{\partial\hat{x}}{\partial v^0}\left(\frac{\partial v^0}{\partial\tau}\frac{\partial\tau}{\partial v^1}+\frac{\partial v^0}{\partial v^1}\right), \tag{6.33}$$

where we have suppressed the points at which partial derivatives are evaluated; thus terms that appear to cancel do not.

In the term $\frac{\partial\hat{x}}{\partial y^0}(\frac{\partial y^0}{\partial\tau}\frac{\partial\tau}{\partial v^1}+\frac{\partial y^0}{\partial x^0}\frac{\partial\check{x}}{\partial v^1}+\frac{\partial y^0}{\partial v^1})$, $\frac{\partial\hat{x}}{\partial y^0}$ is bounded; by Deng's lemma (Theorem 2.2 of [11]), partial derivatives of y^0 are of order $e^{-(\mu_0-\beta)\tau}$; and by Lemma 6.3, partial derivatives of τ and $\check{x}$ are of order $e^{\beta\tau}$. Therefore this term is of order $e^{-(\mu_0-2\beta)\tau}$. One can see inductively that for $2\leqslant i\leqslant r$, if one differentiates this term $i-1$ more times with respect to any combination of the variables, the result is of order $e^{-(\mu_0-2i\beta)\tau}$.

A similar argument applies to the product $\frac{\partial\hat{x}}{\partial v^0}\frac{\partial v^0}{\partial x^0}\frac{\partial\check{x}}{\partial v^1}$ in the middle term. It is important to note that according to Lemma 6.5, any partial derivative of $v^0(0,\tau,x^0,y^1,v^1,w^1,\epsilon)$ of order k, provided we differentiate at least once with respect to x^0 or y^0, is of order $e^{-(\mu_0-k\beta_0)\tau}$.

The remaining terms in (6.33), taking into account where they are evaluated, are

$$\begin{aligned}
&\frac{\partial\hat{x}}{\partial v^0}\big((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon\big)\\
&\quad\times\left(\frac{\partial v^0}{\partial\tau}(0,\tau,x^0,y^1,v^1,w^1,\epsilon)\frac{\partial\tau}{\partial v^1}(v^1,w^1,\epsilon)+\frac{\partial v^0}{\partial v^1}(0,\tau,x^0,y^1,v^1,w^1,\epsilon)\right)\\
&\quad-\frac{\partial\hat{x}}{\partial v^0}\big(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon\big)\\
&\quad\times\left(\frac{\partial v^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon)\frac{\partial\tau}{\partial v^1}(v^1,w^1,\epsilon)+\frac{\partial v^0}{\partial v^1}(0,\tau,0,0,v^1,w^1,\epsilon)\right)\\
&=\frac{\partial\hat{x}}{\partial v^0}\big((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon\big)\\
&\quad\times\left(\left(\frac{\partial v^0}{\partial\tau}(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-\frac{\partial v^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon)\right)\frac{\partial\tau}{\partial v^1}(v^1,w^1,\epsilon)\right.\\
&\quad\left.+\frac{\partial v^0}{\partial v^1}(0,\tau,x^0,y^1,v^1,w^1,\epsilon)-\frac{\partial v^0}{\partial v^1}(0,\tau,0,0,v^1,w^1,\epsilon)\right)\\
&\quad+\left(\frac{\partial\hat{x}}{\partial v^0}\big((y^0,v^0)(0,\tau,x^0,y^1,v^1,w^1,\epsilon),\epsilon\big)-\frac{\partial\hat{x}}{\partial v^0}\big(0,v^0(0,\tau,0,0,v^1,w^1,\epsilon),\epsilon\big)\right)\\
&\quad\times\left(\frac{\partial v^0}{\partial\tau}(0,\tau,0,0,v^1,w^1,\epsilon)\frac{\partial\tau}{\partial v^1}(v^1,w^1,\epsilon)+\frac{\partial v^0}{\partial v^1}(0,\tau,0,0,v^1,w^1,\epsilon)\right).
\end{aligned}\tag{6.34}$$

We consider (6.34) to be the sum of three terms. The first summand,

$$\frac{\partial \hat{x}}{\partial v^0}\big(\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big)$$
$$\times \left(\frac{\partial v^0}{\partial \tau}\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - \frac{\partial v^0}{\partial \tau}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\right)\frac{\partial \tau}{\partial v^1}\big(v^1, w^1, \epsilon\big), \quad (6.35)$$

is itself a product of three terms.

The first is bounded; the second, by Deng's lemma (Theorem 2.2 of [11]), is of order $e^{-(\mu_0-\beta)\tau}$; the third is of order $e^{\beta\tau}$. Thus the product is of order $e^{-(\mu_0-2\beta)\tau}$. Moreover, for $2 \leqslant i \leqslant r$, if one differentiates this product $i-1$ more times with respect to some combination of the variables, the result is a sum of products; each product is a jth derivative of the first term times a kth derivative of the second times an lth derivative of the third, with $j+k+l = i-1$. These derivatives are respectively of order $e^{j\beta\tau}$, $e^{-(\mu-(2k+1)\beta)\tau}$, and $e^{(l+1)\beta\tau}$; hence each product is of order $e^{-(\mu-(j+2k+l+2)\beta)\tau}$. Since $j+2k+l+2 = (j+k+l+1)+(k+1) \leqslant 2i$, each product is of order $e^{-(\mu-2i\beta)\tau}$.

The second summand in (6.34),

$$\frac{\partial \hat{x}}{\partial v^0}\big(\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big)$$
$$\times \left(\frac{\partial v^0}{\partial v^1}\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - \frac{\partial v^0}{\partial v^1}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\right),$$

can be treated similarly.

The third summand in (6.34),

$$\left(\frac{\partial \hat{x}}{\partial v^0}\big(\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big) - \frac{\partial \hat{x}}{\partial v^0}\big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big)\right)$$
$$\times \left(\frac{\partial v^0}{\partial \tau}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\frac{\partial \tau}{\partial v^1}\big(v^1, w^1, \epsilon\big) + \frac{\partial v^0}{\partial v^1}\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\right),$$

is a product of two terms. The first is at most a bound on the second partial derivatives of $\hat{x}$ times

$$\big\|\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big) - \big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big)\big)\big\|,$$

which by Deng's lemma is of order $e^{-\mu_0\tau}$. Using Lemma 6.3, we see that the second is of order $e^{\beta\tau}$. Thus the product is of order $e^{-(\mu_0-\beta)\tau}$.

For $2 \leqslant i \leqslant r$, if one differentiates this product $i-1$ more times with respect to some combination of the variables, the result is a sum of products; each product is a jth derivative of the first term times a kth derivative of the second, with $j+k = i-1$. Any kth derivative of the second term, $2 \leqslant k \leqslant r-1$, with respect to some combination of the variables, it is of order $e^{(k+1)\beta\tau}$. As to jth derivatives of the first term,

$$\frac{\partial \hat{x}}{\partial v^0}\big(\big(y^0, v^0\big)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big) - \frac{\partial \hat{x}}{\partial v^0}\big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big),$$

one must essentially repeat the study performed thus far on derivatives of (6.32). One thus sets up an induction that yields the result. Note that, just as our estimate on the size of this term used a bound on the second partial derivative of $\hat{x}$, a bound on

$$\frac{\partial^l \hat{x}}{\partial (v^0)^l}\big((y^0, v^0)\big(0, \tau, x^0, y^1, v^1, w^1, \epsilon\big), \epsilon\big) - \frac{\partial^l \hat{x}}{\partial (v^0)^l}\big(0, v^0\big(0, \tau, 0, 0, v^1, w^1, \epsilon\big), \epsilon\big)$$

will use a bound on the $(l+1)$st partial derivative of $\hat{x}$. Since $\hat{x}$ is C^{r+1}, we have such bounds through rth partial derivative of $\hat{x}$.

This completes the argument for partial derivatives of (6.32) where we differentiate at least once with respect to v^1 (since the argument began by differentiating (6.32) with respect to v^1). Other partial derivatives of (6.32), and partial derivatives of the other entries of the matrix, are treated similarly. $\square$

Step 6. Note that if a linear operator A is invertible and $\|C - A\| \leqslant \frac{1}{\|A^{-1}\|}$, then C is invertible, and

$$\big\|C^{-1} - A^{-1}\big\| \leqslant \frac{\|A^{-1}\|^2 \|C - A\|}{1 - \|A^{-1}\| \|C - A\|}. \tag{6.36}$$

From (6.26) and our estimate for δ,

$$\begin{aligned} &\big\| D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((\bar{\tau}, \bar{x}^0, w^1), (v^1, \epsilon), y^1\big) - D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((0,0,0), (v^1, \epsilon), 0\big)\big\| \\ &\quad \leqslant K e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 2\beta_1)} + K e^{\frac{2K_1}{\epsilon^d}\beta_1} K e^{-\frac{K_1}{\epsilon^d}(\mu_0 - \beta_1)} \leqslant 2K^2 e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 3\beta_1)}. \end{aligned} \tag{6.37}$$

Since $\mu_0 - 4\beta_1 > 0$, $K_1(\mu_0 - 3\beta_1) > K_1\beta_1 > K_2\beta$. Then from (6.37) and (6.24), for small $\epsilon > 0$,

$$\begin{aligned} &\big\| D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((\bar{\tau}, \bar{x}^0, w^1), (v^1, \epsilon), y^1\big) - D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((0,0,0), (v^1, \epsilon), 0\big)\big\| \\ &\quad \leqslant \frac{1}{K} e^{-\frac{K_2\beta}{\epsilon^d}} \leqslant \frac{1}{\|D_{(\bar{\tau}, \bar{x}^0, w^1)} G((0,0,0), (v^1, \epsilon), 0)^{-1}\|}. \end{aligned}$$

Then from (6.36), (6.24), (6.37), and $\mu_0 - 6\beta_1 > 0$,

$$\begin{aligned} &\big\| D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((\bar{\tau}, \bar{x}^0, w^1), (v^1, \epsilon), y^1\big)^{-1} - D_{(\bar{\tau}, \bar{x}^0, w^1)} G\big((0,0,0), (v^1, \epsilon), 0\big)^{-1}\big\| \\ &\quad \leqslant 2 \cdot K^2 e^{\frac{2K_2\beta}{\epsilon^d}} \cdot 2K^2 e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 3\beta_1)} \leqslant 4K^4 e^{-\frac{K_1}{\epsilon^d}(\mu_0 - 5\beta_1)} \leqslant 2K^4 e^{\frac{K_1}{\epsilon^d}\beta_1}. \end{aligned}$$

The conclusion follows.

Step 7. Let $X = (\bar{\tau}, \bar{x}^0, w^1)$, $Y = ((v^1, \epsilon), y^1)$, and let the solution of $G(X, Y) = 0$ be $X = g(Y)$. Differentiating $G(X, Y) = 0$ yields

$$G_X\big(g(Y), Y\big) g_Y + G_Y\big(g(Y), Y\big) = 0, \tag{6.38}$$

so

$$g_Y = -G_X\big(g(Y), Y\big)^{-1} G_Y\big(g(Y), Y\big).$$

Therefore

$$\|g_Y\| \leqslant \left\|G_X\big(g(Y),Y\big)^{-1}\right\|\left\|G_Y\big(g(Y),Y\big)\right\| \leqslant Ke^{\frac{K_2\beta}{\epsilon^a}} \cdot Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-3\beta_1)} \leqslant K^2e^{-\frac{K_1}{\epsilon^a}(\mu_0-4\beta_1)}.$$

Differentiating (6.38) yields

$$G_X g_{yy} + G_{XX}g_Y^2 + 2G_{XY}g_Y + G_{YY} = 0,$$

so

$$g_{yy} = -G_X^{-1}\big(G_{XX}g_Y^2 + 2G_{XY}g_Y + G_{YY}\big).$$

For small $\epsilon > 0$, the term in parentheses of largest norm is G_{YY}, so

$$\|g_{yy}\| \leqslant 3\left\|G_X^{-1}\right\|\|G_{YY}\| \leqslant Ke^{\frac{K_2\beta}{\epsilon^a}} \cdot 3Ke^{-\frac{K_1}{\epsilon^a}(\mu_0-4\beta_1)} \leqslant 3K^2e^{-\frac{K_1}{\epsilon^a}(\mu_0-5\beta_1)}.$$

The general result follows by induction.

References

[1] P. Brunovský, C^r-inclination theorems for singularly perturbed equations, J. Differential Equations 155 (1999) 133–152.
[2] P. De Maesschalck, C^r transition maps through turning points in singular perturbation problems with one fast variable, University of Hasselt, 2007, preprint.
[3] N. Fenichel, Asymptotic stability with rate conditions II, Indiana Univ. Math. J. 26 (1977) 81–93.
[4] N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, J. Differential Equations 31 (1979) 53–98.
[5] C.K.R.T. Jones, Geometric singular perturbation theory, in: Dynamical Systems, Montecatini Terme, 1994, in: Lecture Notes in Math., vol. 1609, Springer, Berlin, 1995, pp. 44–118.
[6] C.K.R.T. Jones, T. Kaper, A primer on the exchange lemma for fast-slow systems, in: Multiple-Time-Scale Dynamical Systems, Minneapolis, MN, 1997, in: IMA Vol. Math. Appl., vol. 122, Springer, New York, 2001, pp. 85–132.
[7] C.K.R.T. Jones, T. Kaper, N. Kopell, Tracking invariant manifolds up to exponentially small errors, SIAM J. Math. Anal. 27 (2) (1996) 558–577.
[8] C.K.R.T. Jones, N. Kopell, Tracking invariant manifolds with differential forms in singularly perturbed systems, J. Differential Equations 108 (1994) 64–88.
[9] W. Liu, Exchange lemmas for singular perturbation problems with certain turning points, J. Differential Equations 167 (2000) 134–180.
[10] E.F. Mishchenko, Yu.S. Kolesov, A.Yu. Kolesov, N.Kh. Rozov, Asymptotic Methods in Singularly Perturbed Systems, translated from the Russian by Irene Aleksanova, Monogr. Contemp. Math., Consultants Bureau, New York, 1994.
[11] S. Schecter, Exchange lemmas 1: Deng's lemma, J. Differential Equations 245 (2) (2008) 392–410.
[12] S. Schecter, P. Szmolyan, Rarefactions in the Dafermos regularization of a system of conservation laws, North Carolina State University, 2007, preprint.
[13] S.-K. Tin, N. Kopell, C.K.R.T. Jones, Invariant manifolds and singularly perturbed boundary value problems, SIAM J. Numer. Anal. 31 (1994) 1558–1576.