A Review of Multiple-Time-Scale Dynamics

Fundamental Phenomena and Mathematical Methods

Kristiansen, Kristian Uldall

Published in:
Mathematics Online First Collections

Link to article, DOI:
10.1007/16618_2023_75

Publication date:
2023

Document Version
Peer reviewed version

A Review of Multiple Time Scale Dynamics: Fundamental phenomena and mathematical methods

K. U. Kristiansen

September 19, 2023

Department of Applied Mathematics and Computer Science,
Technical University of Denmark,
2800 Kgs. Lyngby,
DK,
krkri@dtu.dk

Abstract

In this paper, we review the most fundamental phenomena in multiple-time scale systems as well as the Geometric Singular Perturbation Theory (GSPT) that can be used to analyze such systems. We put special emphasis on loss of normal hyperbolicity and illustrate how it relates to the different phenomena. Moreover, by working on lower dimensional model problems, we review the main technical tool of GSPT, the blowup method, and demonstrate – in line with recent research in this area – how it may provide detailed insight into problems of this kind.

1 Introduction

Multiple time scale systems are important in many areas of applied science. It is central in neuroscience [1, 30, 59], in biology and chemical reaction networks [15, 21, 27, 57, 62], but also in weather forecasting [58] and many other areas [28, 29, 68, 69]. Oftentimes multiple scales are seen as an obstacle, as finer scales and fast processes require more computational time. Finding ways to circumvent this, without compromising accuracy, through mathematical methods, is the central objective in the mathematical modelling of systems with multiple scales.

In this paper, we will consider models of ordinary differential equations with multiple time scales. These can (at least locally) be written as

$$\begin{aligned} \frac{dx}{dt} &= X(x, y, \epsilon), \\ \epsilon \frac{dy}{dt} &= Y(x, y, \epsilon), \end{aligned} \tag{1}$$

with $0 < \epsilon \ll 1$ a small parameter. Here $x \in \mathbb{R}^k$, $y \in \mathbb{R}^l$ with $k, l \in \mathbb{N}$ and $X : \mathbb{R}^k \times \mathbb{R}^l \times [0, \epsilon_0) \to \mathbb{R}^k$, $Y : \mathbb{R}^k \times \mathbb{R}^l \times [0, \epsilon_0) \to \mathbb{R}^l$, with $\epsilon_0 > 0$, are both assumed to be smooth (C^∞) throughout. Systems of this form, also frequently called slow-fast or singularly perturbed, are multi-scaled since the variables x, at "typical points" (i.e. away from $Y = 0$) vary by a much smaller amount than y for all $0 < \epsilon \ll 1$. For this reason, x and y are called the *slow* and *fast variables*, respectively.

We will not consider stochastic effects (which admittedly are central in many areas of science) and will also ignore models with co-existing spatial and temporal scales as in models described by partial differential equations. Instead, by restricting attention to (1), we will illustrate the different phenomena that can occur in multiple time scale systems and will review the mathematical theory of Geometric Singular Perturbation Theory (GSPT) [34] that – starting from the work of Fenichel [17, 18, 19] – over the past decades have turned multiple scales in systems of the form (1) from being an obstacle to being a desirable setting that can be exploited in the mathematical modelling and the mathematical analysis.[1]

There are already a few review papers of GSPT, see e.g. Jones' [34] and Kaper's [36] and there is also a book by Kuehn on the topic [47]. Recently, Wechselberger in [74] also published a review paper on GSPT, with special emphasis on slow-fast system in nonstandard form (see Section 2.4 below). Our present paper will overlap in some places with these excellent references, but there are also important differences. For example, we will focus on situations with lack of hyperbolicity. In the standard formulation of GSPT [19, 34], these are the cases where the theory breaks down. However, with the aid of blowup [14, 44], GSPT has over the past two decades been extended in various ways to cover these situations. Within this extended formulation of GSPT, the theory also becomes a toolbox and we demonstrate how this toolbox can be used to study some new (and old) examples. Loss of hyperbolicity is important in applications since – as we shall see – it relates to most of the nontrivial phenomena that occur in these type of systems.

2 Geometric Singular Perturbation Theory

The point of departure for the theory of slow-fast systems (1), is the limiting system $\epsilon = 0$:

$$\begin{aligned} \frac{dx}{dt} &= X(x, y, 0), \\ 0 &= Y(x, y, 0). \end{aligned} \tag{2}$$

[1] In fairness, the study of singularly perturbed systems of the form (1) has a long history [49, 55, 67], also before Fenichel's work that led to the development of GSPT. See the appendix of [55] for historical review. However, in contrast to GSPT, these references were not geometric and they were less compatible with dynamical systems theory.

Recall that $x \in \mathbb{R}^k$, $y \in \mathbb{R}^l$. The equation (2) consists of a differential equation for x and an algebraic equation relating x and y. The limit $\epsilon = 0$ is therefore very singular: We replace a differential equation for y in (1) with an equation that does *not* require an initial condition for $\epsilon = 0$. The set defined by the second algebraic equation in (2) (typically) defines a k-dimensional sub-manifold

$$S = \{(x, y) \in \mathbb{R}^{k+l} \, : \, 0 = Y(x, y, 0)\}, \tag{3}$$

and (2) is known as *the reduced problem* associated with (1). In fact, if the Jacobian $D_y Y$ of Y with respect to y is nonsingular, then by the implicit function theorem [50, p. 228] S is indeed a k-dimensional smooth manifold and can (at least locally in some neighborhood U) be written as a graph over x:

$$C = \{(x, y) \in \mathbb{R}^{k+l} \, : \, y = h(x), \quad x \in N\}, \tag{4}$$

where N is some open subset of $\mathbb{R}^k$, such that $C = S \cap U$. In particular, if Y is smooth, then so is the function $h : N \subset \mathbb{R}^k \to \mathbb{R}^l$, and on C, we therefore obtain from (2) a single differential equation for x:

$$\frac{dx}{dt} = X(x, h(x), 0). \tag{5}$$

A solution $x(t)$ of this equation, then gives a solution $(x(t), h(x(t)))$ of (2). The basic question is then:

- Question 1: When, and in what sense, does the solution $(x(t), h(x(t)))$ of (2), requiring only an initial condition on x, relate to or approximate solutions of (1)?

But also:

- Question 2: What happends when $D_y Y$ is singular and near points where S cannot be written as a graph?

In the following, we will address these fundamental questions through our review of GSPT.

For this purpose, we will need to view (1) on a different time scale. Define $\tau = \epsilon^{-1} t$, where t is the time used in (1). The time t is called the *slow time* whereas the time τ is called *fast*; notice that $t = \mathcal{O}(1)$ implies $\tau = \mathcal{O}(\epsilon^{-1})$ for $\epsilon \to 0$. Since $\frac{dt}{d\tau} = \epsilon$, we can write (1) in the following form:

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon X(x, y, \epsilon), \\ \frac{dy}{d\tau} &= Y(x, y, \epsilon), \end{aligned} \tag{6}$$

the two systems being equivalent for all $\epsilon > 0$. The systems (1) and (6) – called *the slow system* and *the fast system*, respectively, in reference to the time used – produce two distinct limiting systems for $\epsilon = 0$: The reduced problem (2) and

$$\begin{aligned} \frac{dx}{d\tau} &= 0, \\ \frac{dy}{d\tau} &= Y(x, y, 0), \end{aligned} \tag{7}$$

respectively. Whereas (2) only involves a differential equation on x, the dynamics of (7) are contained within *layers* of $x = \text{const}$. For this reason, (7) is also referred to as *the layer problem*.[2] In order to bridge the two systems (2) and (7), we realize that equilibria or critical points of (7) are given by S in (3). S is therefore also called *the critical manifold* of the slow-fast systems (1) or (6). Points on S can be attracting for the flow $x(\tau) = \text{const}, y(\tau)$ defined by (7) and this leads to the basic intuition of approximating solutions of either (1) and (6) for $\epsilon > 0$ by concatenating solutions of (7) and (2) on the diverging time scale associated with τ and t, respectively. See Fig. 1 and the associated figure caption.

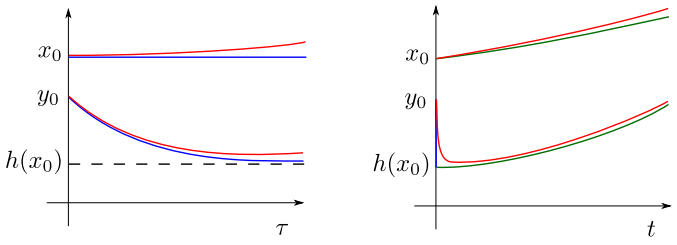
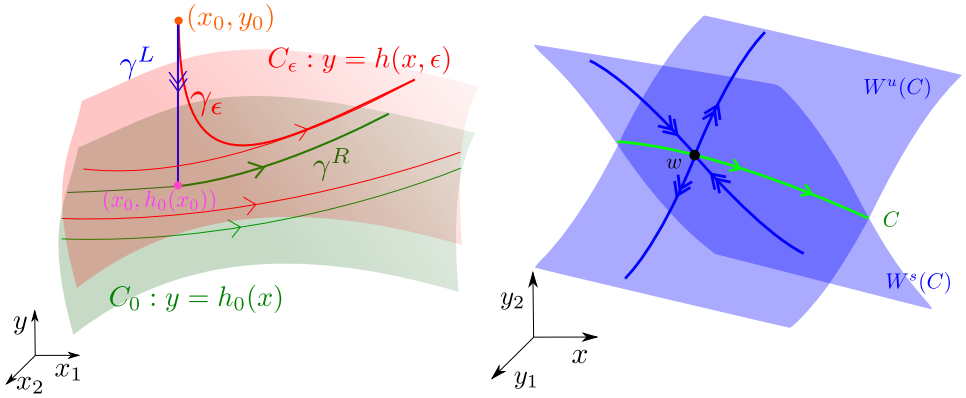


Figure 1: A solution (in red) of (1) or (6) starting at $x(0) = x_0$, $y(0) = y_0$, near an attracting critical manifold $C : y = h(x)$, is approximated for $\epsilon > 0$ small enough by solutions of the layer problem (7) and the reduced problem on the separate time scales defined by τ and t, respectively. In particular, the solution of the layer problem $x(\tau) = x_0$, $y(\tau)$ with $y(0) = y_0$ and $\lim_{\tau\to\infty} y(\tau) = h(x_0)$ is a good approximation of the solution on the time scale associated with the fast time τ, see figure on the left. Next, on the right, we illustrate that the solution $x(t), y(t) = h(x(t))$ of the reduced problem (2), see also (5), with $x(0) = x_0$, $y(0) = h(x_0))$ is a good approximation on the time scale associated with the slow time t. GSPT and Fenichel's theory provides conditions that justify and quantify these approximations in geometric terms using orbits.

In Dynamical Systems Theory, one represent solutions $(x(t), y(t))$ as curves traced out by t, called orbits, in the (x, y)-space. If one restricts

[2]Historically, "layer problem" has its origin in boundary layers in fluid dynamics.

$t \geq 0$ then the curve $(x(t), y(t))$ is also called the forward orbit. In this geometric setting, the approximations of solutions either (1) and (6) through solutions of (7) and (2), can be viewed as follows (see also Fig. 2(a)): Consider a point (x_0, y_0) in a neighborhood of C and let γ^L denote the forward orbit $x(\tau) = \text{const}, y(\tau)$, $\tau \geq 0$, of (7) with initial condition $(x(0), y(0)) = (x_0, y_0)$ and suppose that $(x(\tau), y(\tau))$ converges as $\tau \to \infty$ to $(x_0, h(x_0)) := \lim_{\tau\to\infty}(x(\tau), y(\tau)) \in C$. From $(x_0, h(x_0))$ we can then define another forward orbit $\gamma^R \subset C$ as the solution $(x(t), h(x(t))$, $t \geq 0$, of the reduced problem (2), see also (5), with initial condition $(x_0, h(x_0))$. Together the two orbits $\gamma^L \cup \gamma^R$ is expected to approximate the orbit γ_ϵ of (1). The goal of this section is to quantify this further in more mathematical terms. To prepare the mathematical machinery, we first review some basic concepts from dynamical systems theory.



(a) (b)

Figure 2: In (a): Illustration of an attracting critical manifold C_0. The set γ^L in blue is an orbit of the layer problem, starting at a point (x_0, y_0). It is attracted to $(x_0, h_0(x_0)) \in C$. From there in green we have a forward orbit γ^R of the reduced problem. Taken together, then γ^L and γ^R approximate a true (perturbed) orbit γ_ϵ (in red) of the system for all $0 < \epsilon \ll 1$. In particular, Theorem 2 says that C_0 perturbs to a slow manifold C_ϵ (in red), and that γ_ϵ converges towards a orbit on C_ϵ; as solutions $(x(t), y(t))$ the convergence is exponential $e^{-ct/\epsilon}$. In (b): A normally hyperbolic critical manifold C having stable and unstable manifolds $W^s(C)$ and $W^u(C)$, respectively.

2.1 Basic results from local dynamical systems theory

Consider

$$\frac{dz}{dt} = Z_0(z), \tag{8}$$

with $z \in \mathbb{R}^n$, Z_0 smooth and suppose that $Z_0(0) = 0$ so that $z = 0$ is an equilibrium of (8). We then say that $z = 0$ is *hyperbolic* if the Jacobian $DZ_0(0)$ has no eigenvalues with zero real part. By the implicit function theorem, hyperbolic equilibria are persistent. For example, the perturbed system:

$$\frac{dz}{dt} = Z_0(z) + \alpha Z_1(z, \alpha), \tag{9}$$

where $\alpha \sim 0$ is a parameter and Z_1 is smooth, also has a hyperbolic equilibrium $z = \mathcal{O}(\alpha)$ for all $\alpha \sim 0$. In fact, in the hyperbolic case we have the following:

Let $z(t) = \phi_t(z_0)$ denote the solution of (8) with initial condition $z(0) = z_0$. ϕ_t is called the flow map and it is a diffeomorphism [22, 56]. The flow map ψ_t of the linearization:

$$\frac{dz}{dt} = DZ_0(0)z, \tag{10}$$

is given by the matrix exponential: $\psi_t = e^{DZ_0(0)t}$. *The Hartman-Grobman Theorem* [22, Theorem 1.3.1] then says that (8) and (10) are C^0 *topologically conjugated* insofar that there is a homeomorphism $H : N \to H(N)$ such that

$$\phi_t = H^{-1} \circ \psi_t \circ H. \tag{11}$$

for all t. This in turn implies that if the eigenvalues of $DZ_0(0)$ all have negative real part, then $z = 0$ is *asymptotically stable*, i.e. for any neighborhood N of $z = 0$ there is (a) a neighborhood $M \subset N$ such that $\phi_t(M) \subset N$ for all $t \geq 0$, and (b) $\phi_t(z) \to 0$ for $t \to \infty$ and $z \in M$. Related to this, we also find that in the case where $z = 0$ is hyperbolic, but has eigenvalues with positive and negative real parts, we have stable and unstable spaces of the linearization (10). Through H and (11) this gives stable and unstable sets $W^s(0)$ and $W^u(0)$ of $z = 0$ for (8) where points contract exponentially in forward and backward time. *The stable manifold theorem*, see [22, Theorem 1.3.2], says that $W^s(0)$ and $W^u(0)$ are smooth manifolds that are tangent to the linear spaces E^s and E^u (being the direct sum of all the generalized eigenspaces of $DZ_0(0)$ associated with eigenvalues with negative and positive real parts, respectively) at $z = 0$. Moreover, locally $W^s(0)$ and $W^u(0)$ consists (again, in the hyperbolic case) of all points that remain in a small neighborhood of $z = 0$ for all $t \geq 0$ and $t \leq 0$, respectively.

If $DZ_0(0)$ has eigenvalues with zero real part, then $z = 0$ is said to be *nonhyperbolic*. Specifically, if all eigenvalues have zero real part then $z = 0$ is said to be *fully nonhyperbolic*. Otherwise, when at least one eigenvalue have nonzero real part, we say that $z = 0$ is *partially* (or *semi-*) *hyperbolic*. In this case, we still have smooth stable and unstable manifolds $W^s(0)$ and $W^u(0)$ tangent to the linear spaces E^s and E^u but as opposed to the hyperbolic

case, they do not characterize the local dynamics completely: There may exist points in a neighborhood of $z = 0$ that remain there for all $t \geq 0$ or all $t \leq 0$ without belonging to either $W^s(0)$ or $W^u(0)$. Moreover, points $z \neq 0$ may remain in a neighborhood of $z = 0$ for all $t \in \mathbb{R}$. Such points will instead belong to a *center manifold* according to *the center manifold theorem*:

Theorem 1 *[22, Theorem 3.2.1] Suppose that $z = 0$ is partially hyperbolic and fix any $n \in \mathbb{N}$. Then there is a neighborhood N of $z = 0$ and a center manifold $W^c(0)$ of (8), that is (a) tangent at $z = 0$ to the center space E^c (being the direct sum of all the generalized eigenspaces associated with eigenvalues with zero real parts), (b) a C^n-graph over E^c, and finally (c) invariant in the following sense: If $z \in W^c(0)$ then $\phi_t(z) \in W^c(0)$ for all $t \in \mathbb{R}$ for which $\phi_t(z) \in N$. Moreover, suppose that $\phi_t(z) \in N$ for all $t \in \mathbb{R}$. Then $z \in W^c(0)$.* □

The dynamics on $W^c(0)$ are not captured by the linearization, but will instead depend upon nonlinear terms. The case where $E^u = \emptyset$ is interesting, also because it relates to the situation illustrated in Fig. 2(a). In this case, all solutions that remain in N for all $t \geq 0$ converge to the center manifold at an exponential rate and $W^c(0)$ therefore offers a dimension reduction. Also in the case where Z depends upon parameters μ all *bifurcations* occur on $W^c(0)$. Here bifurcations refer to systems (8) where perturbations e.g. (9) are not topologically conjugated or topologically equivalent [56, p. 107]. The simplest examples of (local) bifurcations are the emergence or disappearance of equilibria or periodic orbits, e.g. through saddle-node bifurcations and Hopf or homoclinic bifurcations, respectively, upon parameter variations, see [22, 56] for further background on the basics of dynamical systems theory.

Remark 1 Whereas the stable and unstable manifolds are C^∞ (or analytic C^ω) if Z is so, the smoothness of $W^c(0)$ is more delicate. In particular, the domain $N(n)$ depends on n in general. Moreover, $W^c(0)$ is not unique in general, see [56, p. 154]. □

2.1.1 Notation

In the following, when illustrating phase portraits of systems of the form (1), we will follow a standard convention and use colours and different arrows on orbits to separate slow and fast directions. In particular, fast orbits are in blue with the corresponding directions indicated by double-headed arrows. On the other hand, slow orbits are in green with the corresponding directions indicated by single-headed ones. We use a similar notation in the context of (8) to separate hyperbolic directions (double-headed arrows) from center directions (single-headed arrows).

2.2 Fenichel's theory of slow-fast systems

Having now introduced the basics of local dynamical systems theory, we turn to (6).

Consider a compact k-dimensional submanifold $C \subset S$. Recall that S is a set of critical points of (7). We then say that C is *normally hyperbolic* if the Jacobian $D_y Y$ at each point in C has no eigenvalues with zero real part. To understand the meaning of this, first notice that the linearization of (7) around $z = (x, y) \in C$ gives

$$\begin{pmatrix} 0_{k\times k} & 0_{k\times l} \\ D_x Y(z,0) & D_y Y(z,0) \end{pmatrix}.$$

The eigenvalues of this block-diagonalized matrix are given by (a) k zero eigenvalues, due to the trivial block $0_{k\times k}$, and (b) the eigenvalues of $D_y Y \in \mathbb{R}^{l\times l}$. Therefore the center space $E^c(z)$ associated with $z \in C$ is at least k-dimensional. In particular, if C is normally hyperbolic then $E^c(z)$ coincides with the tangent space $T_z C$ of C at z. Moreover, each point $z \in C$ has a stable and an unstable manifold $W^s(z)$ and $W^u(z)$. As z ranges of C, we obtain a stable and an unstable manifold $W^s(C)$ and $W^u(C)$ of C, see Fig. 2(b).

Suppose that C is normally hyperbolic. Then by augmenting $\dot{\epsilon} = 0$ to (6), Theorem 1 actually implies that there is a local center manifold $W^c(z, 0)$ of $(z, 0) \in C \times \{0\}$ in the (x, y, ϵ)-space. Theorem 1 is a local result, but as this holds for any $z \in C$ it is clear that in the context of (6), $W^c(z, 0)$ has a more global structure not captured by Theorem 1. This – and more – is laid out by Fenichel's theory on singular perturbations [17, 18, 19, 34]:

Theorem 2 *[34, Theorem 2] Fix any $n \in \mathbb{N}$ and consider a compact k-dimensional submanifold $C_0 \subset S$ of (6) that is normally hyperbolic and suppose for simplicity that C_0 can be written as a graph:*

$$C_0 = \{(x, y) \,:\, y = h_0(x), \quad x \in N\}, \tag{12}$$

with $h_0 : N \to \mathbb{R}^l$ smooth and where $N \subset \mathbb{R}^k$ is a compact domain. Then there exists an $\epsilon_0(n) > 0$ sufficiently small such that for all $\epsilon \in [0, \epsilon_0)$ there exists an invariant manifold C_ϵ of (6) of the following form:

$$C_\epsilon = \{(x, y) \,:\, y = h(x, \epsilon), \quad x \in N\}.$$

Here $h : N \times [0, \epsilon_0) \to \mathbb{R}^l$ is C^n-smooth and satifies $h(\cdot, 0) = h_0$. Also, C_ϵ has stable and unstable manifolds $W^s(C_\epsilon)$ and $W^u(C_\epsilon)$ that are C^n-smooth perturbations of $W^s(C_0)$ and $W^u(C_0)$. Points on $W^s(C_\epsilon)$ converge exponentially at a rate $e^{-ct/\epsilon}$, $c > 0$ uniformly across C, towards C_ϵ as $t \to \infty$ (while remaining in a neighborhood of C_0). □

Remark 2 The assumption (12) that the normally hyperbolic submanifold C_0 can be written as a graph is not a restriction locally. Having said that, it is not always possible to write the set S defined by (3) *globally* as a graph over x in general, not even if S is compact. (A sufficient condition for this would be that the *the projection of S onto the x-space is simply connected.* This is a basic fact of algebraic topology.) □

The invariant manifold C_ϵ is called a slow manifold. As center manifolds, C_ϵ is not unique. However, as with center manifolds, C_ϵ offers a dimension reduction. In particular, by the invariance of C_ϵ, we can substitute $y = h(x, \epsilon)$ into the x-equation in (1):

$$\dot{x} = X(x, h(x, \epsilon), \epsilon), \tag{13}$$

and a solution $x(t)$ of (13) gives a solution $(x(t), h(x(t), \epsilon))$ of (1). Now we notice that (13) is smooth in ϵ and is it therefore a smooth $\mathcal{O}(\epsilon)$-perturbation of the reduced problem (5). We can therefore use standard dynamical systems theory, see Section 2.1, to study (13), and in turn (1), for all $0 < \epsilon \ll 1$.

In the case where the eigenvalues of $D_y Y(z, 0)$ have negative real part, then C_0 is said to be attracting. In this situation, we can through the *fast* exponential contraction along $W^s(C_\epsilon)$ and the *slow* motion along C_ϵ offer a mathematical description of Question 1. Indeed, in the most basic formulation, the orbit γ_ϵ, defined as the solution $(x(t), y(t))$, $t \geq 0$, of (1) with initial condition $(x(0), y(0)) = (x_0, y_0) \in W^s(C_0)$, converges to $\gamma^L \cup \gamma^R$ on compact sets (in the Haussdorff-sense) as $\epsilon \to 0$, see Fig. 2(a). There is in fact a deeper result, because Fenichel showed that there is a well-defined *fiber-projection* π^s [19, 34] from each point q_ϵ on $W^s(C_\epsilon)$ to a base-point $\pi^s(q_\epsilon)$ on C_ϵ, such that $\phi_t(q_\epsilon)$ converges exponentially to the flow of the base point $\pi^s(q_\epsilon)$ as $t \to \infty$ (while remaining in a neighborhood of C_0):

$$|\phi_t(q_\epsilon) - \phi_t(\pi^s(q_\epsilon))| \leq c_1 e^{-c_2 t \epsilon^{-1}},$$

for some $c_1, c_2 > 0$ and all $0 < \epsilon \ll 1$. In this sense, the splitting into fast and slow dynamics for $\epsilon = 0$, through (7) and (2), persists for $\epsilon > 0$. We refer to [34, 36] for further details.

Theorem 2 can be used successively in systems with more than just one time scale separation [7]. As an example, consider a general system having three separate time scales:

$$\begin{aligned}
\frac{du}{d\tau} &= \epsilon\delta U(u, v, w, \epsilon, \delta), \\
\frac{dv}{d\tau} &= \epsilon V(u, v, w, \epsilon, \delta), \\
\frac{dw}{d\tau} &= W(u, v, w, \epsilon, \delta),
\end{aligned}$$

with $0 \leq \epsilon, \delta \ll 1$ being two small parameters for $u \in \mathbb{R}^k$, $v \in \mathbb{R}^l$, $w \in \mathbb{R}^m$. Then u is *super-slow*, v is *slow* and finally w is *fast*, and if the compact $(k+l)$-dimensional critical manifold $C_0 \subset S = \{(u, v, w) : W(u, v, w, 0, \delta) = 0\}$ is normally attracting for all $\delta \geq 0$ sufficiently small, and can be written as a graph

$$C_0 = \{(u, v, w) : w = h_0(u, v, \delta)\},$$

then by Fenichel's theory, see Theorem 2, the reduced problem

$$\begin{aligned} \frac{du}{dt} &= \delta U(u, v, h_0(u, v, \delta), 0, \delta), \\ \frac{dv}{dt} &= V(u, v, h_0(u, v, \delta), 0, \delta), \end{aligned} \tag{14}$$

approximates the flow on the slow time scale t for $0 < \epsilon \ll 1$, uniformly in $\delta \geq 0$. (14) is then a reduced problem of dimension $l + k$ (m-dimensions have been eliminated) and it is also slow-fast, now with respect to the small parameter $\delta > 0$. We can then perform the same analysis, studying the associated layer problem:

$$\begin{aligned} \frac{du}{dt} &= 0, \\ \frac{dv}{dt} &= V(u, v, h_0(u, v, 0), 0, 0), \end{aligned}$$

checking stability of the critical manifold $\{(u, v) : V(u, v, h_0(u, v, 0), 0, 0) = 0\}$, and study the reduced problem:

$$\begin{aligned} \frac{du}{ds} &= U(u, v, h_0(u, v, 0), 0, 0), \\ 0 &= V(u, v, h_0(u, v, 0), 0, 0), \end{aligned} \tag{15}$$

with respect to the *super-slow* time $s = \delta t$. (15) is k-dimensional upon eliminating v from the second equation, and consequently, in the attracting case where all eigenvalues of the Jacobian $(D_v V + D_w V D_v h_0)$ (by the chain rule) have negative real parts, $l + m$ dimensions can be eliminated for all $0 < \epsilon, \delta \ll 1$, see Fig. 3. This process can be repeated for any number of time scales.

2.3 Breakdown of Fenichel's theory and the blowup method

We now turn to Question 2. Fenichel's theory can break down in number of ways. For example, noncompact critical manifold are excluded and $D_y Y$ could have eigenvalues with zero real part, or even be singular. In the latter case, the implicit function theorem fails and the critical manifold may no longer be a graph over x. We now consider two simple examples with these types of breakdown.

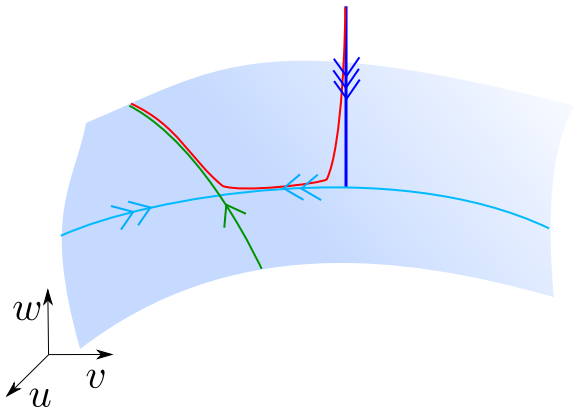


Figure 3: Illustration of the iterative application of Fenichel's theory in the case of three time scales. On the fastest time scale we can reduce to a cyan critical manifold, under the attractivity assumption. On the associated reduced system, in the intermediate time scale, see (14), we then finally reduce to the attracting green critical manifold upon which we have a final reduced problem (15). The red curve illustrates a typical orbit contracting first towards the cyan and then subsequently towards the green manifold.

2.3.1 A fold point

As a way of illustration, suppose

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon, \\ \frac{dy}{d\tau} &= x + y^2. \end{aligned} \tag{16}$$

In this case, the layer problem is given by

$$\frac{dx}{d\tau} = 0, \tag{17}$$

$$\frac{dy}{d\tau} = x + y^2. \tag{18}$$

Consequently, the critical manifold S is given as $x = -y^2$, which is not a graph over x, but rather a graph over the fast variable y. The linearization of (17) around any point in S gives $2y$ as the only nontrivial eigenvalue. The point $(x, y) = (0, 0)$, called a *fold point*, see Fig. 4, is therefore fully nonhyperbolic and it divides S into an attracting subset $S_a = S \cap \{y < 0\}$ and a repelling subset $S_r = S \cap \{y > 0\}$.

S_a and S_r are noncompact. However, we could obviously restrict e.g. S_a to

$$C_a := \{(x, y) \,:\, y = -\sqrt{-x}, x \in I\},$$

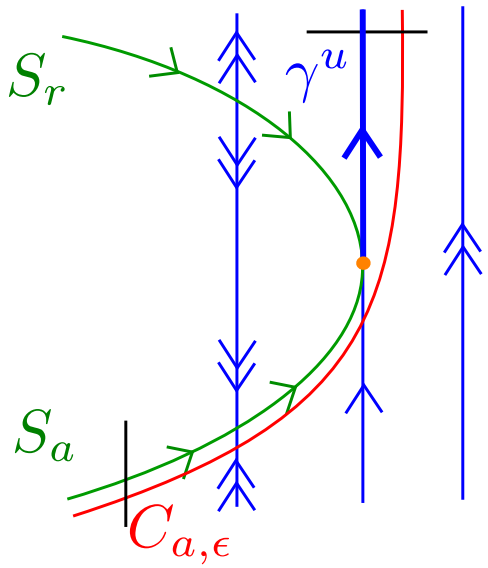


Figure 4: Illustration of a regular fold *jump* point. The reduced problem on the attracting branch S_a points towards the degenerate point, where Fenichel's theory does not apply. Nevertheless, in this simple case, the slow manifold $C_{a,\epsilon}$ (obtained as a perturbation of a compact submanifold $C_a \subset S_a$) is just an orbit segment which can be extended by the flow of (16). In this way, a continuity argument implies that this extended orbit follows the unstable set γ^u (blue, thickened curves) of the fold point for the layer problem. Theorem 3, which can be proven by blowup, provides further details on the transition near a generalized fold (indicated by the two sections in black).

for $I := [x_1, x_2] \subset \mathbb{R}_-$ some compact interval, and then apply Fenichel's theory to this normally hyperbolic and attracting subset. This would give a slow manifold

$$C_{a,\epsilon} := \{(x, y) \, : \, y = -\sqrt{-x} + \mathcal{O}(\epsilon), x \in I\},$$

with slow flow given by $\frac{dx}{dt} = 1$. Consequently, points on $C_{a,\epsilon}$ leave the boundary at $x = x_2$ under the flow, so the description of the dynamics in a uniform neighborhood of $(x, y) = 0$ as $\epsilon \to 0$, requires an extension of the theory.

However, the system (16) is so simple that we can easily describe what happends. In two-dimensions $C_{a,\epsilon}$ is just an orbit segment and we can extend this by applying the flow defined by (16). A continuity argument clearly shows that the orbit converges as $\epsilon \to 0$ to the union of S_a and the orbit $\gamma^u \, : \, x = 0, y > 0$ of the layer problem (17). γ^u is the unstable set of

$(x, y) = (0, 0)$ for (17). For this reason, the point $(x, y) = (0, 0)$ is also called a jump point.

However, the further details (including the asymptotics and the details of contraction) of the transition near the jump point cannot be handled so easy. Also in some situations (including ones in higher dimensions, see also further examples below), the dynamics can be more complicated without having an easy interpretation. For this purpose, *the blowup method* offers unique insight.

The blowup method was first pioneered by Dumortier and Roussarie [14] and later refined in the language of Geometric Singular Perturbation Theory by Krupa and Szmolyan [44]. In particular, [44] offers an excellent introduction to the blowup method in the context of generalized jump points. In particular, they show the following:

Theorem 3 *[44], [47, Theorem 7.4.1] Consider a planar system (6) with*

$$Y(0,0,0) = Y'_y(0,0,0) = 0,\ Y'_x(0,0,0) > 0,\ Y''_{yy}(0,0,0) > 0,\ X(0,0,0) > 0.$$

Then the critical manifold S has a fold point at $(x, y) = (0, 0)$ and locally splits into an attracting branch S_a for $y < 0$ and a repelling branch for $y > 0$. Moreover, by the conditions on X and Y there exists a constant $c > 0$ such that for all $0 < \epsilon \leq \epsilon_0$, with $\epsilon_0 > 0$ sufficiently small, the map from initial conditions transverse to S_a to final conditions at a section transverse to the unstable set γ^u of $(x, y) = (0, 0)$, is a strong contraction with Lipschitz-constant bounded by $e^{-c/\epsilon}$, the image being $\mathcal{O}(\epsilon^{2/3})$ close to γ^u. □

This result is important because it enables the construction of attracting limit cycles (basically, periodic orbits that attract or repel nearby points under the flow) through the contraction mapping theorem [50], see also Section 3.1. We refer the reader to [44] for full details on the blowup method. Here we only give a summary in the context of (16) (which clearly satifies the assumption in Theorem 3):

The point $(x, y, \epsilon) = (0, 0, 0)$ is *fully nonhyperbolic* for the extended system:

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon, \\ \frac{dy}{d\tau} &= x + y^2, \\ \frac{d\epsilon}{d\tau} &= 0. \end{aligned} \tag{19}$$

Indeed, the linearization around $(0, 0, 0)$ only has zero eigenvalues, and consequently the center manifold theory, recall Theorem 1, provides no insight at this point. The goal of the blowup method is essentially to *insert more space* in order to *gain hyperbolicity* so that center manifold theory can be

applied. In further details for (19), one replaces $(x, y, \epsilon) = (0, 0, 0)$ by a unit hemi-sphere

$$S^2_+ := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = 1,\, x_3 \geq 0\},$$

through the non-invertible *blowup* transformation:

$$\Psi: \quad (r, (\bar{x}, \bar{y}, \bar{\epsilon})) \mapsto \begin{cases} x &= r^2\bar{x}, \\ y &= r\bar{y}, \\ \epsilon &= r^3\bar{\epsilon}, \end{cases} \qquad r \geq 0,\, (\bar{x}, \bar{y}, \bar{\epsilon}) \in S^2_+. \tag{20}$$

Notice that the image of $r = 0$ under this transformation is precisely $(x, y, \epsilon) = (0, 0, 0)$ and this is the sense in which the point is blown up to the hemi-sphere $(\bar{x}, \bar{y}, \bar{\epsilon}) \in S^2_+$ by the inverse process. The (hemi-) sphere S^2_+ is referred to as *the blown up (hemi-) sphere*.

If the exponents or *weights* on r, i.e. 2, 1 and 3, had all been equal to 1, then (20) would just correspond to spherical coordinates with r being the radius. The blowup transformation (20) with different weights can therefore be seen as generalized spherical coordinates and this is important. Specifically, notice that there are two important invariant sets of (17): the critical manifold S and $\gamma : x = 0, y \in \mathbb{R}$. These sets have a quadratic tangency at $(x, y) = (0, 0)$ and this intrinsically relates to the breakdown of hyperbolicity. But under (20), these sets (now embedded within $\epsilon = 0$ in the extended space (x, y, ϵ)) separate at $r = 0$. In particular, γ becomes $\bar{\gamma}^{u/s} : (\bar{x}, \bar{y}, \bar{\epsilon}) = (0, \pm 1, 0)$, $r \geq 0$ whereas S becomes two rays $\bar{S}_{r/a}$ of points $(\bar{x}, \bar{y}, \bar{\epsilon}) = (-l^2, \pm l, 0)$, with $r \geq 0$, where $l > 0$ is the unique positive solution of $l^4 + l^2 = 1$. The corresponding base points at $r = 0$, $\bar{g}^{u/s} : (\bar{x}, \bar{y}, \bar{\epsilon}) = (0, \pm 1, 0)$ and $\bar{s}_{r/a} : (\bar{x}, \bar{y}, \bar{\epsilon}) = (-l^2, \pm l, 0)$, are clearly separated.

In general, the geometric objects and the analysis of the dynamics on the blown up sphere should be described in specific directional charts [44]. Consider for example (r_1, x_1, ϵ_1) defined by

$$r_1 = r(-\bar{y}), \quad x_1 = \bar{x}(-\bar{y})^{-2}, \quad \epsilon_1 = \bar{\epsilon}(-\bar{y})^{-3}, \tag{21}$$

for $\bar{y} < 0$. These coordinates produce a chart or coordinate patch, typically referred to as the $\bar{y} = -1$ chart, on the $\bar{y} < 0$ subset of the space $r \geq 0, (\bar{x}, \bar{y}, \bar{\epsilon}) \in S^2_+$, in which (20) takes the local form

$$(r_1, x_1, \epsilon_1) \mapsto \begin{cases} x &= r_1^2 x_1, \\ y &= -r_1, \\ \epsilon &= r_1^3 \epsilon_1. \end{cases} \tag{22}$$

Consequently, $S_a = S \cap \{y < 0\}$ and $\gamma^s := \gamma \cap \{y < 0\}$ become $(r_1, -1, 0)$ and $(r_1, 0, 0)$ for $r_1 > 0$, respectively, in these coordinates. The corresponding

points $s_{a,1} : (r_1, x_1, \epsilon_1) = (0, -1, 0)$ and $g_1^s : (r_1, x_1, \epsilon_1) = (0, 0, 0)$ at $r_1 = 0$, being the local versions of $\bar{s}_a$ and $\bar{g}^s$ under (21), are clearly separated. Using the coordinates (r_3, x_3, ϵ_3) of the corresponding $\bar{y} = 1$ chart, in which (20) takes the local form

$$(r_3, x_3, \epsilon_3) \mapsto \begin{cases} x &= r_3^2 x_3, \\ y &= r_3, \\ \epsilon &= r_3^3 \epsilon_3, \end{cases} \tag{23}$$

we can similarly express $S_r = S \cap \{y > 0\}$ and $\gamma^u := \gamma \cap \{y > 0\}$ as $S_{r,3} : (r_3, x_3, \epsilon_1) = (r_3, -1, 0)$ and

$$\gamma_3^u : (r_3, x_3, \epsilon_1) = (r_3, 0, 0)$$

for $r_3 > 0$, respectively. The corresponding points $s_{r,3} : (r_3, x_3, \epsilon_3) = (0, -1, 0)$ and $g_3^u : (r_3, x_3, \epsilon_3) = (0, 0, 0)$ at $r_3 = 0$ are again separated. It is the splitting of these points $(\bar{s}_{a/r}, \bar{g}^{s/u})$ that makes it makes it possible to gain hyperbolicity. For example, if we apply the local map (22) then (19) becomes

$$\begin{aligned} \frac{dr_1}{dt} &= -r_1^2(x_1 + 1), \\ \frac{dx_1}{dt} &= r_1\epsilon_1 + r_1(x_1 + 1)x_1, \\ \frac{d\epsilon_1}{dt} &= 3r_1(x_1 + 1)\epsilon_1. \end{aligned}$$

Here $r_1 = 0$ is a set of equilibria but $r_1 \geq 0$ is a common factor which can be divided out

$$\begin{aligned} \frac{dr_1}{dt_1} &= -r_1(x_1 + 1), \\ \frac{dx_1}{dt_1} &= \epsilon_1 + r_1(x_1 + 1)x_1, \\ \frac{d\epsilon_1}{dt_1} &= 3(x_1 + 1)\epsilon_1. \end{aligned} \tag{24}$$

For $r_1 > 0$ this division corresponds to a transformation of time defined by $\frac{d}{dt} = r_1 \frac{d}{dt_1}$. For $r_1 = 0$ the two systems are clearly distinct, in particular (24) is nontrivial on this subset. $s_{a,1}$ is partially hyperbolic for (24) whereas g_1^s is fully hyperbolic, the linearization having eigenvalues $-1, 2, 3$. The fact that $s_{a,1}$ is partially hyperbolic for (24) implies that there is a center manifold of this point. This enables an extension of the slow manifold $C_{a,\epsilon}$ as an invariant manifold [44]. The division by $r_1 \geq 0$ is called desingularization and notice how we use dynamical systems theory (e.g. center manifold theory Theorem 1) to infer properties from $r_1 = 0$ to $r_1 > 0$ and therefore to (19) for $\epsilon > 0$ small enough. The analysis in the $\bar{y} = 1$ chart is similar with $s_{r,3}$ being partially hyperbolic, enabling an extension of the repelling slow

manifold $C_{r,\epsilon}$ through center manifold theory, whereas g_3^u is fully hyperbolic, having the set γ_3^u as one-dimensional unstable manifold.

The two charts $\bar{y} = -1$, $\bar{y} = 1$ do not cover the whole sphere. To describe the full dynamics in local coordinates, separate charts are needed. For the purpose of tracking $C_{a,\epsilon}$, the associated $\bar{\epsilon} = 1$ chart with chart-specific coordinates (x_2, y_2, r_2) suffices to complete the analysis. This chart is defined by

$$(x_2, y_2, r_2) \mapsto \begin{cases} x &= r_2^2 x_2, \\ y &= r_2 y_2, \\ \epsilon &= r_2^3. \end{cases}$$

Notice that $r_2 = \sqrt[3]{\epsilon}$ and hence $x = \sqrt[3]{\epsilon^2} x_2, y = \sqrt[3]{\epsilon} y_2$ and the variables are therefore just scaled in this chart. For this reason, the chart $\bar{\epsilon} = 1$ is also referred to as the scaling chart. Here (19) takes the following form

$$\begin{aligned} \frac{dx_2}{dt_2} &= 1, \\ \frac{dy_2}{dt_2} &= x_2 + y_2^2, \end{aligned} \tag{25}$$

and $\frac{dr_2}{dt_2} = 0$ with respect to a new time $t_2 = r_2 \tau$. (Notice that for the simple system (16), $r_2 = \sqrt[3]{\epsilon}$ is completely eliminated upon this transformation.) The center manifold of $s_{a,1}$ in chart $\bar{y} = -1$ can then be extended in the chart $\bar{\epsilon} = 1$ upon change of coordinates defined by

$$r_2 = r_1 \sqrt[3]{\epsilon_1}, \quad x_2 = x_1 \epsilon_1^{-\frac{2}{3}}, \quad y_2 = -\epsilon_1^{-\frac{1}{3}}$$

upon following a (unique) special solution $\phi_2(t_2)$ of (25), see [44, Proposition 2.3], with

$$\phi_2(t_2) \to (\Omega_0, \infty), \quad \text{for} \quad t_2 \to \infty, \tag{26}$$

for some $\Omega_0 > 0$. Transforming this solution to the $\bar{y} = 1$ chart using the change of coordinates defined by

$$r_3 = r_2 y_2, \quad x_2 = x_3 y_2^{-2}, \quad \epsilon_3 = y_2^{-3},$$

it is clear that this solution within $r_2 = 0$ is asymptotic to g_3^u for $t_2 \to \infty$. By working in the $\bar{y} = 1$ chart, an analysis – based upon a linearization of the hyperbolic point g_3^u – allows one to track the extension of the slow manifold $C_{a,\epsilon}$ along γ^u. We again refer to [44] for the full details. We summarize the results of the blowup analysis in Fig. 5.

From one perspective, the blowup method is essentially an approach to find good coordinates to study the dynamics. The geometry and the

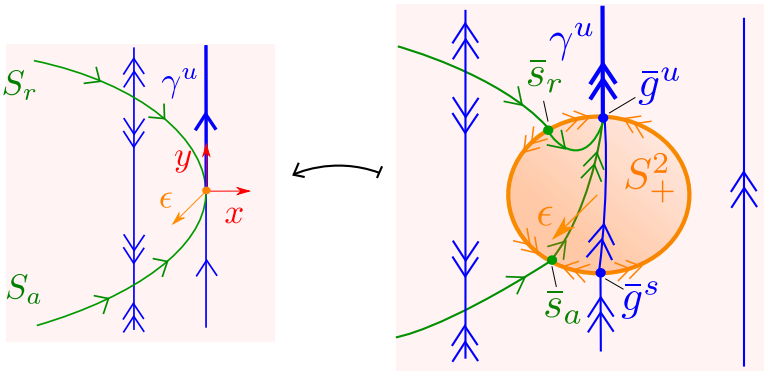


Figure 5: To study degenerate points, like the fold point in Fig. 4, one can use blowup in the extended space, obtained by augmenting ϵ. The figure, shown schematically, illustrates this approach on the fold point. The orange fold point on the left is in the (x, y, ϵ)-space blown up by the blowup transformation (20) to a hemi-sphere S^2_+ (also in orange, shown as a projection) on the right. The blowup transformation (due to the different exponents on r) splits the degenerate sets and upon desingularization (through division by r in this case) we gain hyperbolicity at the points $\bar{s}_{a/r}$ and $\bar{g}^{s/u}$ (indicated by the double-headed arrows, recall Section 2.1.1). This enables the use of center manifold and linearizations, see Section 2.1, to follow the slow manifold close to γ^u. In particular, the green orbit on the blown up sphere, connecting $\bar{s}_a$ and $\bar{g}^u$, is given in the chart $\bar{\epsilon} = 1$ as a special solution $\phi_2(t_2)$ satisfying (26).

abstract blowup transformation organize these coordinates through the separate charts. In fact, the weights of the blowup transformation are often found by first looking for appropriate ϵ-dependent scalings. From these scalings one can then work backwards and define the blowup transformation in such a way that the scaling corresponds to the associated scaling chart.

2.3.2 Intersecting critical manifolds

Next, we consider the following example:

$$\begin{aligned} \frac{du}{d\tau} &= -v^2 u - \epsilon\mu^{-1} u + \epsilon^2, \\ \frac{dv}{d\tau} &= -\epsilon v, \end{aligned} \tag{27}$$

with $\mu \neq 0$. Here u is fast while v is slow for $0 < \epsilon \ll 1$. The system (27) has a unique equilibrium at $(u, v) = (\epsilon\mu, 0)$ for all $\mu \neq 0$, $\epsilon > 0$, and the

associated layer problem is given by

$$\frac{du}{d\tau} = -v^2 u,$$
$$\frac{dv}{d\tau} = 0.$$

Here we find that the two axes $S = \{(u, v) : u = 0\}$ and $M = \{(u, v) : v = 0\}$ are critical manifolds. Their intersection at $(0, 0)$ is transverse. The manifold M is, however, completely degenerate, the linearization about any point in M having only zero eigenvalues. On the other hand, $S_a := S \cap \{v \neq 0\}$ is normally hyperbolic and attracting but it is *noncompact*. We can specifically only apply Fenichel's theory to compact subsets within $v < 0$ or within $v > 0$; the theory does not allow us to cover a full neighborhood of $(u, v) = (0, 0)$ uniformly in $\epsilon > 0$. Nevertheless, notice that S_a carries the following reduced problem

$$\frac{dv}{dt} = -v.$$

This equation has a hyperbolic and attracting node at $v = 0$, but this point does not belong to S_a, but rather it corresponds to the degenerate point $S \cap M$. Consequently, we cannot perturb this equilibrium by Fenichel's theory. We illustrate the dynamics for $\epsilon = 0$ in Fig. 6; notice that this is independent of μ.

At this level, it is tempting to think that $(u, v) = (\epsilon\mu, 0)$ attracts all initial conditions under the forward flow for all $0 < \epsilon \ll 1$. However, we can easily show that this is not case: In fact, M – although being completely degenerate for $\epsilon = 0$ – is clearly invariant for all $\epsilon \geq 0$. It carries the following dynamics

$$\frac{du}{dt} = -\mu^{-1} u + \epsilon,$$

and hence $u = \epsilon\mu$ is attracting for $\mu > 0$ and repelling for $\mu < 0$. Consequently, $(u, v) = (\epsilon\mu, 0)$ is unstable for $\mu < 0$; in fact, it is a saddle in this case with its unstable manifold along M, the u-axis. Therefore for any $c > 0$ we have that any trajectory starting arbitrarily close to S but away from $v = 0$, will eventually be separated from S by a distance that is larger than c. The separation occurs along the unstable manifold M near the degenerate point on S. For a rigorous proof, one can apply a blowup transformation:

$$r \geq 0, (\bar{u}, \bar{v}, \bar{\epsilon}) \in S^2_+ \mapsto \begin{cases} u &= r^2 \bar{u}, \\ v &= r\bar{v}, \\ \epsilon &= r^2 \bar{\epsilon} \end{cases}$$

of the degenerate point $(u, v, \epsilon) = (0, 0, 0)$ within the extended space. In particular, in the associated scaling chart $\bar{\epsilon} = 1$ with chart-specific coordinates (u_2, v_2, r_2) defined by

$$(u_2, v_2, r_2) \mapsto \begin{cases} u &= r_2^2 u_2, \\ v &= r_2 v_2, \\ \epsilon &= r_2^2, \end{cases}$$

with $r_2 = \sqrt{\epsilon}$, we obtain the following equations

$$\begin{aligned} \frac{du_2}{dt_2} &= -v_2^2 u_2 - \mu^{-1} u_2 + 1, \\ \frac{dv_2}{dt_2} &= -v_2, \end{aligned}$$

in terms of $t_2 = \epsilon\tau$. This system only depends upon μ and it is easy to study. In particular, v_2 decouples.

On the other hand, for $\mu > 0$ it can be shown by the Poincaré-Bendixson theorem [22, 56] – using the invariance of M to exclude existence of limit cycles – that $(u, v) = (\epsilon\mu, 0)$ attracts all points for all $\epsilon > 0$. In this case, there is actually an invariant manifold near S but it is not smooth for $\mu > 1$. This rests upon the fact that for $\mu > 1$ the u-axis is a weak direction for the stable node. We leave out further analysis, but show some numerical computations in Fig. 7 for $\mu = 3$, see figure caption for further details.

The example shows that the dynamics near nonhyperbolic points can be hidden. However, blowup and scalings can be used to resolve the hidden dynamics.

2.4 Slow-fast systems in nonstandard form

In Fenichel's original work [17, 18, 19], the variables were not grouped into slow and fast as in (6). Instead, Fenichel considered a more general setting of

$$\frac{dz}{d\tau} = Z(z, \epsilon), \tag{28}$$

with $Z : \mathbb{R}^n \times [0, \epsilon_0) \to \mathbb{R}^n$ smooth and where the critical set $S = \{z : Z(z, 0) = 0\}$ defines a k-dimensional submanifold with $1 \leq k < n$ (again called the critical manifold). These systems are called slow-fast systems in non-standard form [39, 74] and Theorem 2 can also be formulated in terms of (28); notice if $z = (x, y)$, $Z = (\epsilon X, Y)$ then (6) becomes (28) and it may also be possible to go from (28) to (6) but (in general) only locally, see [74].

Consider a compact k-dimensional critical submanifold $C \subset S$. Then C is *normally hyperbolic* if the linearization of the associated layer problem:

$$\frac{dz}{d\tau} = Z(z, 0),$$

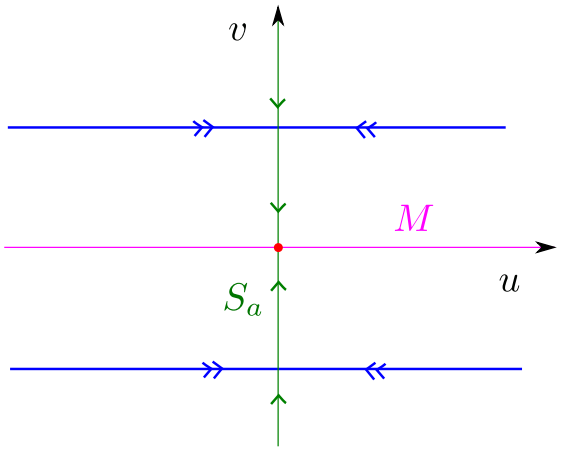


Figure 6: Slow-fast analysis of (27) based on the layer problem and the reduced problem. The set S_a is noncompact and depending on $\mu > 0$ there may be no invariant manifold near S_a for $\epsilon > 0$ or in case there is, this manifold may be nonsmooth; see Fig. 7 and the text for further details.

around any point $q \in C$ only has k eigenvalues with zero real part, i.e. as before if $E^c(q) = T_qC$. As for (6), a normally hyperbolic C has stable and unstable manifolds $W^s(C)$ and $W^u(C)$.

Theorem 4 *[19], [74, Theorem 3.1] A compact k-dimensional normally hyperbolic critical submanifold $C_0 \subset S$ persists as an invariant manifold C_ϵ of (28) for all $\epsilon \in [0, \epsilon_0)$, having stable and unstable manifolds $W^s(C_\epsilon)$ and $W^u(C_\epsilon)$. Each of these objects, C_ϵ, $W^s(C_\epsilon)$ and $W^u(C_\epsilon)$ are smoothly $\mathcal{O}(\epsilon)$ close to their unperturbed versions $C_0, W^s(C_0), W^u(C_0)$, respectively. (Here smoothness is understood in the same sense as in Theorem 2.)* □

As opposed to (6), for (28) no variables are in general acting as parameters for the layer problem $\epsilon = 0$. Moreover, the description of the dynamics on C_ϵ for (28) differs from that of (6). We saw that for (6) the dynamics on C_ϵ are approximated by the reduced problem (2) in terms of the slow time t. The same is true for (28) on C_ϵ, but this limiting system for $\epsilon \to 0$ in terms of the slow time t is more hidden. We refer to [74] for further details.

In [17], Fenichel studied a planar system in nonstandard slow fast form having the unit circle as a critical manifold. Fenichel did not use specific equations but

$$\begin{aligned} \dot{x} &= x(1 - (x^2 + y^2)) - \epsilon y g(x, y), \\ \dot{y} &= y(1 - (x^2 + y^2)) + \epsilon x g(x, y), \end{aligned} \tag{29}$$

provides a concrete example. For $\epsilon = 0$ this system has a (regular/isolated) equilibrium at $(x, y) = (0, 0)$ (it is a hyperbolic repelling node) and the unit

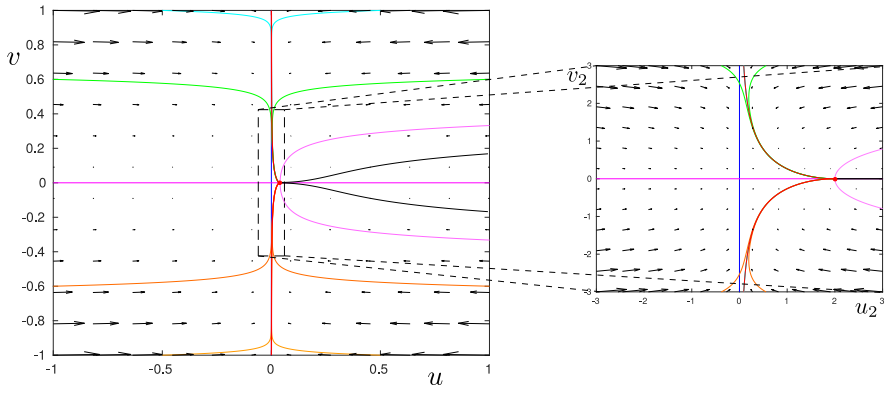


Figure 7: Phase portrait of (27) for $\mu = 3$ and $\epsilon = 0.02$. The inset shows a zoom using the scaled (u_2, v_2)-variables. In this case, the equilibrium indicated by the red point is a stable node and the trajectories in red, yellow, orange, green and cyan that converge towards S_a all have horizontal tangencies at the stable node due to the fact that the u-axis is the weak direction. The pink trajectories are the strong stable manifold of the node and the black trajectories, starting on the other side of the strong stable manifold have tangencies along opposite directions. The red orbits extend the furthest and can be thought of as perturbations of S_a. Due to the lack of hyperbolicity near $v = 0$, and the stable node, the resulting set is not smooth there (it has a cusp). In general however, by Theorem 2 the set will be smooth as a graph over any compact interval contained within either $v < 0$ or $v > 0$ for all $\epsilon > 0$ sufficiently small.

circle $x^2 + y^2 = 1$ as a compact critical manifold C_0. It is attracting and normally hyperbolic since the linearization around any point gives -2 as a nontrivial eigenvalue. By Theorem 4, C_0 perturbs to an invariant smooth slow manifold C_ϵ (diffeomorphic to a circle) for all $0 < \epsilon \ll 1$. Now, Fenichel used his example to demonstrate why C_ϵ is not necessarily smooth. In the context of (29), one could take $g(x, y) = x + \delta y$ with $\delta \neq 0$ small (to break the symmetry $x \mapsto -x$ for $\delta = 0$). Then C_ϵ would have a saddle $q_{s,\epsilon}$ near $(x, y) = (0, -1)$ and a stable node $q_{n,\epsilon}$ near $(x, y) = (1, 0)$. In particular, the linearization around $q_{n,\epsilon}$ produce the following eigenvalues $\lambda_1 \approx -\epsilon$ and $\lambda_2 \approx -2$. Now, C_ϵ is the unstable manifold of $q_{s,\epsilon}$, and therefore locally C^∞ (recall Remark 1), but near $q_{n,\epsilon}$, where

$$\begin{aligned} \dot{x} &\approx \lambda_1 x, \\ \dot{y} &\approx \lambda_2 (y - 1), \end{aligned}$$

so that solutions (upon eliminating time) take the form

$$y - 1 \approx cx^{\frac{\lambda_2}{\lambda_1}}, \tag{30}$$

it will (due to the root in (30)) only have finite smoothness (of degree $\approx \lfloor \frac{2}{\epsilon} \rfloor$). Our description is clearly only qualitative and formal. For further (rigorous) details, we refer to [17].

2.5 Overview

Since Fenichel's theory can be used successively in multi-scaled systems, it is meaningful to restrict attention to lower dimensional systems. In the remainder of the paper, we therefore consider two and three dimensional systems. In Section 3, we first consider planar systems, starting in Section 3.1 with the van der Pol system, perhaps the most classical example of a slow-fast system. We also use this example to demonstrate how to put a general model (with units) into the general form (1) through appropriate scalings and nondimensionalization. Subsequently in Section 3.2, we consider a slightly more complicated planar model of viral blips [75, 76]. This model serves to illustrate that in applications, the full dynamics can often only be analyzed upon patching different regimes together. The blowup method has been shown to be a powerful method to rigorously glue such regimes together [31, 32, 33, 37, 38, 43] and we will briefly demonstrate this in the context of the viral blip model.

In Section 4, we will then turn our attention to three dimensional systems, focusing on new phenomena that occur in the larger space.

3 Models in two dimensions

3.1 Van der Pol

In the 1920s, van der Pol [70, 71], a dutch engineer, studied electrical circuits in radio engineering by modifying the model for a resistor (R), inductor (L) and capacitor (C) system:

$$L\frac{d^2I}{dT^2} + R\frac{dI}{dT} + \frac{1}{C}I = F,$$

where $I(T)$ is the current and F is a forcing (which we will take to be constant). In particular, he argued that the resistor R in some devices depends upon the current, swapping sign when the amplitude is greater than a specific constant value I_0. He therefore replaced R by $R_0(I^2 - I_0^2)$ for some constant R_0:

$$\frac{d^2I}{dT^2} + \frac{R_0(I^2 - I_0^2)}{L}\frac{dI}{dT} + \frac{1}{CL}I = \frac{F}{L}. \tag{31}$$

Let $\omega = \frac{1}{CR_0 I_0^2}$ (having units of s^{-1}), $\frac{d}{dT} = \omega \frac{d}{dt}$, and finally $I = I_0 y$. Then t and y are nondimensional variables (of time and current, respectively). Moreover, if we define the following nondimensionalized numbers

$$\epsilon = \frac{L}{CR_0^2 I_0^4}, \quad \lambda = \frac{CF}{I_0},$$

then (31) becomes

$$\epsilon \frac{d^2 y}{dt^2} + (y^2 - 1)\frac{dy}{dt} + y = \lambda.$$

Upon writing

$$x = \epsilon \frac{dy}{dt} + \frac{1}{3}y^3 - y,$$

we obtain the first order system

$$\begin{aligned} \frac{dx}{dt} &= \lambda - y, \\ \epsilon \frac{dy}{dt} &= x - \frac{1}{3}y^3 + y, \end{aligned} \tag{32}$$

which is frequently called the van der Pol system in the literature. (In fairness, x and y are often swapped in the literature, see e.g. [47], but we prefer to have x slow and y fast for consistency).

The history of (32) is fascinating. Van der Pol, see also [72] in joint work with his colleague van der Mark, demonstrated that the system could sustain different type of oscillations depending upon ϵ and λ. This research stimulated further mathematical interest into nonlinear oscillators. In particular, Cartwright and Littlewood in the 40s and 50s [9, 10], see also [51], studied the system with periodic forcing $\lambda(t)$. An argument can be made that their work marks the birth of chaos theory [64]. Subsequently, the van der Pol system has also been used as the basic model in the development of GSPT and blowup, see [14, 44]. In fact, for $0 < \epsilon \ll 1$ the van der Pol system exemplifies the most basic concepts of planar slow-fast systems: A curved critical manifold, relaxation oscillations and so-called *canards* [44]. Haiduc [24], also proved existence of chaotic dynamics in the forced van der Pol system using GSPT. Haiduc therefore demonstrated a modern and more geometric approach to the results of Cartwright and Littlewood.

The layer problem of (32) is given by

$$\begin{aligned} \frac{dx}{d\tau} &= 0, \\ \frac{dy}{d\tau} &= x - \frac{1}{3}y^3 + y. \end{aligned}$$

The critical manifold $S = \{(x, y) : x = \frac{1}{3}y^3 - y\}$ is a graph over y, not x. In particular, by linearizing the layer problem around any point on S we find that the nontrivial eigenvalue is given by $-y^2 + 1$ and S therefore divides up into an attracting branch $S_{a,1} := S \cap \{y < -1\}$, a repelling branch $S_r := S \cap \{-1 < y < 1\}$ and a separate attracting branch $S_{a,2} := S \cap \{y > 1\}$. On top of that, there are two fold points p_1 and p_2 at $(\frac{2}{3}, -1)$ and $(-\frac{2}{3}, 1)$, respectively, so that $S = S_{a,1} \cup p_1 \cup S_r \cup p_2 \cup S_{a,2}$. See Fig. 8.

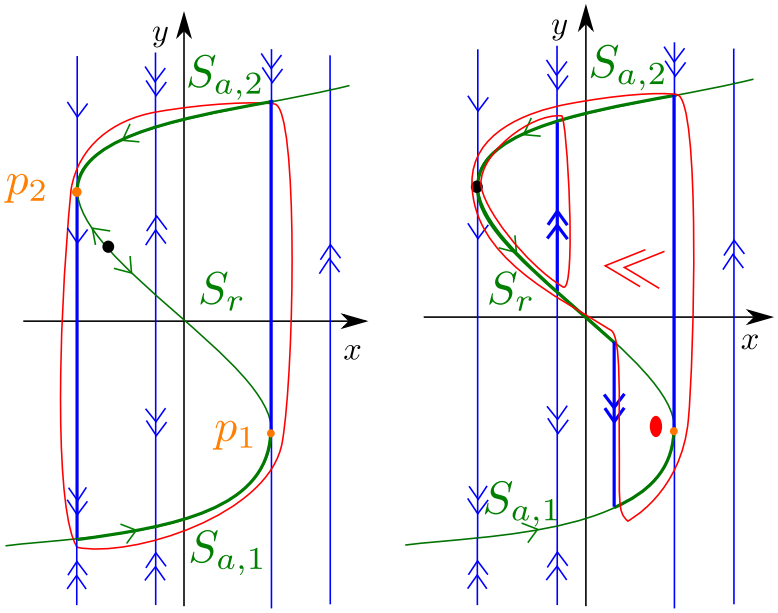


Figure 8: The van der Pol system. (a): The relaxation regime $\lambda \in (-1, 1)$. The singular cycle, consisting of a union of orbit segments of the reduced problem and the layer problem shown in green and blue (thickened curves), respectively, perturbs to a unique attracting limit cycle for all $0 < \epsilon \ll 1$. (b): Canard situation for $\lambda = 1$ where the reduced problem has an orbit that extends through the fold. This gives rise to the canard explosion phenomena [4, 45] where limit cycles grow within a very narrow parameter regime $\lambda \sim 1$ (width of order $\mathcal{O}(e^{-c\epsilon^{-1}})$).

The reduced problem on S is given by $\frac{dx}{dt} = \lambda - y$, but since we cannot write the second equation as a graph over x, we write the reduced problem in terms of y instead. This gives

$$(y^2 - 1)\frac{dy}{dt} = \lambda - y, \tag{33}$$

which is singular at the fold points $y = \mp 1$, corresponding to p_1 and p_2, and has a regular critical point at $y = \lambda$. This point lies on $S_{a,1}$, S_r and

$S_{a,2}$ for $\lambda < -1$, $-1 < \lambda < 1$ and $\lambda > 1$, respectively. In particular, when $\lambda \in (-1, 1)$ then we are in what is known as the "relaxation regime", where both fold points $p_{1,2}$ are regular jump points and this gives to a "singular cycle" as illustrated in Fig. 8(a) using thickened curves. By Theorem 3 and the contraction mapping theorem it follows that this cycle perturbs to an attracting limit cycle (red in Fig. 8(a)) for all $0 < \epsilon \ll 1$. Due to the time scale separation, $y(t)$ along such a cycle looks like copies of Fig. 1(b). Such limit cycle are said to be of relaxation type [71].

For $\lambda = 1$ then (33) becomes

$$(y+1)\frac{dy}{dt} = -1,$$

upon division by the common factor $1 - y$ on both sides. Consequently, for this parameter value, the reduced system is well-defined through the fold point p_2. (A similar result holds for $\lambda = -1$ through p_1.) In this case, p_2 violates the assumption of Theorem 3. It is call a canard point [45]. As illustrated in Fig. 8(b) it gives rise to singular cycles following S_r (thickened curves). For $0 < \epsilon \ll 1$ it can be shown (e.g. using blowup [14, 45]) that the system has a Hopf bifurcation for $\lambda = \lambda_H^+(\sqrt{\epsilon})$, $\lambda_H^+(0) = 1$ where – following the cycles in Fig. 8(b) – a rapid increase of amplitude of limit cycles occurs within an exponentially small neighborhood of $\lambda = \lambda_c^+(\sqrt{\epsilon})$, $\lambda_c(0) = 1$ [14, 45]. Along the branch of limit cycles, some cycles resemble a duck, see the red cycle in Fig. 8(b) with eyes and feathers added for further emphasis. There is also a Hopf bifurcation along $\lambda = \lambda_H^-(\epsilon)$, $\lambda_H^-(0) = -1$, and a similar "canard explosion" near $\lambda = -1$. In fact, for the van der Pol system it can be shown that there is a unique limit cycle for each $\lambda \in (\lambda_H^-(\epsilon), \lambda_H^+(\epsilon))$. For $\lambda \in (-1, 1)$ fixed, the resulting limit cycle is of relaxation type for all $0 < \epsilon \ll 1$.

Remark 3 Following [3], these "duck orbits" are called canards (the french word for ducks). In fact, canards are now used in more broad terms as orbits (not necessarily resembling actual ducks, by any stretch of the imagination) that follow repelling critical manifolds for an $\mathcal{O}(1)$-distance, i.e. on the order of the slow time scale. □

3.2 Viral blip

In this subsection, we consider the following model

$$\begin{aligned} \frac{dx_2}{d\tau} &= 1 - \epsilon x_2 - \left(b + \frac{ay_2}{y_2 + c}\right) x_2 y_2, \\ \frac{dy_2}{d\tau} &= y_2 \left(\left(b + \frac{ay_2}{y_2 + c}\right) x_2 - 1\right). \end{aligned} \tag{34}$$

This model was presented in [76] as a minimal model of viral blips. These are recurrent infections that are characterized by short episodes of high

viral reproduction, separated by long periods of quiescence. Such patterns are observed in many persistent infections, including chronic infection with HIV [76]. x_2 and y_2 represent healthy and infected T cells, respectively, and a, b and c are model parameters used to describe the infection. ϵ is a death rate. Fig. 9 shows a numerical computation of (34) for the parameter values:

$$\epsilon = 0.057, a = 0.364, b = 0.06, c = 0.823, \tag{35}$$

see figure text for further details. [76] demonstrate existence of different bifurcations upon variation of b, including saddle-node, transcritical, Hopf whereas [75] showed existence of a homoclinic. In this section, we will show that all of these bifurcations (except for the Hopf) occur persistently in the regime of (a, b) defined by

$$a = \epsilon a_2, \quad b = \epsilon b_2, \tag{36}$$

considering (a_2, b_2, c) in some fixed compact domain and $0 < \epsilon \ll 1$. For the values in (35), we have $a_2 = 6.386, b_2 = 1.0526$.

Obviously, (36) only defines a particular region of parameter space. It is possible to consider different parameter regimes and (some of these) can be studied by the same methods. The regime (36) also gives rise to a separate slow-fast phenomena, called bifurcation delay [11, 12, 53], which we will review in this context.

Inserting (36) into (34) gives

$$\begin{aligned} \frac{dx_2}{d\tau} &= 1 - \epsilon x_2 - \epsilon \left(b_2 + \frac{a_2 y_2}{y_2 + c} \right) x_2 y_2, \\ \frac{dy_2}{d\tau} &= y_2 \left(\epsilon \left(b_2 + \frac{a_2 y_2}{y_2 + c} \right) x_2 - 1 \right). \end{aligned} \tag{37}$$

It is not slow-fast as (1), not even in the nonstandard form, recall (28). In fact, setting $\epsilon = 0$ gives

$$\begin{aligned} \frac{dx_2}{d\tau} &= 1, \\ \frac{dy_2}{d\tau} &= -y_2, \end{aligned}$$

without any equilibria. Instead, x_2 is increasing unboundedly whereas y_2 decreases monotonically towards $y_2 = 0$. Consequently, for any fixed compact set in the (x_2, y_2)-plane there are can be *no* limit cycles or equilibria contained in this set for all $0 < \epsilon \ll 1$. This indicates that different scaling regimes (where x_2 is large) are important. Looking at the terms in (37), we are led to consider $x = \epsilon x_2$. Notice specifically, how the terms in the equation for y_2 then become independent of ϵ. Inserting this scaling gives

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon \left(1 - x - \left(b_2 + \frac{a_2 y_2}{y_2 + c} \right) x y_2 \right), \\ \frac{dy_2}{d\tau} &= y_2 \left(\left(b_2 + \frac{a_2 y_2}{y_2 + c} \right) x - 1 \right), \end{aligned} \tag{38}$$

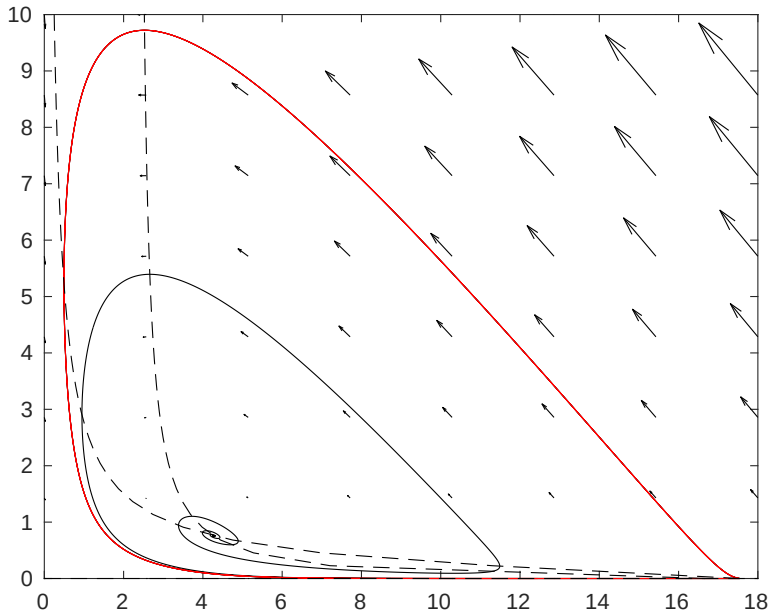


Figure 9: Phase portrait of the system (34) for the parameter values (35). The red orbit is an attracting limit cycle. The dotted lines are nullclines of the system.

which is now slow-fast (in the standard form (6)). Setting $\epsilon = 0$, we obtain the layer problem

$$\begin{aligned} \frac{dx}{d\tau} &= 0, \\ \frac{dy_2}{d\tau} &= y_2 \left(\left(b_2 + \frac{a_2 y_2}{y_2 + c} \right) x - 1 \right), \end{aligned} \tag{39}$$

and two critical manifolds: S_1 defined by $y_2 = 0$ and S_2 defined by

$$x = h_0(y_2) := \frac{1}{b_2 + \frac{a_2 y_2}{y_2 + c}}. \tag{40}$$

Notice that for $y_2 \geq 0$

$$h_0(y_2) > 0, \quad h'(y_2) < 0, \quad h_0(y_2) \to x_c := \frac{1}{a_2 + b_2}, \tag{41}$$

where the last limit is with respect to $y_2 \to \infty$.

S_1 and S_2 intersect transversally in $(x, y_2) = \left(\frac{1}{b_2}, 0 \right)$, which is a degenerate point, in the sense that the linearization gives only zero eigenvalues. Away from this point, we find that S_2 for $y_2 > 0$ is normally hyperbolic and repelling, whereas S_1 is normally hyperbolic and attracting for $x < \frac{1}{b_2}$

and repelling for $x > \frac{1}{b_2}$. In fact, when viewing x as a parameter, the y_2-subsystem undergoes a transcritical bifurcation at $x = \frac{1}{b_2}$, see illustration in Fig. 10. By implicit differentiation, the reduced problem on S_2 can be written as

$$\frac{dy_2}{dt} = \frac{1}{h_0'(y_2)} \left(1 - h_0(y_2) - y_2\right),$$

using (40) and (41). The number of equilibria of the reduced problem on S_2 are therefore given by the roots of the rational function

$$R(y_2) := h_0(y_2) + y_2 - 1,$$

on $y_2 > 0$. The derivative of R at a root also determines its stability ($R'(y_2) < 0$ stable, $R'(y_2) > 0$ unstable). The function R depends upon the parameters a_2, b_2 and c. $R(y_2) = 0$ can in general be solved for b_2 as a function of $a_2, y_2 \neq 1$ and c:

$$b_2 = v(y_2),$$

where $v(y_2) := \frac{1}{1-y_2} - \frac{a_2 y_2}{y_2 + c}$. See sketch in Fig. 10(c) (orange). From this expression, setting $v'(y_2) = 0$ gives $y_2 = y_{2f}$ where

$$y_{2f} := \frac{\sqrt{a_2 c} - c}{\sqrt{a_2 c} + 1},$$

and $v''(y_{2f}) > 0$. Consequently, there is a saddle-node bifurcation at $b_2 = b_{2f} := v(y_{2f})$, see Fig. 10 and the figure caption for further details. For the values in (35), we have $b_{2f} \approx -0.4394$, which is therefore negative. This has no biological meaning, but b_{2f} becomes positive for smaller values of a_2. The general expressions are a bit complicated, but with c fixed as in (35) we find that if $a_2 \lesssim 5.0956$ then $b_{2f} \gtrsim 0$.

Now on S_1 we obtain the following reduced problem:

$$\frac{dx}{dt} = 1 - x, \tag{42}$$

and hence there is always an equilibrium on S_1 at $x = 1$. If $b_2 < 1$ ($b_2 > 1$) then it is on the attracting (repelling) $S_{1,a}$ ($S_{1,r}$, respectively) part of S_1, whereas for $b_2 = 1$ it collides with one of the equilibria on S_2 in a transcritical bifurcation.

By Fenichel's theory most of the previous results, based upon the singular limit $\epsilon = 0$, can be perturbed to all $\epsilon > 0$ sufficiently small. In particular, for fixed $a_2 > c$ there is also a saddle-node bifurcation at $b_2 = b_{2f}$ for all $0 < \epsilon \ll 1$. However, the situation near $b_2 \gtrsim 1$ is special for (39). For $b_2 > 1$, $\frac{dx}{dt} > 0$ on the attracting part $S_{1,a}$ of S_1, and apriori one would therefore expect that trajectories starting near $S_{1,a}$ for all $0 < \epsilon \ll 1$ would

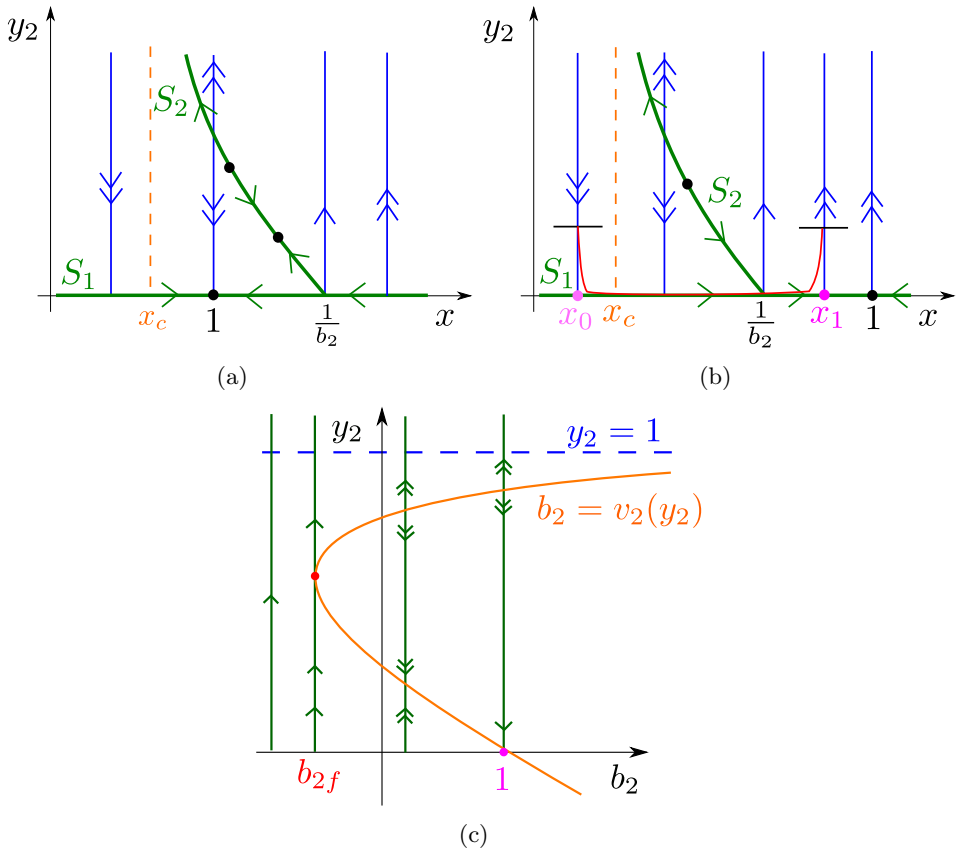


Figure 10: Slow-fast dynamics of (38) for $b_2 \in (b_{2f}, 1)$ in (a) and $b_2 > 1$ in (b). In (b) we indicate a typical orbit in red for $0 < \epsilon \ll 1$ contracting towards S_1 near $x = x_0$ before leaving at $x = x_1$ at an $\mathcal{O}(1)$-distance from the bifurcation point $x = \frac{1}{b_2}$. The orange dotted lines indicate the asymptote of the S_2 manifold, see (41). Figure (c) shows a bifurcation diagram for the reduced problem on the critical manifold S_2. For $b = b_{2f}$ a saddle-node bifurcation occurs (red in (c)) so that there are no equilibria on S_2 for $b_2 < b_{2f}$. At $b_2 = 1$ a transcritical bifurcation (purple in (c)) occurs. (This bifurcation is also associated with a homoclinic bifurcation in the full system (not shown), see Remark 4.) For the values in (35), $b_{2f} < 0$ but by decreasing a_2 one can have $b_{2f} > 0$.

follow $S_{1,a}$ all the way up to $x = \frac{1}{b_2}$ and then follow the fast jump, like in Fig. 5. But this is in fact not what occurs for (39). The point is that $y_2 = 0$ is invariant for all $\epsilon > 0$ and this leads to the *bifurcation delay* [11, 12, 53] where trajectories, originating near $S_{1,a}$ follow the repelling part $S_{1,r}$ for an $\mathcal{O}(1)$-distance before separating for all $0 < \epsilon \ll 1$. Such trajectories are sometimes also called canards, see Remark 3. The theory of bifurcation

delay is well-described for planar systems. We describe it in the context of (39), see Fig. 10(b) for an illustration. Here $\lambda(x) := b_2 x - 1$ is the nontrivial eigenvalue obtained by linearizing the layer problem around $(x, 0) \in S_1$. It changes sign at $x = \frac{1}{b_2}$ since S_1 changes stability there. Consider an initial condition (x_0, δ) with $\delta > 0$ small enough and $x_0 < \frac{1}{b_2}$. Then under the flow of (39) the point gets attracted to $S_{1,a}$ and then x increases according to (42) and eventually the orbit intersects $y_2 = \delta$ in a point with $x = x_1$, say. Here $x_1 = x_1(x_0, \epsilon)$ depends upon x_0 and ϵ, but the theory says the following:

Lemma 1 *[11, 12, 53] Suppose $b_2 > 1$ and fix any $r \in \mathbb{N}$. Then for x_0 in a compact interval contained within $x_0 < \frac{1}{b_2}$, the mapping $(x_0, \epsilon) \mapsto x_1(x_0, \epsilon)$ has a C^r-extension $x_1(x_0, 0)$ to $\epsilon = 0$ given implicitly by*

$$\int_{x_0}^{x_1} \frac{\lambda(x)}{1-x} dx = 0. \tag{43}$$

□

The interpretation is clear: We can write (38) as an equation for $y_2'(x)$ which has $y_2 = 0$ as a solution. Linearizing around this solution gives $y_2'(x) = \epsilon^{-1} \frac{\lambda(x)}{1-x} y_2$, which can be solved

$$y_2(x) = \exp\left(\epsilon^{-1} \int_{x_0}^{x} \frac{\lambda(\tilde{x})}{1-\tilde{x}} d\tilde{x}\right) y(x_0). \tag{44}$$

The quantity $\int_{x_0}^{x} \frac{\lambda(\tilde{x})}{1-\tilde{x}} d\tilde{x}$ measures the accumulative attraction along S_1 and for $x < x_1$ this quantity is negative, so that $y_2(x)$ in (44) is exponentially small. On the other hand, for $x > x_1$ this quantity is positive so that $y_2(x)$ in (44) is exponentially large.

In any case, for $b_2 > 1$ we find no limit cycles in any compact domain of the (x, y_2)-plane for $0 < \epsilon \ll 1$, see Fig. 10(b). This is in contrast to Fig. 9. This suggests that a different scaling capturing y_2 large is important. Define

$$y = \epsilon y_2.$$

Then in terms of (x, y), we obtain

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon(1-x) - \left(b_2 + \frac{a_2 y}{y + \epsilon c}\right) xy, \\ \frac{dy}{d\tau} &= y\left(\left(b_2 + \frac{a_2 y}{y + \epsilon c}\right) x - 1\right). \end{aligned} \tag{45}$$

Setting $\epsilon = 0$ for $y > 0$ gives the limiting system

$$\begin{aligned} \frac{dx}{d\tau} &= -\left(b_2 + a_2\right) xy, \\ \frac{dy}{d\tau} &= y\left(\left(b_2 + a_2\right) x - 1\right). \end{aligned} \tag{46}$$

Viewing this system on $y \geq 0$, we find that $(x, 0)$ defines a set of equilibria. For $y > 0$, upon dividing the right hand side of (46) by y, we obtain a linear system and in this way, we can show that the phaseportrait of (46) consists of heteroclinic connections to $y = 0$, see Fig. 11(a). These connections are organized around the point $(x_c, 0)$, recall (41). In fact, for each $x_1 > x_c$, we can write the connections from $(x_1, 0)$ to $(x_0, 0)$ as graphs over x

$$y(x) = -\left(x_c \log \frac{x_1}{x} - (x_1 - x_0)\right), \quad x \in (x_0, x_1), \tag{47}$$

where $x_0 < x_c$ is such that $y(x_0) = 0$:

$$x_c \log \frac{x_1}{x_0} = x_1 - x_0. \tag{48}$$

This equation defines a mapping $x_1 \mapsto x_0 \in (0, x_c]$ from $x_1 \geq x_c$, having $x_1 = x_c$ as a fixed-point. The correspondence is illustrated in Fig. 12 (dashed) for a new set of parameters $a_2 = 1.5$, $b_2 = 2$. The reason for changing the parameters is that it gives a better plot. In essence there is no difference between these values and those in (35), other than x_c being quite small for the latter set of parameters.

Fig. 12 also shows the correspondence $x_0 \mapsto x_1 \in \left(\frac{1}{b_2}, 1\right)$ given by (43) for $x_0 < \frac{1}{b_2}$ for the same set of parameters. We see that the two curves intersect in a unique point (circle). It is straightforward to show that an intersection occurs for all $b_2 > 1$ (although uniqueness is more involved), and upon combination of (46) with the slow-fast system (38), cf. Lemma 1, see Fig. 11 with $a_2 = 1.5$, $b_2 = 2$ and $c = 0.5$, we therefore see a global picture of the dynamics, including the main mechanism leading to limit cycles for $b_2 > 1$.

Formally, however, there is a technical issue merging the two limit dynamics as the domains do not overlap. Firstly, the analysis in the (x, y_2)-plane is only valid in a compact subset, say $[0, c_1] \times [0, c_2]$, and secondly due to the singularity of the terms $\frac{y}{y+\epsilon c}$ in (45) at $y = \epsilon = 0$ we can only perturb away from (47) within compact subsets of $y > 0$, say $y \in [c_3, c_4]$. Here all $c_i > 0$. There is therefore a gap between $y = \epsilon c_2$ and $y = c_3$ for $0 < \epsilon \ll 1$ which is not captured. To cover this gap, we first have to regularize the system (45) for $y = \epsilon = 0$. We do so by following [39] and use the fact that the system is rational and therefore becomes polynomial when we multiply the right side by the denominator $y + \epsilon c$:

$$\begin{aligned} \frac{dx}{ds} &= (\epsilon(1-x)(y+\epsilon c) - (b_2(y+\epsilon c) + a_2 y)\, xy), \\ \frac{dy}{ds} &= y\left((b_2(y+\epsilon c) + a_2 y)\, x - (y+\epsilon c)\right). \end{aligned} \tag{49}$$

For $y \geq 0$ and any $\epsilon > 0$ this multiplication corresponds to a reparametriza-

tion of time defined implicitly by

$$\frac{d\tau}{ds} = y + \epsilon c.$$

Nevertheless, $y = \epsilon = 0$ is well-defined for (49) as a set of equilibria. It is degenerate, since the linearization only produces zero eigenvalues, and so we cannot perturb this set by using Fenichel. However, for the extended system, obtained by augmenting $\frac{d\epsilon}{ds} = 0$ to (49), we can blowup the degenerate set $(x, 0, 0)$ to a cylinder through a polar blowup transformation:

$$(r, (\bar{y}, \bar{\epsilon})) \mapsto \begin{cases} y &= r\bar{y}, \\ \epsilon &= r\bar{\epsilon}, \end{cases} \tag{50}$$

leaving x fixed, with $r \geq 0$, $(\bar{y}, \bar{\epsilon}) \in S^1_+ = S^1 \cap \{\bar{\epsilon} \geq 0\}$, S^1 being the unit circle in $\mathbb{R}^2$. In this way, we can gain hyperbolicity along $r = 0$ through desingularization. Consider the coordinates (x, r_1, ϵ_1) of the chart $\bar{y} = 1$ associated with (50):

$$(r_1, \epsilon_1) \mapsto \begin{cases} y &= r_1, \\ \epsilon &= r_1\epsilon_1, \end{cases}$$

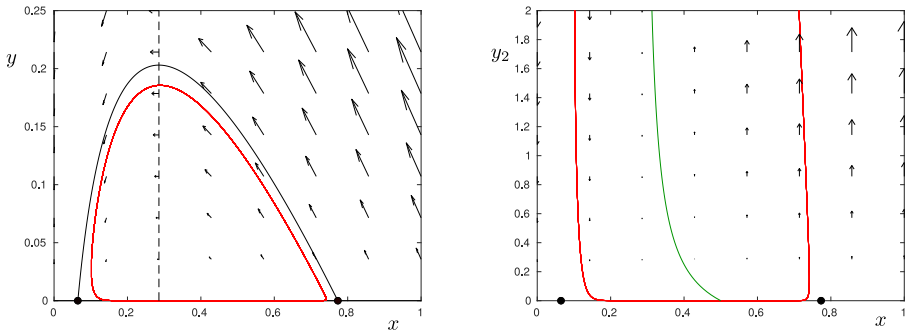
and notice specifically that $\epsilon = r_1\epsilon_1, y_2 = \epsilon_1^{-1}$ or $r_1 = \epsilon y_2, \epsilon_1 = y_2^{-1}$. We can therefore transform coordinates from (x, y_2, ϵ), which capture $y = \epsilon y_2 = \mathcal{O}(\epsilon)$, to (x, r_1, ϵ_1) for $y_2 > 0$. At the same time since $y = r_1$, we cover the regime $y = \mathcal{O}(1)$ by taking $\epsilon_1 = \mathcal{O}(\epsilon)$ small enough. Consequently, the coordinate (x, r_1, ϵ_1) are ideally suited to cover the aforementioned gap and match the two regimes: $y = \mathcal{O}(1)$ and $y = \mathcal{O}(\epsilon)$.

The differential equations in the (x, r_1, ϵ_1) coordinates are obtained from (49) and $\frac{d\epsilon}{ds} = 0$. We find

$$\begin{aligned} \frac{dx}{ds_1} &= r_1 \left(\epsilon_1(1-x)(1+\epsilon_1 c) - (b_2(1+\epsilon_1 c) + a_2)\, x\right), \\ \frac{dr_1}{ds_1} &= r_1 \left((b_2(1+\epsilon_1 c) + a_2)\, x - (1+\epsilon_1 c)\right), \\ \frac{d\epsilon_1}{ds_1} &= -\epsilon_1 \left((b_2(1+\epsilon_1 c) + a_2)\, x - (1+\epsilon_1 c)\right) \end{aligned}$$

upon desingularization through division of the right hand side by r_1. Here $r_1 = \epsilon_1 = 0$ defines a set of equilibria, but it is now partially hyperbolic at every point except $x \neq x_c$. Indeed the linearization around any point $(x, 0, 0)$ gives two nontrivial eigenvalues $\pm(x_c^{-1}x - 1)$. In this way, the connections (47) become heteroclinic orbits connecting partially hyperbolic points whereas the unbounded orbits in the (x, y_2)-plane for $\epsilon = 0$ become trajectories, upon applying the coordinate transformation $\epsilon_1 = y_2^{-1}$, $r_1 = r_2 y_2$,

contained within the stable and unstable manifolds of $r_1 = \epsilon_1 = 0$. We illustrate the findings in Fig. 13. It is possible to describe the transition near $r_1 = \epsilon_1 = 0$ for $x \neq x_c$; they are smooth functions of x, $\epsilon = r_1\epsilon_1$ and $\epsilon \log \epsilon^{-1}$, see [12, Proposition 2.1]. In this way, by working in the (x, y_2, ϵ) coordinates and the (x, r_1, ϵ_1) coordinates one can then describe a return map $y_2 = \delta$, $\delta > 0$, in the (x, y_2) coordinates, with $x \sim x_1$ where x_1 is the value at the intersection point in Fig. 12. It has a smooth extension to $\epsilon = 0$, having $x = x_1$ as an attracting fixed point. This fixed point therefore perturbs for all $0 < \epsilon \ll 1$ and produces the desired limit cycle for any $b_2 > 1$.



(a) (b)

Figure 11: Phase portrait of the two limiting systems (46) (in (a)) and (39) (in (b)) for $a_2 = 1.5, b_2 = 2, c = 0.5$. The circles correspond $x = x_0$ and $x = x_1$; these values are the coordinates of the intersection point in Fig. 12. In red, we illustrate a limit cycle for $\epsilon = 0.01$.

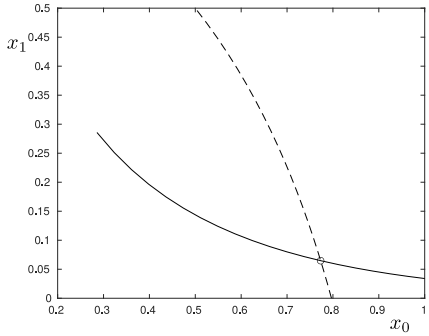


Figure 12: The two curves obtained by (43) (dashed) and (48) for $a_2 = 1.5, b_2 = 2, c = 0.5$. The intersection gives rise to a singular cycle, which perturbs to an attracting limit cycle for all $0 < \epsilon \ll 1$, see Fig. 11.

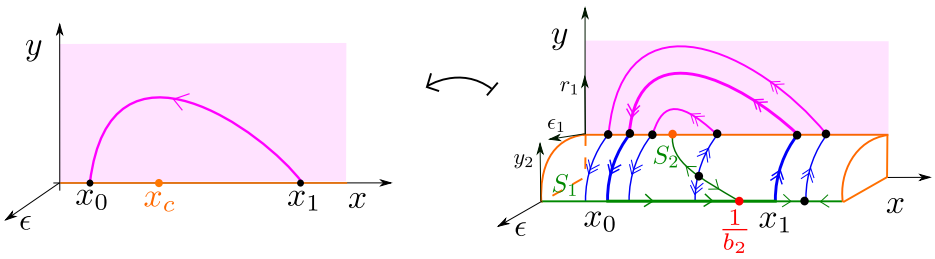


Figure 13: The transformation (50) blows up the degenerate line $y = \epsilon = 0$ of (49) to a cylinder. In this way, we gain hyperbolicity and the associated coordinates (x, r_1, ϵ_1) parametrizes the upper part of the cylinder as indicated in the figure. They enable a rigorous matching between $y = \mathcal{O}(1)$ described by (46) and $y = \mathcal{O}(\epsilon)$, or $y_2 = \mathcal{O}(1)$, described by the slow-fast system (38). The figure focuses on the case $b_2 > 1$ and the thickened curves on the right indicate a singular cycle, with x_0 and x_1 determined by the transverse intersection point shown in Fig. 12.

Remark 4 The emergence of limit cycles in the viral blip model is due to a homoclinic bifurcation around $b_2 \sim 1$. In the present case, this bifurcation is slightly unusual (which in fact erroneously led the authors of [76] to exclude its existence.) In particular, the homoclinic is not a saddle connection to the saddle on S_1 for $b_2 > 1$, but rather to the saddle on S_2, see Fig. 10(a). In fact, this saddle moves towards $(x, y_2) = (1, 0)$ as $b_2 \to 1^-$. For all $b_2 \lesssim 1$ and $0 < \epsilon \ll 1$, we can follow its unstable manifold through (46), see Fig. 11(a), back towards S_1. It will intersect $x = 1$ with $y_2 = \mathcal{O}(e^{-c/\epsilon})$ for some $c > 0$. Using a separate blowup of $(x, y_2, \epsilon, b_2) = (1, 0, 0, 1)$ by augmenting trivial equations for both ϵ and b_2, it is possible to follow the saddle in a neighborhood of $b_2 = 1$ uniformly as $\epsilon \to 0$. From this it then follows that a homoclinic bifurcation occurs at $b_2 = b_{2hom}(\epsilon)$ with $b_{2hom}(\epsilon) = 1 + \mathcal{O}(e^{-c/\epsilon}) < 1$ for some $c > 0$. The homoclinic bifurcation therefore occurs exponentially close to the transcritical bifurcation at $b_2 = 1$ where the saddle for $b_2 \lesssim 1$ collides with the stable node on S_1. □

4 Oscillatory phenomena in three-dimensional systems: Canards and mixed mode oscillations

In this section, we study some separate phenomena that can occur in three dimensional slow-fast systems. First in Section 4.1, we consider the case of two slow variables and a single fast one and then in Section 4.2, we subsequently consider one slow variable and two fast ones. In both cases, focus will again be on canards and their role in applications.

4.1 Folded singularities

Consider the following slow-fast system [65, 73]:

$$\begin{aligned}\frac{dx}{d\tau} &= \epsilon\left(\frac{1}{2}\mu y - (\mu+1)z + \mathcal{O}(x,\epsilon,(y+z)^2)\right),\\ \frac{dy}{d\tau} &= \epsilon,\\ \frac{dz}{d\tau} &= x + z^2 + \mathcal{O}(xz^2, z^3, xyz) + \epsilon\mathcal{O}(x,y,z,\epsilon).\end{aligned} \tag{51}$$

with two slow variables x and y and one fast variable z. $\mu \in \mathbb{R}$ is a separate parameter. For $\epsilon = 0$ we find that the critical manifold S is approximately given by the parabolic cylinder $x = -z^2$, $z < 0$ (S_a) being stable and $z > 0$ (S_r) being unstable. See Fig. 14. The degenerate line $F : x = z = 0$, being a fold line of degenerate points, is in orange whereas S_a and S_r are both in green. Let us for simplicity ignore the $\mathcal{O}$-terms in (51). Then the reduced problem on S is given by

$$\begin{aligned}\frac{dy}{dt} &= 1,\\ 2z\frac{dz}{dt} &= -\frac{1}{2}\mu y + (\mu+1)z,\end{aligned} \tag{52}$$

using $x = -z^2$ and implicit differentiation. Consider S_a where $z < 0$. Then multiplication of the right hand side by $-2z$ gives the topologically equivalent system

$$\begin{aligned}y' &= -2z,\\ z' &= \frac{1}{2}\mu y - (\mu+1)z,\end{aligned} \tag{53}$$

on S_a, see [73], with

$$()' = -2z\frac{d}{dt}, \tag{54}$$

indicating differentiation with respect to a new time. The system (53) is called the desingularized reduced problem and it is well-defined everywhere. In particular, the point p at $(y,z) = (0,0)$ on F, pink in Fig. 14, is a hyperbolic stable node of these equations for $\mu > 0$ with eigenvalues -1 and $-\mu$ and associated eigenvectors:

$$(2,1)^T, \quad (2,\mu)^T, \tag{55}$$

respectively. See illustration of the reduced flow in Fig. 15(a). In this case, p is called a folded node of the slow-fast system (51). For $\mu < 0$, on the

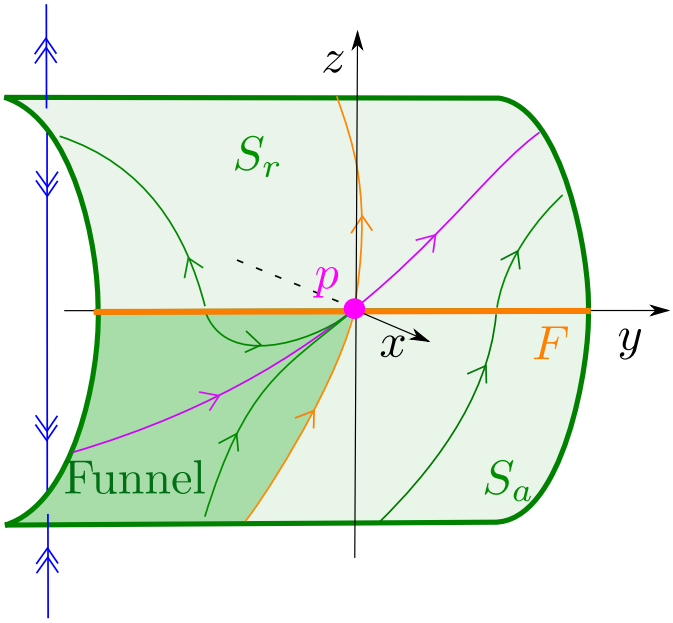


Figure 14: The folded node singularity p. Upon desingularization of the reduced problem, the folded node singularity becomes a stable node. The strong eigenvector associated with the node, gives rise to a strong stable manifold (orange) that forms a boundary of a funnel region [6] (shaded), bounded on the other side by F, where trajectories approach the folded node p (in finite time before desingularization), tangent to a weak eigendirection.

other hand, $(y, z) = (0, 0)$ is a hyperbolic saddle, and it is called a folded saddle for the slow-fast system (51), see [65].

Notice that the orbits of (52) on S_r, where $z > 0$, are also orbits of (53), but – due to the multiplication of a negative quantity there, recall (54) – their directions have to be reversed. In the case $\mu > 0$ there is an open set on S_a (the funnel, see the shaded region in Fig. 14) that reaches $(y, z) = 0$ by following the reduced problem (52). Notice that whereas these orbits are asymptotic for the flow of (53), orbits of (52) reach the point in finite time. We see this clearly from the first equation in (52). Consequently, as we saw for planar systems above, we have orbits (also called canards in the present case) connecting the attracting and the repelling branches of the critical manifold. But in contrast to canards of the van der Pol (32), the canards here are persistent (or generic), insofar that they do not require the presence of an additional parameter. However, the folded node is complicated, and in the author's opinion still not fully understood. The main difficulty lies in the fact that there is no "candidate" for the canard at the (singular) level of

$\epsilon = 0$; all points within the funnel (see Fig. 14) behave in the same way). The folded saddle is easier from this perspective. Here there is a canard for all $\epsilon > 0$ small enough connecting the attracting slow manifold on one side of the fold with the repelling slow manifold on the other side. This perturbed canard remains close to the stable manifold $W^s(q)$ of the desingularized reduced problem.

Nevertheless, for the folded node it is well-established [73] that this singularity is associated with small amplitude oscillations (SAOs), see [65, 73]. This can be explained through blowup. Indeed, the point p is fully nonhyperbolic for $\epsilon = 0$. Consequently, one applies a blowup transformation of the form:

$$r \geq 0, (\bar{x}, \bar{y}, \bar{z}, \bar{\epsilon}) \in S^3_+ \mapsto \begin{cases} x &= r^2 \bar{x}, \\ y &= r\bar{y}, \\ z &= r\bar{z}, \\ \epsilon &= r^2 \bar{\epsilon}, \end{cases}$$

where $S^3_+ = S^3 \cap \{\bar{\epsilon} \geq 0$, S^3 being the unit sphere in $\mathbb{R}^4$, to the extended system obtained by augmenting $\frac{d\epsilon}{d\tau} = 0$ to (51). In this way, one gains hyperbolicity (upon proper desingularization) of the critical manifold S at $r = 0$. We refer to [65, 73] for details and just focus on the associated scaling chart $\bar{\epsilon} = 1$ where

$$(x_2, y_2, z_2, r_2) \mapsto \begin{cases} x &= r_2^2 x_2, \\ y &= r_2 y_2, \\ z &= r_2 z_2, \\ \epsilon &= r_2^2, \end{cases}$$

so that $r_2 = \sqrt{\epsilon}$. Inserting this into (51) we can eliminate time and obtain equations for $x_2'(y_2)$ and $z_2'(y_2)$. Setting $r_2 = 0$ in these equations gives

$$\begin{aligned} \frac{dx_2}{dy_2} &= \frac{1}{2}\mu y_2 - (\mu + 1) z_2, \\ \frac{dz_2}{dy_2} &= x_2 + z_2^2. \end{aligned} \tag{56}$$

For this system we find two solutions:

$$\begin{aligned} x_2(y_2) &= -\frac{\lambda^2}{4} y_2^2 - \frac{\lambda}{2}, \\ z_2(y_2) &= -\frac{\lambda}{2} y_2, \end{aligned} \tag{57}$$

where $\lambda = -1, -\mu$ are the eigenvalues of linearization around p for the desingularized system (53). Notice that upon projecting these solutions onto

the (y_2, z_2)-plane, the associated orbits coincide with the span of the two eigenvectors. For $\mu > 1$, then the solution with $\lambda = -1$ corresponds to the weak direction, whereas the solution with $\lambda = -\mu$ corresponds to the strong direction. These solutions are crucial in the analysis of the perturbation for $0 < \epsilon \ll 1$. Moreover, suppose $\mu > 1$ ($\mu \in (0, 1)$ is equivalent). Then we can linearize (56) around (57) with $\lambda = -1$. This gives

$$x_2''(y_2) - y_2 x_2'(y_2) + (\mu + 1)x_2(y_2) = 0, \tag{58}$$

with $x_2'(y_2) = -(\mu+1)z_2$, when written as a second order equation, see [65]. This equation is known as a Weber equation.

Lemma 2 *[65] For $\mu \in (n, n+1)$ with $n \in \mathbb{N}$ then there exists two linearly independent solutions $u_\pm(y_2)$of (58), having the following asymptotic properties: $u_\pm$ grows exponentially as $y_2 \to \pm\infty$ and polynomially as $y_2 \to \mp\infty$, respectively. Moreover, u_+ has $n+1$ simple zeros.* □

The growth properties of $u_\pm$ allowed [65, 73] to show that (extended versions) of the slow manifolds intersect transversally along perturbed canards whenever $\mu \notin \mathbb{N}$ for all $0 < \epsilon \ll 1$. $\mu \in \mathbb{N}$ are bifurcations where additional intersections of the perturbed manifolds can occur. These are called secondary canards [73]. Regardless, whenever $\mu \notin \mathbb{N}$, then – due to the $n+1$ zeros of u_+ – the tangent space of the attracting slow manifold twists along the perturbed canard. This gives rise to SAOs. In Fig. 15 we illustrate some computations on (51) with the $\mathcal{O}$-terms removed. See figure caption for further details.

The model (51) is a local "normal form" for slow-fast systems (6) – with two slow variable and one fast, i.e. $x \in \mathbb{R}^2$, $y \in \mathbb{R}$ – having a folded critical manifold that intersects the nullcline of one of the slow variables transversally at a point on the fold. Under some further nondegeneracy conditions, see [65], it is possible to bring such a system into the form (51) upon using only scalings, translations and a regular transformation time. It is well-documented [5, 6, 13] that many models have such folded singularities. The main interest comes from the fact that when the system has a global return mechanism to the funnel region (see Fig. 14 and Fig. 15(a)) then – due the attractivity towards the folded node – such a situation allows trajectories to pass through the region of SAOs again and again and this leads to so-called mixed mode oscillations (MMOs). Here the SAOs are interspersed by large amplitude oscillations (LAOs) due to the return mechanism [5, 6, 13]. These references are excellent review papers on the topic, that also lay out the importance of MMOs in different areas of science, including neuroscience, see e.g. [26, 60].

Whereas the folded node relates to an attractor of the system, with an open set of points passing through such a point for $\epsilon \to 0$, the folded saddle

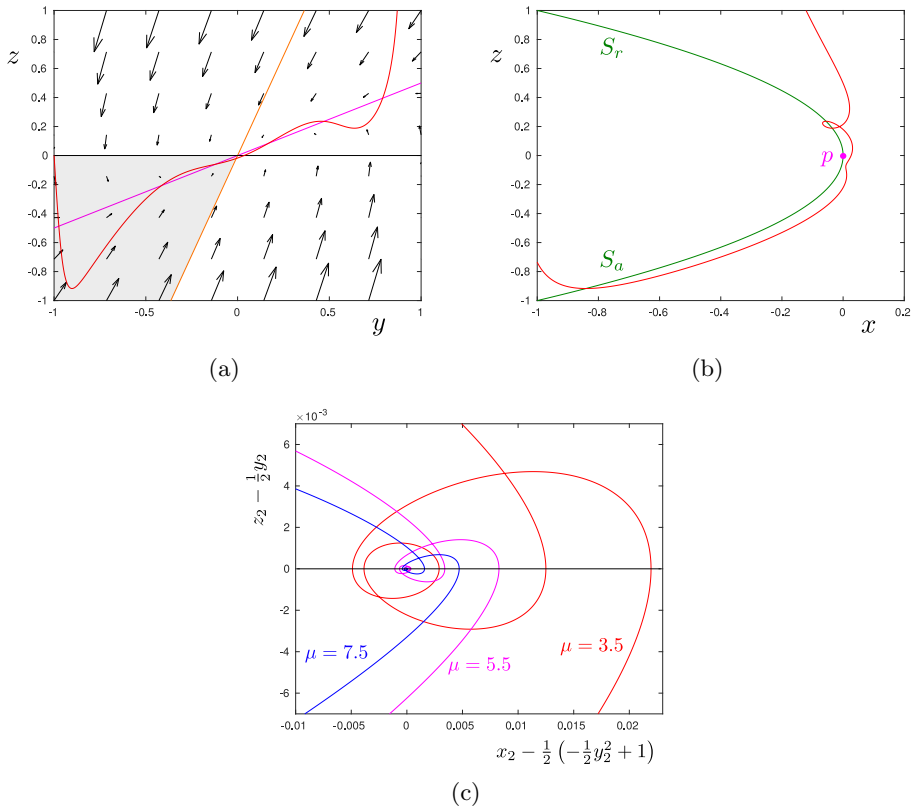


Figure 15: (a): The desingularized reduced problem (53) for $\mu = 5.5$. Curves in black are the invariant spaces of the linearization spanned by the eigenvectors (55). An orbit is shown in red for $\epsilon = 0.05$. (b): Projection onto (x, z). The critical manifold is in green whereas the orbit in red is for $\mu = 5.5$ and $\epsilon = 0.05$. (c): Solutions of (51) for $\epsilon = 0.01$ and three different values of μ. The coordinates on the axis gives a view along the weak canard and we see that the number of twists along this canard increases as μ increases. Specifically, the number of intersections with the black horizontal curve increases by 1 each time μ passes through a positive integer: The red orbit ($\mu = 3.5$) has 5 intersections, the purple ($\mu = 5.5$) has 7 intersections, and finally the blue orbit ($\mu = 7.5$) has 9 intersections.

is clearly different. However, the folded saddle provides an important mechanism for which to prove existence of chaos. It involves contraction, towards the slow manifold, and expansion, due to the saddle, and consequently, if there is a global (twisting) return mechanism, then we have the main ingredients for chaos through a horseshoe [64]. There are a few examples using this approach to prove chaos in models [24, 40, 41], starting from Haiduc's

inspirational proof [24] of chaos in the forced van der Pol.

4.2 Bursting

In the case of two fast variables, we can have Hopf bifurcations and homoclinic bifurcations in the layer problem. This leads to new types of MMOs, also known as bursting [30, 59] where slow phases are interspersed by spike-like oscillations. This type of dynamics is often observed in neural systems and this is where the term bursting comes from. The Morris-Lecar model [52] is possibly the simplest neural model supporting such phenomena, but this model is still fairly complicated, involving highly nonlinear terms. Having said that, it is mainly the geometry that is important and [66] used this to formulate general geometric conditions that lead to bursting.

Here we will consider the following toy model:

$$\begin{aligned} \frac{dx}{d\tau} &= \epsilon, \\ \frac{dy_1}{d\tau} &= y_2, \\ \frac{dy_2}{d\tau} &= x - \beta y_1 + y_1^2 - y_1 y_2 - \delta y_1^3, \end{aligned} \tag{59}$$

that illustrates the bursting phenomenon in an even simpler setting. The parameters $\beta, \delta > 0$ are fixed small enough and $0 \leq \epsilon \ll 1$. The corresponding layer problem of (59) is given by

$$\begin{aligned} \frac{dy_1}{d\tau} &= y_2, \\ \frac{dy_2}{d\tau} &= -x - \beta y_1 + y_1^2 - y_1 y_2 - \delta y_1^3, \end{aligned} \tag{60}$$

with $x = \text{const}$. To quadratic order this system is a Bogdanov-Takens normal form [56, p. 477]. We sketch the phase portrait for different values of x in Fig. 16 for $\beta > 0$ and $\delta > 0$ fixed small enough. In particular, due to the Bogdanov-Takens bifurcation at $x = \beta = 0$, it is possible to show that there is a Hopf bifurcation $x = x_H(\beta, \delta)$ and a homoclinic bifurcation $x = x_{hom}(\beta, \delta)$ both near $x \sim 0$ – in fact, there is also a saddle-node but we will not study this bifurcation – and stable limit cycles exist for all $x \in (x_H, x_{hom})$ (in red in Fig. 16(b)). (There are also Hopf and homoclinic bifurcations in the layer problem of the Morris-Lecar model, but the Hopf plays little role there. By working on (59), we have the opportunity to explain the consequences of this bifurcation on the full system for $0 < \epsilon \ll 1$.)

The cubic term has the effect of introducing a third equilibrium (furthest to the right in Fig. 16) which for all x, $\beta > 0$ and $\delta > 0$ small enough does not undergo bifurcations. Specifically, in this parameter regime it remains a stable node.

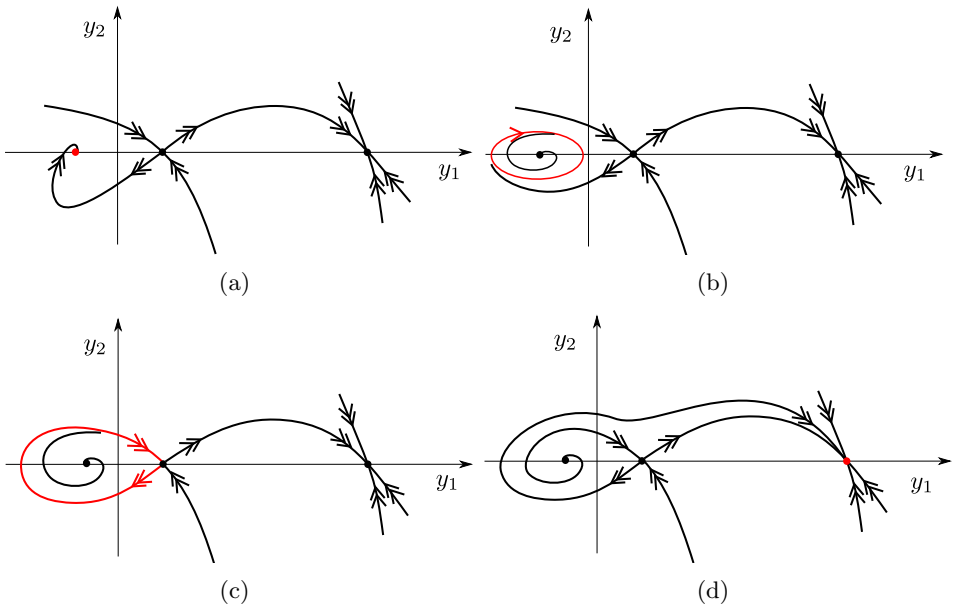


Figure 16: Phase portrait of the layer problem (60) for fixed $\beta > 0$ and $\delta > 0$, both sufficiently small, and different values of x: (a): $x < x_H$, (b): $x \in (x_H, x_{hom})$, (c): $x = x_{hom}$, (d): $x > x_{hom}$, where $x_H = x_H(\beta, \delta)$ and $x_{hom} = x_{hom}(\beta, \delta)$ are the values of x at the Hopf bifurcation and the homoclinic bifurcation, respectively.

The reduced problem is given by $\frac{dx}{dt} = 1$ for this simple system. It is well-defined for $x \lesssim x_H$, as the associated stable equilibrium (red in Fig. 16(a)) produces an attracting critical manifold. At $x = x_H$ this critical manifold loses stability in the (dynamic) Hopf bifurcation [53]. Specifically, for $x \in (x_H, x_{hom})$ the attracting limit cycles of the layer problem form a cylinder. This cylinder is for $x \in I$, with I a compact interval within (x_H, x_{hom}), an example of a normally hyperbolic and attracting cylinder. Fenichel's theory also applies to such sets [17, 18, 19, 47]. Consequently, for all $0 < \epsilon \ll 1$ there is a perturbed invariant cylinder, carrying fast dynamics in the azimuthal direction and slow dynamics in the cylindrical direction (here: the x-direction). This slow flow is in general given to leading order by averaging [53] but in our case it is just $\frac{dx}{dt} = 1$.

At $x = x_{hom}$ the limit cycles of the layer problem further bifurcate in a homoclinic bifurcation, and, in general, points following the invariant cylinder will experience a fast jump and imediately follow the attracting critical manifold that is unfolded by the stable node (red point in Fig. 16(d)) for $x \gtrsim x_{hom}$, [8, 66]. (In exception to this jump behaviour, there is thin set close to the stable manifold of the saddle-type critical manifold, that follows the saddle-type slow manifold for an extended period of time.)

This all seems very straightforward, but there is one important thing: The transition to the stable cylinder, does not occur at the Hopf bifurcation $x = x_H$ for $\epsilon \to 0$. This is evident from the simulations in Fig. 17. Here $\beta = 0.5$, $\delta = 0.3$, and $\epsilon = 10^{-4}$. Fig. 17(a) shows a projection onto (x, y_1) whereas (b) shows a projection onto (y_1, y_2). In Fig. 17(a), we see – as expected – that we first follow the critical manifold (shown in red, dashed), before oscillations start to grow, but well beyond x_H. Consequently, we see a bifurcation delay (as we saw in Section 3.2 on the viral blip in a different context), where the (canard) orbit follows the critical manifold on the repelling side, after the Hopf bifurcation, before moving close to the cylinder. Near $x = x_{hom}$, there is then a fast transition to a separate attracting branch of the critical manifold (around $y_1 \approx 2.5$). This transition is due to the homoclinic bifucation in the layer problem, recall Fig. 16(c).

The bifurcation delay seen in Fig. 17 is generic in (analytic) systems with a dynamic Hopf bifurcation, see [25, 53]. We will illustrate this further on the simplified Shishkova equation [54, 63]:

$$\begin{aligned} \frac{dz}{d\tau} &= (i + x)z + \epsilon h(x), \\ \frac{dx}{d\tau} &= \epsilon, \end{aligned} \tag{61}$$

with $z \in \mathbb{C}$, $x \in \mathbb{R}$ and h analytic. For $\epsilon = 0$, $\frac{dz}{d\tau} = (i + x)z$, $\frac{dx}{d\tau} = 0$ so $z = 0$ is normally attracting for $x < 0$, normally repelling for $x > 0$. $x = 0$ is Hopf bifurcation of $z = 0$; it is degenerate because there are no nonlinear terms. It is possible to include these, but focusing on the delay, these terms are not important and are therefore left out.

Lemma 3 *[25, 54] Consider (61) and suppose that $x(0) = x_0 < 0$, $z(0) = z_0$ and $h(-i) \neq 0$. Fix any $\delta > |z_0|$ and let $x_1(x_0, \epsilon)$ be the value $x(\tau) > 0$ where $\tau > 0$ is the smallest positive value such that $|z(\tau)| = \delta$. Then*

$$\lim_{\epsilon \to 0} x_1(x_0, \epsilon) = \begin{cases} -x_0 & \text{for} \quad x_0 \in (-1, 0), \\ 1 & \text{for} \quad x_0 \leq -1. \end{cases}$$ □

The full details can be found in [25, 54]; the former includes a very nice proof – that also extends to the nonlinear case – using complex time and blowup.

It is easy to get some intuition into the delay in Lemma 3 by considering $h(x) = \text{const.} \neq 0$. In this case, we can eliminate time τ and integrate the resulting equation for $z(x)$:

$$z(x) = e^{\epsilon^{-1} \int_{x_0}^{x} (i+s) ds} dsz(x_0) + h \int_{x_0}^{x} e^{\epsilon^{-1} \int_{u}^{x} (i+s) ds} du. \tag{62}$$

Notice that the first integral is bounded as $\epsilon \to 0$ iff $x \leq -x_0$. The last integral can be evaluated using the erf-function and we can estimate

$$|\int_{x_0}^{x} e^{\epsilon^{-1}\int_u^x (i+s)ds} du| \sim \sqrt{\epsilon} e^{\epsilon^{-1}(x^2-1)},$$

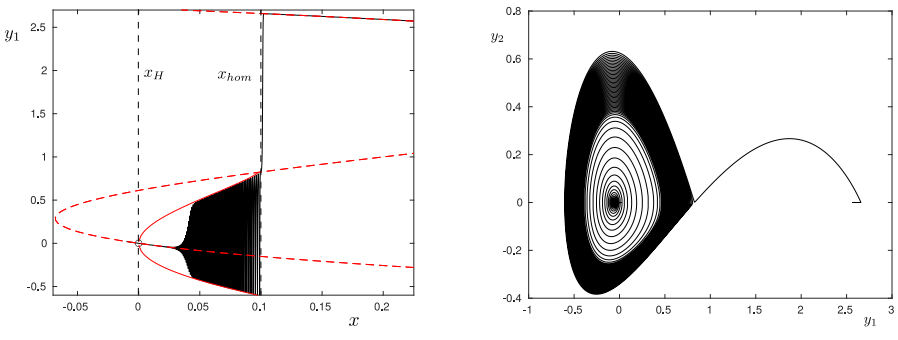
for $x \geq 0$, as $\epsilon \to 0$. (Here we use $\sim$ to indicate that the left hand side can be bounded from below and from above by a constant independent of ϵ multiplying the right hand side.) The right hand side is bounded for $x \geq 0$ iff $x \leq 1$.

In light of Lemma 1, and the bifurcation delay we saw in the model of viral blips, one might expect that $x_1 \approx -x_0$ for all $x_0 < 0$, since this is precisely where (to leading order) the expansion on the repelling side equals the contraction on the attracting side; see the first term in (62). However, the difference between Lemma 1 and Lemma 3 is that the critical manifold is only invariant for the system in Lemma 1. This is reflected by the second term in (62). In general, the reference [25] shows, see also [54, 53], that the attracting and the repelling slow manifolds of (61) split with exponentially small splitting along $x = 0$ if $h(-i) \neq 0$. It is this exponentially small splitting, exemplified in the second term in (62), which grows to order $\mathcal{O}(1)$ at the upper bound $x = 1$, see Lemma 3; this bound is also called a *buffer point* [25, 54, 53].

Remark 5 For non-analytic systems the situation is different [54, 53]. In this case the splitting may not be exponential and consequently there may be no delay at all for $\epsilon \to 0$. As an extreme case of this, we could take $h = -i$ for $x > 0$ and $h = i$ for $x < 0$ in (61). In this case, the attracting (repelling) slow manifold intersects $x = 0$ in a point with $z \approx \epsilon$ ($z \approx -\epsilon$, respectively). The splitting is therefore $\approx 2\epsilon$ in this case and due to the exponential expansion for $x > 0$ there is no delay for $\epsilon \to 0$. However, the infinitely smooth case is to the author's knowledge still not yet fully understood. □

The model (59) could, as (51) above, be understood as a local normal form, which upon coupled with a global return mechanism, could give rise to some recurrent dynamics seen in neural systems, see [59, 30] and [66] for an analysis of the Morris-Lecar models [52]. In fact, Terman in [66] demonstrated the existence of chaotic dynamics in the Morris-Lecar model through a Smale horseshoe [64]. The horseshoe was obtained through a general mechanism associated with the homoclinic bifurcation in the layer problem.

Izhikevich in [30] has an extensive classification of different bursting patterns, also based upon bifurcation of the layer problem and descriptions of the onset and termination of oscillations in bursts. For (59) the bursting occurs along the cylinder of limit cycle, being initiation after the Hopf bifurcation and terminated in the homoclinic bifurcation [56, pp. 375–376].



(a) (b)

Figure 17: Simulations of (59) for $\beta = 0.5$, $\delta = 0.3$, $\epsilon = 10^{-4}$. In (a): Projection onto the (x, y_1)-plane. The dashed red line is the critical manifold of (59). The orbit in black is of canard type, following the repelling part of the critical manifold beyond the Hopf point (dashed line) before jumping to the stable limit cycles (these are indicated using max and min in red). In (b): Projection onto the (y_1, y_2)-space, compare with Fig. 16. For $x \sim x_{hom}$, see also (a), the orbit, due to the homoclinic bifurcation in the layer problem, jumps towards the attracting point near $y_1 \approx 2.7$.

5 Outlook

The theory of multiscale ODE systems is by now a mature theory that has already found application in many areas, with neuroscience as the perhaps most successful one. From the author's point of view, there are only a few important, but technical, open problems of the theory. Firstly, the folded node is not fully understood yet. Similarly, some aspects of the folded saddle-node bifurcation where a true saddle collides with the folded node remain unresolved. This bifurcation, known as the folded saddle-node [46, 23], which bears some resembles with the zero-Hopf bifurcation [2], is believed to be associated with Shilnikov chaos [23], but this is still open in general.

From the application point of view, there are important developments in the area of biochemical reaction networks [16]. These systems are typically complex and suffer from parameter uncertainty, yet they often display slow-fast behaviour when simulated, even without knowledge of an explicit small parameter. Also there is typically no way to group the variables into slow and fast variables. In this application area, Quasi-Steady-State (QSS) is used in an ad-hoc way, perhaps based upon simulation results, to reduce the dimension of the system. By now, it is custom to interpret QSS as a slow-fast phenomenon [35]. Recently, see e.g. [20], there has been some work on the development of methods to identify small parameters in polyno-

mial or rational systems that can then be used for dimension reduction and justification of QSS. The basis for this theory is systems of the form (28) having a manifold of equilibria. Similarly, there is also some recent work using ideas from tropical geometry to identify parameter regimes where dimension reduction takes place, see [61]. The potential of this approach for polynomial or rational models is to map out skeleton dynamics in different parameter regimes, see [42] for results on planar systems, but this needs to be developed further.

Acknowledgement. The author thanks Ilona Kosiuk, Panagiotis Kaklamanos and Peter Szmolyan for providing valuable feedback on earlier versions of the manuscript.

References

[1] R. Amir, M. Michaelis, and M. Devor. Burst discharge in primary sensory neurons: triggered by subthreshold oscillations, maintained by depolarizing afterpotentials. *Journal of Neuroscience*, 22(3):1187–1198, 2002.

[2] I. Baldomá, S. Ibáñez, and T.M. Seara. Hopf-zero singularities truly unfold chaos. *Communications in Nonlinear Science and Numerical Simulation*, 84:105162, 2019.

[3] E. Benoit, J. L. Callot, F. Diener, and M. Diener. Chasse au canards. *Collect. Math.*, 31:37â119, 1981.

[4] M. Brøns and K. Bar-Eli. Canard explosion and excitation in a model of the belousov-zhabotinskii reaction. *Journal of Physical Chemistry*, 95(22):8706–8713, 1991.

[5] M. Brøns, T.J. Kaper, and H.G. Rotstein. Introduction to focus issue: Mixed mode oscillations: Experiment, computation, and analysis. *Chaos*, 18(1):015101, 2008.

[6] M. Brøns, M. Krupa, and M. Wechselberger. Mixed mode oscillations due to the generalized canard phenomenon. In W. Nagata and N. Sri Namachchivaya, editors, *Bifurcation Theory and Spatio-Temporal Pattern Formation*, volume 49 of *Fields Institute Communications*, pages 39–64. American Mathematical Society, 2006.

[7] P. T. Cardin and M. A. Teixeira. Fenichel theory for multiple time scale singular perturbation problems. *SIAM Journal on Applied Dynamical Systems*, 16(3):1425–1452, 2017.

[8] P. Carter. Spike-adding canard explosion in a class of square-wave bursters. *Journal of Nonlinear Science*, 30(6):2613–2669, 2020.

[9] M. L. Cartwright and J. E. Littlewood. On non-linear differential equations of the second order: I. *Journal of the London Mathematical Society*, s1-20(3):180–189, 1945.

[10] M. L. Cartwright and J. E. Littlewood. On non-linear differential equations of the second order: II. *Annals of Mathematics*, 48(2):472, 1947.

[11] P. De Maesschalck. Smoothness of transition maps in singular perturbation problems with one fast variable. *Journal of Differential Equations*, 244(6):1448–1466, 2008.

[12] P. De Maesschalck and S. Schecter. The entry-exit function and geometric singular perturbation theory. *Journal of Differential Equations*, 260(8):6697–6715, 2016.

[13] M. Desroches, J. Guckenheimer, B. Krauskopf, H. M. Osinga, C. Kuehn, and M. Wechselberger. Mixed-mode oscillations with multiple time scales. *SIAM Review*, 54(2):211–288, 2012.

[14] F. Dumortier and R. Roussarie. Canard cycles and center manifolds. *Memoirs of the American Mathematical Society*, 121:1–96, 1996.

[15] I. R. Epstein and K. Showalter. Nonlinear chemical dynamics: Oscillations, patterns, and chaos. *Journal of Physical Chemistry*, 100(31):13132–13147, 1996.

[16] M. Feinberg. *Foundations of chemical reaction network theory.* Springer, 2019.

[17] N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21:193–226, 1971.

[18] N. Fenichel. Asymptotic stability with rate conditions. *Indiana University Mathematics Journal*, 23:1109–1137, 1974.

[19] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31:53–98, 1979.

[20] A. Goeke, S. Walcher, and E. Zerz. Determining "small parameters" for quasi-steady state. *Journal of Differential Equations*, 259(3):1149–1180, 2015.

[21] A. Goryachev, P. Strizhak, and R. Kapral. Slow manifold structure and the emergence of mixed-mode oscillations. *Journal of Chemical Physics*, 107(8):2881–2889, 1997.

[22] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer Verlag, 5th edition, 1997.

[23] J. Guckenheimer and I. Lizarraga. Shilnikov homoclinic bifurcation of mixed-mode oscillations. *SIAM Journal on Applied Dynamical Systems*, 14(2):764–786, 2015.

[24] R. Haiduc. Horseshoes in the forced van der pol system. *Nonlinearity*, 22(1):213–237, 2009.

[25] M. G. Hayes, T. J. Kaper, P. Szmolyan, and M. Wechselberger. Geometric desingularization of degenerate singularities in the presence of fast rotation: A new proof of known results for slow passage through Hopf bifurcations. *Indagationes Mathematicae*, 27(5):1184–1203, 2016.

[26] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Bulletin of Mathematical Biology*, 52(1-2):25–71, 1990.

[27] J. L. Hudson, M. Hart, and D. Marinko. An experimental study of multiple peak periodic and nonperiodic oscillations in the Belousov-Zhabotinskii reaction. *Journal of Chemical Physics*, 71(4):1601–1606, 1979.

[28] T.J.R. Hughes. Multiscale phenomena - greens-functions, the dirichlet-to-neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387–401, NOV 1995.

[29] T.J.R. Hughes, G.R. Feijoo, L. Mazzei, and J.B. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24, 1998.

[30] E. M. Izhikevich. *Dynamical Systems in Neuroscience: The geometry of Excitability and Bursting*. The MIT Press, 2007.

[31] S. Jelbart, K. U. Kristiansen, P. Szmolyan, and M. Wechselberger. Singularly perturbed oscillators with exponential nonlinearities. *Journal of Dynamics and Differential Equations*, pages 1–53, 2021.

[32] S. Jelbart, K. U. Kristiansen, and M. Wechselberger. Singularly perturbed boundary-equilibria bifurcations. *Nonlinearity*, 34:7371–7414, 2021.

[33] S. Jelbart, K. U. Kristiansen, and M. Wechselberger. Singularly perturbed boundary-focus bifurcations. *Journal of Differential Equations*, 296:412–492, 2021.

[34] C. K. R. T. Jones. Geometric singular perturbation theory. In *Dynamical systems (Montecatini Terme, 1994)*, volume 1609 of *Lecture Notes in Math.*, pages 44–118. Springer, Berlin, 1995.

[35] H. G. Kaper and T. J. Kaper. Asymptotic analysis of two reduction methods for systems of chemical reactions. *Physica D: Nonlinear Phenomena*, 165(1-2):66–93, 2002.

[36] T. J. Kaper. An introduction to geometric methods and dynamical systems theory for singular perturbation problems. In *Analyzing multiscale phenomena using singular perturbation methods (Baltimore, MD, 1998)*, volume 56 of *Proceedings of Symposia in Applied Mathematics*, pages 85–131. American Mathematical Society, Providence, RI, 1999.

[37] I. Kosiuk and P. Szmolyan. Geometric singular perturbation analysis of an autocatalator model. *Discrete and Continuous Dynamical Systems - Series S*, 2(4):783–806, 2009.

[38] I. Kosiuk and P. Szmolyan. Scaling in singular perturbation problems: Blowing up a relaxation oscillator. *SIAM Journal on Applied Dynamical Systems*, 10(4):1307–1343, 2011.

[39] I. Kosiuk and P. Szmolyan. Geometric analysis of the goldbeter minimal model for the embryonic cell cycle. *Journal of Mathematical Biology*, 72(5):1337–1368, 2016.

[40] K. U. Kristiansen. A stiction oscillator under slowly varying forcing: uncovering small scale phenomena using blowup. *SIAM Journal on Applied Dynamical Systems*, 20(4):2359–2390, 2021.

[41] K. U. Kristiansen. Blowup analysis of a hysteresis model based upon singular perturbations. *arXiv:2202.05027 v(2)*, 2022.

[42] K. U. Kristiansen. On a tropicalization of planar polynomial odes with finitely many structurally stable phase portraits. *arXiv:2305.18002 v(2)*, 2023.

[43] K. U. Kristiansen and P. Szmolyan. Relaxation oscillations in substrate-depletion oscillators close to the nonsmooth limit. *Nonlinearity*, 34(2):1030–1083, 2021.

[44] M. Krupa and P. Szmolyan. Extending geometric singular perturbation theory to nonhyperbolic points - fold and canard points in two dimensions. *SIAM Journal on Mathematical Analysis*, 33(2):286–314, 2001.

[45] M. Krupa and P. Szmolyan. Relaxation oscillation and canard explosion. *Journal of Differential Equations*, 174(2):312–368, 2001.

[46] M. Krupa and M. Wechselberger. Local analysis near a folded saddle-node singularity. *Journal of Differential Equations*, 248(12):2841–2888, 2008.

[47] C. Kuehn. *Multiple Time Scale Dynamics*. Springer-Verlag, Berlin, 2015.

[48] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[49] N. Levinson. Perturbations of discontinuous solutions of nonlinear systems of differential equations. *Proceedings of the National Academy of Sciences of the United States of America*, 33(7):214–218, 1947.

[50] L. H Loomis and S. Sternberg. *Advanced calculus*. World Scientific Publishing Co., 2014.

[51] S.L. McMurran and J.J. Tattersall. The mathematical collaboration of m. l. cartwright and j. e. littlewood. *American Mathematical Monthly*, 103(10):833–845, 1996.

[52] C. Morris and H. Lecar. Voltage oscillations in the barnacle balanus-nubilus giant muscle fiber. *Biophysical Journal*, 35(1):193–214, 1981.

[53] A. Neishtadt. On stability loss delay for dynamical bifurcations. *Discrete and Continuous Dynamical Systems - Series S*, 2(4):897–909, 2009.

[54] A.I. Neishtadt. Persistence of stability loss for dynamical bifurcations .2. *Differential Equations*, 24(2):171–176, 1988.

[55] Jr. O'Malley. *Singular Perturbation Methods for Ordinary Differential Equations*, volume 89. Springer-Verlag,, 1991.

[56] L. Perko. *Differential equations and dynamical systems*, volume 7. Springer-Verlag, Berlin, 1991.

[57] V. Petrov, S.K. Scott, and K. Showalter. Mixed-mode oscillations in chemical-systems. *Journal of Chemical Physics*, 97(9):6191–6198, 1992.

[58] J. G. et al Powers. The weather research and forecasting model overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, 98(8):1717–1738, 2017.

[59] J. Rinzel. A formal classification of bursting mechanisms in excitable systems. *Mathematical Topics in Population Biology, Morphogenesis and Neurosciences*, pages 267–281, 1987.

[60] J. Rubin and M. Wechselberger. The selection of mixed-mode oscillations in a hodgkin-huxley model with multiple timescales. *Chaos*, 18(1):015105, 2008.

[61] S. S. Samal, D. Grigoriev, H. Fröhlich, and O. Radulescu. Analysis of reaction network systems using tropical geometry. In *Computer Algebra in Scientific Computing - 17th International Workshop, CASC 2015, Aachen, Germany, September 14-18, 2015, Proceedings*, pages 424–439, 2015.

[62] R.A. Schmitz, K.R. Graziani, and J.L. Hudson. Experimental-evidence of chaotic states in belousov-zhabotinskii reaction. *Journal of Chemical Physics*, 67(7):3040–3044, 1977.

[63] M.A. Shishkova. Examination of one system of differential equations with higher derivatives having small parameter. *Doklady Akademii Nauk Sssr*, 209(3):576–579, 1973.

[64] S. Smale. Differentiable dynamical systems. *Mathematics of Time*, pages 1–82, 1980.

[65] P. Szmolyan and M. Wechselberger. Canards in $\mathbb{R}^3$. *Journal of Differential Equations*, 177(2):419–453, December 2001.

[66] D. Terman. Chaotic spikes arising from a model of bursting in excitable-membranes. *SiAM Journal on Applied Mathematics*, 51(5):1418–1450, 1991.

[67] A.N. Tikhonov. On the dependence of the solutions of differential equations on a small parameter. *Matematicheskii Sbornik*, 31:575–586, 1948.

[68] P. Tracqui. Mixed-mode oscillation genealogy in a compartmental model of bone-mineral metabolism. *Journal of Nonlinear Science*, 4(1):69–103, 1994.

[69] P. Tracqui. Organizing centers and symbolic dynamic in the study of mixed-mode oscillations generated by models of biological autocatalytic processes. *Acta Biotheoretica*, 42(2-3):147–166, 1994.

[70] B. van der Pol. A theory of the amplitude of free and forced triode vibrations. *Radio Review*, 1(15):754–762, 1920.

[71] B. van der Pol. On "relaxation oscillations.". *Philosophical Magazine*, 2(11):978–992, 1926.

[72] B. van der Pol and J. van der Mark. The heart beat considered as a relaxation oscillation and an electrical model of the heart. *L'onde Electrique*, 7:365–392, 1928.

[73] M. Wechselberger. Existence and bifurcation of canards in $\mathbb{R}^3$ in the case of a folded node. *SIAM Journal on Applied Dynamical Systems*, 4(1):101–139, January 2005.

[74] M. Wechselberger. Geometric singular perturbation theory beyond the standard form. *Frontiers in Applied Dynamical Systems: Reviews and tutorials 6, Springer International Publishing*, 2020.

[75] P. Yu, W. Zhang, and L. M. Wahl. Dynamical analysis and simulation of a 2-dimensional disease model with convex incidence. *Communications in Nonlinear Science and Numerical Simulation*, 37:163–192, 2016.

[76] W. Zhang, L. M. Wahl, and P. Yu. Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. *SIAM Review*, 56(1):127–155, 2014.