J Nonlinear Sci (2008) 18: 527–550
DOI 10.1007/s00332-008-9021-2

Journal of
Nonlinear
Science

Singularly Perturbed Monotone Systems and an Application to Double Phosphorylation Cycles

Liming Wang · Eduardo D. Sontag

Received: 30 July 2007 / Accepted: 28 February 2008 / Published online: 26 March 2008

Abstract The theory of monotone dynamical systems has been found very useful in the modeling of some gene, protein, and signaling networks. In monotone systems, every net feedback loop is positive. On the other hand, negative feedback loops are important features of many systems, since they are required for adaptation and precision. This paper shows that, provided that these negative loops act at a comparatively fast time scale, the main dynamical property of (strongly) monotone systems, convergence to steady states, is still valid. An application is worked out to a double-phosphorylation "futile cycle" motif which plays a central role in eukaryotic cell signaling.

Keywords Singular perturbation · Monotone systems · Asymptotic stability · MAPK system

Mathematics Subject Classification (2000) 34D15 · 37C65

1 Introduction

Monotone dynamical systems constitute a rich class of models, for which global and almost-global convergence properties can be established. They are particu-

L. Wang (✉) · E.D. Sontag
Department of Mathematics, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA
e-mail: wshwlm@math.rutgers.edu

E.D. Sontag
e-mail: sontag@math.rutgers.edu

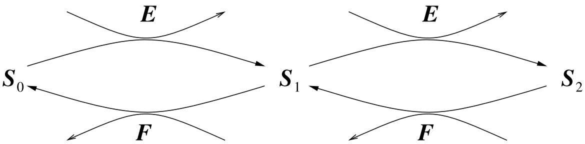

Fig. 1 Dual futile cycle. A substrate S_0 is ultimately converted into a product S_2, in an "activation" reaction triggered or facilitated by an enzyme E and, conversely, S_2 is transformed back (or "deactivated") into the original S_0, helped on by the action of a second enzyme F

larly useful in biochemical applications and also appear in areas like coordination (Moreau 2004) and other problems in control (Chisci and Falugi 2005). One of the fundamental results in monotone systems theory is Hirsch's Generic Convergence Theorem (Hirsch 1983, 1985, 1988; Hirsch and Smith 2005; Smith 1995). 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 (Dancer 1998; Jiang 1994; Smith 1995), to models of biochemical systems. See, for instance, Sontag (2004, 2005) for expositions and many references.

Unfortunately, many models in biology are not monotone, at least with respect to any standard orthant order. This is because in monotone systems (with respect to orthant orders) every net feedback loop should be positive; on the other hand, in many systems negative feedback loops often appear as well, as they are required for adaptation and precision. Intuitively, however, negative loops that act at a comparatively fast time scale should not affect the main characteristics of monotone behavior. The main purpose of this paper is to show that this is indeed the case, in the sense that singularly perturbed strongly monotone systems inherit generic convergence properties. A system that is not monotone may become monotone once fast variables are replaced by their steady-state values. In order to prove a precise time-separation result, we employ tools from geometric singular perturbation theory.

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 (Edelstein-Keshet 1988; Murray 2002). One particular example of great interest in view of current systems biology research is that of dual "futile cycle" motifs, as illustrated in Fig. 1. As discussed in Samoilov et al. (2005), futile cycles (with any number of intermediate steps, and also called substrate cycles, enzymatic cycles, or enzymatic interconversions) underlie signaling processes such as guanosine triphosphatase cycles (Donovan et al. 2002), bacterial two-component systems and phosphorelays (Bijlsma and Groisman 2003; Grossman 1995), actin treadmilling (Chen et al. 2000), and glucose mobilization (Karp 2002), as well as metabolic control (Stryer 1995) and cell division and apoptosis (Sulis and Parsons 2003) and cell-cycle checkpoint control (Lew and Burke 2003). A most important instance is that of Mitogen-Activated Protein Kinase (MAPK) cascades, which regulate primary cellular activities such as proliferation, differentiation, and apoptosis (Asthagiri and Lauffenburger 2001; Chang and Karin 2001; Huang and Ferrell 1996; Widmann et al. 1999) in eukaryotes from yeast to humans.

MAPK cascades usually consist of three tiers of similar structures with multiple feedbacks (Burack and Sturgill 1997; Ferrell and Bhatt 1900; Zhao and Zhang 2001). Here we focus on one individual level of a MAPK cascade, which is a dual futile cycle as depicted in Fig. 1. The precise mathematical model is described later. Numerical simulations of this model have suggested that the system may be monostable or bistable, see Markevich et al. (2004). The latter will give rise to switch-like behavior, which is ubiquitous in cellular pathways (Gardner et al. 2000; Pomerening et al. 2003; Sel'kov 1975; Sha et al. 2003). In either case, the system under meaningful biological parameters shows convergence, not other dynamical properties such as periodic behavior or even chaotic behavior. Analytical studies done for the quasi-steady-state version of the model (slow dynamics), which is a monotone system, indicate that the reduced system is indeed monostable or bistable, see Ortega et al. (2006). Thus, it is of great interest to show that, at least in certain parameter ranges (as required by singular perturbation theory), the full system inherits convergence properties from the reduced system, and this is what we do as an application of our results. We remark that the simplified system, consisting of a unary conversion cycle (no S_2), is known to admit a unique equilibrium (subject to mass conservation constraints) which is a global attractor, see Angeli and Sontag (2008).

A feature of our approach is the use of geometric invariant manifold theory (Fenichel 1979; Jones 1995; Nipp 1992). There is a manifold M_ε, invariant for the full dynamics of a singularly perturbed system, which attracts all near-enough solutions. However, we need to exploit the full power of the theory, and especially the fibration structure and an asymptotic phase property. The system, restricted to the invariant manifold M_ε, is a regular perturbation of the slow ($\varepsilon = 0$) system. As remarked in Theorem 1.2 in Hirsch's early paper (Hirsch 1985), a C^1 regular perturbation of a flow with eventually positive derivatives also has generic convergence properties. So, solutions in the manifold will generally be well-behaved, and asymptotic phase implies that solutions near M_ε 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.

A preliminary version of these results in Wang and Sontag (2006) dealt with the special case of singularly perturbed systems of the form:

$$\begin{aligned} \dot{x} &= f(x, y), \\ \varepsilon \dot{y} &= Ay + h(x), \end{aligned}$$

on a product domain, where A is a constant Hurwitz matrix and the reduced system $\dot{x} = f(x, -A^{-1}h(x))$ is strongly monotone. However, for the application to the above futile cycle, there are two major problems with that formulation: first, the dynamics of the fast system have to be allowed to be nonlinear in y, and second, it is crucial to allow for an ε-dependence on the right-hand side as well as to allow the domain to be a convex polytope depending on ε. We provide a much more general formulation here.

We note that no assumptions are imposed regarding global convergence of the reduced system, which is essential because of the intended application to multistable systems. This seems to rule out the applicability of Lyapunov-theoretic and input-to-state stability tools (Christofides and Teel 1996; Teel et al. 2003).

This paper is organized as follows. The main result is stated in Sect. 2. In Sect. 3, we review some basic definitions and theorems about monotone systems. The detailed proof of the main theorem can be found in Sect. 4, and applications to the MAPK system and another set of ordinary differential equations are discussed in Sect. 5. Finally, in Sect. 6, we summarize the key points of this paper.

2 Statement of the Main Theorem

In this paper, we focus on the dynamics of the following prototypical system in singularly perturbed form:

$$\begin{aligned} \frac{dx}{dt} &= f_0(x, y, \varepsilon), \\ \varepsilon\frac{dy}{dt} &= g_0(x, y, \varepsilon). \end{aligned} \tag{1}$$

We will be interested in the dynamics of this system on an ε-dependent domain D_ε. For $0 < \varepsilon \ll 1$, the variable x changes much slower than y. As long as $\varepsilon \neq 0$, one may also change the time scale to $\tau = t/\varepsilon$, and study the equivalent form:

$$\begin{aligned} \frac{dx}{d\tau} &= \varepsilon f_0(x, y, \varepsilon), \\ \frac{dy}{d\tau} &= g_0(x, y, \varepsilon). \end{aligned} \tag{2}$$

Within this general framework, we will make the following assumptions (some technical terms will be defined later), where the integer $r > 1$ and the positive number ε_0 are fixed from now on:

A1 Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open and bounded. The functions

$$f_0 : U \times V \times [0, \varepsilon_0] \to \mathbb{R}^n$$

and

$$g_0 : U \times V \times [0, \varepsilon_0] \to \mathbb{R}^m$$

are both of class C_b^r, where a function f is in C_b^r if it is in C^r and its derivatives up to order r as well as f itself are bounded.

A2 There is a function

$$m_0 : U \to V$$

of class C_b^r, such that $g_0(x, m_0(x), 0) = 0$ for all x in U.

It is often helpful to consider $z = y - m_0(x)$, and the fast system (2) in the new coordinates becomes:

$$\begin{aligned}\frac{dx}{d\tau} &= \varepsilon f_1(x,z,\varepsilon),\\ \frac{dz}{d\tau} &= g_1(x,z,\varepsilon),\end{aligned}\tag{3}$$

where

$$\begin{aligned}f_1(x,z,\varepsilon) &= f_0\bigl(x, z+m_0(x),\varepsilon\bigr),\\ g_1(x,z,\varepsilon) &= g_0\bigl(x, z+m_0(x),\varepsilon\bigr) - \varepsilon\bigl[D_x m_0(x)\bigr] f_1(x,z,\varepsilon).\end{aligned}$$

When $\varepsilon = 0$, the system (3) degenerates to

$$\frac{dz}{d\tau} = g_1(x,z,0), \qquad x(\tau) \equiv x_0 \in U,\tag{4}$$

seen as equations on $\{z \mid z + m_0(x_0) \in V\}$.

A3 The steady state $z = 0$ of (4) is globally asymptotically stable on $\{z \mid z + m_0(x_0) \in V\}$ for all $x_0 \in U$.

A4 All eigenvalues of the matrix $D_y g_0(x, m_0(x), 0)$ have negative real parts for every $x \in U$, i.e., the matrix $D_y g_0(x, m_0(x), 0)$ is Hurwitz on U. (The notation $D_y g_0(x, m_0(x), 0)$ means the partial derivatives of $g_0(x, y, \varepsilon)$ with respect to y evaluated at the point $(x, m_0(x), 0)$.)

A5 There exists a family of convex compact sets $D_\varepsilon \subset U \times V$, which depend continuously on $\varepsilon \in [0, \varepsilon_0]$, such that (1) is positively invariant on D_ε for $\varepsilon \in (0, \varepsilon_0]$.

A6 The flow ψ_t^0 of the limiting system (set $\varepsilon = 0$ in (1)):

$$\frac{dx}{dt} = f_0\bigl(x, m_0(x), 0\bigr)\tag{5}$$

has eventually positive derivatives on K_0 with respect to some cone, where K_0 is the projection of

$$D_0 \cap \bigl\{(x,y) \mid y = m_0(x), x \in U\bigr\}$$

onto the x-axis.

A7 The set of equilibria of (1) on D_ε is totally disconnected.

Remark 1 Assumption A3 implies that $y = m_0(x)$ is a unique solution of $g_0(x, y, 0) = 0$ on U.

Continuity in A5 is understood with respect to the Hausdorff metric.

In mass-action chemical kinetics, the vector fields are polynomials. So, A1 follows naturally.

Our main theorem is:

Theorem 1 *Under assumptions* A1 *to* A7, *there exists a positive constant $\varepsilon^* < \varepsilon_0$ such that for each $\varepsilon \in (0, \varepsilon^*)$, the forward trajectory of* (1) *starting from almost every point in D_ε converges to some equilibrium.*

3 Monotone Systems of Ordinary Differential Equations

In this section, we review several useful definitions and theorems regarding monotone systems. As we wish to provide results valid for arbitrary orders, not merely orthants, and some of these results, though well-known, are not readily available in a form needed for reference, we provide some technical proofs.

Definition 1 A nonempty, closed set $C \subset \mathbb{R}^N$ is a cone if

1. $C + C \subset C$
2. $\mathbb{R}_+ C \subset C$
3. $C \cap (-C) = \{0\}$.

We always assume $C \neq \{0\}$. Associated to a cone C is a partial order on $\mathbb{R}^N$. For any $x, y \in \mathbb{R}^N$, we define

$$\begin{aligned} x \geq y \quad &\Longleftrightarrow \quad x - y \in C, \\ x > y \quad &\Longleftrightarrow \quad x - y \in C, \quad x \neq y. \end{aligned}$$

When $\operatorname{Int} C$ is not empty, we can define

$$x \gg y \quad \Longleftrightarrow \quad x - y \in \operatorname{Int} C.$$

Definition 2 The dual cone of C is defined as

$$C^* = \left\{\lambda \in \left(\mathbb{R}^N\right)^* \mid \lambda(C) \geq 0\right\}.$$

An immediate consequence is

$$\begin{aligned} x \in C \quad &\Longleftrightarrow \quad \lambda(x) \geq 0, \quad \forall \lambda \in C^*, \\ x \in \operatorname{Int} C \quad &\Longleftrightarrow \quad \lambda(x) > 0, \quad \forall \lambda \in C^* \setminus \{0\}. \end{aligned}$$

With this partial ordering on $\mathbb{R}^N$, we analyze certain features of the dynamics of an ordinary differential equation:

$$\frac{dz}{dt} = F(z), \tag{6}$$

where $F : \mathbb{R}^N \to \mathbb{R}^N$ is a C^1 vector field. We are interested in a special class of equations that preserve the ordering along the trajectories. For simplicity, the solutions of (6) are assumed to exist for all $t \geq 0$ in the sets considered in the following.

Definition 3 The flow ϕ_t of (6) is said to have (eventually) positive derivatives on a set $V \subseteq \mathbb{R}^N$, if $[D_z\phi_t(z)]x \in \operatorname{Int} C$ for all $x \in C \setminus \{0\}$, $z \in V$, and $t \geq 0$ ($t \geq t_0$ for some $t_0 > 0$ independent of z).

It is worth noticing that $[D_z\phi_t(z)]x \in \operatorname{Int} C$ is equivalent to $\lambda([D_z\phi_t(z)]x) > 0$ for all $\lambda \in C^*$ with norm one. We will use this fact in the proof of the next lemma, which deals with "regular" perturbations in the dynamics. The proof is in the same spirit as in Theorem 1.2 of Hirsch (1983), but generalized to the arbitrary cone C.

Lemma 1 *Assume $V \subset \mathbb{R}^N$ is a compact set in which the flow ϕ_t of* (6) *has eventually positive derivatives. Then there exists $\delta > 0$ with the following property. Let ψ_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 $\psi_s(z) \in V$ for all $s \in [0, t]$ where $t \geq t_*$, then $[D_z\psi_t(z)]x \in \operatorname{Int} C$ for all $x \in C \setminus \{0\}$.*

Proof Pick $t^* = t_0 > 0$ so that $\lambda([D_z\phi_t(z)]x) > 0$ for all $t \geq t_0$, $z \in V$, $\lambda \in C^*$, $x \in C$ with $|\lambda| = 1$, $|x| = 1$. Then there exists $\delta > 0$ with the property that when the C^1 norm of $F(z) - G(z)$ is less than δ, we have $\lambda([D_z\psi_t(z)]x) > 0$ for $t_0 \leq t \leq 2t_0$.

When $t > 2t_0$, we write $t = r + kt_0$, where $t_0 \leq r < 2t_0$ and $k \in \mathbb{N}$. If $\psi_s(z) \in V$ for all $s \in [0, t]$, we can define $z_j := \psi_{jt_0}(z)$ for $j = 0, \ldots, k$. For any $x \in C \setminus \{0\}$, using the chain rule, we have:

$$\big[D_z\psi_t(z)\big]x = \big[D_z\psi_r(z_k)\big]\big[D_z\psi_{t_0}(z_{k-1})\big]\cdots\big[D_z\psi_{t_0}(z_0)\big]x.$$

It is easy to see that $[D_z\psi_t(z)]x \in \operatorname{Int} C$. □

Corollary 1 *If V is positively invariant under the flow ψ_t, then ψ_t has eventually positive derivatives in V.*

Proof If V is positively invariant under the flow ψ_t, then for any $z \in V$ the condition $\psi_s(z) \in V$ for $s \in [0, t]$ is satisfied for all $t \geq 0$. By the previous lemma, ψ_t has eventually positive derivatives in V. □

Definition 4 The system (6) or the flow ϕ_t of (6) is called monotone (resp. strongly monotone) in a set $W \subseteq \mathbb{R}^N$, if for all $t > 0$ and $z_1, z_2 \in W$,

$$z_1 \geq z_2 \quad \Rightarrow \quad \phi_t(z_1) \geq \phi_t(z_2) \quad \big(\text{resp. } \phi_t(z_1) \gg \phi_t(z_2) \text{ when } z_1 \neq z_2\big).$$

It is eventually (strongly) monotone if there exists $t_0 > 0$ such that ϕ_t is (strongly) monotone for all $t \geq t_0$.

Definition 5 An set $W \subseteq \mathbb{R}^N$ is called p-convex, if W contains the entire line segment joining x and y whenever $x \leq y$, $x, y \in W$.

Proposition 1 *Let $W \subseteq \mathbb{R}^N$ be p-convex. If the flow ϕ_t has* (*eventually*) *positive derivatives in W, then it is* (*eventually*) *strongly monotone in W.*

Proof For any $z_1 > z_2 \in W, \lambda \in C^* \setminus \{0\}$ and $t \geq 0$ ($t \geq t_0$ for some $t_0 > 0$), we have that $\lambda(\phi_t(z_1) - \phi_t(z_2))$ equals

$$\int_0^1 \lambda\big(\big[D_z\phi_t(sz_1 + (1-s)z_2)\big](z_1 - z_2)\big)\,ds > 0.$$

Therefore, ϕ_t is (eventually) strongly monotone in W. □

The following two lemmas are variations of Hirsch's Generic Convergence Theorem.

Lemma 2 *Suppose that the flow ϕ_t of* (6) *has eventually positive derivatives in a p-convex open set $W \subseteq \mathbb{R}^N$. Let $W^c \subseteq W$ be the set of points whose forward orbit has compact closure in W. If the set of equilibria is totally disconnected* (*e.g.*, *countable*), *then the forward trajectory starting from almost every point in W^c converges to an equilibrium.*

This result follows from a generalization of Theorem 4.1 in Hirsch (1985) to an arbitrary cone.

Definition 6 A point x in a set $W \subseteq \mathbb{R}^N$ is called strongly accessible from below (resp., above) if there exists a sequence $\{y_n\}$ in W converging to x such that $y_n < y_{n+1} < x$ (resp., $y_n > y_{n+1} > x$).

In our motivating example, as well as in most biochemical systems after reduction by elimination of stoichiometric constraints, the set of equilibria is discrete, and thus Lemma 2 will apply. However, the following more general result is also true, and applies even when the set of equilibria is not discrete. This follows as a direct application of Theorem 2.26 in Hirsch and Smith (2005).

Lemma 3 *Suppose that the flow ϕ_t of* (6) *has compact closure and eventually positive derivatives in a p-convex open set $W \subseteq \mathbb{R}^N$. If any point in W can be strongly accessible either from above or from below in W, then the forward trajectory from every point*, *except for initial conditions in a nowhere dense set*, *converges to an equilibrium.*

4 Details of the Proof

Our approach to solve the varying domain problem is motivated by Nipp (1992). The idea is to extend the vector fields from $U \times V \times [0, \varepsilon]$ to $\mathbb{R}^n \times \mathbb{R}^m \times [0, \varepsilon_0]$, then apply geometric singular perturbation theorems (Sakamoto 1990) on $\mathbb{R}^n \times \mathbb{R}^m \times [0, \varepsilon_0]$, and finally restrict the flows to D_ε for the generic convergence result.

4.1 Extensions of the Vector Fields

For a given compact set $K \subset \mathbb{R}^n$ ($K_0 \subseteq K \subset U$), the following procedure is adopted from Nipp (1992) to extend a C_b^r function with respect to the x coordinate from U to $\mathbb{R}^n$, such that the extended function is C_b^r and agrees with the old one on K. This is a routine "smooth patching" argument.

Let U_1 be an open subset of U with C^r boundary and such that $K \subset U_1 \subseteq U$. For $\Theta_0 > 0$ sufficiently small, define

$$U_1^{\Theta_0} := \{x \in U_1 \mid \Theta(x) \geq \Theta_0\}, \quad \text{where } \Theta(x) := \min_{u \in \partial U_1} |x - u|,$$

such that K is contained in $U_1^{\Theta_0}$. Consider the scalar C^∞ function ρ:

$$\rho(a) := \begin{cases} 0, & a \leq 0, \\ \exp(1 - \exp(a-1)/a), & 0 < a < 1, \\ 1, & a \geq 1. \end{cases}$$

Define

$$\hat{\Theta}(x) := \begin{cases} 0, & x \in \mathbb{R}^n \setminus U_1, \\ \Theta(x), & x \in U_1 \setminus U_1^{\Theta_0}, \\ \Theta_0, & x \in U_1^{\Theta_0}, \end{cases}$$

and

$$\bar{\Theta}(x) := \rho\left(\frac{\hat{\Theta}(x)}{\Theta_0}\right).$$

For any $q \in C_b^r(U)$, let

$$\bar{\bar{q}}(x) := \begin{cases} q(x), & x \in U_1, \\ 0, & x \in \mathbb{R}^n \setminus U_1, \end{cases} \quad \text{and} \quad \bar{q}(x) := \bar{\Theta}(x)\bar{\bar{q}}(x).$$

Then $\bar{q}(x) \in C_b^r(\mathbb{R}^n)$ and $\bar{q}(x) \equiv q(x)$ on K.

We fix some $d_0 > 0$ such that

$$L_{d_0} := \{z \in \mathbb{R}^m \mid |z| \leq d_0\} \subset \bigcap_{x \in K} \{z \mid z + m_0(x) \in V\}.$$

Then we extend the functions f_1 and m_0 to $\bar{f}_1$ and $\bar{m}_0$, respectively, with respect to x in the above way. To extend g_1, let us first rewrite the differential equation for z as

$$\frac{dz}{d\tau} = [B(x) + C(x,z)]z + \varepsilon H(x,z,\varepsilon) - \varepsilon[D_x m_0(x)] f_1(x,z,\varepsilon),$$

where

$$B(x) = D_y g_0(x, m_0(x), 0) \quad \text{and} \quad C(x, 0) = 0.$$

Following the above procedures, we extend the functions C and H to $\bar{C}$ and $\bar{H}$, but the extension of B is defined as

$$\bar{B}(x) := \bar{\Theta}(x)\bar{\bar{B}}(x) - \mu\big(1 - \bar{\Theta}(x)\big)I_n,$$

where μ is the positive constant such that the real parts of all eigenvalues of $B(x)$ is less than $-\mu$ for every $x \in K$. According to the definition of $\bar{B}(x)$, all eigenvalues of $\bar{B}(x)$ will have negative real parts less than $-\mu$ *for every* $x \in \mathbb{R}^n$. The extension $\bar{g}_1$, defined as

$$\big[\bar{B}(x) + \bar{C}(x,z)\big]z + \varepsilon\bar{H}(x,z,\varepsilon) - \varepsilon\big[D_x\bar{m}_0(x)\big]\bar{f}_1(x,z,\varepsilon),$$

is then $C_b^{r-1}(\mathbb{R}^n \times L_{d_0} \times [0,\varepsilon_0])$ and agrees with g_1 on $K \times L_{d_0} \times [0,\varepsilon_0]$.

To extend functions $\bar{f}_1$ and $\bar{g}_1$ in the z direction from L_{d_0} to $\mathbb{R}^m$, we use the same extension technique but with respect to z. Let us denote the extensions of $\bar{f}_1, \bar{C}, \bar{H}$ and the function $z = z$ by $\tilde{f}_1, \tilde{C}, \tilde{H}$ and $\tilde{z}$, respectively, then define $\tilde{g}_1$ as

$$\big[\bar{B}(x) + \tilde{C}(x,z)\big]\tilde{z}(z) + \varepsilon\tilde{H}(x,z,\varepsilon) - \varepsilon\big[D_x\bar{m}_0(x)\big]\tilde{f}_1(x,z,\varepsilon),$$

which is now $C_b^{r-1}(\mathbb{R}^n \times \mathbb{R}^m \times [0,\varepsilon_0])$ and agrees with g_1 on $K \times L_{d_1} \times [0,\varepsilon_0]$ for some d_1 slightly less than d_0. Notice that $z = 0$ is a solution of $\tilde{g}_1(x,z,0) = 0$, which guarantees that for the extended system in (x,y) coordinates $(y = z + \bar{m}_0(x))$

$$\begin{aligned} \frac{dx}{d\tau} &= \varepsilon f(x,y,\varepsilon), \\ \frac{dy}{d\tau} &= g(x,y,\varepsilon), \end{aligned} \tag{7}$$

$y = \bar{m}_0(x)$ is the solution of $g(x,y,0) = 0$. To summarize, (7) satisfies

E1 The functions

$$\begin{aligned} f &\in C_b^r\big(\mathbb{R}^n \times \mathbb{R}^m \times [0,\varepsilon_0]\big), \\ g &\in C_b^{r-1}\big(\mathbb{R}^n \times \mathbb{R}^m \times [0,\varepsilon_0]\big), \\ \bar{m}_0 &\in C_b^r\big(\mathbb{R}^n\big), \quad g\big(x,\bar{m}_0(x),0\big) = 0, \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

E2 All eigenvalues of the matrix $D_y g(x,\bar{m}_0(x),0)$ have negative real parts less than $-\mu$ for every $x \in \mathbb{R}^n$.

E3 The function $\bar{m}_0$ coincides with m_0 on K, and the functions f and g coincide with f_0 and g_0, respectively, on

$$\Omega_{d_1} := \big\{(x,y) \mid x \in K,\ |y - m_0(x)| \le d_1\big\}.$$

Conditions E1 and E2 are the assumptions for geometric singular perturbation theorems, and condition E3 ensures that on Ω_{d_1} the flow of (2) coincides with the flow of (7). If we apply geometric singular perturbation theorems to (7) on $\mathbb{R}^n \times \mathbb{R}^m \times [0,\varepsilon_0]$, the exact same results are true for (2) on Ω_{d_1}. For the rest of the paper, we identify the flow of (7) and the flow of (2) on Ω_{d_1} without further mentioning this fact. (Later, in Lemmas 5–8, when globalizing the results, we consider again the original system.)

4.2 Geometric Singular Perturbation Theory

The theory of geometric singular perturbation can be traced back to the work of Fenichel (1979), which first revealed the geometric aspects of singular perturbation problems. Later on, the works by Knobloch and Aulbach (1984), Nipp (1992), and Sakamoto (1990) also presented results similar to Fenichel (1979). By now, the theory is fairly standard, and there have been enormous applications to traveling waves of partial differential equations, see Jones (1995) and the references there.

To apply geometric singular perturbation theorems to the vector fields on $\mathbb{R}^n \times \mathbb{R}^m \times [0, \varepsilon_0]$, we use the theorems stated in Sakamoto (1990). The following lemma is a restatement of the theorems in Sakamoto (1990), and we refer to Sakamoto (1990) for the proof.

Lemma 4 *Under conditions* E1 *and* E2, *there exists a positive* $\varepsilon_1 < \varepsilon_0$ *such that for every* $\varepsilon \in (0, \varepsilon_1]$:

1. *There is a* C_b^{r-1} *function*

$$m : \mathbb{R}^n \times [0, \varepsilon_1] \to \mathbb{R}^m$$

such that the set M_ε *defined by*

$$M_\varepsilon := \left\{(x, m(x, \varepsilon)) \mid x \in \mathbb{R}^n\right\}$$

is invariant under the flow generated by (7). *Moreover,*

$$\sup_{x\in\mathbb{R}^n} \left|m(x, \varepsilon) - \bar{m}_0(x)\right| = O(\varepsilon), \quad \text{as } \varepsilon \to 0.$$

In particular, we have $m(x, 0) = \bar{m}_0(x)$ *for all* $x \in \mathbb{R}^n$.

2. *The set consisting of all the points* (x_0, y_0) *such that*

$$\sup_{\tau\geq 0}\left|y(\tau; x_0, y_0) - m(x(\tau; x_0, y_0), \varepsilon)\right|e^{\frac{\mu\tau}{4}} < \infty,$$

where $(x(\tau; x_0, y_0), y(\tau; x_0, y_0))$ *is the solution of* (7) *passing through* (x_0, y_0) *at* $\tau = 0$, *is a* C^{r-1}*-immersed submanifold in* $\mathbb{R}^n \times \mathbb{R}^m$ *of dimension* $n + m$, *denoted by* $W^s(M_\varepsilon)$, *the stable manifold of* M_ε.

3. *There is a positive constant* δ_0 *such that if*

$$\sup_{\tau\geq 0}\left|y(\tau; x_0, y_0) - m(x(\tau; x_0, y_0), \varepsilon)\right| < \delta_0,$$

then $(x_0, y_0) \in W^s(M_\varepsilon)$.

4. *The manifold* $W^s(M_\varepsilon)$ *is a disjoint union of* C^{r-1}*-immersed manifolds* $W_\varepsilon^s(\xi)$ *of dimension* m:

$$W^s(M_\varepsilon) = \bigcup_{\xi\in\mathbb{R}^n} W_\varepsilon^s(\xi).$$

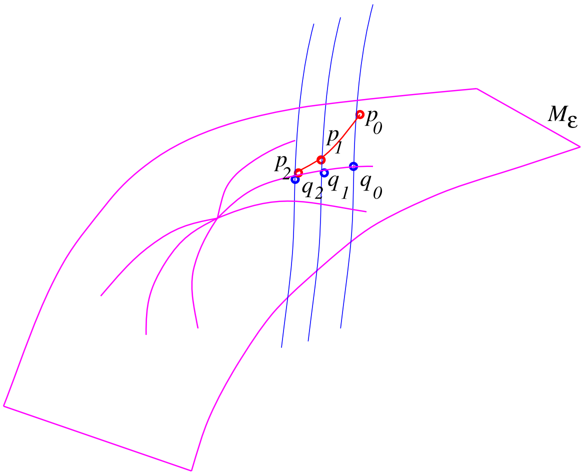

Fig. 2 An illustration of the "positive invariant" and "asymptotic phase" properties. Let p_0 be a point on the fiber $W^s_\varepsilon(q_0)$ (*vertical curve*). Suppose the solution of (7) starting from $q_0 \in M_\varepsilon$ evolves to $q_1 \in M_\varepsilon$ at time τ_1, then the solution of (7) starting from p_0 will evolve to $p_1 \in W^s_\varepsilon(q_1)$ at time τ_1. At time τ_2, they evolve to q_2, p_2 respectively. These two solutions are always on the same fiber. If we know that the one starting from q_0 converges to an equilibrium, then the one starting from p_0 also converges to an equilibrium

For each $\xi \in \mathbb{R}^n$, let $H_\varepsilon(\xi)(\tau)$ be the solution for $\tau \geq 0$ of

$$\frac{dx}{d\tau} = \varepsilon f(x, m(x,\varepsilon), \varepsilon), \qquad x(0) = \xi \in \mathbb{R}^n.$$

Then, the manifold $W^s_\varepsilon(\xi)$ is the set

$$\left\{(x_0, y_0) \mid \sup_{\tau\geq 0}\big|\tilde{x}(\tau)\big|e^{\frac{\mu\tau}{4}} < \infty, \sup_{\tau\geq 0}\big|\tilde{y}(\tau)\big|e^{\frac{\mu\tau}{4}} < \infty\right\},$$

where

$$\tilde{x}(\tau) = x(\tau; x_0, y_0) - H_\varepsilon(\xi)(\tau),$$
$$\tilde{y}(\tau) = y(\tau; x_0, y_0) - m\big(H_\varepsilon(\xi)(\tau), \varepsilon\big).$$

5. *The fibers are "positively invariant" in the sense that $W^s_\varepsilon(H_\varepsilon(\xi)(\tau))$ is the set*

$$\big\{\big(x(\tau; x_0, y_0), y(\tau; x_0, y_0)\big) \mid (x_0, y_0) \in W^s_\varepsilon(\xi)\big\}$$

for each $\tau \geq 0$, see Fig. 2.

6. *The fibers restricted to the δ_0 neighborhood of M_ε, denoted by $W^s_{\varepsilon,\delta_0}$, can be parameterized as follows. There are two C^{r-1}_b functions*

$$P_{\varepsilon,\delta_0} : \mathbb{R}^n \times L_{\delta_0} \to \mathbb{R}^n,$$
$$Q_{\varepsilon,\delta_0} : \mathbb{R}^m \times L_{\delta_0} \to \mathbb{R}^m,$$

and a map

$$T_{\varepsilon,\delta_0} : \mathbb{R}^n \times L_{\delta_0} \to \mathbb{R}^n \times \mathbb{R}^m$$

mapping (ξ, η) to (x, y), where

$$x = \xi + P_{\varepsilon,\delta_0}(\xi, \eta), \qquad y = m(x, \varepsilon) + Q_{\varepsilon,\delta_0}(\xi, \eta)$$

such that

$$W^s_{\varepsilon,\delta_0}(\xi) = T_{\varepsilon,\delta_0}(\xi, L_{\delta_0}).$$

Remark 2 The δ_0 in property 3 can be chosen uniformly for $\varepsilon \in (0, \varepsilon_0]$. Without loss of generality, we assume that $\delta_0 < d_1$.

Notice that property 4 ensures that for each $(x_0, y_0) \in W^s(M_\varepsilon)$, there exists a ξ such that

$$\begin{aligned}\bigl|x(\tau; x_0, y_0) - H_\varepsilon(\xi)(\tau)\bigr| &\to 0,\\ \bigl|y\bigl(\tau; x_0, y_0\bigr) - m\bigl(H_\varepsilon(\xi)(\tau), \varepsilon\bigr)\bigr| &\to 0.\end{aligned}$$

as $\tau \to \infty$. This is often referred to as the "asymptotic phase" property, see Fig. 2.

4.3 Further Analysis of the Dynamics

The first property of Lemma 4 concludes the existence of an invariant manifold M_ε. There are two reasons to introduce M_ε. First, on M_ε the x-equation is decoupled from the y-equation:

$$\begin{aligned}\frac{dx}{dt} &= f\bigl(x, m(x, \varepsilon), \varepsilon\bigr),\\ y(t) &= m\bigl(x(t), \varepsilon\bigr).\end{aligned} \tag{8}$$

This reduction allows us to analyze a lower dimensional system, whose dynamics may have been well studied. Second, when ε approaches zero, the limit of (8) is (5) on K_0. If (5) has some desirable property, it is natural to expect that this property is inherited by (8). An example of this principle is provided by the following lemma:

Lemma 5 *There exists a positive constant $\varepsilon_2 < \varepsilon_1$, such that for each $\varepsilon \in (0, \varepsilon_2)$, the flow ψ_t^ε of (8) has eventually positive derivatives on K_ε, which is the projection of $M_\varepsilon \cap D_\varepsilon$ to the x-axis.*

Proof Assumption A6 states that the flow ψ_t^0 of the limiting system (5) has eventually positive derivatives on K_0. By the continuity of $m(x, \varepsilon)$ and D_ε at $\varepsilon = 0$, we can pick ε_2 small enough such that the flow ψ_t^0 has eventually positive derivatives on K_ε for all $\varepsilon \in (0, \varepsilon_2)$. Applying Corollary 1, we conclude that the flow ψ_t^ε of (8) has eventually positive derivatives on K_ε provided K_ε is positively invariant under (8), which follows easily from the facts that (7) is positively invariant on D_ε and M_ε is an invariant manifold. □

The next lemma asserts that the generic convergence property is preserved for (8), see Fig. 3.

Lemma 6 *For each $\varepsilon \in (0, \varepsilon_2)$, there exists a set $C_\varepsilon \subseteq K_\varepsilon$ such that the forward trajectory of (8) starting from any point of C_ε converges to some equilibrium, and the Lebesgue measure of $K_\varepsilon \setminus C_\varepsilon$ is zero.*

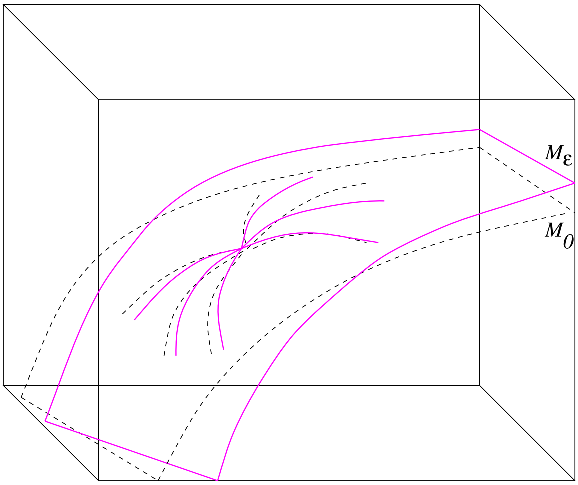

Fig. 3 This is a sketch of the manifolds M_0 (surface bounded by *dashed curves*), M_ε (surface bounded by *dotted curves*), and D_ε (the *cube*). It highlights two major characteristics of M_ε. First, M_ε is close to M_0. Second, the trajectories on M_ε converge to equilibria if those on M_0 do

Proof There exists a convex open set W_ε containing K_ε such that flow ψ_t^ε of (8) has eventually positive derivatives on W_ε. Assumption A5 assures that $K_\varepsilon \subseteq W_\varepsilon^c$. The proof is completed by applying Lemma 2 under the assumption A7. □

By now, we have discussed flows restricted to the invariant manifold M_ε. Next, we will explore the conditions for a point to be on $W^s(M_\varepsilon)$, the stable manifold of M_ε. Property 3 of Lemma 4 provides a sufficient condition, namely, any point (x_0, y_0) such that

$$\sup_{\tau\geq 0}\left|y(\tau; x_0, y_0) - m\bigl(x(\tau; x_0, y_0), \varepsilon\bigr)\right| < \delta_0 \tag{9}$$

is on $W^s(M_\varepsilon)$. In fact, if we know that the difference between y_0 and $m(x_0, \varepsilon)$ is sufficiently small, then the above condition is always satisfied. More precisely, we have:

Lemma 7 *There exists* $\varepsilon_3 > 0$, $\delta_0 > d > 0$, *such that for each* $\varepsilon \in (0, \varepsilon_3)$, *if the initial condition satisfies* $|y_0 - m(x_0, \varepsilon)| < d$, *then* (9) *holds, i.e.*, $(x_0, y_0) \in W^s(M_\varepsilon)$.

Proof Follows from the proof of Claim 1 in Nipp (1992). □

Before we get further into the technical details, let us give an outline of the proof of the main theorem. The proof can be decomposed into three steps. First, we show that almost every trajectory on $D_\varepsilon \cap M_\varepsilon$ converges to some equilibrium. This is precisely Lemma 6. Second, we show that almost every trajectory starting from $W^s(M_\varepsilon)$ converges to some equilibrium. This follows from Lemma 6 and the "asymptotic phase" property in Lemma 4, but we still need to show that the set of nonconvergent initial conditions is of measure zero. The last step is to show that all trajectories in D_ε will eventually stay in $W^s(M_\varepsilon)$, which is our next lemma:

Lemma 8 *There exist positive* τ_0 *and* $\varepsilon_4 < \varepsilon_3$, *such that* $(x(\tau_0), y(\tau_0)) \in W^s(M_\varepsilon)$ *for all* $\varepsilon \in (0, \varepsilon_4)$, *where* $(x(\tau), y(\tau))$ *is the solution to* (2) *with the initial condition* $(x_0, y_0) \in D_\varepsilon$.

Proof It is convenient to consider the problem in (x, z) coordinates. Let $(x(\tau), z(\tau))$ be the solution to (3) with initial condition (x_0, z_0), where $z_0 = y_0 - m(x_0, 0)$. We first show that there exists a τ_0 such that $|z(\tau_0)| \le d/2$.

Expanding $g_1(x, z, \varepsilon)$ at the point $(x_0, z, 0)$, the equation of z becomes

$$\frac{dz}{d\tau} = g_1(x_0, z, 0) + \frac{\partial g_1}{\partial x}(\xi, z, 0)(x - x_0) + \varepsilon R(x, z, \varepsilon)$$

for some $\xi(\tau)$ between x_0 and $x(\tau)$ (where $\xi(\tau)$ can be picked continuously in τ). Let us write

$$z(\tau) = z^0(\tau) + w(\tau),$$

where $z^0(\tau)$ is the solution to (4) with initial the condition $z^0(0) = z_0$, and $w(\tau)$ satisfies

$$\begin{aligned}\frac{dw}{d\tau} &= g_1(x_0, z, 0) - g_1(x_0, z^0, 0) + \frac{\partial g_1}{\partial x}(\xi, z, 0)(x - x_0) + \varepsilon R(x, z, \varepsilon)\\ &= \frac{\partial g_1}{\partial z}(x_0, \zeta, 0)w + \varepsilon \frac{\partial g_1}{\partial x}(\xi, z, 0)\int_0^\tau f_1(x(s), z(s), \varepsilon)\,ds + \varepsilon R(x, z, \varepsilon), \quad (10)\end{aligned}$$

with the initial condition $w(0) = 0$ and some $\zeta(\tau)$ between $z^0(\tau)$ and $z(\tau)$ (where $\zeta(\tau)$ can be picked continuously in τ).

By assumption A3, there exist a positive τ_0 such that $|z^0(\tau)| \le d/4$ for all $\tau \ge \tau_0$. Notice that we are working on the compact set D_ε, so τ_0 can be chosen uniformly for all initial conditions in D_ε.

We write the solution of (10) as

$$\begin{aligned}w(\tau) = &\int_0^\tau \frac{\partial g_1}{\partial z}(x_0, \zeta, 0)w\,ds\\ &+ \varepsilon\int_0^\tau \left(\frac{\partial g_1}{\partial x}(\xi, z, 0)\int_0^{s'} f_1(x, z, \varepsilon)\,ds' + R(x, z, \varepsilon)\right)ds.\end{aligned}$$

Since the functions f_1, R and the derivatives of g_1 are bounded on D_ε, we have

$$\big|w(\tau)\big| \le \int_0^\tau L|w|\,ds + \varepsilon\int_0^\tau \left(M_1\int_0^{s'} M_2\,ds' + M_3\right)ds,$$

for some positive constants L, M_i, $i = 1, 2, 3$. The notation $|w|$ means the Euclidean norm of $w \in \mathbb{R}^m$. Moreover, if we define

$$\alpha(\tau) = \int_0^\tau \left(M_1\int_0^{s'} M_2\,ds' + M_3\right)ds,$$

then

$$\left|w(\tau)\right| \leq \int_0^\tau L|w|\,ds + \varepsilon\alpha(\tau_0),$$

for all $\tau \in [0, \tau_0]$ as α is increasing in τ. Applying Gronwall's inequality (Sontag 1990), we have

$$\left|w(\tau)\right| \leq \varepsilon\alpha(\tau_0)e^{L\tau},$$

which holds in particular at $\tau = \tau_0$. Finally, we choose ε_4 small enough such that $\varepsilon\alpha(\tau_0)e^{L\tau_0} < d/4$ and $|m(x,\varepsilon) - m(x,0)| < d/2$ for all $\varepsilon \in (0, \varepsilon_4)$. Then we have

$$\begin{aligned}\left|y(\tau_0) - m(x(\tau_0),\varepsilon)\right| &\leq \left|y(\tau_0) - m(x(\tau_0),0)\right| + \left|m(x(\tau_0),\varepsilon) - m(x(\tau_0),0)\right| \\ &< \left|z(\tau_0)\right| + d/2 \\ &< d/2 + d/2 = d.\end{aligned}$$

That is, $(x(\tau_0), y(\tau_0)) \in W^s(M_\varepsilon)$ by Lemma 7. □

By now, we have completed all three steps, and are ready to prove Theorem 1.

4.4 Proof of Theorem 1

Proof Let $\varepsilon^* = \min\{\varepsilon_2, \varepsilon_4\}$. For $\varepsilon \in (0, \varepsilon^*)$, it is equivalent to prove the result for the fast system (2). Pick an arbitrary point (x_0, y_0) in D_ε, and there are three cases:

1. $y_0 = m(x_0, \varepsilon)$, that is, $(x_0, y_0) \in M_\varepsilon \cap D_\varepsilon$. By Lemma 6, the forward trajectory converges to an equilibrium except for a set of measure zero.
2. $0 < |y_0 - m(x_0, \varepsilon)| < d$. By Lemma 7, we know that (x_0, y_0) is in $W^s(M_\varepsilon)$. Then, property 4 of Lemma 4 guarantees that the point (x_0, y_0) is on some fiber $W^s_{\varepsilon,d}(\xi)$, where $\xi \in K_\varepsilon$. If $\xi \in C_\varepsilon$, that is, the forward trajectory of ξ converges to some equilibrium, then by the "asymptotic phase" property of Lemma 4, the forward trajectory of (x_0, y_0) also converges to an equilibrium. To deal with the case when ξ is not in C_ε, it is enough to show that the set

   $$B_{\varepsilon,d} = \bigcup_{\xi \in K_\varepsilon \setminus C_\varepsilon} W^s_{\varepsilon,d}(\xi)$$

   has measure zero in $\mathbb{R}^{m+n}$. Define

   $$S_{\varepsilon,d} = (K_\varepsilon \setminus C_\varepsilon) \times L_d.$$

   By Lemma 6, $K_\varepsilon \setminus C_\varepsilon$ has measure zero in $\mathbb{R}^n$, thus $S_{\varepsilon,d}$ has measure zero in $\mathbb{R}^n \times \mathbb{R}^m$. On the other hand, property 6 in Lemma 4 implies $B_{\varepsilon,d} = T_{\varepsilon,d}(S_{\varepsilon,d})$. Since Lipschitz maps send measure zero sets to measure zero sets, $B_{\varepsilon,d}$ is of measure zero.
3. $|y_0 - m(x_0, \varepsilon)| \geq d$. By Lemma 8, the point $(x(\tau_0), y(\tau_0))$ is in $W^s(M_\varepsilon)$ and we are back to case 2. The proof is completed if the set $\phi^\varepsilon_{-\tau_0}(B_{\varepsilon,d})$ has measure zero, where ϕ^ε_τ is the flow of (2). This is true because ϕ^ε_τ is a diffeomorphism for any finite τ. □

5 Applications

5.1 An Application to the Dual Futile Cycle

Our structure of futile cycles in Fig. 1 implicitly assumes a sequential instead of a random mechanism. By a sequential mechanism, we mean that the kinase phosphorylates the substrates in a specific order, and the phosphatase works in the reverse order. A few kinases are known to be sequential, for example, the auto-phosphorylation of FGF-receptor-1 kinase (Furdui et al. 2006). This assumption dramatically reduces the number of different phospho-forms and simplifies our analysis. In a special case when the kinetic constants of each phosphorylation are the same and the kinetic constants of each dephosphorylation are the same, the random mechanism can be easily included in the sequential case. We therefore write down the chemical reactions in Fig. 1 as follows:

$$S_0 + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C_1 \overset{k_2}{\to} S_1 + E \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} C_2 \overset{k_4}{\to} S_2 + E,$$

$$S_2 + F \underset{h_{-1}}{\overset{h_1}{\rightleftarrows}} C_3 \overset{h_2}{\to} S_1 + F \underset{h_{-3}}{\overset{h_3}{\rightleftarrows}} C_4 \overset{h_4}{\to} S_0 + F.$$

There are three conservation relations:

$$\begin{aligned}
S_{\text{tot}} &= [S_0] + [S_1] + [S_2] + [C_1] + [C_2] + [C_4] + [C_3],\\
E_{\text{tot}} &= [E] + [C_1] + [C_2],\\
F_{\text{tot}} &= [F] + [C_4] + [C_3],
\end{aligned}$$

where brackets indicate concentrations. Based on mass action kinetics, we have the following set of ordinary differential equations:

$$\begin{aligned}
\frac{d[S_0]}{d\tau} &= h_4[C_4] - k_1[S_0][E] + k_{-1}[C_1],\\
\frac{d[S_2]}{d\tau} &= k_4[C_2] - h_1[S_2][F] + h_{-1}[C_3],\\
\frac{d[C_1]}{d\tau} &= k_1[S_0][E] - (k_{-1} + k_2)[C_1],\\
\frac{d[C_2]}{d\tau} &= k_3[S_1][E] - (k_{-3} + k_4)[C_2],\\
\frac{d[C_4]}{d\tau} &= h_3[S_1][F] - (h_{-3} + h_4)[C_4],\\
\frac{d[C_3]}{d\tau} &= h_1[S_2][F] - (h_{-1} + h_2)[C_3].
\end{aligned} \tag{11}$$

After rescaling the concentrations and time, (11) becomes

$$
\begin{aligned}
\frac{dx_1}{dt} &= -k_1 S_{\text{tot}} x_1(1-y_1-y_2) + k_{-1}y_1 + h_4 c y_3,\\
\frac{dx_2}{dt} &= -h_1 S_{\text{tot}} c x_2(1-y_3-y_4) + h_{-1} c y_4 + k_4 y_2,\\
\varepsilon\frac{dy_1}{dt} &= k_1 S_{\text{tot}} x_1(1-y_1-y_2) - (k_{-1}+k_2)y_1,\\
\varepsilon\frac{dy_2}{dt} &= k_3 S_{\text{tot}}(1-x_1-x_2-\varepsilon y_1-\varepsilon y_2-\varepsilon c y_3-\varepsilon c y_4)\\
&\quad\times(1-y_1-y_2)-(k_{-3}+k_4)y_2,\\
\varepsilon\frac{dy_3}{dt} &= h_3 S_{\text{tot}}(1-x_1-x_2-\varepsilon y_1-\varepsilon y_2-\varepsilon c y_3-\varepsilon c y_4)\\
&\quad\times(1-y_3-y_4)-(h_{-3}+h_4)y_3,\\
\varepsilon\frac{dy_4}{dt} &= h_1 S_{\text{tot}} x_2(1-y_3-y_4)-(h_{-1}+h_2)y_4,
\end{aligned}
\tag{12}
$$

where

$$
\begin{aligned}
&x_1=\frac{[S_0]}{S_{\text{tot}}},\qquad x_2=\frac{[S_2]}{S_{\text{tot}}},\qquad y_1=\frac{[C_1]}{E_{\text{tot}}},\qquad y_2=\frac{[C_2]}{E_{\text{tot}}},\\
&y_3=\frac{[C_4]}{F_{\text{tot}}},\qquad y_4=\frac{[C_3]}{F_{\text{tot}}},\qquad \varepsilon=\frac{E_{\text{tot}}}{S_{\text{tot}}},\qquad c=\frac{F_{\text{tot}}}{E_{\text{tot}}},\qquad t=\tau\varepsilon.
\end{aligned}
$$

These equations are in the form of (1). The conservation laws suggest taking $\varepsilon_0 = 1/(1+c)$ and

$$
\begin{aligned}
D_\varepsilon = \big\{&(x_1,x_2,y_1,y_2,y_3,y_4)\mid 0\le y_1+y_2\le 1,\\
&0\le y_3+y_4\le 1, x_1,x_2,y_1,y_2,y_3,y_4\ge 0,\\
&0\le x_1+x_2+\varepsilon(y_1+y_2+cy_3+cy_4)\le 1\big\}.
\end{aligned}
$$

For $\varepsilon\in(0,\varepsilon_0]$, taking the inner product of the normal of ∂D_ε and the vector fields, it is easy to check that (12) is positively invariant on D_ε, so A5 holds. We want to emphasize that, in this example, the domain D_ε is a convex polytope varying with ε.

It can be proved that on D_ϵ system (12) has at most a finite number of steady states, and thus A7 holds. This is a consequence of a more general result, proved using some of the ideas given in Gunawardena (2005), concerning the number of steady states of more general systems of phosphorylation/dephosphorylation reactions, see Wang and Sontag (2008).

Solving $g_0(x,y,0)=0$, we get

$$
\begin{aligned}
y_1 &= \frac{x_1}{\frac{K_{m1}}{S_{\text{tot}}}+\frac{K_{m1}(1-x_1-x_2)}{K_{m2}}+x_1},\\
y_2 &= \frac{\frac{K_{m1}(1-x_1-x_2)}{K_{m2}}}{\frac{K_{m1}}{S_{\text{tot}}}+\frac{K_{m1}(1-x_1-x_2)}{K_{m2}}+x_1},
\end{aligned}
$$

$$y_3 = \frac{\frac{K_{m3}(1-x_1-x_2)}{K_{m4}}}{\frac{K_{m3}}{S_{\mathrm{tot}}} + \frac{K_{m3}(1-x_1-x_2)}{K_{m4}} + x_2},$$

$$y_4 = \frac{x_2}{\frac{K_{m3}}{S_{\mathrm{tot}}} + \frac{K_{m3}(1-x_1-x_2)}{K_{m4}} + x_2},$$

where K_{m1}, K_{m2}, K_{m3} and K_{m4} are the Michaelis–Menten constants defined as

$$K_{m1} = \frac{k_{-1}+k_2}{k_1}, \qquad K_{m2} = \frac{k_{-3}+k_4}{k_3},$$
$$K_{m3} = \frac{h_{-1}+h_2}{h_1}, \qquad K_{m4} = \frac{h_{-3}+h_4}{h_3}.$$

Now, we need to find a proper set $U \subset \mathbb{R}^2$ satisfying assumptions A1–A4. Suppose that U has the form

$$U = \left\{(x_1, x_2) \mid x_1 > -\sigma,\ x_2 > -\sigma,\ x_1 + x_2 < 1 + \sigma\right\},$$

for some positive σ, and V is any bounded open set such that D_ε is contained in $U \times V$, then A1 follows naturally. Moreover, if

$$\sigma \le \sigma_0 := \min\left\{\frac{K_{m1}K_{m2}}{S_{\mathrm{tot}}(K_{m1}+K_{m2})}, \frac{K_{m3}K_{m4}}{S_{\mathrm{tot}}(K_{m3}+K_{m4})}\right\},$$

A2 also holds. To check A4, let us look at the matrix:

$$B(x) := D_y g_0(x, m_0(x), 0) = \begin{pmatrix} B_1(x) & 0 \\ 0 & B_2(x) \end{pmatrix},$$

where the column vectors of $B_1(x)$ are

$$B_1^1(x) = \begin{pmatrix} -k_1 S_{\mathrm{tot}} x_1 - (k_{-1}+k_2) \\ -k_3 S_{\mathrm{tot}}(1-x_1-x_2) \end{pmatrix},$$

$$B_1^2(x) = \begin{pmatrix} -k_1 S_{\mathrm{tot}} x_1 \\ -k_3 S_{\mathrm{tot}}(1-x_1-x_2) - (k_{-3}+k_4) \end{pmatrix},$$

and the column vectors of $B_2(x)$ are

$$B_2^1(x) = \begin{pmatrix} -h_3 S_{\mathrm{tot}}(1-x_1-x_2) - (h_{-3}+h_4) \\ -h_1 S_{\mathrm{tot}} x_2 \end{pmatrix},$$

$$B_2^2(x) = \begin{pmatrix} -h_3 S_{\mathrm{tot}}(1-x_1-x_2) \\ -h_1 S_{\mathrm{tot}} x_2 - (h_{-1}+h_2) \end{pmatrix}.$$

If both matrices B_1 and B_2 have negative traces and positive determinants, then A4 holds. The trace of B_1 is

$$-k_1 S_{\mathrm{tot}} x_1 - (k_{-1}+k_2) - k_3 S_{\mathrm{tot}}(1-x_1-x_2) - (k_{-3}+k_4).$$

It is negative provided that

$$\sigma \le \frac{k_{-1}+k_2+k_{-3}+k_4}{S_{\text{tot}}(k_1+k_3)}.$$

The determinant of B_1 is

$$k_1(k_{-3}+k_4)S_{\text{tot}}x_1 + k_3(k_{-1}+k_2)S_{\text{tot}}(1-x_1-x_2) + (k_{-1}+k_2)(k_{-3}+k_4).$$

It is positive if

$$\sigma \le \frac{(k_{-1}+k_2)(k_{-3}+k_4)}{S_{\text{tot}}(k_1(k_{-3}+k_4)+k_3(k_{-1}+k_2))}.$$

The condition for B_2 can be derived similarly. To summarize, if we take

$$\sigma = \min\left\{\sigma_0, \frac{k_{-1}+k_2+k_{-3}+k_4}{S_{\text{tot}}(k_1+k_3)}, \frac{(k_{-1}+k_2)(k_{-3}+k_4)}{S_{\text{tot}}(k_1(k_{-3}+k_4)+k_3(k_{-1}+k_2))}, \right.$$
$$\left. \frac{h_{-1}+h_2+h_{-3}+h_4}{S_{\text{tot}}(h_1+h_3)}, \frac{(h_{-1}+h_2)(h_{-3}+h_4)}{S_{\text{tot}}(h_1(h_{-3}+h_4)+h_3(h_{-1}+h_2))}\right\},$$

then the assumptions A1, A2 and A4 will hold.

Notice that dy/dt in (12) is linear in y when $\varepsilon = 0$, so g_1 (defined as in (3)) is linear in z, and hence the equation for z can be written as

$$\frac{dz}{d\tau} = B(x_0)z, \quad x_0 \in U,$$

where the matrix $B(x_0)$ is Hurwitz for every $x_0 \in U$. Therefore, A3 also holds.

It remains to show that assumption A6 is satisfied. Let us look at the reduced system ($\varepsilon = 0$ in (12)):

$$\begin{aligned}
\frac{dx_1}{dt} &= -\frac{k_2 x_1}{\frac{K_{m1}}{S_{\text{tot}}} + \frac{K_{m1}(1-x_1-x_2)}{K_{m2}} + x_1} + \frac{h_4 c \frac{K_{m3}(1-x_1-x_2)}{K_{m4}}}{\frac{K_{m3}}{S_{\text{tot}}} + \frac{K_{m3}(1-x_1-x_2)}{K_{m4}} + x_2} := F_1(x_1,x_2) \\
\frac{dx_2}{dt} &= -\frac{h_2 c x_2}{\frac{K_{m3}}{S_{\text{tot}}} + \frac{K_{m3}(1-x_1-x_2)}{K_{m4}} + x_2} + \frac{k_4 \frac{K_{m1}(1-x_1-x_2)}{K_{m2}}}{\frac{K_{m1}}{S_{\text{tot}}} + \frac{K_{m1}(1-x_1-x_2)}{K_{m2}} + x_1} := F_2(x_1,x_2).
\end{aligned} \tag{13}$$

It is easy to see that F_1 is strictly decreasing in x_2, and F_2 is strictly decreasing in x_1 on

$$K_0 = \big\{(x_1,x_2) \mid x_1 \ge 0, x_2 \ge 0, x_1+x_2 \le 1\big\}.$$

The reduced system (13) is a strictly competitive system. By Theorem 1.1 of Hirsch (1985), flows of (13) have positive derivatives with respect to the cone

$$\big\{(x_1,x_2) \mid x_1 \le 0, x_2 \ge 0\big\},$$

and thus assumption A6 is satisfied. Applying Theorem 1, we have:

Theorem 2 *There exist a positive $\varepsilon^* < \varepsilon_0$ such that for each $\varepsilon \in (0, \varepsilon^*)$, the forward trajectory of* (12) *starting from almost every point in D_ε converges to some equilibrium.*

It is worth pointing out that the conclusion we obtained from Theorem 2 is valid only for small enough ε; that is, the concentration of the enzyme should be much smaller than the concentration of the substrate. Unfortunately, this is not always true in biological systems, especially when feedbacks are present. However, if the sum of the Michaelis–Menten constants and the total concentration of the substrate are much larger than the concentration of enzyme, a different scaling,

$$x_1 = \frac{[S_0]}{A}, \qquad x_2 = \frac{[S_2]}{A}, \qquad \varepsilon' = \frac{E_{\text{tot}}}{A}, \qquad t = \tau\varepsilon',$$

where $A = S_{\text{tot}} + K_{m1} + K_{m2} + K_{m3} + K_{m4}$ allows us to obtain the same convergence result.

5.2 Another Example

The following example demonstrates the importance of the smallness of ε. Consider an $m+1$-dimensional system,

$$\begin{aligned} \frac{dx}{dt} &= \gamma(y_1, \dots, y_m) - \beta(x) \\ \varepsilon\frac{dy_i}{dt} &= -d_i y_i - \alpha_i(x), \quad d_i > 0,\ i = 1, \dots, m, \end{aligned} \tag{14}$$

under the following assumptions:

1. There exists an integer $r > 1$ such that the derivatives of γ, β, and α_i are of class C_b^r for sufficiently large bounded sets.
2. The function $\beta(x)$ is odd, and it approaches infinity as x approaches infinity.
3. The function $\alpha_i(x)$ $(i = 1, \dots, m)$ is bounded by positive constant M_i for all $x \in \mathbb{R}$.
4. The number of roots to the equation

$$\gamma(\alpha_1(x), \dots, \alpha_m(x)) = \beta(x)$$

is countable.

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 $\varepsilon > 1$, limit cycles may appear.

Assumption 4 implies A7, and because of the form of (14), A3 and A4 follow naturally. A6 also holds, as every one-dimensional system is strongly monotone. For A5, we take

$$D_\varepsilon = \left\{(x, y) \mid |x| \le a, |y_i| \le b_i, i = 1, \dots, m\right\},$$

where b_i is an arbitrary positive number greater than $\frac{M_i}{d_i}$ and a can be any positive number such that

$$\beta(a) > N_b := \max_{|y_i| \le b_i} \gamma(y_1, \ldots, y_m).$$

Picking such b_i and a assures

$$x\frac{dx}{dt} < 0, \qquad y_i\frac{dy_i}{dt} < 0,$$

i.e., the vector fields point transversely inside on the boundary of D_ε. Let U and V be some bounded open sets such that $D_\varepsilon \subset U \times V$, and assumption 1 holds on U and V. Then A1 and A2 follow naturally. By our main theorem, for sufficiently small ε, the forward trajectory of (14) starting from almost every point in D_ε converges to some equilibrium.

On the other hand, convergence does not hold for large ε. Let

$$\beta(x) = \frac{x^3}{3} - x, \qquad \alpha_1(x) = 2\tanh x,$$
$$m = 1, \qquad \gamma(y_1) = y_1, \qquad d_1 = 1.$$

It is easy to verify that $(0, 0)$ is the only equilibrium, and the Jacobian matrix at $(0, 0)$ is

$$\begin{pmatrix} 1 & 1 \\ -2/\varepsilon & -1/\varepsilon \end{pmatrix}.$$

When $\varepsilon > 1$, the trace of the above matrix is $1 - 1/\varepsilon > 0$, its determinant is $1/\varepsilon > 0$, so the (only) equilibrium in D_ε is repelling. On the other hand, the set D_ε is chosen such that the vector fields point transversely inside on the boundary of D_ε. By the Poincaré–Bendixson Theorem, there exists a limit cycle in D_ε.

6 Conclusions

Singular perturbation techniques are routinely used in the analysis of biological systems. The *geometric* approach is a powerful tool for global analysis, since it permits one to study the behavior for finite ε on a manifold in which the dynamics are “close” to the slow dynamics. Moreover, and most relevant to us, a suitable fibration structure allows the “tracking” of trajectories and hence the lifting to the full system of the exceptional set of nonconvergent trajectories, if the slow system satisfies the conditions of Hirsch’s Theorem. Using the geometric approach, we were able to provide a global convergence theorem for singularly perturbed strongly monotone systems, in a form that makes it applicable to the study of double futile cycles.

Acknowledgements We have benefited greatly from correspondence with Christopher Jones and Kaspar Nipp about geometric singular perturbation theory. We also wish to thank Alexander van Oudenaarden for questions that triggered much of this research, and David Angeli, Thomas Gedeon, and Hal Smith for helpful discussions.

This work was supported in part by NSF Grant DMS-0614371.

References

Angeli, D., Sontag, E.D.: Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles. Nonlinear Anal. Ser. B Real World Appl. **9**, 128–140 (2008)

Asthagiri, A.R., Lauffenburger, D.A.: A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. Biotechnol. Prog. **17**, 227–239 (2001)

Bijlsma, J.J., Groisman, E.A.: Making informed decisions: regulatory interactions between two-component systems. Trends Microbiol. **11**, 359–366 (2003)

Burack, W.R., Sturgill, T.W.: The Activating dual phosphorylation of MAPK by MEK is nonprocessive. Biochemistry **36**, 5929–5933 (1997)

Chang, L., Karin, M.: Mammalian MAP kinase signaling cascades. Nature **410**, 37–40 (2001)

Chen, H., Bernstein, B.W., Bamburg, J.R.: Regulating actin filament dynamics in vivo. Trends Biochem. Sci. **25**, 19–23 (2000)

Chisci, L., Falugi, P.: Asymptotic tracking for state-constrained monotone systems. In: Proc. 44th IEEE Conf. Decision and Control, Seville, paper ThC17.5 (2005)

Christofides, P.D., Teel, A.R.: Singular perturbations and input-to-state stability. IEEE Trans. Automat. Contr. **41**, 1645–1650 (1996)

Dancer, E.N.: Some remarks on a boundedness assumption for monotone dynamical systems. Proc. Am. Math. Soc. **126**, 801–807 (1998)

Donovan, S., Shannon, K.M., Bollag, G.: GTPase activating proteins: critical regulators of intracellular signaling. Biochim. Biophys. Acta **1602**, 23–45 (2002)

Edelstein-Keshet, L.: Mathematical Models in Biology. McGraw-Hill, New York (1988)

Fenichel, N.: Geometric singular perturbation theory for ordinary differential equations. J. Diff. Equ. **31**, 53–98 (1979)

Ferrell, J.E., Bhatt, R.R.: Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. J. Biol. Chem. **272**, 19008–19016 (1997)

Furdui, C.M., Lew, E.D., Schlessinger, J., Anderson, K.S.: Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. Molec. Cell **21**, 711–717 (2006)

Gardner, T.S., Cantor, C.R., Collins, J.J.: Construction of a genetic toggle switch in Escherichia coli. Nature **403**, 339–342 (2000)

Grossman, A.D.: Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis. Annu. Rev. Genet. **29**, 477–508 (1995)

Gunawardena, J.: Multisite protein phosphorylation makes a good threshold but can be a poor switch. Proc. Natl. Acad. Sci. **102**, 14617–14622 (2005)

Hirsch, M.: Differential equations and convergence almost everywhere in strongly monotone flows. Contemp. Math. **17**, 267–285 (1983)

Hirsch, M.: Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere. SIAM J. Math. Anal. **16**, 423–439 (1985)

Hirsch, M.: Stability and convergence in strongly monotone dynamical systems. J. Reine Angew. Math. **383**, 1–53 (1988)

Hirsch, M., Smith, H.L.: Monotone dynamical systems. In: Handbook of Differential Equations, Ordinary Differential Equations, vol. 2. Elsevier, Amsterdam (2005)

Huang, C.-Y.F., Ferrell, J.E.: Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Natl. Acad. Sci. USA **93**, 10078–10083 (1996)

Jiang, J.F.: On the global stability of cooperative systems. Bull. Lond. Math. Soc. **6**, 455–458 (1994)

Jones, C.K.R.T.: Geometric singular perturbation theory. In: Dynamical Systems (Montecatini Terme, 1994). Lect. Notes in Math., vol. 1609. Springer, Berlin (1995)

Karp, G.: Cell and Molecular Biology. Wiley, New York (2002)

Knobloch, K.W., Aulbach, B.: Singular perturbations and integral manifolds. J. Math. Phys. Sci. **13**, 415–424 (1984)

Lew, D.J., Burke, D.J.: The spindle assembly and spindle position checkpoints. Annu. Rev. Genet. **37**, 251–282 (2003)

Markevich, N.I., Hoek, J.B., Kholodenko, B.N.: Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. J. Cell Biol. **164**, 353–359 (2004)

Moreau, L.: Stability of continuous-time distributed consensus algorithms. In: Proc. 43rd IEEE Conf. Decision and Control, Paradise Island, Bahamas, paper ThC10.4 (2004)

Murray, J.D.: Mathematical Biology, 3rd edn. Springer, New York (2002)

Nipp, K.: Smooth attractive invariant manifolds of singularly perturbed ODE's. Research Report, No. 92–13 (1992)

Ortega, F., Garcés, J., Mas, F., Kholodenko, B.N., Cascante, M.: Bistability from double phosphorylation in signal transduction: kinetic and structural requirements. FEBS J. **273**, 3915–3926 (2006)

Pomerening, J.R., Sontag, E.D., Ferrell, J.E. Jr.: Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. Nature Cell Biology **5**, 346–351 (2003)

Sakamoto, K.: Invariant manifolds in singular perturbation problems for ordinary differential equations. Proc. R. Soc. Edinb. A **116**, 45–78 (1990)

Samoilov, M., Plyasunov, S., Arkin, A.P.: Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. Proc. Natl. Acad. Sci. USA **102**, 2310–2315 (2005)

Sel'kov, E.E.: Stabilization of energy charge, generation of oscillation and multiple steady states in energy metabolism as a result of purely stoichiometric regulation. Eur. J. Biochem. **59**(1), 151–157 (1975)

Sha, W., Moore, J., Chen, K., Lassaletta, A.D., Yi, C.S., Tyson, J.J., Sible, J.C.: Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts. Proc. Natl. Acad. Sci. **100**, 975–980 (2003)

Smith, H.L.: Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems. Mathematical Surveys and Monographs, vol. 41. Am. Math. Soc., Providence (1995)

Sontag, E.D.: Mathematical Control Theory: Deterministic Finite Dimensional Systems. Springer, New York (1990). 2nd edn. (1998)

Sontag, E.D.: Some new directions in control theory inspired by systems biology. Syst. Biol. **1**, 9–18 (2004)

Sontag, E.D.: Molecular systems biology and control. Eur. J. Control **11**, 396–435 (2005)

Stryer, L.: Biochemistry. Freeman, New York (1995)

Sulis, M.L., Parsons, R.: PTEN: from pathology to biology. Trends Cell Biol. **13**, 478–483 (2003)

Teel, A.R., Moreau, L., Nesic, D.: A unification of time-scale methods for systems with disturbances. IEEE Trans. Automat. Contr. **48**, 1526–1544 (2003)

Wang, L., Sontag, E.D.: A remark on singular perturbations of strongly monotone systems. In: Proc. IEEE Conf. Decision and Control. San Diego, p. WeB10.5 (2006)

Wang, L., Sontag, E.D.: On the number of steady states in a multiple futile cycle. J. Math. Biol. (2008). doi:10.1007/s00285-007-0145-z

Widmann, C., Spencer, G., Jarpe, M.B., Johnson, G.L.: Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. Physiol. Rev. **79**, 143–180 (1999)

Zhao, Y., Zhang, Z.Y.: The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3. J. Biol. Chem. **276**, 32382–32391 (2001)