A Remark on Singular Perturbations of Strongly Monotone Systems

Liming Wang and Eduardo D. Sontag

***Abstract*— This paper extends to singular perturbations of strongly monotone systems a result of Hirsch's on generic convergence to equilibria.**

I. INTRODUCTION

Monotone systems constitute a rich class of models for which global and almost-global convergence properties can be established. They are particularly useful in biochemical models (see discussion and references in [13], [14], and also appear in areas like coordination ([10]) and other problems in control theory ([1]). This paper studies extensions, using geometric singular perturbation theory, of Hirsch's generic convergence theorem for monotone systems ([4], [5], [6], [12]). Informally stated, Hirsch's result says that almost every bounded solution of a strongly monotone system converges to the set of equilibria. There is a rich literature regarding the application of this powerful theorem, as well as of other results dealing with everywhere convergence when equilibria are unique ([12], [2], [7]), to models of biochemical systems. Unfortunately, many models in biology are not monotone. In order to address this drawback (as well as to study properties of large systems which are monotone but which are hard to analyze in their entirety), a recent line of work introduced an input/output approach that is based on the analysis of interconnections of monotone systems. For example, the approach allows one to view a *non*-monotone system as a "negative" feedback loop of monotone open-loop systems, thus leading to results on global stability and the emergence of oscillations under transmission delays, and to the construction of relaxation oscillators by slow adaptation rules on feedback gains. See [13], [14] for expositions and many references. The present paper is in the same character.

Our motivation arose from the observation that time-scale separation may also lead to monotonicity. This point of view is of special interest in the context of biochemical systems; for example, Michaelis Menten kinetics are mathematically justified as singularly perturbed versions of mass action kinetics. A system that is not monotone may become monotone once that fast variables are replaced by their steady-state values. A trivial linear example that illustrates this point is $\dot{x}=-x-y$, $\varepsilon\dot{y}=-y+x$, with $\varepsilon>0$. This system is not monotone with respect to any orthant cone. On the other hand, for $\varepsilon \ll 1$, the fast variable y tracks x, so the slow dynamics is well-approximated by $\dot{x} = -2x$ (which is strongly monotone, because every scalar system is).

We consider systems $\dot{x} = f(x,y)$, $\varepsilon\dot{y} = g(x,y)$ for which the reduced system $\dot{x} = f(x,h(x))$ is (strongly) monotone and the fast system $\dot{y} = g(x,y)$ has a unique globally asymptotically stable steady state $y = h(x)$ for each x, and satisfies an input to state stability type of property with respect to x. One may expect that the original system inherits global (generic) convergence properties, at least for all $\varepsilon>0$ small enough, and this is indeed the object of our study. This question may be approached in several ways. One may view $y-h(x)$ as an input to the slow system, and appeal to the theory of asymptotically autonomous systems. Another approach, the one that we develop here, is through geometric invariant manifold theory ([3], [8], [11]). There is a manifold M_ε, invariant for the full dynamics, which attracts all near-enough solutions, with an asymptotic phase property. The system restricted to the invariant manifold M_ε is a regular perturbation of the fast (ε=0) system. As remarked in Theorem 1.2 in Hirsch's early paper [4], a C^1 regular perturbation of a flow with eventually positive derivatives also has generic convergence. So, solutions in the manifold will be generally well-behaved, and asymptotic phase implies that solutions track solutions in M_ε, and hence also converge to equilibria if solutions on M_ε do. A key technical detail is to establish that the tracking solutions also start from the "good" set of initial conditions, for generic solutions of the large system.

For simplicity, we discuss here only the case of cooperative systems (monotonicity with respect to the main orthant), but proofs in the case of general cones are similar and will be discussed in a paper under preparation.

II. STATEMENT OF MAIN RESULT

We are interested in systems in singularly perturbed form:

$$\begin{aligned} \frac{dx}{dt} &= f(x,y) \\ \varepsilon\frac{dy}{dt} &= g(x,y), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $0 < \varepsilon \ll 1$, and f and g are smooth functions. We will present some preliminary results in general, but for our main theorem we will restrict attention to the case when g has the special form $g(x,y) = Ay + h(x)$, where A is a Hurwitz matrix (all eigenvalues have negative real part) and h is a smooth function. That is, we

This research was supported in part by NSF grant DMS-0504557

Liming Wang is with the Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA, wshwlm@math.rutgers.edu

Eduardo D. Sontag is with the Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA, sontag@math.rutgers.edu

will specialize to systems of the following form:

$$\begin{aligned} \frac{dx}{dt} &= f(x,y) \\ \varepsilon \frac{dy}{dt} &= Ay + h(x). \end{aligned} \tag{2}$$

(We remark later how our results may be extended to a broader class of systems.) Setting ε to zero, we have:

$$\frac{dx}{dt} = f(x, m_0(x)), \tag{3}$$

where $m_0(x) = -A^{-1}h(x)$. As usual in singular perturbation theory, our goal is to use properties of the limiting system (3) in order to derive conclusions about the full system (2) when $0 < \varepsilon \ll 1$. We will assume given three sets K, $\widetilde{K}$, and L which satisfy the following hypotheses:

H1 The set $\widetilde{K}$ is an n-dimensional C^∞ simply connected compact manifold with boundary.
H2 The set $\widetilde{K}$ is convex.
H3 The set L is a bounded open subset of $\mathbb{R}^m$, and $M_0 = \{(x,y) \mid y = m_0(x),\ x \in \widetilde{K}\}$, the graph of m_0, is contained in $\widetilde{K} \times L$.
H4 The flow $\{\psi_t\}$ of the limiting system (3) has eventually positive derivatives on $\widetilde{K}$.
H5 For each $\varepsilon > 0$ sufficiently small, the forward trajectory under (2) of each point in $\widetilde{D} = \mathrm{Int}\widetilde{K} \times L$ is precompact in $\widetilde{D}$.
H6 The equilibrium set of (2) is countable.
H7 The set $K \subset \mathrm{Int}\widetilde{K}$ is compact, and for each $\varepsilon > 0$ sufficiently small, the set $D = K \times L$ is positively invariant.

The main theorem is:

Theorem 1: Under assumptions **H1-H7**, there exists $\varepsilon^* > 0$ such that for each $0 < \varepsilon < \varepsilon^*$, the forward trajectory of (2) starting from almost every point in D converges to some equilibrium.

Remark: A variant of this result is to assume that the reduced system (3) has a unique equilibrium. In this case, one may improve the conclusions of the theorem to global (not just generic) convergence, by appealing to results of Hirsch and others that apply when equilibria are unique. The proof is simpler in that case, since the foliation structure given by Fenichel's theory (see below) is not required. In the opposite direction, one could drop the assumption of countability and instead provide theorems on generic convergence to the set of equilibria, or even to equilibria if hyperbolicity conditions are satisfied, in the spirit of what is done in the theory of strongly monotone systems.

III. Terminology

The following standard terminology is defined for a general ordinary differential equation:

$$\frac{dz}{dt} = F(z), \tag{4}$$

where $F : \mathbb{R}^N \to \mathbb{R}^N$ is a C^1 vector field. For any $z \in \mathbb{R}^N$, we denote the maximally defined solution of (4) with initial condition z by $t \to \phi_t(z), t \in I(z)$, where $I(z)$ is an open interval in $\mathbb{R}$ that contains zero. For each $t \in \mathbb{R}$, the set of $z \in \mathbb{R}^N$ for which $\phi_t(z)$ is defined is an open set $W(t) \subseteq \mathbb{R}^N$, and $\phi_t : W(t) \to W(-t)$ is a diffeomorphism. The collection of maps $\{\phi_t\}_{t\in\mathbb{R}}$ is called the flow of (4). We also write just $z(t)$ for the solution of (4), if the initial condition $z(0)$ is clear from the context. The forward trajectory of $z \in \mathbb{R}^N$ is a parametrized curve $t \to \phi_t(z)$. Its image is the forward orbit of z, denoted as $O_+(z)$. The backward trajectory and the backward orbit $O_-(z)$ are defined analogously. A set $U \subseteq \mathbb{R}^N$ is *positively (respectively, negatively) invariant* if $O_+(U) \subseteq U$. It is *invariant* if it is both positively and negatively invariant.

We borrow the notation from [8] for the forward evolution of a set $U \subseteq V \subseteq \mathbb{R}^N$ restricted to V:

$$U \cdot_V t = \{\phi_t(p) :\ p \in U \text{ and } \phi_s(p) \in V \text{ for all } 0 \le s \le t\}.$$

Let us denote the interior and the closure of a set U as Int$U$ and $\overline{U}$ respectively.

Definition 1: The flow $\{\phi_t\}$ of (4) is said to have *eventually positive derivatives on a set* $V \subseteq \mathbb{R}^N$ if there exists t_0 such that the matrix $D_z\phi_t(z)$ has only positive entries (also called the matrix is positive) for every $t \ge t_0$, $z \in V$.

The next lemma is a restatement of theorem 4.4 in [5]:

Lemma 1: Suppose that the open set $W \subseteq R^n$ is convex and the flow $\{\phi_t\}$ of (4) has eventually positive derivatives on W. Let $W^c \subseteq W$ be the set of points whose forward orbit has compact closure in W. If the set of equilibrium points is countable, then $z(t)$ converges to a equilibrium as t goes to infinity, for almost every $z \in W^c$.

The following fact follows from differentiability of solutions with respect to "regular" perturbations in the dynamics; see [5], Theorem 1.2:

Lemma 2: Assume $V \subset W$ is a compact set in which the flow $\{\phi_t\}$ has eventually positive derivatives. Then, there exists $\delta > 0$ with the following property. Let $\{\psi_t\}$ denote the flow of a C^1 vector field G such that the C^1 norm of $F(z) - G(z)$ is less than δ for all z in V. Then there exists $t_* > 0$ such that if $t \ge t_*$ and $\psi_s(z) \in V$ for all $s \in [0,t]$, then the matrix $D_z\psi_t(z)$ is positive.

Definition 2: A compact, connected C^r manifold $M \subset \mathbb{R}^N$ with boundary is said to be *locally invariant* under the flow of (1), if for each $p \in \text{Int}M$, there exists a time interval $I_p = (t_1, t_2), t_1 < 0 < t_2$, such that $\phi_t(p) \in M$ for all $t \in I_p$.

When $\varepsilon \neq 0$, we can set $\tau = t/\varepsilon$, and (1) is equivalent to its fast system:

$$\frac{dx}{d\tau} = \varepsilon f(x, y) \quad (5)$$
$$\frac{dy}{d\tau} = g(x, y).$$

Definition 3: Let M be an n-dimensional manifold (possibly with boundary) which is contained in $\{(x, y) \mid g(x, y) = 0\}$. We say that M is normally hyperbolic relative to (5) if all eigenvalues of the matrix $D_y g(p)$ have nonzero real part for every $p \in M$.

IV. Proof of the Main Theorem

Our proofs are based on Fenichel's theorems [3], in the forms presented and developed by Jones in [8].

Fenichel's First Theorem *Under assumption H1, if M_0 is normally hyperbolic relative to the fast system of (2), then there exists $\varepsilon_0 > 0$, such that for every $0 < \varepsilon < \varepsilon_0$ and $r > 0$, there is a function $y = m_\varepsilon(x)$, defined on $\widetilde{K}$, of class C^r jointly in x and ε, such that*

$$M_\varepsilon = \{(x, y) \mid y = m_\varepsilon(x),\ x \in \widetilde{K}\}$$

is locally invariant under (2), see Figure 1.

The requirement that M_0 be normally hyperbolic is satisfied in our case, as $g(x, y) = Ay + h(x)$ and therefore $D_y g(p) = A$, which is invertible, for each $p \in M_0$.

We will pick a particular $r > 1$ in the above theorem from now on.

Let us interpret local invariance in terms of equations. Let $(x(t), y(t))$ be the solution to (2) with initial condition (x_0, y_0), such that $x_0 \in \text{Int}\widetilde{K}$ and $y_0 = m_\varepsilon(x_0)$. Local invariance implies that $(x(t), y(t))$ satisfies

$$\frac{dx(t)}{dt} = f\big(x(t), m_\varepsilon(x(t))\big) \quad (6)$$
$$y(t) = m_\varepsilon(x(t)), \quad (7)$$

for all t small enough. Actually, this is also true for all $t \geq 0$. The argument is as follows. By **H5**, $(x(t), y(t))$ is well-defined and remains in $\widetilde{D}$ for all $t \geq 0$. Let $T = \{t \geq 0 \mid y(t) = m_\varepsilon(x(t))\}$. Then, T is not empty, and T is closed by the continuity of $m_\varepsilon(x(t))$ and $y(t)$. Also, T is open, since M_ε is locally invariant. So $T = \{t \geq 0\}$, that is, $x(t)$ is a solution to (6) and $y(t) = m_\varepsilon(x(t))$ for all $t \geq 0$.

In (6)-(7), the x-equation is decoupled from the y-equation, which allows us to reduce to studying a lower-dimension system. Another advantage is that, as ε approaches

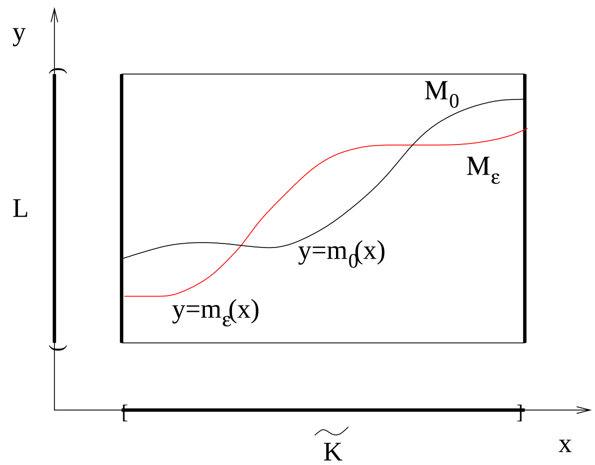

Fig. 1. For simplicity, we sketch manifolds M_ε and M_0 of a system where $n = m = 1$. The set $\widetilde{K}$ is a compact set in x, and L is an open set in y. The red curve denotes the locally invariant manifold M_ε and the black curve denotes M_0.

zero, the limit of system (6) is system (3), which describes the flow on M_0. If M_0 has some desirable property, it is natural to expect this property is inherited by the perturbed manifold M_ε. An example of this principle is provided by the following lemma.

Lemma 3: Under assumptions **H1**-**H4**, for each $0 < \varepsilon < \varepsilon_0$, the flow $\{\psi_t\}$ of (6) has eventually positive derivatives on $\text{Int}\widetilde{K}$.

Proof. Applying Lemma 2, there exist $\delta > 0$ such that when the C^1 norm of $m_0(x) - m_\varepsilon(x)$ is less than δ for all $x \in \widetilde{K}$, there exists $t_* > 0$ with the property that: if $t \geq t_*$ and $\psi_s(x) \in \widetilde{K}$ for all $s \in [0, t]$, then the matrix $D_x\psi_t(x)$ is positive. Since m_ε is of class C^r, jointly in x and ε, we can pick $\varepsilon > 0$ small enough to control δ. If we then prove $\text{Int}\widetilde{K}$ is invariant under (6), we are done. To see this, pick any $x_0 \in \text{Int}\widetilde{K}$, and let $y_0 = m_\varepsilon(x_0), y(t) = m_\varepsilon(x(t))$. Then, $(x(t), y(t))$ is the solution to (2) with initial condition $(x_0, y_0) \in \widetilde{D}$. By **H5**, $(x(t), y(t))$ stays in $\widetilde{D}$ for all $t \geq 0$, and therefore $x(t) \in \text{Int}\widetilde{K}$ for all $t \geq 0$.

Flows with eventually positive derivatives have particularly appealing properties, as in Lemma 1. To apply that lemma, we need to check two conditions. First, for every point in $\text{Int}\widetilde{K}$, its forward trajectory under (6) has compact closure in $\text{Int}\widetilde{K}$. Second, the number of equilibria of (6) is countable. Suppose that the first property does not hold, and let $x(t)$ be a solution to (6) with $x(0) \in \text{Int}\widetilde{K}$ but $\lim_{j\to\infty} x(t_j) \notin \text{Int}\widetilde{K}$ for some sequence $\{t_j\}$. So, $\big(x(t), m_\varepsilon(x(t))\big)$ is a solution for (2), and its forward orbit is not precompact in $\widetilde{D}$. This violates **H5**. To check the second condition, we notice that (2) is reduced to (6) on M_ε, so the equilibrium set of (6) is a subset of the x-coordinate of the equilibrium set of (2), which is countable. Therefore, the second condition also holds. Applying Lemma 1 we have:

Lemma 4: Under assumptions **H1-H6**, for each $0 < \varepsilon < \varepsilon_0$, there exists a set $C_\varepsilon \subseteq \mathrm{Int}\widetilde{K}$ such that the forward trajectory of (6) for every point of C_ε converges to some equilibrium, and the measure of $\mathrm{Int}\widetilde{K} \setminus C_\varepsilon$ is zero.

Until now, we have discussed the flow only when restricted to the locally invariant manifold M_ε. The next theorem, stated in the form given by [8], deals with more global behavior. In [8], the theorem is stated for $\varepsilon > 0$, but it also holds for $\varepsilon = 0$ ([9]). (We will apply this result again with a fixed $r > 1$.)

Fenichel's Third Theorem *Under assumption H1, if M_0 is normally hyperbolic relative to (2), then there exists $0 < \varepsilon_1 < \varepsilon_0$, $\delta_0 > 0$, such that for every $0 \le \varepsilon < \varepsilon_1$, $0 < \delta < \delta_0$ and $r > 0$, there is a function*

$$h_{\varepsilon,\delta} : M_\varepsilon \times [-\delta, \delta] \to \mathbb{R}^n$$

such that the following properties hold:

1) *For each $\bar{p} = (\bar{x}, m_\varepsilon(\bar{x})) \in M_\varepsilon$, $h_{\varepsilon,\delta}(\bar{p}, 0) = \bar{x}$.*
2) *The stable fibers $W^s_{\varepsilon,\delta}(\bar{p})$, defined as*

$$\{(x,y) \mid x = h_{\varepsilon,\delta}(\bar{p},\lambda),\ y = \lambda + m_\varepsilon(x),\ |\lambda| \le \delta\},$$

form a "positively invariant" family when $\varepsilon \neq 0$ in the sense that

$$W^s_{\varepsilon,\delta}(\bar{p}) \cdot_{W^s_{\varepsilon,\delta}(M_\varepsilon)} t \subseteq W^s_{\varepsilon,\delta}(\phi_t(\bar{p})),$$

where $W^s_{\varepsilon,\delta}(M_\varepsilon) = \bigcup_{\bar{p} \in M_\varepsilon} W^s_{\varepsilon,\delta}(\bar{p})$.

3) *"Asymptotic Phase". There are positive constants k and α such that for any p and $\bar{p}$, if $p \in W^s_{\varepsilon,\delta}(\bar{p})$, $\varepsilon \neq 0$, then*

$$|\phi_t(p) - \phi_t(\bar{p})| \le k e^{-\alpha t}$$

for all $t \ge 0$ as long as $\phi_t(p)$ and $\phi_t(\bar{p})$ stay in $W^s_{\varepsilon,\delta}(M_\varepsilon)$.

4) *The stable fibers are disjoint, i.e., for $p_i \in W^s_{\varepsilon,\delta}(\bar{p}_i)$, $i = 1, 2$, either $W^s_{\varepsilon,\delta}(\bar{p}_1) \bigcap W^s_{\varepsilon,\delta}(\bar{p}_2) = \emptyset$ or $W^s_{\varepsilon,\delta}(\bar{p}_1) = W^s_{\varepsilon,\delta}(\bar{p}_2)$.*
5) *The function $h_{\varepsilon,\delta}(\bar{p}, \lambda)$ is C^r jointly in ε, $\bar{p}$ and λ. When $\varepsilon = 0$, $h_{0,\delta}(\bar{p}, \lambda) = \bar{x}$, where $\bar{p} = (\bar{x}, m_0(\bar{x})) \in M_0$.*

The next lemma gives a sufficient condition to guarantee that a point is on some fiber.

Lemma 5: There exists $0 < \varepsilon_2 < \varepsilon_1$, such that for every $0 < \varepsilon < \varepsilon_2$, $0 < \delta < \delta_0$, the set $\mathcal{A}_\delta := \{(x, y) \mid x \in K,\ |y - m_0(x)| \le \frac{\delta}{2}\} \subset W^s_{\varepsilon,\delta}(M_\varepsilon)$.

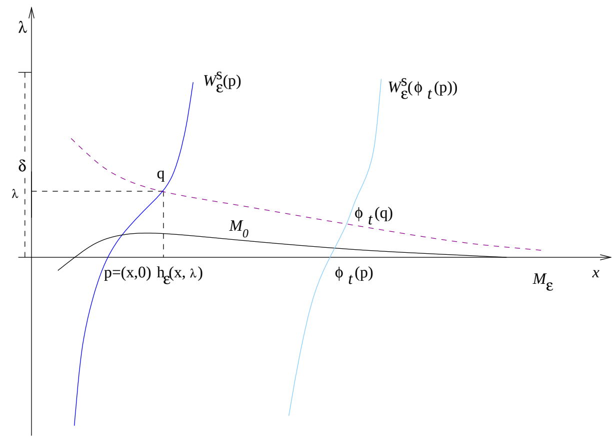

Fig. 2. To illustrate the geometric meaning of Fenichel's Third Theorem, we sketch the locally invariant manifold and stable fibers of a system, in the case n=m=1. The dimension of the manifolds M_ε, M_0, and stable fibers is one. M_ε is the graph of $\lambda = 0$, and M_0 is the graph of $m_0(x) - m_\varepsilon(x)$ (black curve). These manifolds may intersect at some equilibrium points. Through each point $\bar{p}$ in the manifold M_ε, there is a stable fiber $W^s_{\varepsilon,\delta}(\bar{p})$ (blue curve). We call $\bar{p}$ the "base point" of the fiber. The fiber consists of the pairs $(x, \lambda) = (h_{\varepsilon,\delta}(\bar{p}, \lambda), \lambda)$, where $|\lambda| \le \delta$. If a solution (purple dashed curve) starts on fiber $W^s_{\varepsilon,\delta}(\bar{p})$, after a small time t, it evolves to a point on another stable fiber $W^s_{\varepsilon,\delta}(\phi_t(\bar{p}))$ (light blue curve); this is the "positive invariance" property.

To prove this lemma, we need the following result:

Lemma 6: Let U and V be compact, convex subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Given a continuous function h:

$$\begin{array}{rcl} U \times V & \to & \mathbb{R}^n \times \mathbb{R}^m \\ (x, y) & \mapsto & (h_1(x, y), h_2(x, y)) \end{array}$$

satisfying $\|h_1(x, y) - x\| \le \rho_1$, $\|h_2(x, y) - y\| \le \rho_2$ for all $(x, y) \in U \times V$. Then every $(\alpha, \beta) \in U \times V$ with $\mathrm{dist}(\alpha, \partial U) \ge \rho_1$ and $\mathrm{dist}(\beta, \partial V) \ge \rho_2$ is in the image of h.

Proof. For such a point (α, β), consider the map $H(x, y) = (H_1(x, y), H_2(x, y)) := (x, y) - (h_1(x, y), h_2(x, y)) + (\alpha, \beta)$. Thus H maps $U \times V$ into itself. If not, say $H_1(x, y)$ is not in U, that is, $x - h_1(x, y) + \alpha$ is not in U. Since $\|x - h_1(x, y)\| \le \rho_1$, so $\mathrm{dist}(\alpha, \partial U) < \rho_1$, contradiction. The case when $H_2(x, y)$ is not in V follows similarly. Since H maps $U \times V$ into itself, and the product of convex sets is still convex, by Brouwer's Fixed Point Theorem, there is some $(\bar{x}, \bar{y}) \in U \times V$ so that $H(\bar{x}, \bar{y}) = (\bar{x}, \bar{y})$, which means that $(h_1(\bar{x}, \bar{y}), h_2(\bar{x}, \bar{y})) = (\alpha, \beta)$, as we wanted to prove.

Proof of Lemma 5. Define the map D_ε, for each $0 \le \varepsilon < \varepsilon_1$:

$$\begin{array}{rcl} W^s_{0,\delta}(M_0) & \to & W^s_{\varepsilon,\delta}(M_\varepsilon) \\ (x, \lambda) & \mapsto & (h_{\varepsilon,\delta}((x, m_\varepsilon(x)), \lambda), \lambda + m_\varepsilon(x) - m_0(x)) \end{array}$$

In this proof, $(\cdot,\cdot)$ denote $x = x$, $\lambda = y - m_0(x)$ coordinates. The map D_ε is continuous, and D_0 is identity. It satisfies $\|h_{\varepsilon,\delta}((x, m_\varepsilon(x)), \lambda) - x\| \leq C_1(\varepsilon)$ and $\|\lambda + m_\varepsilon(x) - m_0(x) - \lambda\| \leq C_2(\varepsilon)$ for some positive function C_i of ε, and $C_i \to 0$ as $\varepsilon \to 0$, $i = 1, 2$. Apply Lemma 6 with $U = \widetilde{K}$, $V = [-\delta, \delta]$, $\rho_1 = C_1(\varepsilon)$ and $\rho_2 = C_2(\varepsilon)$. Since $\text{dist}(\partial K, \partial\widetilde{K})$ and $\frac{\delta}{2}$ are fixed, we can pick $0 < \varepsilon_2 < \varepsilon_1$ such that $\text{dist}(\partial K, \partial\widetilde{K}) > C_1(\varepsilon)$ and $\frac{\delta}{2} > C_2(\varepsilon)$ for all $\varepsilon \in (0, \varepsilon_2)$. By Lemma 6, all points in $K \times [-\frac{\delta}{2}, \frac{\delta}{2}] = \mathcal{A}_\delta$ are also in $W^s_{\varepsilon,\delta}(M_\varepsilon)$.

Lemma 7: Under assumption **H7**, for any given $\delta > 0$, there exists $0 < \varepsilon_3 < \varepsilon_2$ such that for each $0 < \varepsilon < \varepsilon_3$, if $p \in D$, then there exists $T_0 \geq 0$, and $\phi_t(p) \in \mathcal{A}_\delta$ for all $t \geq T_0$.

Proof. Setting $z = y - m_0(x)$ and $\tau = t/\varepsilon$, (2) becomes

$$\begin{aligned} \frac{dx}{d\tau} &= \varepsilon f(x, z + m_0(x)) \\ \frac{dz}{d\tau} &= Az - \varepsilon m_0'(x) f(x, z + m_0(x)). \end{aligned}$$

So

$$z(\tau) = z(0)e^{A\tau} - \varepsilon \int_0^\tau e^{A(\tau-s)} m_0'(x) f\big(x, z + m_0(x)\big)\, ds.$$

Notice that $|e^{At}| \leq Ce^{-\beta t}$, for some positive constants C and β. So,

$$\left|\varepsilon \int_0^\tau e^{A(\tau-s)} m_0'(x) f\big(x, z + m_0(x)\big)\, ds\right| \leq \frac{2\varepsilon MC}{\beta},$$

where M is an upper bound of the function $|m_0'(x) f(x, y)|$ on $\overline{D}$. Let

$$\varepsilon = \frac{\delta\beta}{4MC} \quad \text{and} \quad T_0' = \max\{\frac{1}{\beta} \ln \frac{2C|z(0)|}{\delta}, 0\}\,.$$

Then, we have $|z(\tau)| \leq \delta$ for all $\tau \geq T_0'$. Back to the slow time scale, we let $T_0 = \varepsilon T_0'$. Therefore, $\phi_t(p) \in \mathcal{A}_\delta$ for all $t \geq T_0$, if the x-coordinate of $\phi_t(p)$ stays in K, but this can be easily derived from **H7**.

Remark: Except for the normal hyperbolicity assumption, Lemma 7 is the only place where the special structure (2) was used. Consider a more general system as in (1), and assume that $g(x, m_0(x)) = 0$ on $\widetilde{K}$ for some smooth function m_0. By the same change of variables as in the above proof, (1) is equivalent to

$$\begin{aligned} \frac{dx}{d\tau} &= \varepsilon f(x, z + m_0(x)) \\ \frac{dz}{d\tau} &= g(x, z + m_0(x)) - \varepsilon m_0'(x) f(x, z + m_0(x)). \end{aligned}$$

The only property that we need in the lemma is that for any initial condition $(x(0), z(0))$, the solution $(x(t), z(t))$ satisfies

$$\limsup_{t\to\infty} |z(t)| \leq \gamma\left(\limsup_{t\to\infty} d(t)\right)$$

where γ is a function of class $\mathcal{K}$, that is to say, a continuous function $[0, \infty) \to [0, \infty)$ with $\gamma(0) = 0$, and $d(t) = \varepsilon m_0'(x(t)) f\big(x(t), z(t) + m_0(x(t))\big)$. In terms of the functions m_0 and g, we may introduce the control system $dz/dt = G(d(t), z) + u(t)$, where d is a compact-valued "disturbance" function and u is an input, and $G(d, z) = g(d, z + m_0(d))$. Then, the property of input-to-state stability with input u (uniformly on d), which can be characterized in several different manners, including by means of Lyapunov functions, provides the desired condition.

Lemma 7 proves that every trajectory in D is attracted to A_δ and therefore is also attracted to M_ε. This will lead to our proof of the main theorem.

Proof of the main theorem.: Choose $\varepsilon^* = \varepsilon_3$ and some $0 < \delta < \delta_0$. For any $p \in D$, there are three cases:

1) $p \in M_\varepsilon$. By Lemma 4, the forward trajectory converges to an equilibrium except for a set of measure zero.
2) $p \in \mathcal{A}_\delta \subset W^s_{\varepsilon,\delta}(M_\varepsilon)$. Then p is on some fiber, say $W^s_{\varepsilon,\delta}(\bar{p})$. If the x-coordinate of $\bar{p}$, denoted as $\bar{x}$, is in C_ε (defined in Lemma 4), then $\phi_t(\bar{p}) \to q$, some equilibrium point of (2). By the "asymptotic phase" property of Fenichel's Third Theorem, $\phi_t(p)$ also converges to q. To deal with the case when $\bar{x} \notin C_\varepsilon$, it is enough to show that the set
$$\mathcal{B}_{\varepsilon,\delta} = \bigcup_{\bar{x} \in \text{Int}\widetilde{K} \setminus C_\varepsilon} W^s_{\varepsilon,\delta}(\bar{p})$$
as a subset of $\mathbb{R}^{m+n}$ has measure zero. Define
$$\mathcal{C}_{\varepsilon,\delta} = \Big(\text{Int}\widetilde{K} \setminus C_\varepsilon\Big) \times [-\delta, \delta].$$
Since $\text{Int}K \setminus C_\varepsilon$ has measure zero in $\mathbb{R}^n$, also $\mathcal{C}_{\varepsilon,\delta}$ has measure zero. The map γ:
$$\begin{aligned} \mathcal{C}_{\varepsilon,\delta} &\to \mathcal{B}_{\varepsilon,\delta} \\ (\bar{x}, \lambda) &\mapsto (h_{\varepsilon,\delta}(\bar{p}, \lambda), \lambda + m_\varepsilon(h_{\varepsilon,\delta}(\bar{p}, \lambda))) \end{aligned}$$
is Lipschitz, and $\gamma(\mathcal{C}_{\varepsilon,\delta}) = \mathcal{B}_{\varepsilon,\delta}$. We are done, because Lipschitz maps send measure zero sets to measure zero sets.
3) $p \in D \setminus \mathcal{A}_\delta$. By Lemma 7 $\phi_t(p) \in \mathcal{A}_\delta$ for all $t \geq T_0$. Without loss of generality, we assume that T_0 is an integer. If $\phi_{T_0}(p) \in \mathcal{A}_\delta \setminus \mathcal{B}_{\varepsilon,\delta}$, then $\phi_t(p)$ converges to an equilibrium. Otherwise, $p \in \bigcup_{k\geq 0, k\in\mathbb{Z}} \phi_{-k}(\mathcal{B}_{\varepsilon,\delta})$. Since the set $\mathcal{B}_{\varepsilon,\delta}$ has measure zero and ϕ_{-k} is

Lipschitz, $\phi_{-k}(\mathcal{B}_{\varepsilon,\delta})$ has measure zero for all k, and the countable union of them still has measure zero.

V. An Example

Consider the following system:

$$\begin{aligned} \frac{dx_i}{dt} &= \gamma_i(y_1, \dots, y_m) - \beta_i(x_1, \dots, x_n) \\ \varepsilon \frac{dy_j}{dt} &= -d_j y_j - \alpha_j(x_1, \dots, x_n), \quad d_j > 0, \end{aligned} \tag{8}$$

where α_j, β_i and γ_i ($i = 1, \dots, n, j = 1, \dots, m$) are smooth functions. We assume that

1) When $n > 1$, for all $i, k = 1, \dots, n$, $i \neq k$, and all $x \in \mathbb{R}^n$, the partial derivatives $\frac{\partial \beta_i}{\partial x_k}(x) < 0$ and $\sum_{l=1}^{m} \frac{\partial \gamma_i}{\partial y_l}(x) \frac{\partial \alpha_l}{\partial x_k}(x) \leq 0$.
2) The function β_i satisfies that that $\beta_i(x_1, \dots, x_n) = +\infty$ as all $x_i \to +\infty$ and $\beta_i(x_1, \dots, x_n) = -\infty$ as all $x_i \to -\infty$.
3) There exists a positive constant M_j such that $|\alpha_j(x)| \leq M_j$ for all $x \in \mathbb{R}^n$.
4) The number of roots of the system of equations $\gamma_i(\alpha_1(x), \dots, \alpha_m(x)) = \beta_i(x), \quad i = 1, \dots, m$, is countable.

The conditions are very natural. The condition on the β_i's is satisfied, for example, if there is a linear decay term $-x_i$ in the differential equation for x_i, and all other variables appear saturated in this rate, see an more interesting example in [15].

We are going to show that on any large enough region, and provided that ε is sufficiently small, almost every trajectory converges to an equilibrium. To emphasize the need for small ε, we also show that when $\varepsilon > 1$, a limit cycle could appear.

To apply our main theorem, we take $L = \{ y \in \mathbb{R}^m \mid |y_j| < b_j, \ j = 1, \dots, m \}$, where b_j is an arbitrary positive number greater than $\frac{M_j}{d_j}$. Picking such b_j assures $y_j \frac{dy_j}{dt} < 0$ for all $x \in \mathbb{R}$ and $|y_j| = b_j$, i.e. the vector field points transversely inside on the boundary of L. Let $K = \{ x \in \mathbb{R}^n \mid -a_{i,2} \leq x_i \leq a_{i,1}, \ i = 1, \dots, n \}$, where $a_{i,1}$ and $a_{i,2}$ can be any positive numbers such that $\beta_i(x) > N_i := \max_{|y_j| \leq b_j} |\gamma_i(y_1, \dots, y_m)|$, whenever $x \in \mathbb{R}^n$ satisfies that its ith coordinate $x_i \geq a_{i,1}$; and $\beta_i(x) < -N_i$, whenever $x \in \mathbb{R}^n$ satisfies that its ith coordinate $x_i \leq -a_{i,2}$. All large enough $a_{i,j}$'s satisfy this condition, because of the assumption made on β. So, we have $x_i \frac{dx_i}{dt} < 0$ for all $y \in L$, $x_i = a_{i,1}$ and $x_i = -a_{i,2}$. We then take $\widetilde{K} = \{ x \in \mathbb{R}^n \mid -a_{i,2} - 1 \leq x_i \leq a_{i,1} + 1, \ i = 1, \dots n \}$, $D = K \times L$ and $\widetilde{D} = \mathrm{Int}\widetilde{K} \times L$. Thus, the vector field will point into the interior of D and $\widetilde{D}$. Hypotheses **H5** and **H7** follow directly from this fact. It is easy to see the other hypotheses also hold. By our main theorem, for sufficiently small ε, the forward trajectory of (8) starting from almost every point in D converges to some equilibrium.

On the other hand, convergence does not hold for large ε. Let $n = 1$, $\beta_1(x_1) = \frac{x_1^3}{3} - x_1$, $m = 1$, $\alpha_1(x_1) = 4 \tanh x_1$, $\gamma(y_1) = y_1$, $d_1 = 1$. It is easy to verify that $(0, 0)$ is the only equilibrium. When $\varepsilon > 1$, the trace of the Jacobian at $(0, 0)$ is $1 - \frac{1}{\varepsilon} > 0$, its determinant is $\frac{15}{\varepsilon} > 0$, so the (only) equilibrium in D is repelling. By the Poincaré-Bendixson Theorem, there exists a limit cycle in D.

References

[1] L. Chisci, P. Falugi, *Asymptotic tracking for state-constrained monotone systems*, Proc. 44th IEEE Conf. Decision and Control, Seville, 2005, paper ThC17.5.

[2] Dancer, E.N., "Some remarks on a boundedness assumption for monotone dynamical systems," *Proc. of the AMS* **126**(1998), pp. 801–807.

[3] Fenichel, N., "Geometric singular perturbation theory for ordinary differntial equations," *J. of Differential Equations* **31**(1979), pp. 53–98.

[4] Hirsch, M., "Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere," *SIAM J. Mathematical Analysis* **16**(1985), pp. 423–439.

[5] Hirsch, M., "Differential equations and convergence almost everywhere in strongly monotone flows," *Contemporary Mathematics* **17**(1983), pp. 267–285.

[6] Hirsch, M., H.L. Smith, "Monotone dynamical systems," in *Handbook of Differential Equations, Ordinary Differential Equations: Volume 2*, Elsevier, Amsterdam, 2005.

[7] Jiang, J.F., "On the global stability of cooperative systems," *Bulletin of the London Math Soc* **6**(1994), pp. 455–458.

[8] Jones, C.K.R.T., "Geometric singular perturbation theory," in *Dynamical Systems (Montecatini. Terme, 1994), Lect. Notes in Math. 1609*, Springer-Verlag, Berlin.

[9] Jones, C.K.R.T., personal communication.

[10] L. Moreau, *Stability of continuous-time distributed consensus algorithms*, Proc. 43rd IEEE Conf. Decision and Control, Paradise Island, Bahamas, 2004, paper ThC10.4.

[11] Nipp, K., "Smooth attractive invariant manifolds of singularly perturbed ODE's," Research Report, No. **92-13** (1992),

[12] Smith, H.L., *Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems, Mathematical Surveys and Monographs, vol. 41*, (AMS, Providence, RI, 1995).

[13] Sontag, E.D. "Some new directions in control theory inspired by systems biology," *Systems Biology* **1**(2004), pp. 9–18.

[14] Sontag, E.D., "Molecular systems biology and control," *European J. Control* **11** (2005), pp. 396-435.

[15] Wang, L. and Sontag, E.D. "Almost global convergence in singular perturbations of strongly monotone systems," arXiv, math.DS/0604369, 2006.