

Geometric singular perturbation theory in biological practice

Geertje Hek

Received: 10 August 2008 / Revised: 30 December 2008 / Published online: 5 April 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Geometric singular perturbation theory is a useful tool in the analysis of problems with a clear separation in time scales. It uses invariant manifolds in phase space in order to understand the global structure of the phase space or to construct orbits with desired properties. This paper explains and explores geometric singular perturbation theory and its use in (biological) practice. The three main theorems due to Fenichel are the fundamental tools in the analysis, so the strategy is to state these theorems and explain their significance and applications. The theory is illustrated by many examples.

Mathematics Subject Classification (2000) 34e15 · 34c30 · 34c60 · 92b09

1 Introduction

Systems in nature often evolve on time scales differing several orders of magnitude, or take place on various length scales. When studying such systems, simplifying assumptions may be of great help; if not to understand the full system, then at least to get a first insight in the system's behaviour. In particular, one can make assumptions with respect to processes that have much slower or faster time scales than those of interest. A simplifying assumption may be to assume slow processes to stand still, or fast processes to adjust instantaneously to changing circumstances. The phase space may as well be adjusted to the analysis' needs, for example by 'blowing up' parts in phase space where small length scales are of importance.

The well-known Hodgkin–Huxley equations (Hodgkin and Huxley 1952) serve as an important example. The typical spiking behaviour of its stable periodic orbit has been analysed and understood through simplifications of the types sketched here, see Keener and Sneyd (1998), Rinzel (1985) and Carpenter (1979).

G. Hek (✉)
Universiteit van Amsterdam, Amsterdam, The Netherlands
e-mail: G.M.Hek@uva.nl

If a natural system is modelled by differential equations, such simplifications are often the results of various different scalings, and the different approximations will all have limited validity in space or time. Especially if one is interested in solutions or orbits that themselves depend on more than one time or length scale, matching between the various (partial) solutions is then a next step towards a global understanding.

Around 1980, geometric singular perturbation theory was introduced. This is a geometric approach to problems with a clear separation in time scales. It uses invariant manifolds in phase space in order to understand the global structure of the phase space or to construct orbits with desired properties. For the slow–fast systems of concern in this paper, the foundation of this approach was given by Fenichel (1979). The ideas are based on previous work by Fenichel (1971, 1974, 1977) and, independently, by Hirsch et al. (1977). Since then, the methods have evolved and found their way towards applications, of which many have a biological background. Recent work on the Hodgkin–Huxley equations (Moehlis 2006; Rubin and Wechselberger 2007) again serves as an example.

However, there are several reasons to write an introduction to these methods for a more or less biologically inclined public: (i) in applications, authors often only use part of the theory, (ii) it is not always clear why and how the methods are applied, and (iii) by nature, biological systems involve many time scales and since the complexity of models increases, this kind of approaches can be more and more useful. The goal of this paper is to explain and explore the full geometric singular perturbation theory and its use in (biological) practice. The three main theorems due to Fenichel are the fundamental tools in the analysis, so the strategy will mainly be to explain these theorems and their significance to applications. For a more mathematical introduction, including proofs of the theorems and many references, we refer to the surveys of Jones (1995) and Kaper (1999); for an extensive exposition and proofs to Wiggins (1994).

1.1 Basic set-up and ideas

The basic equations we consider are singularly perturbed systems of ODEs with two different time-scales, that are of the form

$$\begin{aligned}\dot{u} &= f(u, v, \varepsilon), \\ \dot{v} &= \varepsilon g(u, v, \varepsilon),\end{aligned}\tag{1.1}$$

where $\dot{} = \frac{d}{dt}$, $u \in \mathbf{R}^k$ and $v \in \mathbf{R}^l$ with $k, l \geq 1$ in general. Our examples are restricted to the cases $k = 1, 2$ and $l = 1, 2$. The parameter ε is a small parameter ($0 < \varepsilon \ll 1$), which gives the system a singular character. The functions f and g are assumed to be sufficiently smooth. ‘Sufficiently smooth’ here means at least C^1 in u, v and ε . In general, to obtain C^r invariant manifolds, f and g should be C^{r+1} functions of u, v, ε , and the subset \mathcal{M}_0 of $\{f(u, v, 0) = 0\}$ we consider below should be a C^{r+1} submanifold of the phase space \mathbf{R}^{k+l} ; see Fenichel (1979).

With a change of time scale, system (1.1) can be reformulated as

$$\begin{aligned}\varepsilon u' &= f(u, v, \varepsilon), \\ v' &= g(u, v, \varepsilon),\end{aligned}\tag{1.2}$$

with $' = \frac{d}{d\tau}$ and $\tau = \varepsilon t$. The time scale given by t is said to be fast whereas that for τ is slow. Thus (1.1) is called the *fast* system and (1.2) the *slow* system. Both systems are equivalent as long as $\varepsilon \neq 0$. Each of the scalings is naturally associated with a limit as $\varepsilon \rightarrow 0$. These limits are respectively given by

$$\begin{aligned} \dot{u} &= f(u, v, 0), \\ \dot{v} &= 0, \end{aligned} \tag{1.3}$$

and

$$\begin{aligned} 0 &= f(u, v, 0), \\ v' &= g(u, v, 0). \end{aligned} \tag{1.4}$$

The latter is called the *reduced system*. It is a differential algebraic system, that describes the evolution of the slow variable constrained to the set $f(u, v, 0) = 0$. That set is exactly the set of critical points for (1.3). The two limits are two different approximations of the full $\varepsilon > 0$ system, but in either formulation one pays a price. Under (1.3) the flow is defined in \mathbf{R}^{k+l} , but is in fact an l -parameter family of k -dimensional systems. Moreover, the flow under (1.3) on the l -dimensional set $f(u, v, 0) = 0$ is trivial. On the other hand, (1.4) does prescribe a nontrivial flow on $f(u, v, 0) = 0$, but at the same time its validity is limited to only this set. The goal of geometric singular perturbation theory is now to analyse the dynamics of system (1.1) with ε nonzero but small by suitably combining the dynamics of these two limits.

The basic ingredient is illustrated in Fig. 1 and is as follows. Suppose we are given an l -dimensional manifold \mathcal{M}_0 , possibly with boundary, which is contained in the set $f(u, v, 0) = 0$. Suppose it is compact and *normally hyperbolic*, that is, the eigenvalues λ of the Jacobian $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ are uniformly bounded away from the imaginary axis. Then this so-called *critical manifold* persists as a locally invariant *slow manifold* \mathcal{M}_ε of the full problem (1.1) that is $\mathcal{O}(\varepsilon)$ close to \mathcal{M}_0 . The restriction of the flow (1.1) to \mathcal{M}_ε is a small perturbation of the flow of the limiting problem (1.4). Moreover, the stable and unstable manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ of \mathcal{M}_0 (that correspond to eigenvalues λ of $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ with respectively $Re(\lambda) < 0$ and $Re(\lambda) > 0$) persist as manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ too. They lie within $\mathcal{O}(\varepsilon)$ distance of, and are diffeomorphic to, $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ respectively.

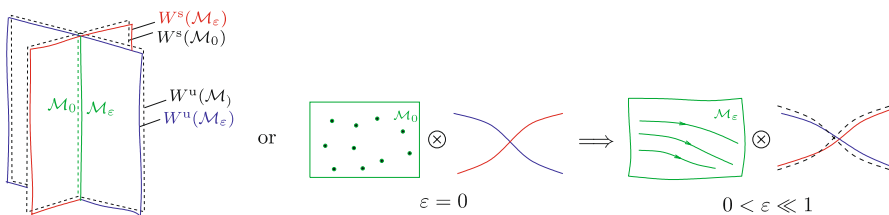


Fig. 1 Unperturbed critical manifolds \mathcal{M}_0 consisting of saddle-type fixed points and their local stable and unstable manifolds $W^{s,u}(\mathcal{M}_0)$. The manifolds persist for $0 < \varepsilon \ll 1$ as perturbed manifolds \mathcal{M}_ε with a slow flow on them, and $W^{s,u}(\mathcal{M}_\varepsilon)$

The problems (1.3) and (1.4) are lower dimensional and can often be analysed in sufficient detail. By ‘gluing’ together fast and slow pieces of orbits, respectively obtained in the fast and slow limits, one can formally construct global *singular* structures, such as singular periodic orbits, and singular homoclinic orbits. Under appropriate conditions one can verify the persistence of these global structures for small $\varepsilon \neq 0$, and reveal their bifurcations as parameters in the system vary. The essential idea behind geometric singular perturbation theory is that this persistence can be established by showing that these singular structures correspond to *transversal intersections* of a pair of stable and unstable manifolds.

We will pay special attention to a subclass of (1.1) that is often met in the literature, and for which geometric singular perturbation theory can be particularly useful. It is the class of *nearly integrable* systems

$$\begin{aligned}\dot{x} &= \frac{\partial H(x, y, v)}{\partial y} + \varepsilon g_1(x, y, v), \\ \dot{y} &= -\frac{\partial H(x, y, v)}{\partial x} + \varepsilon g_2(x, y, v), \\ \dot{v} &= \varepsilon h(x, y, v),\end{aligned}\tag{1.5}$$

in which the variable $v \in \mathbf{R}^n$, $n \geq 1$, may be seen as a slowly varying coefficient in the equations for x and y . For $\varepsilon = 0$ this coefficient is fixed to some v_0 , and (1.5) possesses an integral or Hamiltonian $H(x, y, v_0)$. The almost integrable structure serves as a skeleton on which solutions with certain properties can be constructed. Biological examples of this type of work are, e.g. found in activator-inhibitor (Gierer–Meinhardt type) models (Doelman et al. 2001b; Doelman and van der Ploeg 2002; Doelman and Kaper 2003) or models for mixed mode oscillations in neurons (Krupa et al. 2008a,b). Other examples include models derived from Ginzburg–Landau equations (Doelman and Holmes 1996; Doelman 1996) or nonlinear Schrödinger equations (Rottschäfer and Kaper 2003), models in optics (Kapitula 1998), for gas discharge dynamics (Doelman et al. 2009) or for Josephson junctions (Derks et al. 2003), to name but a few.

Although classical singular perturbation techniques only aim at two different time-scales, in biological literature one also meets systems with *three* different time-scales. These are in general written in the form

$$\begin{aligned}\varepsilon u' &= f(u, v, w, \varepsilon, \delta), \\ v' &= g(u, v, w, \varepsilon, \delta), \\ w' &= \delta h(u, v, w, \varepsilon, \delta),\end{aligned}\tag{1.6}$$

where δ is a second parameter that may or may not depend on ε . To allow straightforward application of singular perturbation techniques, it is very important that ε and δ are two *independent* small parameters ($0 < \varepsilon, \delta \ll 1$). Examples of this type are for instance found in food chain models with a third class of so-called super or top-predators (Muratori and Rinaldi 1992; Deng 2001; Deng and Hines 2002) or in hormone secretion models (Kunpasuruang et al. 2002). Models with three or more

time-scales are also used to study neuronal behaviour, in particular to explain firing of neurons or so-called mixed mode oscillations. Various papers in this field show different ways to deal with, or even to benefit from the time-scale structure. In [Medvedev and Cisternas \(2004\)](#) and [Krupa et al. \(2008a,b\)](#) the three time-scales *are related*, but geometric singular perturbation theory could still be applied after (local) rescalings that brought the system into a new one with only two time-scales. And the classical example of Hodgkin–Huxley was analysed in [Moehlis \(2006\)](#) using that a particularly fast gating variable may be assumed to always be at its saturation value, so that only two time-scales are left.

Remark 1 Although this overview is written for a biological inclined public, it is necessarily mostly of a mathematical nature. However, in every section, the theory is illustrated with examples, of which some are biologically inspired. Others are just chosen because of their mathematical features. In several examples, the biological motivation and interpretation of the results is explicitly clarified and discussed. This is the case in the illustrative Example 1.1 below, as well as in Examples 2.1 and 3.3. In other examples, some of the biological background of the model is described.

Example 1.1 The Hydra, a type of fresh-water polyp, is considered as a model organism for morphogenesis. Small pieces of Hydra can regenerate to a complete new animal: after removing the original head, a new head will be formed, but the orientation of the animal will always be preserved. [Gierer and Meinhardt \(1972\)](#) introduced a mathematical model in order to better understand its growth and regeneration. They realised that the local formation of a head was in fact an example of pattern formation, and they identified the crucial conditions for such pattern formation: local self-activation and long-range inhibition. Hence they modelled the growth of a new head on a piece of Hydra using an activator-inhibitor model; see, e.g. [Murray \(1989\)](#). A slowly diffusing activator stimulates growth of a new head, whereas a rapidly diffusing inhibitor is produced by the activator and prevents the formation of a second head. At a position with a peak in activator concentration, a head will be formed.

In [Doelman et al. \(2001a,b\)](#) and [van der Ploeg \(2005\)](#) a generalised or ‘normal form’ version that no longer models the orientation-preservation is presented:

$$\begin{cases} \varepsilon^2 U_t = U_{xx} - \varepsilon^2 \mu U + U^{\alpha_1} V^{\beta_1}, \\ V_t = \varepsilon^2 V_{xx} - V + U^{\alpha_2} V^{\beta_2}. \end{cases} \quad (1.7)$$

Here $U(x, t)$ and $V(x, t)$ represent inhibitor and activator concentrations, respectively. Although in biological systems the domain is necessarily bounded by dimensions of the tissue, here an unbounded domain is considered: $x \in \mathbb{R}$. This is reflecting the choice to study patterns with a much smaller spatial scale than the length scales of the domain. Apart from this, the scaling of the equations should be mentioned, too: it is suitable for finding *large* pulse solutions. Such solutions correspond to very high activator concentrations on very narrow regions, and almost zero activator concentrations elsewhere. The small parameter $0 < \varepsilon^2 \ll 1$ expresses the ratio of the diffusion rates. Here we will describe the case $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (0, -1, 2, 2)$, corresponding to the classical Gierer–Meinhardt model.

Since the two components have completely different diffusion rates, they also have different natural length scales, which makes the problem suitable for Fenichel theory. Here we focus on the use of Fenichel's methods to construct a periodic, stationary pulse solution to (1.7). In terms of the polyp, such solution would be an infinite sequence of Hydra, each with a single head being formed (or perhaps more realistically, a single Hydra with periodic boundary conditions). For proofs and a detailed construction and stability analysis of this and other periodic and aperiodic patterns we refer to Doelman et al. (2001b). Alternative methods to study similar solutions are presented in Iron et al. (2001) and Ward and Wei (2002, 2003).

For stationary solutions, we set $U_t = V_t = 0$. In (1.7) one then observes that V varies over $\mathcal{O}(\varepsilon)$ length scales, so that its natural spatial coordinate is $\xi = x/\varepsilon$. The inhibitor U varies over much longer scales. Although the independent variables are spatial rather than temporal variables, it is still common to use the terminology 'fast' and 'slow': V is called the 'fast component' and U the 'slow component'. In terms of ξ , (1.7) is written as

$$\begin{cases} U_{\xi\xi} = \varepsilon^4 \mu U - \varepsilon^2 V^2, \\ V_{\xi\xi} = V - \frac{V^2}{U}. \end{cases}$$

By taking $\varepsilon \rightarrow 0$, the equation for U reduces to $U_{\xi\xi} = 0$, with $U = \text{const.}$ as the only possible bounded (and therefore biologically relevant) solutions. Thus we take

$$U(\xi) = U_0 = \text{const.} \quad \text{and} \quad V_{\xi\xi} = V - \frac{V^2}{U_0}. \quad (1.8)$$

as the fast reduced limit. For any fixed inhibitor concentration U_0 , the V -equation now has a homoclinic solution

$$V_h(\xi; U_0) = \frac{3}{2} U_0 \operatorname{sech} \left(\frac{\xi}{2} \right), \quad (1.9)$$

a local peak in activator concentration that decays exponentially to $V = 0$ as $\xi \rightarrow \pm\infty$. At $\mathcal{O}(\frac{1}{\varepsilon})$ -distances in ξ (so $\mathcal{O}(1)$ in x) from the centre of the pulse, V_h is exponentially small in ε , so that it can be approximated by $V = 0$. In these regions the slow reduced limit, describing the limiting inhibitor behaviour,

$$V(x) = 0 \quad \text{and} \quad U_{xx} = \varepsilon^2 \mu U \quad (1.10)$$

gives a leading order approximation for the pulse solution. The idea is now to construct a singular periodic pulse solution as an alternate concatenation of two parts. The first part is approximated by a slow solution $V(x) = 0$, $U(x) = A e^{\varepsilon \sqrt{\mu} x} + B e^{-\varepsilon \sqrt{\mu} x}$ to (1.10) that tends to $U = U_0$ as x tends to some x_0 . The second part is a connecting narrow pulse described by (1.8), with suitably chosen A , B , U_0 .

System (1.7) with $U_t = V_t = 0$ is written as a system of ODEs of the form (1.2) (with reverse roles for u and v though) so that it is suitable for direct application of Fenichel’s theory:

$$\begin{aligned}
 u' &= p, \\
 p' &= -v^2 + \varepsilon^2 \mu u, \\
 \varepsilon v' &= q, \\
 \varepsilon q' &= v - \frac{v^2}{u}
 \end{aligned}
 \tag{1.11}$$

with $' = \frac{d}{dx}$. The set $\{v = q = 0, u > 0\}$ being a set of normally hyperbolic fixed points of the fast (v, q) subsystem, one can take a compact critical manifold $\mathcal{M}_0 \subset \{v = q = 0, u > 0\}$, as large as one wishes. It is invariant under the $\varepsilon = 0$ and the $\varepsilon > 0$ flow, so that the persisting slow manifold \mathcal{M}_ε is in fact equal to \mathcal{M}_0 . Since for $\varepsilon = 0$ every saddle point in \mathcal{M}_0 has 1-dimensional stable and unstable manifolds, the 2-parameter collection of all saddle-points in \mathcal{M}_0 together have a 3D stable manifold $W^s(\mathcal{M}_0)$ and unstable manifold $W^u(\mathcal{M}_0)$. Fenichel’s second theorem now guarantees that these manifolds persist for small nonzero ε , see Sect. 3; the third theorem gives an idea about how individual orbits in these manifolds tend to \mathcal{M}_ε as $x \rightarrow \pm\infty$, see Sect. 6.

The individual stable and unstable manifolds with $u, v > 0$ of each saddle point in \mathcal{M}_0 coincide in a homoclinic orbit $(v = V_h(\xi), q = \frac{dv}{d\xi})$, so that the complete unperturbed manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ coincide for $u, v > 0$. A question now is, whether this is still true for small nonzero ε . Melnikov-type integrals give the answer, as explained in Sect. 5. The intersection $W^s(\mathcal{M}_\varepsilon) \cap W^u(\mathcal{M}_\varepsilon)$ turns out to be a 2D set, which is in fact a 1-parameter family of orbits that are biasymptotic to \mathcal{M}_ε . This is what one would generically expect as the intersection of two 3D manifolds in a 4D phase space. The intersection must consist of orbits, since it is a subset of two manifolds that are both invariant under the flow (1.11).

For $\varepsilon > 0$, the flow on $\mathcal{M}_\varepsilon = \mathcal{M}_0$ is described by (1.10), or $u' = p, p' = \varepsilon^2 \mu u$, and clearly has a saddle structure for $\mu > 0$. The last question is now, whether or not one of the orbits in \mathcal{M}_ε can be connected to one of the ‘surviving’ biasymptotic orbits to \mathcal{M}_ε , to form a skeleton for a one-pulse periodic pattern as in Fig. 2. The take-off and touch-down points of the $\varepsilon > 0$ biasymptotic orbits should be calculated to find the answer, which is yes (see Corollary 9 and Example 6.2).

The big advantage of the geometrical approach sketched here, is that it allows to determine the basic geometrical features leading to certain orbits or patterns. Only using the geometry of the limiting pulse solution (1.9) and of the curves of possible take-off and touch-down points within the slow manifold \mathcal{M}_ε , it was possible to construct a whole plethora of periodic and aperiodic patterns to (1.7) (see Doelman et al. (2001b)). This is not unique for this example: especially in existence proofs for arbitrary periodic and aperiodic solutions, or for chaotic behaviour, the key ingredients are mostly of a geometrical nature. See, e.g. Carpenter (1979) for a classical example in excitable membranes and Deng and Hines (2002) for an example of food chain model chaos.

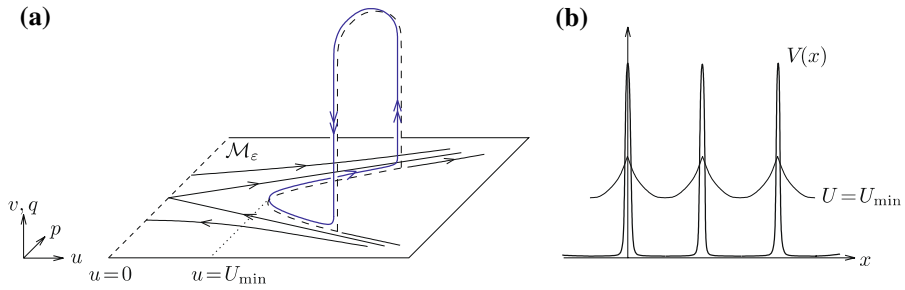


Fig. 2 **a** Schematic illustration of the constructed periodic orbit and the singular skeleton structure on which it is based (*dashed lines*). A *single arrow* means slow flow, a *double arrow* means fast flow. Note that $\mathcal{M}_\varepsilon = \mathcal{M}_0$ can be chosen as large as one wishes, so that it will contain the orbit in $\{v = q = 0\} \supset \mathcal{M}_\varepsilon$ that is part of the singular orbit. **b** Sketch of the corresponding periodic pattern $(U(x), V(x))$

We end this example with a remark that the point $(U, V) = (0, 0)$ is a singularity of the nonlinearities in (1.7). This being fairly nonbiological, it may be seen as a modelling mistake. But anyway, it is possible to control these singularities and extend \mathcal{M} , so that amongst others a pulse-solution that is homoclinic to $(U, V) = (0, 0)$ can be constructed, similar to the one in Example 6.2. See Doelman et al. (2001a) for the details, or Doelman and van der Ploeg (2002) for constructions in two spatial dimensions.

2 Fenichel’s first theorem

In general it is natural to expect that the set $f(u, v, 0) = 0$ of critical points of (1.3) is, at least locally, an l -dimensional manifold in \mathbf{R}^{k+l} , since it is obtained by solving k equations and f was assumed to be sufficiently smooth. If indeed \mathcal{M}_0 is an l -dimensional manifold contained in $f(u, v, 0) = 0$, and \mathcal{M}_0 is normally hyperbolic, then Fenichel’s first theorem says that this manifold persists for small nonzero ε as a manifold \mathcal{M}_ε with a slow flow on it. The theorem is as follows.

Theorem 2 (Fenichel) *Suppose $\mathcal{M}_0 \subset \{f(u, v, 0) = 0\}$ is compact, possibly with boundary, and normally hyperbolic, that is, the eigenvalues λ of the Jacobian $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ all satisfy $\text{Re}(\lambda) \neq 0$. Suppose f and g are smooth. Then for $\varepsilon > 0$ and sufficiently small, there exists a manifold \mathcal{M}_ε , $\mathcal{O}(\varepsilon)$ close and diffeomorphic to \mathcal{M}_0 , that is locally invariant under the flow of the full problem (1.1).*

The statement that \mathcal{M}_ε is locally invariant means in practice that it is ‘invariant’ in the sense that one wishes: orbits on \mathcal{M}_ε can only leave this manifold in the slow direction, via the boundary of \mathcal{M}_ε , and not via the directions ‘perpendicular’ to \mathcal{M}_ε .

Theorem 2 implies that the restriction of the flow (1.1) to \mathcal{M}_ε is a small perturbation of the flow of the limiting problem (1.4). This can be easily seen in the case that \mathcal{M}_0 is given by a graph of a function $p_0(v)$ (which is at least locally true since the Jacobian $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ is invertible by the normal hyperbolicity assumption). If

$$\mathcal{M}_0 = \{(u, v) \in \mathbf{R}^{k+l} \mid u = p_0(v)\}$$

with $v \in C$, $C \subset \mathbf{R}^l$ compact, then the perturbed manifold \mathcal{M}_ε is described by a perturbation p_ε of p_0 :

$$\mathcal{M}_\varepsilon = \{(u, v) \in \mathbf{R}^{k+l} \mid u = p_\varepsilon(v)\}.$$

Substituting this into the slow scaling (1.2), one sees that the flow on \mathcal{M}_ε is given by

$$v' = g(p_\varepsilon(v), v, \varepsilon). \tag{2.1}$$

This is a slow flow, since it is $\mathcal{O}(1)$ with respect to the slow time τ . Moreover, if $\varepsilon \rightarrow 0$ this flow indeed has the limit $v' = g(p_0(v), v, \varepsilon)$.

For more precise statements about the smoothness of \mathcal{M}_ε and p_ε and about local invariance we refer to Fenichel (1979) and Jones (1995).

2.1 Application of Fenichel’s first theorem

A 2-dimensional predator–prey model, that can easily be analysed by standard phase plane analysis, serves as a first illustration of Theorem 2.

Example 2.1 Consider a classical Rosenzweig–MacArthur predator–prey model (May 1972; Shimazu et al. 1972) as in Rinaldi and Muratori (1992) that is in a rescaled form given by

$$\begin{aligned} \dot{u} &= u \left(1 - u - \frac{av}{u+d} \right), \\ \dot{v} &= \varepsilon v \left(\frac{au}{u+d} - 1 \right). \end{aligned} \tag{2.2}$$

Here the numbers of prey u and predators v are both non-negative. They are scaled with the constant predator-free carrying capacity of the prey. The parameter $\varepsilon > 0$ is the ratio between the linear death rate of the predator and the linear growth rate of the prey, and the positive parameters a and d determine the impact of predation on the prey.

If the prey reproduce much faster than the predators, and the predator is rather aggressive but is in comparison not so efficient, then the ratio ε becomes a small parameter, see Rinaldi and Muratori (1992). This means that (2.2) is of the form (1.1), with $f(u, v, \varepsilon) = u(1 - u) - \frac{auv}{u+d}$ and $g(u, v, \varepsilon) = v \left(\frac{u}{u+d} - 1 \right)$. Note that g is not well-defined in $u = -d$. Although $u = -d$ is a nonbiological value this is illustrative: g does not satisfy the smoothness assumptions in Theorem 2. In the $\varepsilon = 0$ limit, the nullcline $\{(u, v) \mid f(u, v, 0) = 0, u \geq 0, v \geq 0\}$ consists of the parts $\mathcal{M}_0^0 := \{(u, v) \mid u = 0, v \geq 0\}$ and $\mathcal{M}_0^1 := \{(u, v) \mid v = \frac{1}{a}(1 - u)(u + d), u, v \geq 0\}$. These critical manifolds consist of all possible prey equilibria in case of a constant,

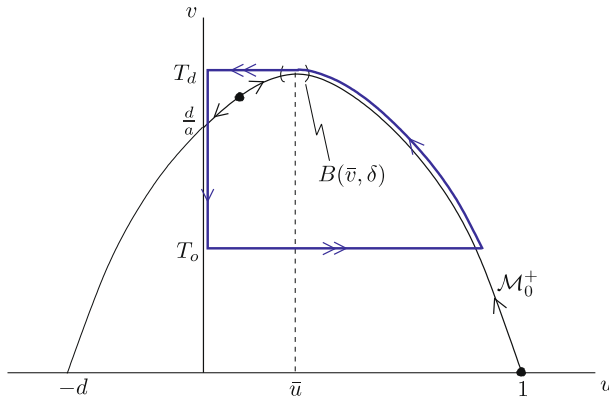


Fig. 3 Configuration of the slow manifolds for system (2.2) with $0 < d < 1$ and $a > (1 + d)/(1 - d)$. The periodic orbit is the one constructed in Rinaldi and Muratori (1992): the touch down point T_d is determined by the v -coordinate of the fold (\bar{u}, \bar{v}) and the take off point T_o is calculated in Rinaldi and Muratori (1992); see Sect. 4

but arbitrary predator population. The manifolds are normally hyperbolic everywhere but in $(0, \frac{d}{a}) \in \mathcal{M}_0^0 \cap \mathcal{M}_0^1$ and $(\bar{u}, \bar{v}) = (\frac{1-d}{2}, \frac{(1+d)^2}{4a}) \in \mathcal{M}_0^1$ for $d < 1$.

On \mathcal{M}_0^1 the flow with respect to time $\tau = \epsilon t$ can be obtained by writing \mathcal{M}_0^1 as a graph of a function in the slow variable v . If $0 < d < 1$, this cannot be done in a global way, so we write \mathcal{M}_0^1 as the union of two hyperbolic parts and a third, small part around the nonhyperbolic fold point (\bar{u}, \bar{v}) , i.e. as $\mathcal{M}_0^1 = \mathcal{M}_0^+ \cup \mathcal{M}_0^- \cup B(\bar{v}, \delta)$. Here $0 < \delta \ll 1$,

$$\mathcal{M}_0^\pm := \left\{ (u, v) \mid u = u_\pm(v) := \frac{1}{2}(1-d \pm \sqrt{(1+d)^2 - 4av}), u \geq 0, 0 \leq v \leq \bar{v} - \frac{\delta}{2} \right\}$$

and $B(\bar{v}, \delta)$ is an open δ -neighbourhood within \mathcal{M}_0^1 of (\bar{u}, \bar{v}) . If $d \geq 1$ we write $\mathcal{M}_0^1 = \mathcal{M}_0^+$, so that the flow can for all $d > 0$ be written as $v' = v \left(\frac{u(a-1)-d}{u_\pm(v)+d} \right)$ on \mathcal{M}_0^\pm .

The flow on \mathcal{M}_0^1 has critical points $v = 0$ (repelling, on the branch \mathcal{M}_0^+) and $v = \frac{d(a-1-d)}{(a-1)^2}$, which is a repeller on \mathcal{M}_0^- if $a > \frac{1+d}{1-d}$ and an attractor on \mathcal{M}_0^+ if $1 < a < \frac{1+d}{1-d}$ (see Fig. 3).

The biological meaning of this flow is the following. Assume that for a constant predator population $v_1 \neq \bar{v}$ the prey is in an equilibrium state $(u_1, v_1) \in \mathcal{M}_0^\pm$. If the size of the predator population now starts to change slowly to a new value v_2 , a new equilibrium $(u_2, v_2) \in \mathcal{M}_0^\pm$ will form, according to the predator–prey interaction on the nullcline \mathcal{M}_0^\pm . Only if v passes the value \bar{v} , this will not be a continuous process.

Theorem 2 now assures that for sufficiently small $\epsilon \neq 0$ the critical manifolds \mathcal{M}_0^\pm persist as perturbed manifolds \mathcal{M}_ϵ^\pm that are invariant for the flow with $\epsilon \neq 0$. The slow flow on \mathcal{M}_ϵ^\pm is approximated by the limiting slow flow $v' = v \left(\frac{u(a-1)-d}{u_\pm(v)+d} \right)$. In terms

of biology, if the predator population is not constant but varies slowly, the dynamics for $0 < \varepsilon \ll 1$ is still predominantly determined by the predator–prey interaction on the perturbed versions $\mathcal{M}_\varepsilon^\pm$ of nullclines \mathcal{M}_0^\pm .

Remark 3 A fold, like (\bar{u}, \bar{v}) in the above example, is nonhyperbolic and therefore needs special attention. In many applications it is not mentioned that Theorem 2 does not apply in such points. Often without any consequences since in general fold points that are important for the dynamics are *jump points*. There the flow jumps off the slow manifold and starts to follow the fast vector field. Such jumping behaviour is part of the mechanics behind the classical relaxation oscillations (see, e.g. Murray 1989). However, it must not be overlooked that less-intuitive behaviour can occur as soon as the slow flow (1.4) on $\{f = 0\}$ has a critical point at or very close to the fold: so-called canards can exist. For a thorough understanding of the (sometimes rather nonintuitive) flow for small nonzero ε , geometric singular perturbation theory can be extended using blow-up or geometric desingularisation techniques, which have been applied successfully in various situations where hyperbolicity is lost. See Dumortier (1993), Dumortier and Roussarie (1996), Krupa and Szmolyan (2001), Popović and Szmolyan (2004), Huber and Szmolyan (2005) and references therein. Clear explanations, including an application to the Hodgkin–Huxley based FitzHugh–Nagumo model (FitzHugh 1960; Nagumo et al. 1962) (see also Murray 1989), are provided in Wechselberger (2005). Earlier successful attempts to understand this behaviour used nonstandard analysis (Bénoît et al. 1981; Diener 1984) or matched asymptotic expansions (Eckhaus 1983).

Example 2.2 If the parameter a in (2.2) is also small and of order $\mathcal{O}(\varepsilon)$, the influence of predation on the prey is small and the system becomes

$$\begin{aligned} \dot{u} &= u(1 - u) - \varepsilon \frac{\tilde{a}uv}{u + d}, \\ \dot{v} &= \varepsilon v \left(\frac{au}{u + d} - 1 \right). \end{aligned} \tag{2.3}$$

In the $\varepsilon = 0$ limit the set $\{(u, v) | f(u, v, 0) = 0, u \geq 0, v \geq 0\}$ consists of the parts $\mathcal{M}_0^0 := \{(u, v) | u = 0, v \geq 0\}$ and $\mathcal{M}_0^1 := \{(u, v) | u = 1, v \geq 0\}$. Both critical manifolds are normally hyperbolic and all conditions in Theorem 2 are satisfied. Thus for $0 < \varepsilon \ll 1$ there are locally invariant slow manifolds $\mathcal{M}_\varepsilon^0$ and $\mathcal{M}_\varepsilon^1$ that are $\mathcal{O}(\varepsilon)$ close and diffeomorphic to \mathcal{M}_0^0 and \mathcal{M}_0^1 , respectively.

Informally, it is immediately clear that $\mathcal{M}_\varepsilon^0 = \mathcal{M}_0^0$, since the set $\{u = 0, v \geq 0\}$ is still locally invariant for $\varepsilon > 0$ (in fact invariant under the forward flow). However, the set $\{u = 1\}$ is no longer invariant for $\varepsilon > 0$, so that $\mathcal{M}_\varepsilon^1 \neq \mathcal{M}_0^1$. Using its invariance, the perturbed manifold $\mathcal{M}_\varepsilon^1$ can be approximated by asymptotic expansions as follows.

Under the assumption that the perturbed manifold $\mathcal{M}_\varepsilon^1$ can be described as a graph $\{(u, v) | u = p_\varepsilon(v), u \geq 0, v \geq 0\}$, this manifold is invariant under the flow of (2.3) if

$$\frac{dp_\varepsilon}{dv} \dot{v} = p_\varepsilon(v)(1 - p_\varepsilon(v)) - \varepsilon \frac{\tilde{a}vp_\varepsilon(v)}{p_\varepsilon(v) + d},$$

so, by filling in \dot{v} , if $\varepsilon \frac{dp_\varepsilon}{dv} v \left(\frac{p_\varepsilon(v)}{p_\varepsilon(v) + d} - 1 \right) = p_\varepsilon(v)(1 - p_\varepsilon(v)) - \varepsilon \frac{\tilde{a}vp_\varepsilon(v)}{p_\varepsilon(v) + d}$.

Now expand

$$p_\varepsilon(v) = p_0(v) + \varepsilon p_1(v) + \varepsilon^2 p_2(v) + \dots,$$

where $p_0(v) = 1$ describes \mathcal{M}_0^1 . Then gathering orders of ε results in

$$\begin{aligned} \mathcal{O}(1) : & \quad 0 = p_0(1 - p_0) \\ \mathcal{O}(\varepsilon) : & \quad 0 = -p_0 p_1 - \frac{\tilde{a}vp_0}{p_0 + d} \\ \mathcal{O}(\varepsilon^2) : & \quad \frac{dp_1}{dv} v \left(\frac{ap_0}{p_0 + d} - 1 \right) = -p_0 p_2 - p_1^2 - \frac{\tilde{a}vp_1}{p_0 + d} + \frac{\tilde{a}vp_0 p_1}{(p_0 + d)^2}, \text{ etc.} \end{aligned}$$

This yields the approximation

$$p_\varepsilon(v) = 1 - \varepsilon \frac{\tilde{a}v}{1 + d} + \varepsilon^2 \left(\frac{\tilde{a}(a - 1 - d)v}{(1 + d)^2} - \frac{\tilde{a}^2 v^2}{(1 + d)^3} \right) + \dots \tag{2.4}$$

Clearly, this approximation is only valid for $v < \mathcal{O}(1/\varepsilon)$, since otherwise the perturbation terms become of order $\mathcal{O}(1)$. Following this same procedure for \mathcal{M}^0 , one would indeed find $\mathcal{M}_\varepsilon^0 = \mathcal{M}_0^0$ up to any order in ε . For $\varepsilon = 0$ the limit (1.4) prescribes a the singular slow flow on \mathcal{M}_0^1 : $v' = \frac{v(a-1-d)}{1+d}$. For sufficiently small nonzero ε , the flow on $\mathcal{M}_\varepsilon^1$ is a perturbation of this flow, that can be approximated by inserting $u = p_\varepsilon(v)$ with $p_\varepsilon(v)$ given by (2.4) into the equation for v' .

The above example is very simple, so that the benefit of an expansion like (2.4) is not immediately clear. For many (biological) problems, the dynamics on \mathcal{M} is however of great importance. Already when the dimensions are slightly higher, e.g. with a 2 or 3-dimensional slow manifold, it can be very useful to describe \mathcal{M}_ε as a graph $\{(u, v) | u = p_\varepsilon(v), u \geq 0, v \geq 0\}$ and find an approximation as in (2.4). In Beck et al. (2006) for instance, a travelling wave is sought as a solution to a bioremediation model. This wave corresponds to the motion of a biologically active zone in which microorganisms consume contaminated soil and an added nutrient. The corresponding phase space is 5-dimensional and comprises a 3-dimensional slow manifold, within which even slower behaviour occurs. For the construction of the wave, it was necessary to find an expansion for \mathcal{M}_ε and even to apply Fenichel theory a second time within \mathcal{M}_ε .

On the other hand, in many cases the $\varepsilon = 0$ limiting flow (1.4) on \mathcal{M}_0 combined with global geometrical properties of the fast limiting system (1.3) suffice to draw conclusions about orbits for $0 < \varepsilon \ll 1$.

2.2 More general notion of slow manifold

It is worth mentioning here, that Fenichel’s original work (Fenichel 1971) covered more general invariant manifolds than only critical manifolds of fixed points of (1.3).

The theory for instance also applies to continuous families of periodic orbits of (1.3), parameterised by the slow variable or parameter v . Such family is often also called a *slow manifold*. In practice it may be a cylindrical manifold \mathcal{P}_0 filled with periodic orbits, that perturbs for $0 < \varepsilon \ll 1$ to a nearby manifold \mathcal{P}_ε with a still almost periodic flow on it, now with a slow drift determined by (2.1).

Bursting phenomena in excitable media (Keener and Sneyd 1998; Rubin and Terman 2002; Terman 1991) are important examples in which this is heavily used: the spiking behaviour results from a solution that winds several times around such manifold \mathcal{P}_ε . The fast ‘jumps’ are inherited from the $\varepsilon = 0$ fast periodic orbits, whereas the slow drift along \mathcal{P}_ε causes the spikes to be separated.

See, e.g. Best et al. (2005) for an application to a pacemaker network, or Higuera et al. (2005) for an application to Faraday waves.

3 Fenichel’s second theorem: stable and unstable manifolds

Fenichel’s first theorem only gives a very local picture of the system (1.1) for small nonzero ε . It guarantees under some conditions the existence of the slow manifold, and gives an approximation for the flow on this slow manifold. We however aim for a more global picture, that in particular addresses the interplay between the slow manifold and the surrounding phase space. In general, such interplay takes place via stable and unstable manifolds. Those are the objects of concern in Fenichel’s second theorem.

It is well-known (see for instance Guckenheimer and Holmes 1983, sec. 1.7 or Wiggins 1990, sec 1.2C) that hyperbolic fixed points of ODEs persist under small perturbations, together with their stable and unstable manifolds. The theorem below states that, likewise, a normally hyperbolic critical manifold \mathcal{M}_0 also persists under a small perturbation *together with its stable and unstable manifolds*.

Consider the Eq. (1.1). Suppose that for $\varepsilon = 0$ the normally hyperbolic critical manifold $\mathcal{M}_0 \subset \{f(u, v, 0) = 0\}$ has an $l + m$ -dimensional stable manifold $W^s(\mathcal{M}_0)$ and an $l + n$ -dimensional unstable manifold $W^u(\mathcal{M}_0)$, with $m + n = k$. In other words, suppose that the Jacobian $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ has m eigenvalues λ with $Re(\lambda) < 0$ and n eigenvalues with $Re(\lambda) > 0$. Then the following holds:

Theorem 4 (Fenichel) *Suppose $\mathcal{M}_0 \subset \{f(u, v, 0) = 0\}$ is compact, possibly with boundary, and normally hyperbolic, and suppose f and g are smooth. Then for $\varepsilon > 0$ and sufficiently small, there exist manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$, that are $\mathcal{O}(\varepsilon)$ close and diffeomorphic to $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$, respectively, and that are locally invariant under the flow of (1.1).*

The manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ are still ‘stable’ and ‘unstable’ manifolds, as the notation already suggests, but in a different sense, as \mathcal{M}_ε is no longer a set of fixed points. They have the property that solutions in $W^s(\mathcal{M}_\varepsilon)$ decay to \mathcal{M}_ε at an exponential rate in forward time, and solutions in $W^u(\mathcal{M}_\varepsilon)$ decay to \mathcal{M}_ε at an exponential rate in backward time (see Sect. 6). However, one should be careful here, since again the *local invariance* implies that the solutions only decay to \mathcal{M}_ε as long as they

stay in a neighbourhood of the *compact, possibly bounded* \mathcal{M}_ε . We refer to the precise formulation and proof of decay estimates by Jones (1995).

The manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ have respective dimensions $l + m$ and $l + n$, so that indeed the stability properties of \mathcal{M}_0 are inherited by \mathcal{M}_ε . When $mn > 0$, Theorem 2 can be concluded from this theorem by taking the intersections of $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$; see Jones (1995).

3.1 Application of Fenichel’s second theorem

In the 2-dimensional Examples 2.1 and 2.2, Theorem 4 in fact only gives information that could as well be deduced by simply looking at the sign of f at either side of the curves $f = 0$. For $d < 1$, $\varepsilon = 0$ the 1-dimensional critical manifolds \mathcal{M}_0^+ is attracting, with 2-dimensional stable manifold $W^s(\mathcal{M}_0^+)$, and \mathcal{M}_0^- is repelling, with 2-dimensional unstable manifold $W^u(\mathcal{M}_0^-)$. By Theorems 2 and 4 a normally hyperbolic, attracting (repelling) manifold is perturbed to an attracting (repelling) manifold invariant for the flow with $\varepsilon > 0$. More precisely, Theorem 4 states that, for $\varepsilon > 0$ and sufficiently small, the manifolds $W^{u,s}(\mathcal{M}_0^\pm)$ persist as $\mathcal{O}(\varepsilon)$ perturbed manifolds $W^s(\mathcal{M}_\varepsilon^+)$ and $W^u(\mathcal{M}_\varepsilon^-)$.

However, since the phase space is also 2-dimensional, the manifolds are of codimension-0 and can therefore not really change position: the dimensions guarantee that for $\varepsilon > 0$ the stable and unstable manifolds $W^s(\mathcal{M}_\varepsilon^+)$ and $W^u(\mathcal{M}_\varepsilon^-)$ still coincide for $\frac{b}{a} < v < \bar{v}$, as was the case for $\varepsilon = 0$.

In higher dimensions there is of course more freedom. In the sequel we sketch what can be concluded by Theorems 2 and 4 about manifolds with nonzero codimension, and address some questions that remain open.

The first two examples are not directly modelling a (biological) system, but are however related to the Ginzburg–Landau equation that is used in a whole plethora of applications; see Aranson and Kramer (2002) and Mielke (2002). As described in Doelman and Holmes (1996), use of a travelling wave Ansatz in seeking solutions to that equation, followed by a reduction based on $SO(2)$ (phase) equivariance, leads to the study of the following 3-dimensional ODEs.

Example 3.1 Consider the 3-dimensional problem

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= x - x^2 + f_i(x, y, z; \varepsilon, \mu), \\ \dot{z} &= \varepsilon g(x, y, z; \varepsilon, \mu), \end{aligned} \tag{3.1}$$

where μ denotes a vector of $\mathcal{O}(1)$ parameters and $0 \leq \varepsilon \ll 1$. We make the explicit choices $\mu = (a, d)$ with $d > 0$ and $g = 1 - d^2 z^2$ with either $f_1 = y(\varepsilon + az)$ or $f_2 = y(z^2 + \varepsilon z - a)$ and are interested in what happens in the part of the phase space that is bounded by the invariant planes $z = \pm \frac{1}{d}$. The system is of the form (1.1) with $u = (x, y)$, $v = z$. For $\varepsilon = 0$ both systems contain two manifolds of equilibria, $\mathcal{M}_0 = \{x = y = 0, |z| \leq \frac{1}{d}\}$ and $\mathcal{N}_0 = \{x = 1, y = 0, |z| \leq \frac{1}{d}\}$. We focus on \mathcal{M}_0 , which is normally hyperbolic. According to Theorem 2 the compact \mathcal{M}_0 persists as

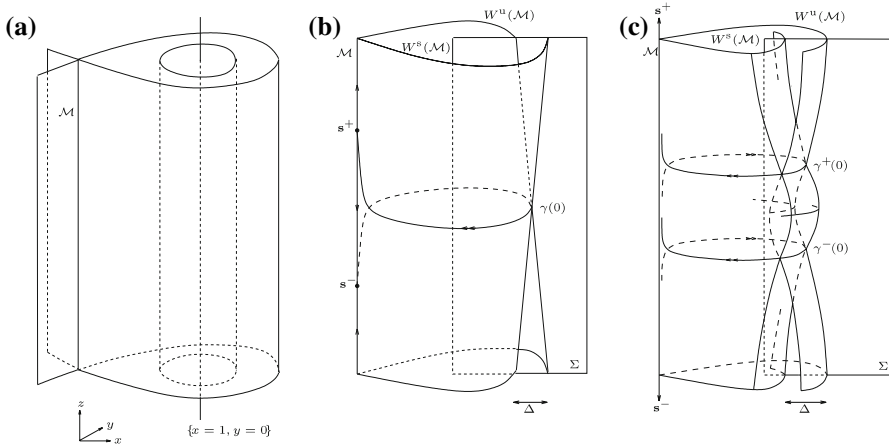


Fig. 4 Phase portraits illustrating Examples 3.1, 3.2 and 5.1, with $\mathcal{M} = \mathcal{M}_0$ or $\mathcal{M} = \mathcal{M}_\varepsilon$ as appropriate. The distances Δ are $\mathcal{O}(1)$ for system (3.1), whereas they are $\mathcal{O}(\varepsilon)$ for system (3.2) or (5.6)

a normally hyperbolic slow manifold \mathcal{M}_ε that is invariant under the flow of (3.1). In theory this manifold is $\mathcal{O}(\varepsilon)$ close to \mathcal{M}_0^1 , but again the set $\{x = y = 0, |z| \leq \frac{1}{d}\}$ is still invariant under the flow of (3.1) for nonzero ε , so that $\mathcal{M}_\varepsilon = \mathcal{M}_0$. The flow on \mathcal{M}_ε is given by $\dot{z} = \varepsilon(1 - d^2z^2)$. The matrix $\frac{\partial f}{\partial u}(u, v, \varepsilon = 0)|_{\mathcal{M}_0}$ has eigenvalues $\lambda_\pm(z, a)$ that satisfy $\lambda_- < 0 < \lambda_+$, so \mathcal{M}_0 has 2-dimensional stable and unstable manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$.

The simple form of (3.1) permits to analyse the global behaviour of $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ by studying the $\varepsilon = 0$ planar subsystems in planes $\{z = \bar{z}\}$ with $|\bar{z}| \leq \frac{1}{d}$. We are interested in those pieces of $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ that lie within the region $x > 0$.

Depending on the sign of $f_i(x, y, \bar{z}; 0, a)$, a planar subsystem is damped or forced. When $f_i(x, y, \bar{z}; 0, a) = 0$, the subsystem contains a homoclinic orbit connecting the point $(0, 0, \bar{z})$ to itself. In terms of the manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ this yields the following.

1. For $f_1 = y(\varepsilon + az)$, $\varepsilon = 0$ and $a \neq 0$ the manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ intersect in the plane $z = 0$ and do not intersect outside this plane. The relative positions of $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ depend on the sign of a , and manifolds intersect in a transversal way (the angle between them can for instance be measured in the plane $\Sigma := \{x > 1, y = 0\}$; see Fig. 4b). If $a = 0$, then the manifolds coincide, as in Fig. 4a.
2. For $f_2 = y(z^2 + \varepsilon z - a)$, $\varepsilon = 0$ and $0 < a < \frac{1}{d^2}$ the manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ intersect transversally in the planes $z = \pm\sqrt{a}$ and do not intersect outside these planes. For $a < 0$ the manifolds do not intersect and $W^u(\mathcal{M}_0)$ winds around $W^s(\mathcal{M}_0)$. Finally, for $a = 0$ the manifolds are tangent in the plane $z = 0$.

By Theorem 4 the manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ persist as perturbed manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$, that are $\mathcal{O}(\varepsilon)$ close and diffeomorphic (but not equal) to $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$. Taking the flow on \mathcal{M}_ε into consideration, we see that all

orbits in $W^s(\mathcal{M}_\varepsilon)$, except the one in the plane $z = -\frac{1}{d}$, lie in fact in the stable manifold $W^s(0, 0, \frac{1}{d})$, and all orbits in $W^u(\mathcal{M}_\varepsilon)$ except the one in $z = \frac{1}{d}$ lie in $W^u(0, 0, -\frac{1}{d})$. The two remaining orbits are the (strong) stable manifold of $(0, 0, -\frac{1}{d})$ and unstable manifold of $(0, 0, \frac{1}{d})$. Concerning the intersections of $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ we conclude:

1. Since the intersection for $a \neq 0$ between $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ at $z = 0$ was transversal and the perturbed manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ are $\mathcal{O}(\varepsilon)$ close and diffeomorphic to $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$, one can use the implicit function theorem to conclude that the manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ still intersect precisely once, in a transversal way, for small enough nonzero ε . The intersection must necessarily be a heteroclinic orbit $\gamma(t)$ connecting the fixed points $s^\pm := (0, 0, \pm\frac{1}{d})$. If $a = 0$, no conclusions can be drawn solely from Fenichel’s theorem.
2. For $a = \mathcal{O}(\varepsilon^2) > 0$, the zeroes of the unperturbed function f_2 are only $\mathcal{O}(\varepsilon)$ apart, so the implicit function theorem cannot guarantee persistence of the intersections $W^s(\mathcal{M}_0) \cap W^u(\mathcal{M}_0)$ under $\mathcal{O}(\varepsilon)$ perturbations of f_2 . For $a = \mathcal{O}(\varepsilon^2) < 0$, one can neither conclude that there are no intersections $W^s(\mathcal{M}_\varepsilon) \cap W^u(\mathcal{M}_\varepsilon)$ for small $\varepsilon > 0$.

For larger $0 < a < \frac{1}{d^2}$, persistence can however be concluded for sufficiently small $\varepsilon > 0$: there are two transversal intersections $W^s(\mathcal{M}_\varepsilon) \cap W^u(\mathcal{M}_\varepsilon)$ that are both heteroclinic orbits $\gamma^\pm(t)$ connecting the fixed points s^\pm , as illustrated in Fig. 4c. For larger negative a there are still no intersections. Somewhere between these cases there must necessarily be a transitional situation at some $a = a^*$, for which the manifolds are tangent in one single heteroclinic orbit.

Example 3.2 Consider again (3.1), but now with a function f_i that is of order $\mathcal{O}(\varepsilon)$, which makes it in a sense even simpler:

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= x - x^2 + \varepsilon f(x, y, z; \mu), \\ \dot{z} &= \varepsilon g(x, y, z; \mu), \end{aligned} \tag{3.2}$$

with μ again a vector of $\mathcal{O}(1)$ parameters. It is of the form (1.5): for $\varepsilon = 0$ Eq. (3.2) is completely integrable, and solutions lie on the intersections of planes $z = z_0 = \text{const.}$ with level sets of the function

$$H(x, y) = \frac{y^2}{2} - \frac{x^2}{2} + \frac{x^3}{3}, \tag{3.3}$$

which is the Hamiltonian for the first two components of (3.2) with $\varepsilon = 0$; see Fig. 4a.

Independent of the choice of f and g , (3.2) has for $\varepsilon = 0$ a 1-parameter family of homoclinic orbits

$$x_0(t) = \frac{3}{2} \operatorname{sech}^2\left(\frac{t}{2}\right), \quad y_0(t) = \dot{x}_0(t), \quad z = z_0 \tag{3.4}$$

to fixed (saddle) points on $\mathcal{M}_0 = \{x = y = 0\}$. This normally hyperbolic invariant manifold persists for sufficiently small ε , along with its stable and unstable manifolds $W^{s,u}(\mathcal{M}_0)$, by Theorems 2 and 4. The structure of (3.2) again implies that $\mathcal{M}_\varepsilon = \mathcal{M}_0$. However, no matter what choices for f and g are made, nothing may be concluded about the structure of $W^{s,u}(\mathcal{M}_\varepsilon)$ and possible intersections solely by application of Fenichel’s second theorem. Conclusions *can* be drawn via adiabatic Melnikov theory (Robinson 1983; Wiggins 1990; Guckenheimer and Holmes 1983), see Sect. 5.

Example 3.3 Consider a tritrophic variant of the Rosenzweig–MacArthur food chain model (Rosenzweig and MacArthur 1963) as proposed by Deng (2001). The food chain model is of the type (1.6) and comprises a logistic prey x , a Holling type II predator y and a Holling type II top-predator z (Holling 1959).

$$\begin{aligned} \varepsilon \frac{dx}{d\tau_1} &= x \left(1 - x - \frac{y}{\beta_1 + x} \right) := xf(x, y), \\ \frac{dy}{d\tau_1} &= y \left(\frac{x}{\beta_1 + x} - \delta_1 - \frac{z}{\beta_2 + y} \right) := yg(x, y, z), \\ \frac{dz}{d\tau_1} &= \delta z \left(\frac{y}{\beta_2 + y} - \delta_2 \right) := \delta zh(y), \end{aligned} \tag{3.5}$$

where ε and δ are two *independent* small parameters ($0 < \varepsilon, \delta \ll 1$), reflecting a drastic trophic time diversification. We do not intend to study this system thoroughly, but only to demonstrate the use of Fenichel’s first two theorems in this example. For a much completer analysis, including existence proofs for various types of chaotic behaviour in both the $0 < \varepsilon, \delta \ll 1$ regime and in a regime with larger δ , we refer to Deng (2001) and Deng and Hines (2002, 2003). Muratori and Rinaldi (1992) essentially studied the same model, but they imposed the three time-scales by inserting small parameters in appropriate places, whereas Deng really rescaled the system so that it is clear what the assumption $0 < \varepsilon, \delta \ll 1$ means for the original parameters in Rosenzweig and MacArthur (1963).

Dynamics in system (3.5) occur at three different time-scales t , $\tau_1 = \varepsilon t$ and $\tau_2 = \varepsilon \delta t$. The system (3.5) is written with respect to the *intermediate* time-scale τ_1 . Evolution on time-scales t and τ_2 is respectively called *fast* and *slow*.

Without any predefined relation between ε and δ , it is *a priori* clear that $0 < \varepsilon \delta < \varepsilon$, so that x is the fast variable, y the intermediate one, and z the slow one. But one cannot analyse the system (3.5) by asymptotic expansions in the small parameters. For example, one does not know whether $\delta^2 < \varepsilon$ or not. The system is however suitable for application of geometric singular perturbation techniques. The system above may as long as $\varepsilon, \delta \neq 0$ equivalently be written as a fast or a slow system:

$$\begin{aligned} \frac{dx}{dt} &= xf(x, y), & \varepsilon \delta \frac{dx}{d\tau_2} &= xf(x, y), \\ \frac{dy}{dt} &= \varepsilon yg(x, y, z), & \delta \frac{dy}{d\tau_2} &= yg(x, y, z), \\ \frac{dz}{dt} &= \varepsilon \delta zh(y), & \frac{dz}{d\tau_2} &= zh(y). \end{aligned} \tag{3.6}$$

Putting either one of the small parameters equal to zero, while keeping the other fixed but nonzero, two different sets of limiting equations are obtained, both fitting immediately in the framework of Theorems 2 and 4. The systems

$$\begin{aligned} \varepsilon = 0 : \quad & \frac{dx}{dt} = xf(x, y), & 0 = xf(x, y), \\ & \frac{dy}{dt} = 0, & \text{(fast) and } \frac{dy}{d\tau_1} = yg(x, y, z), \text{ (interm.)} \\ & \frac{dz}{dt} = 0, & \frac{dz}{d\tau_1} = \delta zh(y), \end{aligned} \tag{3.7}$$

respectively, describe the limiting fast flow towards or away from the equilibrium surface $\{xf(x, y) = 0\}$ and the limiting interaction between top-predators and predators on this surface. The systems

$$\begin{aligned} \varepsilon \frac{dx}{d\tau_1} = xf(x, y), & \quad 0 = xf(x, y), \\ \delta = 0 : \quad \frac{dy}{d\tau_1} = yg(x, y, z), & \text{(interm.) and } 0 = yg(x, y, z), \text{ (slow)} \\ \frac{dz}{d\tau_1} = 0, & \quad \frac{dz}{d\tau_2} = zh(y) \end{aligned} \tag{3.8}$$

describe the predator–prey interaction for a constant top-predator population, with equilibrium curves $\{xf(x, y) = 0, yg(x, y, z) = 0\}$, and the limiting dynamics of the top-predator on these curves. When both $\varepsilon, \delta \rightarrow 0$, the two intermediate limits become the same limit that is naturally associated to the intermediate scaling in (3.5).

The dynamics of (3.5) is predominantly determined by the interaction of predators and top-predators on the (attracting parts) of the nullclines $\{x = 0\}$ and $\{f = 0\}$. The $\varepsilon = 0$ limits (3.7) are most suitable to study both the fast motion to (or from) these nullclines and the interaction on them. Therefore we start our analysis with the $\varepsilon = 0$ fast limiting system, in which the sets $\{x = 0\}$ and $\{f = 0\}$ are 2-dimensional critical manifolds. Both are normally hyperbolic everywhere but in their intersection line $\{x = 0, y = \beta_1\}$ and in the fold $\{x = \frac{1-\beta_1}{2}, y = \frac{(1+\beta_1)^2}{4}\}$ on $\{f = 0\}$.

To study the dynamics in the positive octant, Theorems 2 and 4 can be applied to compact pieces $\mathcal{M}_0^\pm \subset \{f = 0\}$ and $\mathcal{N}_0^\pm \subset \{x = 0\}$, that avoid the non-hyperbolic lines above, and can be taken as large as one wants. Take for instance $\mathcal{M}_0^- := \{f = 0, \Delta \leq x \leq \frac{1-\beta_1}{2} - \Delta, 0 \leq z \leq M\}$, $\mathcal{M}_0^+ := \{f = 0, \frac{1-\beta_1}{2} + \Delta \leq x \leq 1, 0 \leq z \leq M\}$ and $\mathcal{N}_0^- := \{x = 0, 0 \leq y \leq \beta_1 - \Delta, 0 \leq z \leq M\}$, $\mathcal{N}_0^+ := \{x = 0, \beta_1 + \Delta \leq y \leq M, 0 \leq z \leq M\}$, where Δ is a small arbitrary constant and M a large one, independent of ε (see Fig. 5a).

The intermediate manifolds \mathcal{M}_0^+ and \mathcal{N}_0^+ are attracting with 3D stable manifolds, and \mathcal{M}_0^- and \mathcal{N}_0^- are repelling with 3D unstable manifolds. Theorems 2 and 4 imply that for small $\varepsilon > 0$ there exist manifolds $\mathcal{M}_\varepsilon^\pm$ and $\mathcal{N}_\varepsilon^\pm$, $\mathcal{O}(\varepsilon)$ close to \mathcal{M}_0^\pm and \mathcal{N}_0^\pm , again with 3D (un)stable manifolds. This gives the opportunity to

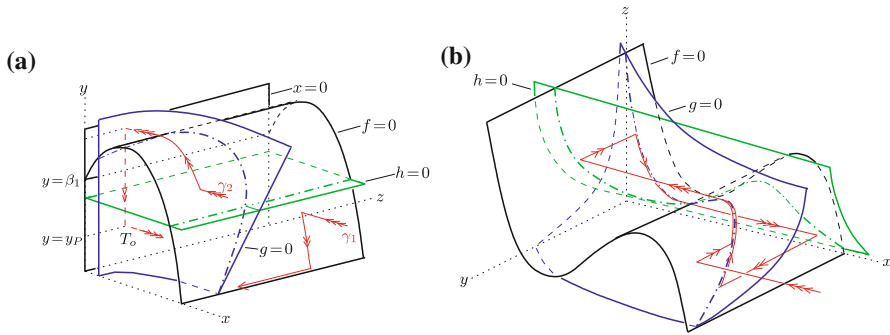


Fig. 5 The configuration of equilibrium manifolds determines the flow in two systems with three time-scales. *Three arrows* indicate a fast trajectory, *two arrows* an intermediate one, and a *single arrow* means a slow trajectory. **a** The situation for (3.5) for certain parameter regions. The orbit γ_2 keeps following the intermediate manifold $\{x = 0\}$ after it loses stability at $y = \beta_1$. The jump position $y = y_P$ is calculated using Pontryagin’s delayed loss of stability. **b** The configuration for (1.6) in Kunpasuruang et al. (2002), with a limit cycle that is totally determined by the three manifolds and their stability types

construct singular orbits for $\varepsilon = 0$, consisting of fast parts off $\mathcal{M}_0^\pm, \mathcal{N}_0^\pm$, and slow or intermediate parts on $\mathcal{M}_0^\pm, \mathcal{N}_0^\pm$. The fast parts are described by the fast limit in (3.7), the parts on $\mathcal{M}_0^\pm, \mathcal{N}_0^\pm$ by the intermediate limit in (3.7), which is again a singularly perturbed system since $0 < \delta \ll 1$. As soon as an $\varepsilon = 0$ fast orbit reaches \mathcal{M}_0^+ or \mathcal{N}_0^+ , the intermediate dynamics on either one of these manifolds, determined by the equilibrium states $y = 0$ and $g = 0$, comes into play. To proceed with the construction of a singular orbit, now the limit $\delta = 0$ can be considered, and Fenichel’s theorems can be applied to the flow restricted to \mathcal{M}_0^+ or \mathcal{N}_0^+ . Everywhere but in their intersection point, the curves $\{f = 0, g = 0\}$ and $\{f = 0, y = 0\}$ are normally hyperbolic for the $\delta = 0$ flow restricted to $\{f = 0\}$. For this restricted flow, $\{f = 0, g = 0\}$ is repelling and $\{f = 0, y = 0\}$ is attracting.

Now, if an orbit of interest hits \mathcal{M}_0^+ , it will either tend to $\{f = 0, y = 0\}$ and then follow the slow flow along that line, or to the fold line $\{x = \frac{1-\beta_1}{2}, y = \frac{(1+\beta_1)^2}{4}\}$, where \mathcal{M}_0^+ loses its normal hyperbolicity. See orbits γ_1 and γ_2 in Fig. 5a, respectively. Likewise, if an orbit hits \mathcal{N}_0^+ , it will tend to $\{x = 0, y = \beta_1\}$ where \mathcal{N}_0^+ loses its normal hyperbolicity.

From a biological modelling and analysis point of view, the following can be concluded from this example. Firstly, it shows that Fenichel’s Theorems 2 and 4 provide a solid basis for the nullcline analysis that is commonly used to understand the dynamics in this kind of systems. Secondly, the normal-hyperbolicity criterium tells when and where such analysis should not be used. But thirdly, it becomes clear that for most (biologically interesting) orbits, more than just Theorems 2 and 4 is needed to know their fate: Unless the parameters in (3.5) are such that $h\left(y = \frac{(1+\beta_1)^2}{4}\right) = 0$, the fold of $\{f = 0\}$ consists of jump points only, so that the flow jumps off \mathcal{M}_0^+ as soon as it reaches the fold, and starts to follow the fast vector field until it hits \mathcal{N}_0^+ . See Remark 3. When the intermediate part of an orbit reaches the intersection (or transcritical) line

$\{x = 0, y = \beta_1\}$, it will keep following $\{x = 0\}$ although that manifold has become unstable as soon as $y < \beta_1$. See again orbit γ_2 in Fig. 5a. This is related to Pontryagin's delayed loss of stability and will be shortly discussed in Sect. 4.3.

Remark 5 In Kunpasuruang et al. (2002) a hormone secretion model of the type (1.6) is presented. It is proved that in some parameter regimes the configuration of the equilibrium manifolds $f = 0$, $g = 0$ and $h = 0$ is as in Fig. 5b. The manifold $f = 0$ is bistable, $g = 0$ is attracting, and $h = 0$ is repelling, so that a construction as above in Example 3.3 immediately yields the illustrated limit cycle.

Remark 6 In Muratori and Rinaldi (1992) many figures illustrate possible singular orbits that can be constructed for (3.5), depending on the parameters and initial conditions. Although they are very clear, it may be confusing that in that paper $\{f = 0\}$ is called the fast manifold and $\{g = 0\}$ is called the intermediate manifold. The names chosen above in Example 3.3 better reflect the standard name 'slow manifold': on $\{f = 0\}$ the intermediate time-scale is dominating, hence the name *intermediate manifold*; on $\{g = 0\}$ the flow is slow, hence its name *slow manifold*.

4 Construction of global orbits

As we saw, Fenichel's first two theorems give information about the nature and position of certain lower-dimensional submanifolds of the phase space. In certain situations this information in itself is very useful to understand the dynamics of systems of type (1.1). This is for instance the case if the critical manifold \mathcal{M} is attracting so that the dynamics will after a certain transition time settle on \mathcal{M} . Especially when the dimensions are high, one can imagine that such knowledge can help to reduce the number of 'important' dimensions tremendously: the system can be understood by just studying the flow on \mathcal{M} (Gear et al. 2005). In general however, the orbits that are of (biological) interest are of a different nature. One can think of periodic solutions to (1.1) that consist of repeated fast 'spikes' with slow rest phases in between. The corresponding orbits in (u, v) -phase space do not live solely on a single slow manifold \mathcal{M} and its stable and unstable manifolds, but can however often be constructed using Fenichel's theorems together with additional information. The next two sections are dedicated to two useful tools to gain such additional knowledge: (i) Melnikov methods to compute intersections of manifolds, and (ii) Fenichel's third theorem to compute so-called base points on a slow manifold.

Here we briefly mention a couple of other geometrical and analytical tools that are related to Fenichel's theory and to the examples we present. Some other important methods are discussed in Sect. 7.

4.1 Fenichel normal form

The Fenichel normal form is particularly useful when estimates of the flow near a slow manifold \mathcal{M} are needed. If one for instance aims to construct a Poincaré mapping for the flow in order to prove existence of a periodic orbit that passes close to

\mathcal{M} , then this mapping will probably be a composition of two or more maps: one that describes the essentially linear flow near \mathcal{M} and one that describes the global return behaviour imposed by, e.g. an $\varepsilon = 0$ periodic orbit. See various constructions in Guckenheimer and Holmes (1983), Kuznetsov (1995), or Terman (1991) for an example.

The flow near a normally hyperbolic manifold \mathcal{M} is necessarily almost linear, and can within a certain neighbourhood be linearised by straightening out \mathcal{M} and its stable and unstable manifolds. The manifolds can then be transformed to coordinate planes. The stable and unstable manifolds consist of Fenichel fibers (see Sect. 6) which in turn may be ‘normalised’ in some sense, too. See Jones (1995) for a very clear explanation.

4.2 Transversally intersecting manifolds

Suppose that one aims to construct a solution to (1.1) of which the corresponding $\varepsilon = 0$ limiting singular orbit consists of several slow and fast pieces, as in Example 1.1. One says that such singular orbit persists for small nonzero ε , if there is an ε_0 such that the construction is valid for any $0 < \varepsilon < \varepsilon_0$. Transversality will often be part of a persistence proof: whenever intersecting manifolds are used in the construction, it is necessary that their mutual intersections are transversal, so that the Implicit Function Theorem implies persistence of the intersections under small perturbations. Intuitively it is clear why this is important: the manifolds used are only known up to $\mathcal{O}(\varepsilon)$ from Fenichel’s theorems, and indeed one obtains validity of the construction for an open ε -set and not just for a single small ε .

We will not discuss this issue in more detail here, but refer to Examples 5.1, 6.1 and 6.2 for applications. Another good example is the construction of the FitzHugh–Nagumo pulse in Jones (1995). There the speed c of the pulse is considered as a fourth variable, and this allows a completely geometrical construction that includes the determination of jump positions at the slow manifolds. The validity of the construction heavily builds on transversal intersections of manifolds.

4.3 Calculation of time spent near a slow manifold

If part of an orbit is a subset of the stable manifold $W^s(\mathcal{M})$ of a slow manifold \mathcal{M} , then in forward time, the orbit will either get exponentially close to the manifold and never leave its neighbourhood again, or it will finally leave \mathcal{M} via its boundary or a nonhyperbolic subset. In predator–prey models with logistic prey, the latter often occurs at the trivial ‘no prey’ equilibrium manifold. See the Examples 2.1 and 3.3 and the corresponding Figs. 3 and 5. In Example 2.1, the manifold $\{u = 0\}$ is stable above, and unstable below the intersection or transcritical point $v = \frac{d}{a}$. In an ε -neighbourhood of $(0, \frac{d}{a})$, the derivative \dot{u} of the *fast* variable u is however $\mathcal{O}(\varepsilon)$, whereas the derivative \dot{v} of the *slow* variable is $\mathcal{O}(1)$, so that an orbit that passes $(0, \frac{d}{a})$ will first continue in a predominantly downward direction until \dot{u} becomes $\mathcal{O}(1)$ and \dot{v} becomes $\mathcal{O}(\varepsilon)$. This is called Pontryagin’s delayed loss of stability. The singular limiting orbit jumps off $\{u = 0\}$ at a take off point T_o that is closely related to the touch down point T_d : since the flow near $\{u = 0\}$ is almost linear, both

above and below $v = \frac{d}{a}$, the amount of time spent to get exponentially close to $\{u = 0\}$ determines the amount of time to leave the neighbourhood of $\{u = 0\}$ again. T_o and T_d are related by an integral, see [Rinaldi and Muratori \(1992\)](#).

Similarly, in [Example 3.3](#), the intermediate manifold $\{x = 0\}$ loses stability at $y = \beta_1$. The orbit γ_2 keeps following this intermediate manifold until the fast flow becomes dominant again at $y = y_P$. The Pontryagin value y_P is calculated in [Deng \(2001\)](#).

In the above examples, the slow manifolds are either fully stable, or fully unstable, so that the (local) phase space is completely covered by their stable or unstable manifold. If a slow manifold \mathcal{M} is of saddle type though, then most orbits in phase space do not lie on its stable or unstable manifolds $W^{s,u}(\mathcal{M})$. Still, it is important to know these manifolds, since any orbit that is $\mathcal{O}(\varepsilon)$ close to a stable manifold $W^s(\mathcal{M}_0)$ does in general not limit on \mathcal{M} , but will reach an $\mathcal{O}(\varepsilon)$ neighbourhood of \mathcal{M} within finite time. To know the fate of orbits and manifolds as they pass *near* such slow manifold, geometric methods exist that do not depend on a calculation like Pontryagin’s one. See the discussion in [Sect. 7](#) and the FitzHugh–Nagumo example in [Jones \(1995\)](#).

5 Intersections of manifolds: Melnikov methods

Above we treated two examples in which the problem (1.1) of our interest is a small perturbation of an integrable system that contains a *homoclinic manifold* in the limiting case (1.3), consisting of a continuous family of homoclinic orbits. The perturbed system ($\varepsilon > 0$) no longer contains such a manifold, but one or more homoclinic orbits may ‘survive’ the perturbation. Melnikov introduced an integral to calculate whether such surviving homoclinic orbits exist ([Melnikov 1963](#)). Classically, the ideas concern 2-dimensional systems

$$\dot{u} = f(u) + \varepsilon g(u, t),$$

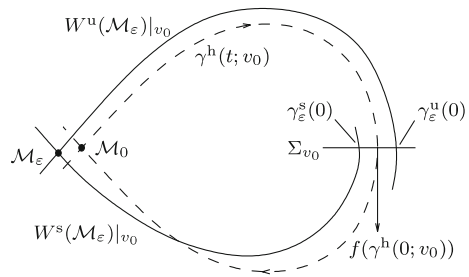
where $\dot{} = \frac{d}{dt}$ and g is a periodic function in t with period 2π . With $u = (x, y) \in \mathbf{R}^2$, a Hamiltonian $f = (\frac{\partial H}{\partial y}, -\frac{\partial H}{\partial x})^T$ and $g = (g_1, g_2)^T$, the systems are of the specific form

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial y} + \varepsilon g_1(x, y, t), \\ \dot{y} &= -\frac{\partial H}{\partial x} + \varepsilon g_2(x, y, t), \end{aligned} \tag{5.1}$$

which may, with $v \in S^1$, equivalently be written as

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial y} + \varepsilon g_1(x, y, v), \\ \dot{y} &= -\frac{\partial H}{\partial x} + \varepsilon g_2(x, y, v), \\ \dot{v} &= 1. \end{aligned} \tag{5.2}$$

Fig. 6 Melnikov method in the plane $\{v = v_0\}$



For $\varepsilon = 0$ the system (5.1) is integrable and is assumed to possess a homoclinic orbit $\gamma^h(t)$ to a saddle p_0 as in Fig. 6, filled with periodic orbits. For $\varepsilon > 0$, the variable v may be seen as a periodically varying coefficient in the equation for $u = (x, y)$. This classical problem and its Melnikov integral are often associated with the study of so-called homoclinic tangles for maps, and corresponding to it, with chaos. See for instance Guckenheimer and Holmes (1983, Sect. 4.6–4.8) or Wiggins (1990, Sect. 4.5–4.6).

However, in the setting of (1.1) the system (5.2) is of the form (1.5):

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial y} + \varepsilon g_1(x, y, v), \\ \dot{y} &= -\frac{\partial H}{\partial x} + \varepsilon g_2(x, y, v), \\ \dot{v} &= \varepsilon h(x, y, v), \end{aligned} \tag{5.3}$$

in which the variable $v \in \mathbf{R}$ is no longer periodic, but may be seen as a slowly varying coefficient in the equations for x and y . By the above assumption on (5.1), the $\varepsilon = 0$ limit of this system possesses a homoclinic orbit $\gamma^h(t; v_0)$ filled with periodic orbits in every plane $\{v = v_0\}$. The homoclinic orbits form a 2-dimensional homoclinic manifold to a curve \mathcal{M}_0 ; see Fig. 4a. This homoclinic manifold is part of both the stable manifold $W^s(\mathcal{M}_0)$ and the unstable manifold $W^u(\mathcal{M}_0)$ of \mathcal{M}_0 : the bounded parts of $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ coincide. In other words, in this degenerate case the distance between both manifolds is zero in every plane $\{v = v_0\}$.

If H , g_i and h are smooth enough and \mathcal{M}_0 is a compact, normally hyperbolic manifold, Fenichel’s first and second theorems apply and guarantee existence of a normally hyperbolic slow manifold \mathcal{M}_ε for $0 < \varepsilon \ll 1$, and stable and unstable manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$, which in general no longer coincide but are $\mathcal{O}(\varepsilon)$ apart. If an unperturbed orbit $\gamma^h(t; v_0)$ survives as a biasymptotic orbit to \mathcal{M}_ε (for example a heteroclinic orbit between two fixed points on \mathcal{M}_ε), it should be subset of $W^s(\mathcal{M}_\varepsilon) \cap W^u(\mathcal{M}_\varepsilon)$, and thus the distance between $W^s(\Gamma_\varepsilon)$ and $W^u(\Gamma_\varepsilon)$ measured in a certain cross-section Σ_{v_0} should be zero. The Melnikov method basically calculates this distance.

In a fixed plane $\{v = v_0\}$ the situations for $\varepsilon = 0$ and $\varepsilon > 0$ are sketched in Fig. 6. Here $\gamma^h(t) = (t; v_0)$ is defined as the homoclinic solution in $\{v = v_0\}$ to the $\varepsilon = 0$ system (5.3) with $\gamma^h(0) = (x_0, y_0, v_0)$, and Σ_{v_0} is the line through

$u_0 = (x_0, y_0)$ perpendicular to $f(\gamma^h(0; v_0))$. The solutions γ_ε^u in $W^u(\mathcal{M}_\varepsilon)$ and γ_ε^s in $W^s(\mathcal{M}_\varepsilon)$ to the $\varepsilon > 0$ system (5.3) are determined by the initial condition $\gamma_\varepsilon^{u,s}(0) \in \Sigma_{v_0}$. In $\{v = v_0\}$, the separation of the manifolds $W^s(\mathcal{M}_\varepsilon)|_{v_0}$ and $W^u(\mathcal{M}_\varepsilon)|_{v_0}$ can be measured along Σ_{v_0} , or, equivalently, along the normal $f^\perp(\gamma^h(0; v_0)) = \left(\frac{\partial H}{\partial x}(\gamma^h(0; v_0)), \frac{\partial H}{\partial y}(\gamma^h(0; v_0))\right)^T$. This yields a distance

$$d(t = 0, v_0, \varepsilon) = f(\gamma^h(0; v_0)) \wedge [\gamma_\varepsilon^u(0; v_0) - \gamma_\varepsilon^s(0; v_0)],$$

where the wedge product is defined by $a \wedge b = a_1 b_2 - a_2 b_1$ and $f \wedge [\gamma_\varepsilon^u - \gamma_\varepsilon^s]$ is the projection of $\gamma_\varepsilon^u - \gamma_\varepsilon^s$ on f^\perp . The distance $d(t = 0, v_0, \varepsilon)$ however depends on Σ_{v_0} ; we therefore take

$$\Delta(0, v_0) = \frac{\partial}{\partial \varepsilon} d(t = 0, v_0, \varepsilon)|_{\varepsilon=0} = f(\gamma^h(0; v_0)) \wedge \left[\frac{\partial}{\partial \varepsilon} \gamma_\varepsilon^u(0; v_0) - \frac{\partial}{\partial \varepsilon} \gamma_\varepsilon^s(0; v_0) \right], \tag{5.4}$$

which is uniquely defined and measures the $\mathcal{O}(\varepsilon)$ distance between the manifolds in $\{v = v_0\}$. By Robinson (1983) (or Wiggins 1990, sec. 4.5) this distance is equal to the integral

$$\Delta W(v_0) = \int_{-\infty}^{\infty} f(\gamma^h(t; v_0)) \wedge \left[g(\gamma^h(t; v_0)) + \frac{\partial f}{\partial v}(\gamma^h(t; v_0)) \frac{\partial v}{\partial \varepsilon} \right] dt, \tag{5.5}$$

where $\frac{\partial v}{\partial \varepsilon}$ satisfies $\frac{d}{dt} \left(\frac{\partial v}{\partial \varepsilon} \right) = h(\gamma^h(0; v_0))$ and $\frac{\partial v}{\partial \varepsilon} = 0$ at $t = 0$. Simple zeroes v_0 of this so-called *adiabatic Melnikov integral* correspond to transverse intersections of $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$, and hence to biasymptotic orbits to \mathcal{M}_ε . These orbits that are traced by this integral make a *single loop* through phase space, before they return to \mathcal{M}_ε .

A very important feature of this measure for the distance between the two manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ is, that it satisfies to explicitly know the *unperturbed* $\varepsilon = 0$ homoclinic orbits to calculate the integral. Although it uses the perturbed $\varepsilon \neq 0$ differential equations and describes the distance between the perturbed manifolds, the integration is along solutions of the unperturbed system.

Remark 7 In fact this Melnikov method applies to more general systems than (5.3). The backbone structure need not be Hamiltonian and the ‘slowly varying parameter’ v can be more-dimensional. In case of a Hamiltonian structure (5.3) the method is equivalent to the following, more intuitive method if H is constant on the perturbed manifold \mathcal{M}_ε . For $\varepsilon = 0$, H is by construction constant along orbits, but this is in general no longer the case for $\varepsilon > 0$. A connecting orbit $\gamma^h(t)$ to \mathcal{M}_ε must however satisfy $\lim_{t \rightarrow \infty} H(\gamma^h(t)) = \lim_{t \rightarrow -\infty} H(\gamma^h(t))$ if H is constant on \mathcal{M}_ε , so that in fact $\Delta H|_{\gamma^h} = \int_{-\infty}^{\infty} \frac{dH}{dt}(\gamma^h(t)) dt = 0$ must hold.

Example 5.1 We return to Example 3.2, in which we treated the system

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= x - x^2 + \varepsilon f(x, y, z; \mu), \\ \dot{z} &= \varepsilon g(x, y, z; \mu), \end{aligned} \tag{5.6}$$

We focus on the fate of the family of homoclinic loops to the vertical line $\mathcal{M}_0 = \mathcal{M}_\varepsilon = \{x = y = 0\}$ that exist for $\varepsilon = 0$. As we mentioned in Example 3.2, based solely on Theorem 4 the fate of this family for small $\varepsilon > 0$ is still unclear. However, since the bounded parts of $W^u(\mathcal{M}_0)$ and $W^s(\mathcal{M}_0)$ coincide, the distance between the perturbed manifolds $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ is of order $\mathcal{O}(\varepsilon)$ and may thus be computed via the adiabatic Melnikov integral.

It is convenient to measure this distance in the cross-section $\Sigma = \{y = 0, x > 1\}$. The solutions $\gamma_\varepsilon^u = (x_\varepsilon^u, y_\varepsilon^u, z_\varepsilon^u)$ in $W^u(\mathcal{M}_\varepsilon)$ and $\gamma_\varepsilon^s = (x_\varepsilon^s, y_\varepsilon^s, z_\varepsilon^s)$ in $W^s(\mathcal{M}_\varepsilon)$ of Eq. (5.6) are determined by their initial conditions in Σ : $\gamma_\varepsilon^{u,s}(0) = (x_\varepsilon^{u,s}, 0, z_0)$. The solution $\gamma_0(t) = (x_0(t), y_0(t), z_0)$ is defined as the homoclinic solution to the unperturbed system ($\varepsilon = 0$) with $\gamma_0(0) = (\frac{3}{2}, 0, z_0)$; see Eq. (3.4). The distance (5.4) is then written as

$$\Delta W(t, z_0) = \begin{pmatrix} y_0(t) \\ x_0(t) - x_0^2(t) \end{pmatrix} \wedge \begin{pmatrix} \frac{\partial}{\partial \varepsilon} [x_\varepsilon^u(t) - x_\varepsilon^s(t)]|_{\varepsilon=0} \\ \frac{\partial}{\partial \varepsilon} [y_\varepsilon^u(t) - y_\varepsilon^s(t)]|_{\varepsilon=0} \end{pmatrix},$$

and the resulting adiabatic Melnikov integral (5.5) is equal to

$$\begin{aligned} \Delta W(z_0) &= \int_{-\infty}^{\infty} \begin{pmatrix} y_0(t) \\ x_0(t) - x_0^2(t) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ f(x_0(t), y_0(t), z_0; \mu) \end{pmatrix} dt \\ &= \int_{-\infty}^{\infty} y_0(t) f(x_0(t), y_0(t), z_0; \mu) dt. \end{aligned} \tag{5.7}$$

If this integral has a simple zero at z_0^* , then $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ intersect transversally and a (transversal) homoclinic orbit to \mathcal{M}_ε persists, which makes a fast loop near the plane z_0^* . No such intersections exist where ΔW remains bounded away from 0. For the specific cases $f_1 = y(b + cz)$, $g_1 = (v + \rho x + z^2)$ and $f_2 = y(z^2 - a)$, $g_2 = (1 + bx - cz^2)$, respectively treated in Doelman and Holmes (1996), Doelman and Hek (2000) and Hek et al. (1998), the integral (5.7) is computed as

$$\Delta W_1(z_0; b, c) = \frac{6}{5}(b + cz_0); \quad \Delta W_2(z_0; a) = \frac{6}{5}(z_0^2 - a). \tag{5.8}$$

Hence the following conclusions can be drawn.

1. The integral $\Delta W_1(z; b, c)$ has a single zero $z = -\frac{b}{c}$, corresponding to a transversal homoclinic orbit to \mathcal{M}_ε . Depending on the parameters, this orbit may connect

two fixed points on \mathcal{M}_ε as in Fig. 4b, may be a homoclinic orbit to a single fixed point, or connect $z = \pm\infty$ to each other or to a fixed point; see Doelman and Holmes (1996).

We remark that in this case one should be careful with the application of Fenichel’s theorems, since these require a compact manifold \mathcal{M}_ε . If there are two fixed points $(0, 0, \pm \frac{1}{\sqrt{-v}})$ for $\varepsilon > 0$, the piece bounded by these fixed points can be taken as $\mathcal{M}_0 = \mathcal{M}_\varepsilon$. The conclusions about homoclinic or heteroclinic orbits to fixed points can then be drawn. For the orbits tending to $z = \pm\infty$ the flow should be compactified in one way or another in order to conclude their existence. We refer again to Doelman and Holmes (1996).

2. The integral $\Delta W_2(z; a)$ has zeroes $z = \pm\sqrt{a}$ if $a \geq 0$, while $\Delta W_2(0, z) \neq 0$ for all z if $a < 0$.

Since $\Delta W(0; 0) = \frac{\partial}{\partial z} \Delta W(0; 0) = 0$ and $\frac{\partial^2}{\partial z^2} \Delta W(0; 0) \neq 0$, $\frac{\partial}{\partial a} \Delta W(0; 0) \neq 0$, there is a unique value $a = a^* = 0 + \mathcal{O}(\varepsilon)$, for which $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ have quadratic contact (cf. Theorem 4.5.4 of Guckenheimer and Holmes 1983). Therefore, for $a > a^*$ there are two orbits $\gamma^\pm(t)$ that are biasymptotic to $\mathcal{M}_\varepsilon = \{x = y = 0, |z| \leq \frac{1}{\sqrt{c}}\}$ and make one fast loop through phase space, while for $a < a^*$ there are none. Varying the parameter a thus gives a global saddle-node bifurcation at $a = a^*$. See Fig. 4c and Doelman and Hek (2000) and Hek et al. (1998).

The constructed orbits are heteroclinic orbits between the fixed points $s^\pm := (0, 0, \pm \frac{1}{\sqrt{c}})$ on \mathcal{M}_ε .

For $\varepsilon \downarrow 0$, these orbits have a singular limiting structure, that consists of a fast loop (the unperturbed homoclinic orbit (3.4) with $z_0 = -\frac{b}{c}$ or $z_0 = \pm\sqrt{a}$), in general concatenated to two slow pieces of orbit on \mathcal{M}_0 . See the illustration in Fig. 4. In the setting with f_1, g_1 it is possible that the fast loop leaves or enters a neighbourhood of \mathcal{M}_ε via the strong stable or unstable manifold of either one of the saddle points. In such case the limiting structure will have only one, or even no slow pieces.

6 Fenichel’s third theorem: fibering and base points

A normally hyperbolic critical manifold \mathcal{M}_0 is by definition filled with critical points v_0 , each of which has its own stable and/or unstable manifolds $W^s(v_0)$ and $W^u(v_0)$. Consider again the Eq. (1.1) and suppose that the $\varepsilon = 0$ Jacobian $\frac{\partial f}{\partial u}(u, v, 0)|_{\mathcal{M}_0}$ has m eigenvalues λ with $Re(\lambda) < 0$ and n eigenvalues with $Re(\lambda) > 0$, as in the setting of Theorem 4. Then $W^s(v_0)$ is m -dimensional and $W^u(v_0)$ is n -dimensional, and the manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$ are the unions

$$W^s(\mathcal{M}_0) = \bigcup_{v_0 \in \mathcal{M}_0} W^s(v_0), \quad W^u(\mathcal{M}_0) = \bigcup_{v_0 \in \mathcal{M}_0} W^u(v_0).$$

In other words, the manifolds $W^s(v_0)$ and $W^u(v_0)$ form collections of fibers for $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$, respectively, with base points $v_0 \in \mathcal{M}_0$.

In our study of (1.1) we have so far seen that a compact critical manifold \mathcal{M}_0 and its stable and unstable manifolds perturb to analogous objects \mathcal{M}_ε , $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ when ε is sufficiently small. A natural question to ask now, is whether the individual stable and unstable manifolds $W^s(v_0)$ and $W^u(v_0)$ also perturb to analogous objects. Fenichel’s third theorem answers this question. Before we state this theorem we first treat an example to introduce the type of problems that may be solved using it. This example is a real Ginzburg–Landau equation coupled to a reaction–diffusion equation. A second, biological example is the 1-dimensional Gierer–Meinhardt problem presented in Example 1.1.

Example 6.1 In Doelman et al. (2004) the following set of a Ginzburg–Landau equation coupled to a diffusion equation was proposed as a prototype equation for certain systems with two competing instability mechanisms. Such a competition of instability mechanisms is a very natural phenomenon that for instance occurs in binary fluid convection (Riecke 1992, 1996) or in biological models for the density of ion channels embedded in a biomembrane (Peter and Zimmerman 2006). Other examples include biochemical systems (Dewel et al. 1995), geophysical morphodynamics (Komarova and Newell 2000) and systems with symmetries (Coullet and Fauve 1985). The equations considered are

$$\begin{cases} A_t = A_{xx} - A + A^3 + \mu AB, \\ \varepsilon^2 \tau B_t = \varepsilon^{-2} B_{xx} - \alpha \varepsilon^2 B + \nu A^2 + \beta A^2 B \end{cases} \tag{6.1}$$

with parameters $\mu, \nu, \alpha, \beta \in \mathbf{R}$, a small parameter $0 < \varepsilon \ll 1$ and $A(x, t) : \mathbf{R} \times \mathbf{R}^+ \rightarrow \mathbf{R}$, $B(x, t) : \mathbf{R} \times \mathbf{R}^+ \rightarrow \mathbf{R}$ real amplitudes. The subscripts t and x respectively denote partial differentiation with respect to the time variable t and the spatial variable x .

The model was developed to study the influence of an additional (slow) diffusion equation on both the existence and the stability of the pulse solutions of the Ginzburg–Landau equation, see Doelman et al. (2004, 2007). Such pulse solutions represent in case of (binary) fluid convection a localised convective pattern within an otherwise quiet fluid. In the ion channel example (Peter and Zimmerman 2006) a pulse describes the amplitude of an ion channel pattern within in otherwise homogeneous distribution of such channels.

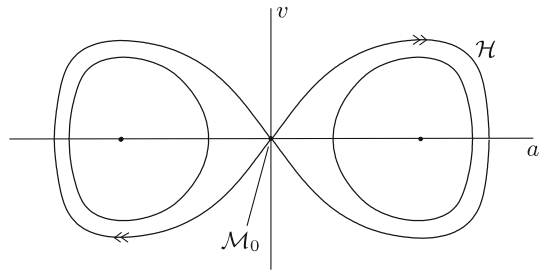
Here we focus on the existence question: “Does there exist a standing single-pulse solution to (6.1)?”

In searching a standing pulse, the time derivatives A_t and B_t are set zero. With new coordinates $(a, v, b, d) = (A, A_x, B, B_x/\varepsilon)$ the system is rewritten as

$$\begin{cases} \dot{a} = v, \\ \dot{v} = a - a^3 - \mu ab, \\ \dot{b} = \varepsilon d, \\ \dot{d} = \varepsilon[\alpha \varepsilon^2 b - \nu a^2 - \beta ba^2], \end{cases} \tag{6.2}$$

where the dot represents differentiation with respect to x . Note that the system is of the form (1.5). A physically relevant pulse solution satisfies $\lim_{x \rightarrow \pm\infty} |A(x, t)| = \lim_{x \rightarrow \pm\infty} |B(x, t)| = 0$. Therefore, we look for homoclinic solutions $\gamma_h(x)$ to (6.2)

Fig. 7 The (a_0, v_0) phase space of system (6.2) with $\varepsilon = 0$



that satisfy $\lim_{x \rightarrow \pm\infty} \gamma_h(x) = (0, 0, 0, 0)$, where $S = (0, 0, 0, 0)$ is a fixed point of (6.2).

In system (6.2) the differentiation is with respect to the *long length scale* x . Alternatively, the *short length scale* $\xi = \varepsilon x$ can be chosen as the independent variable, with resulting system

$$\begin{cases} \varepsilon a' = v, \\ \varepsilon v' = a - a^3 - \mu ab, \\ b' = d, \\ d' = \varepsilon^2 \alpha b - va^2 - \beta ba^2. \end{cases} \tag{6.3}$$

where $' = \frac{d}{d\xi}$. We remark that, with some abuse of terminology, system (6.2) still tends to be called the fast system, and system (6.3) the slow one. Putting $\varepsilon = 0$ in (6.2) hence yields the associated fast reduced system, which possesses 2-dimensional invariant manifolds $\{a_0 = 0, v_0 = 0\}$, $\{a_0 = -\sqrt{1 - \mu b_0}, v_0 = 0\}$ and $\{a_0 = +\sqrt{1 - \mu b_0}, v_0 = 0\}$. Only part of the first one, $\mathcal{M}_0 := \{(a_0, v_0, b_0, d_0) \mid a_0 = v_0 = 0, 1 - \mu b_0 > 0\}$, is filled with hyperbolic equilibria. Thus \mathcal{M}_0 is a normally hyperbolic invariant manifold; it has 3-dimensional stable and unstable manifolds $W^s(\mathcal{M}_0)$ and $W^u(\mathcal{M}_0)$, which are the unions of the two-parameter families of 1-dimensional stable and unstable manifolds of the saddle points $(a, v, b_0, d_0) = (0, 0, b_0, d_0) \in \mathcal{M}_0$. For each (b_0, d_0) , \mathcal{M}_0 is connected to itself by two homoclinic orbits $(\pm a_0(x; b_0), \pm v_0(x; b_0))$ with

$$a_0(x; b_0) = \sqrt{2(1 - \mu b_0)} \operatorname{sech}(\sqrt{1 - \mu b_0} x), \tag{6.4}$$

$$v_0(x; b_0) = a_0'(x; b_0). \tag{6.5}$$

The 2-parameter families of homoclinic orbits form homoclinic manifolds $\mathcal{H} = W^s(\mathcal{M}_0) \cap W^u(\mathcal{M}_0)$, filled with periodic orbits (see Fig. 7). From now on, we focus on the homoclinic orbits with positive a -coordinate. The orbits $(a_0(x; b_0), v_0(x; b_0))$ correspond to pulse solutions $A(x, t) = a_0(x; b_0)$, with $\lim_{x \rightarrow \pm\infty} A = 0$, in the uncoupled Ginzburg–Landau equation $A_t = A_{xx} - (1 - \mu b_0)A + A^3$.

Let us return to the question whether (6.1) admits single-pulse solutions with $\lim_{x \rightarrow \pm\infty} |A(x, t)| = \lim_{x \rightarrow \pm\infty} |B(x, t)| = 0$. Since the ODE system (6.2) is nearly integrable, and the integrable limit possesses two families of homoclinic orbits that correspond to pulse solutions to the unperturbed Ginzburg–Landau equation, one expects

that possible pulse solutions to the full equation (6.1) have one of the unperturbed pulses as ‘building block’. Therefore, the first question to answer is whether (6.1) still has solutions that satisfy $\lim_{x \rightarrow \pm\infty} A = 0$, or, in terms of a and v , whether (6.2) has solutions with $\lim_{x \rightarrow \pm\infty} (a, v) = (0, 0)$. We apply Fenichel’s theorems and the Melnikov integral to deduce that of the 2-parameter family (6.4) a 1-parameter family of orbits asymptotic to \mathcal{M}_ε remains.

If ε is sufficiently small, Theorem 2 guarantees that for any compact subset \mathcal{M}_0^c of \mathcal{M}_0 there exists a locally invariant manifold $\mathcal{M}_\varepsilon^c$ for the perturbed system (6.2) which is diffeomorphic to and $\mathcal{O}(\varepsilon)$ C^1 -close to \mathcal{M}_0^c . Clearly, $\mathcal{M}_\varepsilon^c = \mathcal{M}_0^c$ in the current setting. In addition, Theorem 4 states that for $0 < \varepsilon \ll 1$ the manifold $\mathcal{M}_\varepsilon^c$ has 3-dimensional stable and unstable manifolds $W^s(\mathcal{M}_\varepsilon^c)$ and $W^u(\mathcal{M}_\varepsilon^c)$ that are again $\mathcal{O}(\varepsilon)$ close and diffeomorphic to their counterparts $W^s(\mathcal{M}_0^c)$ and $W^u(\mathcal{M}_0^c)$. Since this is true for any compact subset of \mathcal{M}_ε , we will be somewhat sloppy and speak about \mathcal{M}_ε in the sequel.

For $0 < \varepsilon \ll 1$, the stable and unstable manifolds $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ will no longer merge in a homoclinic manifold \mathcal{H} , but may intersect in one or more 2-dimensional surfaces. Since the system (6.2) is an $\mathcal{O}(\varepsilon)$ perturbation of an integrable system with periodic orbits inside \mathcal{H} , components of $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ inside \mathcal{H} wind around $(a, v) = (\sqrt{1 - \mu b_0}, 0)$ and intersect the hyperplane $\{v = 0\}$ several times.

Adiabatic Melnikov theory (see Sect. 5) now provides the measure (5.5) as a function of b_0 and d_0 that determines the $\mathcal{O}(\varepsilon)$ distance between the first intersections of $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ with $\{v = 0\}$:

$$\Delta(b_0, d_0) = \int_{-\infty}^{\infty} \mu d_0 a_0(x; b_0) v_0(x; b_0) x \, dx. \tag{6.6}$$

Zeroes of this measure $\Delta(b_0, d_0)$ give the first intersections of $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ in leading order, so assuming $\mu \neq 0$ and using that the integrand in (6.6) is an even function of x , the intersections are located at

$$d_0 = 0. \tag{6.7}$$

A priori this is only a leading order result. Since system (6.1) is symmetric under $\{x \rightarrow -x\}$, a reversibility symmetry inherited by (6.2) as

$$x \rightarrow -x, \quad v \rightarrow -v, \quad d \rightarrow -d \tag{6.8}$$

the result is however exact:

$$W^u(\mathcal{M}_\varepsilon) \cap W^s(\mathcal{M}_\varepsilon) \cap \{v = 0\} \subset \{d = 0\}.$$

In other words, any orbit that is biasymptotic to \mathcal{M}_ε with initial conditions $v(0) = d(0) = 0$ is reversible itself.

This means, that there is a one-parameter family of orbits biasymptotic to \mathcal{M}_ε . As pointed out in Sect. 3, solutions in $W^s(\mathcal{M}_\varepsilon)$ decay to \mathcal{M}_ε at an exponential rate as $x \rightarrow \infty$, solutions in $W^u(\mathcal{M}_\varepsilon)$ decay to \mathcal{M}_ε at an exponential rate as $x \rightarrow -\infty$. Thus the slow segments of the biasymptotic orbits should be close to some orbit in \mathcal{M}_ε . However, without further knowledge about the fate of the biasymptotic orbits in a neighbourhood of \mathcal{M}_ε , one cannot conclude whether or not there is an orbit among them that is homoclinic to the fixed point S . Fenichel’s third theorem will serve to answer this question.

6.1 Fenichel’s third theorem

Although the critical points $v_0 \in \mathcal{M}_0$ do in general not perturb to fixed points, the answer to the above stated question is positive. Indeed, the individual stable and unstable manifolds $W^s(v_0)$ and $W^u(v_0)$ perturb to analogous objects, as is formulated in the theorem below. Readers who find it too technical may want to read Corollary 9 first, as this is the useful consequence for applications.

Whereas the manifolds $W^{u,s}(v_0)$ are invariant, their counterparts $W^{u,s}(v_\varepsilon)$ are not. This is intuitively clear since their ‘base point’ v_ε itself is not invariant under the flow of (1.1). However, the whole families $\{W^u(v_\varepsilon) \mid v_\varepsilon \in \mathcal{M}_\varepsilon\}$ and $\{W^s(v_\varepsilon) \mid v_\varepsilon \in \mathcal{M}_\varepsilon\}$ are invariant in a certain sense.

To state this invariance we use the notation $x \cdot t$ to denote the application of a flow after time t to an initial point x . Similarly, $V \cdot t$ denotes the application of the flow after time t to a set V , and $x \cdot [t_1, t_2]$ is the resulting trajectory if the flow is applied over the interval $[t_1, t_2]$. However, in order to avoid difficulties we restrict ourselves to a neighbourhood Δ of \mathcal{M}_ε in which the linear terms of (1.1) are dominant, and consider only trajectories in $W^u(\mathcal{M}_\varepsilon)$ that have not left Δ in forward time (yet), and trajectories in $W^s(\mathcal{M}_\varepsilon)$ that have not left Δ in backward time. This is stated more precisely following the definition from Jones (1995):

Definition The forward evolution of a set $V \subset \Delta$ restricted to Δ is given by the set

$$V \cdot_\Delta t := \{x \cdot t \mid x \in V \text{ and } x \cdot [0, t] \subset \Delta\}.$$

We are now ready to state

Theorem 8 (Fenichel) *Suppose $\mathcal{M}_0 \subset \{f(u, v, 0) = 0\}$ is compact, possibly with boundary, and normally hyperbolic, and suppose f and g are smooth. Then for every $v_\varepsilon \in \mathcal{M}_\varepsilon$, $\varepsilon > 0$ and sufficiently small, there are an m -dimensional manifold $W^s(v_\varepsilon) \subset W^s(\mathcal{M}_\varepsilon)$ and an n -dimensional manifold $W^u(v_\varepsilon) \subset W^u(\mathcal{M}_\varepsilon)$, that are $\mathcal{O}(\varepsilon)$ close and diffeomorphic to $W^s(v_0)$ and $W^u(v_0)$, respectively. The families $\{W^{u,s}(v_\varepsilon) \mid v_\varepsilon \in \mathcal{M}_\varepsilon\}$ are invariant in the sense that*

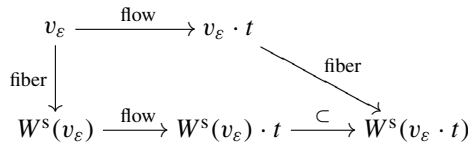
$$W^s(v_\varepsilon) \cdot_\Delta t \subset W^s(v_\varepsilon \cdot t)$$

if $v_\varepsilon \cdot s \in \Delta$ for all $s \in [0, t]$, and

$$W^u(v_\varepsilon) \cdot_\Delta t \subset W^u(v_\varepsilon \cdot t)$$

if $v_\varepsilon \cdot s \in \Delta$ for all $s \in [t, 0]$.

The idea is, that the order in which the flow after time t is applied to a base point and the fiber of a base point is constructed does not matter, as depicted in this diagram.



In the unperturbed setting (1.1) with $\varepsilon = 0$, the decay in forward time of points in $W^s(\mathcal{M}_0)$ to \mathcal{M}_0 is clearly to the base point v_0 of their fiber, where the decay rate as $t \rightarrow \infty$ is exponential, since all associated eigenvalues have nonzero real part.

The Fenichel fibers of Theorem 8 offer an analogous matching between points in $W^s(\mathcal{M}_\varepsilon)$ and \mathcal{M}_ε (and of course similarly between points in $W^u(\mathcal{M}_\varepsilon)$ and \mathcal{M}_ε). If associated to a point $x \in W^s(\mathcal{M}_\varepsilon)$ there is a base point $x^+ \in \mathcal{M}_\varepsilon$, i.e. $x \in W^s(x^+)$, then the exponential decay is inherited from the unperturbed case as follows, and as illustrated in Fig. 8.

Corollary 9 *There are constants $\kappa_s, \alpha_s > 0$ so that if $x \in W^s(x^+) \cap \Delta$, then*

$$\|x \cdot t - x^+ \cdot t\| \leq \kappa_s e^{-\alpha_s t}$$

for all $t \geq 0$ for which $x \cdot [0, t] \subset \Delta$ and $x^+ \cdot [0, t] \subset \Delta$.

Similarly, there are constants $\kappa_u, \alpha_u > 0$ so that if $x \in W^u(x^-) \cap \Delta$, then

$$\|x \cdot t - x^- \cdot t\| \leq \kappa_u e^{\alpha_u t}$$

for all $t \leq 0$ for which $x \cdot [t, 0] \subset \Delta$ and $x^- \cdot [t, 0] \subset \Delta$.

Although this is only formulated in the neighbourhood Δ of \mathcal{M}_ε where the linear behaviour is dominant, this has of course a consequence for points ‘further away’ on

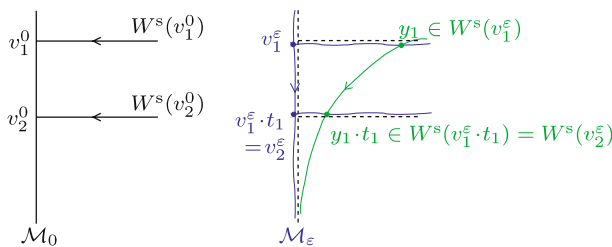


Fig. 8 Fibering of $W^s(\mathcal{M}_0)$ for $\varepsilon = 0$ (left panel) and $W^s(\mathcal{M}_\varepsilon)$ for small $\varepsilon > 0$ (right panel)

$W^u(\mathcal{M}_\varepsilon)$ or $W^s(\mathcal{M}_\varepsilon)$. In general, if a point $x \in W^s(\mathcal{M}_\varepsilon)$ has base point $x^+ \in \mathcal{M}_\varepsilon$, then this corollary ensures that there are constants $C_1, C_2, \kappa > 0$ such that

$$\|x \cdot T - x^+ \cdot T\| \leq C_1 e^{-\kappa/\varepsilon} \quad \forall T \geq \frac{C_2}{\varepsilon}.$$

Consequently, associated to an orbit $\gamma(t; x) \subset W^s(\mathcal{M}_\varepsilon)$ there is an orbit $\gamma^+(t; x^+) \subset \mathcal{M}_\varepsilon$, with $\gamma(0; x) = x, \gamma^+(0; x^+) = x^+$, such that $\|\gamma(t; x) - \gamma^+(t; x^+)\| \leq C_1 e^{-\kappa/\varepsilon}$ for $t \geq \frac{C_2}{\varepsilon}$ (assuming that $\gamma^+(t; x^+)$ stays in \mathcal{M}_ε for all $t \geq \frac{C_2}{\varepsilon}$). We say that $\|\gamma(t; x) - \gamma^+(t; x^+)\|$ is exponentially small in ε for $t \geq \frac{C_2}{\varepsilon}$. Analogous statements can be formulated for $W^u(\mathcal{M}_\varepsilon)$.

6.2 Application of Fenichel’s third theorem

The fact that any point in $W^u(\mathcal{M}_\varepsilon)$ or $W^s(\mathcal{M}_\varepsilon)$ now has a base point in \mathcal{M}_ε associated to it, can be used to concatenate fast pieces of orbits to slow pieces of orbits that follow an l -dimensional slow manifold with $l > 1$. If one has sufficient information about the fast flow, one can construct so-called *take-off* and *touch-down* sets (often curves), that are in fact sets of base points associated to sets of orbits in $W^u(\mathcal{M}_\varepsilon)$ or $W^s(\mathcal{M}_\varepsilon)$. We will do so in the following example, which is a continuation of Example 6.1.

Example 6.2 Consider again Eq. (6.2). We are now in a position to decide whether or not the slow and fast segments of one of the solutions with $\lim_{x \rightarrow \pm\infty}(a, v) = (0, 0)$ can be concatenated to form a homoclinic orbit to $S = (0, 0, 0, 0)$.

The (slow) flow on \mathcal{M}_ε for the full system (6.2) can be determined by substituting $(a, v) = (0, 0)$. The resulting 2-dimensional system has one equilibrium point $(b, d) = (0, 0)$, which is S in the four-dimensional system. It is given by

$$\begin{aligned} \dot{b} &= \varepsilon d, \\ \dot{d} &= \varepsilon^3 \alpha b. \end{aligned} \tag{6.9}$$

Note here, that the flow on \mathcal{M}_ε is in fact ‘superslow’ (with $\dot{d} = \mathcal{O}(\varepsilon^3)$), which is the reason why we use (6.2) rather than the slow limiting problem (Eq. 6.3 with $\varepsilon = 0$) to analyse it.

If the restriction of $S = (0, 0, 0, 0)$ to \mathcal{M}_ε is a saddle point, there exist orbits within \mathcal{M}_ε that satisfy $\lim_{x \rightarrow \infty}(b, d) = (0, 0)$ and orbits that satisfy $\lim_{x \rightarrow -\infty}(b, d) = (0, 0)$. This is the case if $\alpha > 0$, which is assumed for the remainder of this example. The stable and unstable manifolds of S (restricted to \mathcal{M}_ε) are given by

$$l^{u,s} = W^{u,s}(0, 0) |_{\mathcal{M}_\varepsilon} = \{(b, d) \mid d = \pm \varepsilon \sqrt{\alpha b}\}. \tag{6.10}$$

In Example 6.1 the distance between $W^u(\mathcal{M}_\varepsilon)$ and $W^s(\mathcal{M}_\varepsilon)$ was measured in the hyperplane $\{v = 0\}$. This means that the biasymptotic orbits $\gamma(x)$ to \mathcal{M}_ε found there can be parameterised by their initial condition, that we can chose as $y_0 = \gamma(0, y_0) \in W^u(\mathcal{M}_\varepsilon) \cap W^s(\mathcal{M}_\varepsilon) \cap \{v = 0\}$. For these orbits $\gamma(x; y_0)$, Theorem 8

and Corollary 9 imply that there exist two orbits $\gamma^+ = \gamma^+(x; y_0^+) \subset \mathcal{M}_\varepsilon$ and $\gamma^- = \gamma^-(x; y_0^-) \subset \mathcal{M}_\varepsilon$, such that $\|\gamma(x; y_0) - \gamma^+(x; y_0^+)\|$ is exponentially small in ε when $x \geq \mathcal{O}(\frac{1}{\varepsilon})$ and $\|\gamma(x; y_0) - \gamma^-(x; y_0^-)\|$ is exponentially small in ε when $-x \geq \mathcal{O}(\frac{1}{\varepsilon})$. To be precise, this statement should be restricted to a compact subset $\mathcal{M}_\varepsilon^c \subset \mathcal{M}_\varepsilon$: the exponential closeness is only guaranteed as long as $\gamma^\pm(x; y_0^\pm)$ stays in $\mathcal{M}_\varepsilon^c$. There are however two reasons why the compactness of \mathcal{M}_ε is not of importance here. The first is, that we aim at constructing a homoclinic orbit, that has by definition compact slow parts (bounded by the take off or touch down point and the point S). The second is, that the existence of the saddle point S on \mathcal{M}_ε gives this slow manifold sufficient structure to be uniquely defined (possible nonuniqueness is the reason why the invariant manifold theorems are restricted to compact manifolds).

We return to the homoclinic orbit. If a homoclinic orbit $\gamma_h(x)$ to S exists, it must either tend to S via a strong stable or unstable manifold, or via its stable and unstable manifolds l^u and l^s within \mathcal{M}_ε . In the latter, generic case $\|\gamma_h(x; y_0) - l^{u,s}\|$ is exponentially small in ε for $|x| \geq \mathcal{O}(\frac{1}{\varepsilon})$. Whether such an orbit γ_h exists depends on the positions of the base points $y_0^\pm = \gamma^\pm(0; y_0^\pm) \in \mathcal{M}_\varepsilon$.

We define the curves $T_o \subset \mathcal{M}_\varepsilon$ (take off) and $T_d \subset \mathcal{M}_\varepsilon$ (touch down) as

$$T_o := \bigcup_{y_0} \{y_0^- = \gamma^-(0; y_0^-)\} \quad \text{and} \quad T_d := \bigcup_{y_0} \{y_0^+ = \gamma^+(0; y_0^+)\},$$

where the unions are over all $y_0 \in W^s(\mathcal{M}_\varepsilon) \cap W^u(\mathcal{M}_\varepsilon) \cap \{v = 0\}$. The take off set T_o represents the collection of base points of all of the Fenichel fibers in $W^u(\mathcal{M}_\varepsilon)$ that are asymptotic to \mathcal{M}_ε as $x \rightarrow \infty$. Similarly, T_d represents the set of base points of the fibers in $W^s(\mathcal{M}_\varepsilon)$ that are asymptotic to \mathcal{M}_ε as $x \rightarrow -\infty$.

Detailed information about the positions of T_o and T_d is given by the relation between $y_0 = (a_0, 0, b_0, d_0)$ and its base points $y_0^- = (a_0^-, v_0^-, b_0^-, d_0^-)$ and y_0^+ , respectively, or in fact by the relation between b_0 and b_0^\pm, d_0 and d_0^\pm . Thus we determine the sets T_o and T_d by measuring the change in b and d of $\gamma(x, y_0) \subset W^u(\mathcal{M}_\varepsilon) \cap W^s(\mathcal{M}_\varepsilon)$ during half a circuit through the fast field. The accumulated change in d during a circuit can be measured by integrating d' along the orbit $\gamma(x, y_0)$ until the orbit settles down near \mathcal{M}_ε . It takes $\mathcal{O}(|\log \varepsilon|)$ ‘time’ to leave an $\mathcal{O}(\varepsilon)$ neighbourhood of \mathcal{M}_ε , make a fast loop and return to the neighbourhood of \mathcal{M}_ε ; see for instance Lemma 3.2 in Doelman and Hek (2000). We therefore integrate up to $\mathcal{O}(|\log \varepsilon|)$, and take for the positive and negative half circuits

$$\Delta^- d = - \int_{k \log \varepsilon}^0 d'|_{\gamma(x; y_0)} dx \quad \text{and} \quad \Delta^+ d = \int_0^{-k \log \varepsilon} d'|_{\gamma(x; y_0)} dx,$$

respectively. Here $k > 0$ is some arbitrary constant independent of ε . For $\gamma(x; y_0)$ with $y_0 = (a_0, 0, b_0, d_0)$ we find, by (6.2),

$$\begin{aligned} \Delta d &= \Delta^+ d - \Delta^- d \\ &= \int_{k \log \varepsilon}^{-k \log \varepsilon} d' |_{\gamma(x; y_0)} dx = -\varepsilon \int_{k \log \varepsilon}^{-k \log \varepsilon} (v + \beta b_0) a_0^2(x; b_0) dx + \mathcal{O}(\varepsilon^2 |\log \varepsilon|), \end{aligned} \tag{6.11}$$

where we have used that $\gamma(x, y_0)$ can be approximated by the unperturbed homoclinic orbit $\gamma_0(x, y_0) = (a_0(x; b_0), v_0(x; b_0), b_0, d_0)$ (6.4) during the fast circuit. Since the $a_0(x)$ and $v_0(x)$ terms of γ_0 converge exponentially to zero as $x \rightarrow \pm\infty$, this integral is equal to

$$\Delta d = -\varepsilon \int_{-\infty}^{\infty} (v + \beta b_0) a_0^2(x; b_0) dx + \mathcal{O}(\varepsilon^{1+2K}) + \mathcal{O}(\varepsilon^2 |\log \varepsilon|), \tag{6.12}$$

and the improper integral exists for the same reason. It follows from (6.4) that $K = k\sqrt{1 - \mu b_0} > 0$. Note that the symmetry of $\gamma_0(x, y_0)$ also implies that $\Delta^+ d = -\Delta^- d = \frac{1}{2}\Delta d$. We conclude by (6.4), and by choosing k large enough, that

$$\begin{aligned} \Delta d &= -2\varepsilon(v + \beta b_0)\sqrt{1 - \mu b_0} \int_{-\infty}^{\infty} \operatorname{sech}^2(x) dx + \mathcal{O}(\varepsilon^2 |\log \varepsilon|) \\ &= -4\varepsilon(v + \beta b_0)\sqrt{1 - \mu b_0} + \mathcal{O}(\varepsilon^2 |\log \varepsilon|). \end{aligned} \tag{6.13}$$

Similarly we could calculate Δb , but the reversibility symmetry (6.8 immediately gives $\Delta b = 0$ over a jump. Since correspondingly $d_0 = 0$ at a jump, $-\Delta^- b = \Delta^+ b = \int_0^\infty b' |_{\gamma(x, y_0)} dx + \mathcal{O}(\varepsilon^2 |\log \varepsilon|) = \mathcal{O}(\varepsilon^2 |\log \varepsilon|)$.

The take off and touch down curves are now obtained by correcting (6.7) for the change in b and d during half an excursion through the fast field. At leading order only the highest order terms of $\Delta^+ d$ and $\Delta^- d$ play a role:

$$\begin{aligned} T_o &:= \left\{ (b_0^-, d_0^-) = (b_0, d_0) \mid d_0 = 2\varepsilon(v + \beta b_0)\sqrt{1 - \mu b_0} \right\}, \\ T_d &:= \left\{ (b_0^+, d_0^+) = (b_0, d_0) \mid d_0 = -2\varepsilon(v + \beta b_0)\sqrt{1 - \mu b_0} \right\} \end{aligned}$$

up to corrections of $\mathcal{O}(\varepsilon^2 |\log \varepsilon|)$.

We recall that a generic homoclinic orbit to S satisfies $\|\gamma_h(x, y_0) - l^{u,s}\| = \mathcal{O}(e^{-\kappa/\varepsilon})$ for $|x| > \mathcal{O}(1/\varepsilon)$ and some $\kappa > 0$. Therefore, its take off and touch down points y_0^- and y_0^+ must lie on l^u and l^s , respectively, and thus in the intersections $T_o \cap l^u$ and $T_d \cap l^s$ (see Fig. 9). The leading order term b_0 of the homoclinic orbit hence has to fulfil the equation that is satisfied in $T_o \cap l^u$ and $T_d \cap l^s$:

$$\sqrt{\alpha} b_0 = 2(v + \beta b_0)\sqrt{1 - \mu b_0}. \tag{6.14}$$

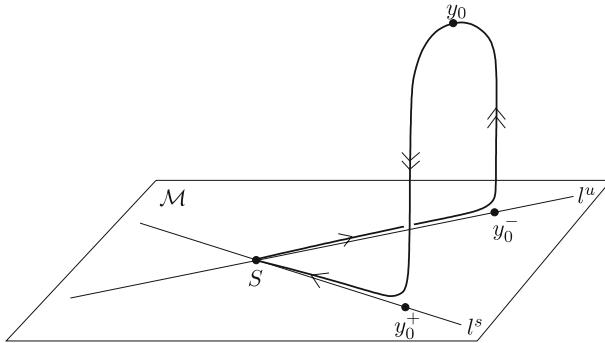


Fig. 9 Sketch of the homoclinic orbit with its take off point $y_0^- \in T_o \cap l^u$ and touch down point $y_0^+ \in T_d \cap l^s$. The constructed singular orbit is concatenated in y_0^\pm . A single arrow indicates slow flow, a double arrow fast flow

Note here that the reversibility symmetry (6.8) of system (6.2) maps l^u to l^s and T_o to T_d (and vice versa), so there is only one equation to satisfy here. In general, it is however important that the size of Δb fits with the geometry of the problem, i.e. that it supports a jump from l^u to l^s .

Depending on the parameters there are 0, 1, or 2 intersections $T_o \cap l^u$ and $T_d \cap l^s$ in this example. It can be shown that these intersections are generically transversal, but there are codimension-1 manifolds in $(\alpha, \beta, \mu, \nu)$ on which the curves are tangent.

The idea of the construction of a homoclinic orbit to S is to formally construct a singular homoclinic orbit that consists of two slow parts in \mathcal{M}_0 (parts of l^u and l^s) with a middle fast part given by the unperturbed homoclinic orbit (6.4) with $d_0 = 0$ and b_0 given by (6.14) (see Fig. 9). The essential idea behind singular perturbation theory is that persistence of this singular orbit for small $\varepsilon > 0$ is guaranteed if the singular structure corresponds to transversal intersections of pairs of manifolds.

In general one needs to check transversality of the intersections $T_o \cap l^u$ and $T_d \cap l^s$ and of the intersection $W^u(\mathcal{M}_\varepsilon) \cap W^s(\mathcal{M}_\varepsilon)$ (which follows in general from the explicit $\mathcal{O}(\varepsilon)$ expression (6.6)), but here the symmetry (6.8) ensures that any transversal intersection $l^u \cap T_o$ automatically corresponds to a (symmetric) homoclinic orbit $\gamma_h \in W^u(S) \cap W^s(S)$. For parameter values at which $T_o \cap l^u$ and $T_d \cap l^s$ are tangent, no precise conclusions can be drawn. However, close to these parameter values, the full system (6.2) undergoes a saddle-node bifurcation of homoclinic orbits to S , as in Example 5.1.

In the setting of system (6.1) the orbit $\gamma_h(x)$ corresponds to a pulse solution $(A_h(x, t), B_h(x, t)) = (a_h(x), b_h(x))$ that satisfies $\lim_{x \rightarrow \pm\infty} (A_h(x, t), B_h(x, t)) = (0, 0)$.

To conclude, this means that after coupling of the Ginzburg–Landau equation with the second B -equation, the resulting system (6.1) still admits a single-pulse solution. In this specific example, the physical meaning of this mathematical conclusion is, that the coupled system (6.1) may indeed explain the mechanisms in binary fluid convection or in similar experiments. For similar examples, such as the ion channel dynamics

in Peter and Zimmerman (2006), similar analysis can predict or confirm existence of localised patterns of ion channels within an otherwise homogeneous state. For more details, and for a construction when the reversibility symmetry (6.8) is broken, we refer to Doelman et al. (2004).

7 Discussion

This overview is of course not complete. It covers only a part of the literature on geometric singular perturbation techniques, and it does not mention other widely used methods to study singularly perturbed systems. Classically, many singularly perturbed problems were analysed by asymptotic analysis. Different scalings, asymptotic expansions in the small parameter, and matching of the obtained different local approximations formed the basic concepts of asymptotic analysis. We refer to the survey by Eckhaus (1979) for an overview and rigorous proofs of different aspects of the theory. A recent text covering a wide variety of the methods and their applications is Verhulst (2005). This type of analysis often gives explicit expansions and validity results, and is, by its nature, quite laborious. A group of French people developed a different analytical approach that simplifies the study of multiple-time scale problems. Their method uses nonstandard analysis (Robinson 1974; Robert 1988), as was first suggested by Reeb. This analysis has a different axiomatic basis, that allows to consider the small perturbation parameter ε as an infinitesimal number. Singularly perturbed differential equations are then considered as infinitesimally small perturbations of standard or limited expressions. Attached to these nonstandard methods is the rather well-known terminology of ‘canards’, used to indicate certain types of ‘slow–fast’ solutions. The name ‘canard’ originates in the study of relaxation oscillations in the Van der Pol equation and was introduced by Diener and Diener (1981). It refers to the shape of transitional limit cycles that appear in an exponentially small neighbourhood of a certain critical parameter value (Diener 1984; Eckhaus 1983; Braaksma 1993).

Other important issues to address here concern various instruments with a geometric flavour that can be combined with geometric singular perturbation theory to gain insight in the behaviour of orbits passing near a slow manifold or to construct n -pulse homoclinic orbits or n -periodic orbits.

Fenichel’s theorems only consider slow manifolds and their *local* stable and unstable manifolds. To unravel the *global* geometry of the stable and unstable manifolds and to keep track of their intersections it is important to understand the behaviour of orbits and manifolds as they pass near a slow manifold. The basic idea in studying such behaviour in forward time, is to take a disk \mathcal{D} that transversally intersects the stable manifold $W^s(\mathcal{M})$ and use the fact that any point $q = q(0)$ in the intersection $W^s(\mathcal{M}) \cap \mathcal{D}$ satisfies $\lim_{t \rightarrow \infty} \|q(t) - \mathcal{M}\| = 0$ by the definition of the stable manifold. Conclusions about the fate of other points in \mathcal{D} as $t \rightarrow \infty$ can then be drawn. Similarly, the fate of a disk intersecting $W^u(\mathcal{M})$ as $t \rightarrow -\infty$ can be understood. Two well-known lemmas based on this idea are the Lambda Lemma for maps (see for instance Guckenheimer and Holmes 1983, p. 247, Wiggins 1990, p. 473) and the Exchange Lemma (Jones and Kopell 1994; Jones et al. 1996; Kaper and Jones 2001) for flows. The Lambda Lemma can be applied by considering the flow after time T as a map and concerns the accumulation of images of \mathcal{D} on $W^u(\mathcal{M})$.

The Exchange Lemma concerns the passage of invariant sets (manifolds or orbits) near a slow manifold and has in a sense proven more suitable for the study of singularly perturbed systems. A series of applications, in which the disk \mathcal{D} is a subset of $W^u(\mathcal{M})$, can be found in [Doelman and Holmes \(1996\)](#), [Doelman \(1996\)](#) and [Doelman and Hek \(2000\)](#). In [Doelman and Holmes \(1996\)](#) both approaches were used, in the other papers the more intricate behaviour was studied using the Exchange Lemma.

In nearly integrable systems (1.5), there is often a region in phase space in which the fast reduced problem contains a continuous family of periodic orbits. This means that the flow in this region is still almost periodic for $0 < \varepsilon \ll 1$, which may be exploited to reduce the system (1.5) to another, lower dimensional problem.

The idea of reducing the study of continuous time systems to the study of an associated discrete time system (map) is due to [Poincaré \(1899\)](#), who first utilised it in his studies of the three body problem in celestial mechanics. The conceptually clear construction of a Poincaré map involves the elimination of a variable of the problem, and hence offers dimensional reduction, see again [Guckenheimer and Holmes \(1983, sec. 1.5\)](#) or [Wiggins \(1990, sec. 1.2\)](#). The explicit construction of a Poincaré map $\mathcal{P} : \Sigma \rightarrow \Sigma$, where Σ is a plane transversal to the flow, requires an extensive knowledge of the geometrical structure of the phase space of the particular ODE one wants to reduce. It is in general not possible, but in nearly integrable systems (1.5), the periodicity in the limiting $\varepsilon = 0$ problem suffices to construct an approximation of the map. This approximation can for instance be used to find n -periodic orbits with slow/fast structure or successive images of $W^u(\mathcal{M}) \cap \Sigma$ or $W^s(\mathcal{M}) \cap \Sigma$. The latter can in turn be used to construct higher order intersections $W^u(\mathcal{M}) \cap W^s(\mathcal{M})$, or n -pulse homoclinic orbits to \mathcal{M} . See for instance ([Doelman and Holmes 1996](#); [Hek 2000](#); [Doelman et al. 2001a,b](#)).

Apart from the above mentioned geometrical methods, higher order Melnikov methods have also successfully been used to find n -pulse homoclinic orbits. Examples can for instance be found in [Kaper and Kovačič \(1996\)](#), [Soto-Treviño and Kaper \(1996\)](#) and [Camassa et al. \(1998\)](#). Another strong, topological method to draw conclusions about (higher order) orbits that connect \mathcal{M} to itself uses the Conley index. See for instance [Kokubu et al. \(1996\)](#), in which a system similar to (3.1) is treated.

Acknowledgments This work was supported by NWO grant MEERVOUD 632000.002, and very much stimulated by Odo Diekmann.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aranson I, Kramer L (2002) The world of the Ginzburg–Landau equation. *Rev Modern Phys* 74:99–143
- Beck M, Doelman A, Kaper TJ (2006) A geometric construction of traveling waves in a bioremediation model. *J Nonlinear Sci* 16:329–349
- Benoît E, Callot JF, Diener F, Diener M (1981) Chasse au canard. *Collect Math* 31(31):37–119
- Best J, Borisyuk A, Rubin J, Terman D, Wechselberger M (2005) The dynamic range of bursting in a model respiratory pacemaker network. *SIAM J Appl Dyn Syst* 4(4):1107–1139

- Braaksma B (1993) Critical phenomena in dynamical systems of Van der Pol type. Ph.D. thesis, Rijksuniversiteit Utrecht
- Camassa R, Kovačič G, Tin S-K (1998) A Melnikov method for homoclinic orbits with many pulses. *Arch Rat Mech Anal* 143(2):105–195
- Carpenter GA (1979) Bursting phenomena in excitable membranes. *SIAM J Appl Math* 36:334–372
- Couillet P, Fauve S (1985) Large-scale oscillatory instability for systems with translational and Galilean invariances. In: *Macroscopic modelling of turbulent flows* (Nice, 1984). *Lecture Notes in Physics*, vol 230. Springer, Berlin, pp 290–295
- Deng B (2001) Food chain chaos due to junction-fold point. *Chaos* 11(3):514–525
- Deng B, Hines G (2002) Food chain chaos due to Shilnikov's orbit. *Chaos* 12(3):533–538
- Deng B, Hines G (2003) Food chain chaos due to transcritical point. *Chaos* 13(2):578–585
- Derks G, Doelman A, van Gils SA, Visser AT (2003) Travelling waves in a singularly perturbed sine-Gordon equation. *Physica D* 180(1–2):40–70
- Dewel G, Métens S, Hilali MF, Borckmans P, Price CB (1995) Resonant patterns through coupling with a zero mode. *Phys Rev Lett* 74:4647–4650
- Diener M (1984) The canard unchained or how fast/slow dynamical systems bifurcate. *Math Intell* 6:38–49
- Diener F, Diener M (1981) Chasse au canard. I. Les canards. *Collect Math* 32(1):37–74
- Doelman A (1996) Breaking the hidden symmetry in the Ginzburg–Landau equation. *Physica D* 97:398–428
- Doelman A, Hek GM (2000) Homoclinic saddle-node bifurcations in singularly perturbed systems. *J Dyn Differ Equ* 12(1):169–216
- Doelman A, Holmes P (1996) Homoclinic explosions and implosions. *Phil Trans R Soc Lond A* 354:845–893
- Doelman A, Kaper TJ (2003) Semi-strong pulse interactions in a class of coupled reaction–diffusion equations. *SIAM Appl Dyn Syst* 2(1):53–96
- Doelman A, van der Ploeg H (2002) Homoclinic stripe patterns. *SIAM Appl Dyn Syst* 1(1):65–104
- Doelman A, Gardner RA, Kaper TJ (2001a) Large stable pulse solutions in reaction–diffusion equations. *Indiana Univ Math J* 50(1):443–507
- Doelman A, Kaper TJ, van der Ploeg H (2001b) Spatially periodic and aperiodic multi-pulse patterns in the one-dimensional Gierer–Meinhardt equation. *Meth Appl Anal* 8(3):387–414
- Doelman A, Hek G, Valkhoff N (2004) Stabilization by slow diffusion in a real Ginzburg–Landau system. *J Nonlinear Sci* 14:237–278
- Doelman A, Hek G, Valkhoff NJM (2007) Algebraically decaying pulses in a Ginzburg–Landau system with a neutrally stable mode. *Nonlinearity* 20:357–389
- Doelman A, van Heijster PJA, Kaper TJ (2009) Pulse dynamics in a three-component system: existence analysis. *J Dyn Differ Equ* (to appear)
- Dumortier F (1993) Techniques in the theory of local bifurcations: blow up, normal forms, nilpotent bifurcations, singular perturbations. In: Szolmiuk D (ed) *Bifurcations and periodic orbits of vector fields*. Kluwer, Dordrecht, pp 19–73
- Dumortier F, Roussarie R (1996) Canard cycles and center manifolds. In: *Memoires of the AMS* 557
- Eckhaus W (1979) *Asymptotic analysis of singular perturbations*. North-Holland, Amsterdam
- Eckhaus W (1983) Relaxation oscillations including a standard chase on french ducks. In: Verhulst F (ed) *Asymptotic analysis II. Lecture Notes in Mathematics*, vol 985. Springer, New York, pp 449–494
- Fenichel N (1971) Persistence and smoothness of invariant manifolds for flows. *Ind Univ Math J* 21:193–225
- Fenichel N (1974) Asymptotic stability with rate conditions. *Ind Univ Math J* 23:1109–1137
- Fenichel N (1977) Asymptotic stability with rate conditions, II. *Ind Univ Math J* 26:81–93
- Fenichel N (1979) Geometric singular perturbation theory for ordinary differential equations. *J Differ Equ* 31:53–98
- FitzHugh R (1960) Thresholds and plateaus in the Hodgkin–Huxley nerve equations. *J Gen Physiol* 43:867–896
- Gear WC, Kaper TJ, Kevrekidis IG, Zagaris A (2005) Projecting to a slow manifold: singularly perturbed systems and legacy codes. *SIAM Appl Dyn Syst* 4:711–732
- Gierer A, Meinhardt H (1972) A theory of biological pattern formation. *Kybernetik* 12:30–39
- Guckenheimer J, Holmes PJ (1983) *Nonlinear oscillations, dynamical systems and bifurcations of vector fields*. Springer, New York
- Hek GM (2000) Bifurcations of homoclinic orbits in singularly perturbed flows. Ph.D. thesis, Universiteit Utrecht

- Hek GM, Doelman A, Holmes PJ (1998) Homoclinic saddle-node bifurcations and subshifts in a three-dimensional flow. *Arch Rat Mech Anal* 145:291–329
- Higuera M, Knobloch E, Vega JM (2005) Dynamics of nearly inviscid Faraday waves in almost circular containers. *Physica D* 201:83–120
- Hirsch MW, Pugh CC, Shub M (1977) Invariant manifolds. *Lecture Notes in Mathematics*, vol 583. Springer, Heidelberg
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol (Lond)* 117:500–544
- Holling CS (1959) Some characteristics of simple types of predation and parasitism. *Can Entomologist* 91:385–398
- Huber A, Szmolyan P (2005) Geometric singular perturbation analysis of the Yamada model. *SIAM J Appl Dyn Syst* 4(3):607–648
- Iron D, Ward MJ, Wei J (2001) The stability of spike solutions of the one-dimensional Gierer–Meinhardt model. *Physica D* 150(1–2):25–62
- Jones CKRT (1995) Geometric singular perturbation theory. In: Johnson R (ed) *Dynamical systems*, Montecatini Terme, *Lecture Notes in Mathematics*, vol 1609. Springer, Berlin, pp 44–118
- Jones CKRT, Kopell N (1994) Tracking invariant manifolds with differential forms in singularly perturbed systems. *J Differ Equ* 108(1):64–88
- Jones CKRT, Kaper TJ, Kopell N (1996) Tracking invariant manifolds up to exponentially small errors. *SIAM J Math Anal* 27(2):558–577
- Kaper TJ (1999) An introduction to geometric methods and dynamical systems theory for singular perturbation problems. In: Cronin J, O'Malley RE Jr (eds) *Analyzing multiscale phenomena using singular perturbation methods*. In: *Proc Symposia Appl Math*, vol 56. American Mathematical Society, Providence, pp 85–132
- Kaper TJ, Jones CKRT (2001) A primer on the exchange lemma for fast–slow systems. In: Jones CKRT, Khibnik AI (eds) *Multiple-time-scale dynamical systems*. IMA volumes in Math and its Applications 122. Springer, New York
- Kaper TJ, Kovačič G (1996) Multi-bump orbits homoclinic to resonance bands. *Trans Am Math Soc* 348(10):3835–3887
- Kapitula T (1998) Bifurcating bright and dark solitary waves of the nearly nonlinear cubic–quintic Schrödinger equation. *Proc R Soc Edin A* 128(3):585–629
- Keener J, Sneyd J (1998) *Mathematical physiology*. Springer, New York
- Kokubu H, Mischaikow K, Oka H (1996) Existence of infinitely many connecting orbits in a singularly perturbed ordinary differential equation. *Nonlinearity* 9(5):1263–1280
- Komarova NL, Newell AC (2000) Nonlinear dynamics of sand banks and sand waves. *J Fluid Mech* 415:285–321
- Krupa M, Szmolyan P (2001) Extending geometric singular perturbation theory to nonhyperbolic points—fold and canard points in two dimensions. *SIAM J Math Anal* 33(2):286–314
- Krupa M, Popović N, Kopell N (2008a) Mixed-mode oscillations in three time-scale systems: a prototypical example. in preparation. *SIAM Appl Dyn Syst* 7(2):361–420
- Krupa M, Popović N, Kopell N, Rotstein HG (2008b) Mixed-mode oscillations in a three time-scale model for the dopaminergic neuron. *Chaos* 18:015106
- Kunpasuruang W, Lenbury Y, Hek G (2002) A nonlinear mathematical model for pulsatile discharges of luteinizing hormone mediated by hypothalamic and extra-hypothalamic pathways. *Math Models Methods Appl Sci* 12(5):607–624
- Kuznetsov YuA (1995) *Elements of applied bifurcation theory*. Springer, New York
- May RM (1972) Limit cycles in predator–prey communities. *Science* 177:900–902
- Medvedev GS, Cisternas JE (2004) Multimodal regimes in a compartmental model of the dopamine neuron. *Physica D* 194(3–4):333–356
- Melnikov VK (1963) On the stability of the center for time periodic perturbations. *Trans Moscow Math Soc* 12:1–57
- Mielke A (2002) The Ginzburg–Landau equation in its role as a modulation equation. In: Fiedler B (ed) *Handbook of dynamical systems*, vol 2. North-Holland, Amsterdam, pp 759–834
- Moehlis J (2006) Canards for a reduction of the Hodgkin–Huxley equations. *J Math Biol* 52(2):141–153
- Muratori S, Rinaldi S (1992) Low- and high-frequency oscillations in three-dimensional food chain systems. *SIAM J Appl Math* 52:1688–1706
- Murray JD (1989) *Mathematical biology*. Springer, New York

- Nagumo JS, Arimoto S, Yoshizawa S (1962) An active pulse transmission line simulating nerve axon. *Proc IRE* 50:6191–6198
- Peter R, Zimmerman W (2006) Traveling ion channel density waves affected by a conservation law. *Phys Rev E* 74(1):016–206
- Poincaré H (1899) *Les Méthodes Nouvelles de la Mécanique Céleste*. Gauthier-Villars, Paris, 3 vols
- Popović N, Szmoljan P (2004) A geometric analysis of the Lagerstrom model problem. *J Differ Equ* 199(2):290–325
- Riecke H (1992) Self-trapping of traveling-wave pulses in binary mixture convection. *Phys Rev Lett* 68:301–304
- Riecke H (1996) Solitary waves under the influence of a long-wave mode. *Physica D* 92(1–2):69–94
- Rinaldi S, Muratori S (1992) Slow–fast limit cycles in predator–prey models. *Ecol Model* 61:287–308
- Rinzel J (1985) Excitation dynamics: insights from simplified membrane models. *Federation Proc* 44:2944–2946
- Robert A (1988) *Nonstandard analysis*. Wiley, New York (Translated from the French by the author)
- Robinson A (1974) *Nonstandard analysis*. American Elsevier, New York
- Robinson C (1983) Sustained resonance for a nonlinear system with slowly varying coefficients. *SIAM Math An* 14:847–860
- Rosenzweig ML, MacArthur RH (1963) Graphical representation and stability conditions of predator–prey interactions. *Am Nat* 97:209–223
- Rottschäfer V, Kaper TJ (2003) Geometric theory for multi-bump, self-similar, blowup solutions of the cubic nonlinear Schrödinger equation. *Nonlinearity* 16(3):929–961
- Rubin J, Terman D (2002) Geometric singular perturbation analysis of neuronal dynamics. In: Fiedler B (ed) *Handbook of dynamical systems*, vol 2. Elsevier, Amsterdam
- Rubin J, Wechselberger M (2007) Giant squid—hidden canard: the 3D geometry of the Hodgkin–Huxley model. *Biol Cybern* 97(1):5–32
- Shimazu Y, Sugiyama K, Kojima T, Tomida E (1972) Some problems in ecology oriented environmentology. *J Earth Sci Nagoya Univ* 20:31–89
- Soto-Treviño C, Kaper TJ (1996) Higher-order Melnikov theory for adiabatic systems. *J Math Phys* 37:6220–6249
- Terman D (1991) Chaotic spikes arising from a model of bursting in excitable membranes. *SIAM J Appl Math* 51(5):1418–1450
- van der Ploeg H (2005) Singular pulse patterns in the Gierer–Meinhardt equation. Ph.D. thesis, Universiteit van Amsterdam
- Verhulst F (2005) *Methods and applications of singular perturbations*. Springer, Berlin
- Ward MJ, Wei J (2002) Asymmetric spike patterns for the one-dimensional Gierer–Meinhardt model: equilibria and stability. *Eur J Appl Math* 13(3):283–320
- Ward MJ, Wei J (2003) Hopf bifurcations and oscillatory instabilities of spike solutions for the one-dimensional Gierer–Meinhardt model. *J Nonlinear Sc* 13(2):209–264
- Wechselberger M (2005) Existence and bifurcation of Canards in \mathbf{R}^3 in the case of a folded node. *SIAM J Appl Dyn Sys* 4:101–139
- Wiggins S (1990) *Introduction to applied nonlinear dynamical systems and chaos*. Springer, New York
- Wiggins S (1994) *Normally hyperbolic invariant manifolds in dynamical systems*. Springer, New York