A geometric singular perturbation analysis of generalised shock selection rules in reaction-nonlinear diffusion models

Bronwyn Bradshaw-Hajek[1], Ian Lizarraga[2], Robert Marangell[2], and Martin Wechselberger[2]

[1]UniSA STEM, University of South Australia, Mawson Lakes, South Australia, Australia
[2]School of Mathematics and Statistics, University of Sydney, NSW, Australia

Abstract

Reaction-nonlinear diffusion (RND) partial differential equations are a fruitful playground to model the formation of sharp travelling fronts, a fundamental pattern in nature. In this work, we demonstrate the utility and scope of regularisation as a technique to investigate shock-fronted solutions of RND PDEs, using geometric singular perturbation theory (GSPT) as the mathematical framework. In particular, we show that *composite* regularisations can be used to construct families of shock-fronted travelling waves sweeping out distinct generalised area rules, which interpolate between the equal area and extremal area (i.e. algebraic decay) rules that are well-known in the shockwave literature. Our analysis blends Melnikov methods with GSPT techniques applied to the PDE over distinct spatiotemporal scales.

We also consider the spectral stability of these new interpolated shockwaves. Using techniques from geometric spectral stability theory, we determine that our RND PDE admits nonlinearly stable shock-fronted travelling waves. The multiple-scale nature of the regularised RND PDE continues to play an important role in the analysis of the spatial eigenvalue problem.

1 Introduction

Continuum transport models of coupled systems of cell populations have for the most part used standard linear diffusion to model population spread [23, 35]. Reaction-diffusion equations, such as the extensively studied Fisher equation [10], are used to model population growth dynamics combined with a simple linear Fickian diffusion process, and are typically capable of exhibiting travelling wave solutions.

In cell migration, advection (or transport) is another source of pattern formation. It may represent, e.g., tactically-driven movement, where cells migrate in a directed manner in response to a concentration gradient [24, 35]. Such a concentration gradient develops, for example, in a soluble fluid (chemotaxis) or as a gradient of cellular adhesion sites or of substrate-bound chemoattractants (haptotaxis). Well studied examples of individual cells exhibiting directed motion in response to a chemical gradient include bacteria chemotactically migrating towards a food source. Wound healing, angiogenesis or malignant tumor invasion are just a few examples of chemotactic and/or haptotactic cell movement where the migrating cells form part of a dense population of cells as may be found in tissues. Such migrating cell populations not only form travelling waves but may

also develop sharp interfaces in the wave form. [26, 27, 34, 17, 45, 14, 15].

Another important experimental observation is that motility varies with population density [40, 8, 4]. Such density-dependent nonlinear diffusion processes are also implicated in the formation of sharp interfaces. In the context of population dynamics, living cells make informed decisions through, e.g., sensing the local cell density, and they perform a 'biased walk'. This could lead to, e.g., the tendency to cluster or aggregate with other nearby cells–think of flocking or swarming. Such aggregation mechanisms can be achieved through, e.g., density dependent negative (or backward) diffusion, and such nonlinear diffusion models with subdomains of backward diffusion are known to admit shock-type solutions [19]. Our goal in this paper is to consider the emergence of shock-fronted travelling waves in *reaction-nonlinear diffusion* (RND) models that arise as continuum limits of discrete motile processes made by aggregations of cells or other biological agents [39, 20, 4].

In general, shocks are problematic because as the wave front steepens (and a shock forms) the solution becomes multivalued and physically nonsensical. The model breaks down and it becomes impossible to compute the temporal evolution of the solution [32, 33]. To deal with such ill-posed problems, shock solutions of PDEs are mathematically formalized as weak solutions in the appropriate function space, and it is generally understood that such solutions are nonunique—this is related to the issue of where exactly the shock discontinuity develops in the domain. Nonunique families of weak solutions are known to arise in models of nonlinear diffusion processes. In physical and chemical problems, shock selection is typically enforced by physical constraints, such as conservation laws, entropy and energy conditions, just to name a few. For example, in advection-reaction models they may represent hyperbolic balance laws, i.e., hyperbolic conservation laws with source terms, where the formation of shock fronts is well-known. These shock selection rules are usually referred to as admissibility criteria. We refer the reader to [41, 46] for a comprehensive development of shock structure theory from this point-of-view.

One mathematical approach to deal with ill-posed problems and shock formation is to employ *regularisation*, i.e. adding small perturbative higher order terms to these models to give rise to well-posed problems and, hence, introduce smoothing effects. In the context of hyperbolic conservation/balance laws, these are usually small viscous (diffusive) regularisations, e.g., the well-known viscous Burgers equation. Due to dissipative mechanisms, these physical shocks are observed as narrow transition regions with steep gradients of field variables. Mathematically, questions of existence and uniqueness of such viscous shock profiles are fundamental.[1]

Regularisation techniques have been employed in physical and chemical problems having nonlinear diffusion processes. These regularised models are usually referred to as phase separation problems [9]. The Cahn-Hilliard equation modelling nucleation in a binary alloy is probably the most famous of these phase separation problems [5, 18, 38], while Sobolev regularisation of phase separation models is another technique [36]. These regularisation techniques are not so well-known [37, 8, 39, 6] within the population dynamics modelling community.

[1] Another option is dispersive regularisation, e.g., the Korteweg-de Vries (KdV) equation. Note that both regularisations (viscous and dispersive) deal with the same problem (inviscid Burgers equation) but create very different outcomes.

Possible shock formation in such regularised RND models is the main focus of this article, and we will use tools from geometric singular perturbation theory and dynamical systems theory to tackle this problem. Again, questions of existence and uniqueness of such regularised shock profiles are fundamental. In particular, we focus on the shock selection criteria based on different composite regularisations. It is worth noting that Witelski considered shock formation in regularised advective nonlinear diffusion models in [47, 48] by means of singular perturbation theory, i.e., he also included advection or transport phenomena in his study. We, on the other hand, focus on nonlinear diffusion as the sole shock formation mechanism. Furthermore, we also consider the spectral stability of these regularised shock waves, and we even construct new kinds of regularised waves, including nonmonotone shockwaves as well as shockwaves containing slow passage through regions of negative diffusion—all of which extends Witelski's approach significantly.

The manuscript is structured as follows: in section 2 we introduce our RND model and its composite regularisations. In sections 3 and 4 we show the existence of travelling and standing waves in such models using the GSPT machinary. In section 5 we then show spectral stability results for regularised shock waves, and we conclude in section 6.

2 The setup for RND Models

We start by considering a dimensionless reaction–nonlinear diffusion model of the form

$$u_t = (D(u)u_x)_x + f(u) = \Phi(u)_{xx} + f(u) \tag{1}$$

where $x \in \mathbb{R}$ denotes the spatial domain, $t \in \mathbb{R}_+$ denotes the time domain, $u \in \mathbb{R}_+$ denotes a (population/agent) density, $D(u)$ models a (population/agent) density dependent diffusivity. $\Phi(u)$ is an anti-derivative of $D(u)$, i.e. $\Phi'(u) = D(u)$, referred to as the *potential*. The (dimensionless) population/agent density u is scaled such that $u \in [0, 1]$ forms the domain of interest where $u = 1$ is the carrying capacity of the population/agent density. This domain of interest is also reflected in the reaction term $f(u)$ which is often modelled either as *logistic growth* or as *bistable growth*. We consider the latter in this paper:

$$f(u) = \kappa u(u - \alpha)(1 - u)\,,\ \kappa > 0,\ 0 < \alpha < 1. \tag{2}$$

We focus on RND models where not only diffusion is present but also *aggregation* (or *backward diffusion*) [5, 18, 8, 35, 37]. By imposing different motility rates for agents that are isolated compared to other agents, one obtains density dependent nonlinear diffusion [20]. Aggregation will manifest itself in these models in sign changes of the density dependent diffusion coefficient $D(u)$. The simplest density-dependent nonlinear diffusion coefficient model that we consider is of the polynomial form

$$D(u) = \beta(u - \gamma_1)(u - \gamma_2) \tag{3}$$

with $0 < \gamma_1 < \gamma_2 < 1$, i.e., we model diffusion-aggregation-diffusion (DAD) in the domain of interest. For sparse population density diffusive behaviour is assumed, while for intermediate population density aggregation will happen which again turns into diffusive behaviour for large population densities (close to carrying capacity). This DAD model assumption leads to a non-monotone cubic potential $\Phi(u)$ as sketched in Figure 1. It is this non-monotonicity of the potential

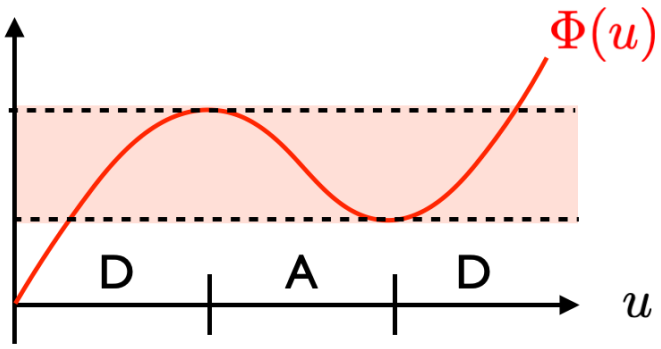


Figure 1: The graph of the potential Φ and the admissible jump zone (shaded) that allows for possible shock connections. We denote the graph $(u, \Phi(u))$ by $S = S_s^l \cup F_l \cup S_m \cup F_r \cup S_s^r$ which is referred to as a critical manifold; see Sec. 3.1 for details. Possible shocks are confined to a fixed potential value $\Phi(u) = const$, i.e. jumps must occur over the middle branch of the potential, S_m, connecting the outer two branches of the potential, S_s^r and S_s^l (shaded region).

Φ which creates a bistability zone of diffusive states which can lead to *phase-separation*, i.e., shock formation.

2.1 RND dynamics and shock formation in travelling wave coordinates

Let us look for one of the simplest possible coherent structures in such RND models (1), travelling waves with wave speed $c \in \mathbb{R}$ that connect the asymptotic end states $u_- = 1 \to u_+ = 0$ or vice versa, i.e., population/agents invade or evade the unoccupied domain with constant speed. A travelling wave analysis introduces a co-moving frame $z = x - ct$ in (1), $c \in \mathbb{R}$. Stationary solutions, i.e., $u_t = 0$, in this co-moving frame include travelling waves/fronts, and they are found as special (heteroclinic) solutions of the corresponding ODE problem

$$-cu_z - (D(u)u_z)_z = f(u)\,.$$

Define the variable $v := -cu - D(u)u_z$ to obtain the corresponding 2D dynamical system

$$\begin{aligned} D(u)u_z &= -(v + cu) \\ v_z &= f(u)\,. \end{aligned} \tag{4}$$

Note that this dynamical system is singular where $D(u) = 0$, i.e., wherever the diffusion-aggregation transition happens. To be able to study this problem (4) on the whole domain of interest including these transition zones near $D(u) = 0$, we make an auxiliary state-dependent transformation $dz = D(u)d\zeta$ which gives the so-called *desingularised problem*

$$\begin{aligned} u_\zeta &= -(v + cu) \\ v_\zeta &= D(u)f(u)\,. \end{aligned} \tag{5}$$

This problem is topologically equivalent to (4) in the diffusion regime $D(u) > 0$ while one has to reverse the orientation in the aggregation regime $D(u) < 0$ to obtain the equivalent flow.

Remark 2.1 *We emphasize that the auxiliary system is only a proxy system to study the problem near $D(u) = 0$. To completely understand the original flow near $D(u) = 0$, one has to use additional techniques such as the blow-up method; see, e.g., [42].*

Remark 2.2 *The desingularised system* (5) *is Hamiltonian when $c = 0$, with generating function*

$$\tilde{H}(u, v) = -\frac{v^2}{2} - \int D(u) f(u)\, du. \tag{6}$$

In this case, local segments of the (un)stable manifolds of the saddle points at $u = 0$ and $u = 1$ lie inside contours of $\tilde{H}(u, v)$.

The asymptotic end states of the travelling waves form equilibrium states of the desingularised (and the original) problem defined by $f(u_\pm) = 0$, and $v_\pm = -cu_\pm$. Our focus is on these asymptotic end states given by the equilibria

$$(u_\mp, v_\mp) = (1, -c)\,, \qquad (u_\pm, v_\pm) = (0, 0)$$

Remark 2.3 *In the case of a bistable reaction term* (2)*, there exists an additional equilibrium in the domain of interest defined by $f(u_b = \alpha) = 0$ which gives $(u_b, v_b) = (\alpha, -c\alpha)$. Its importance/relevance will be discussed later on.*

In our setup of the RND model, travelling wave solutions connecting u_- and u_+ (if they exist) allow for nonsmooth solutions, because the zeroes of the diffusion coefficient $D(u)$ in the relevant domain of interest $u \in [0, 1]$ define singularities in this problem. Discontinuous jumps (shocks) can occur anywhere within the admissible jump zone; see Figure 1. In the absence of an obvious integral conservation law, our approach is to define *geometric* criteria for shock selection.

One strategy is to formally select a shock height from the admissible jump zone in Fig. 1, i.e. we specify the endpoints of the shock $u = u_l$ and $u = u_r$ such that $\Phi(u_l) = \Phi(u_r)$. In terms of system (5), let us denote by $W^u(u_+, c)$ the unstable manifold of u_+ and by $W^s(u_-, c)$ the stable manifold of u_-, and let $v_+(c)$ denote the v-coordinate of the first intersection of $W^u(u_+, c)$ with the section $\{(u, v) : u = u_r\}$ and similarly, denote by $v_-(c)$ the v-coordinate of the first intersection of $W^s(u_-, c)$ with the section $\{(u, v) : u = u_l\}$. We can then attempt to locate a wavespeed $c = c_*$ such that

$$v_+(c_*) = v_-(c_*).$$

If such a wavespeed exists, then we are able to construct a *formal* (nonsmooth) shock connecting u_- and u_+ according to a given shock selection rule.

Shock selection can also be enforced from other geometric conditions. Let us consider a symmetric setup with $\gamma_2 = 1 - \gamma_1$ and $\alpha = 1/2$, i.e., the roots of the diffusion $D(u)$ are placed symmetrically about the midpoint $u = 1/2$ in the interval $u \in (0, 1)$ and the middle root of the reaction term $f(u)$ is placed exactly in the middle. In view of Remark 2.2 and the symmetry, the (un)stable manifolds of the saddle points at $u_\pm = 0, 1$ happen to lie on the same contour of $\tilde{H}(1, 0) = \tilde{H}(0, 0) = 0$ when $c = 0$, i.e.,

$$\tilde{H}(1, 0) = \int_0^1 D(u) f(u) = 0\,.$$

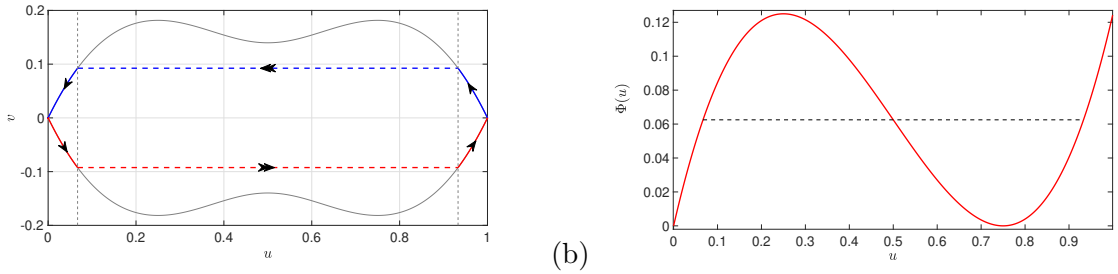


Figure 2: A pair of symmetric standing waves (solid curve segments connected by dashed lines representing shock discontinuities of the solutions, red and blue online) of the desingularised travelling wave equations (5), subject to the equal area shock selection rule. Vertical dashed lines (gray online) denote shock selection values $u = u_{l/r}$. Corresponding contour $\tilde{H}(u,v) = \tilde{H}(0,0) = \tilde{H}(1,0) = 0$ underlaid (gray solid curve). (b) Corresponding shock selection with $u_l \approx 0.06699$ and $u_r \approx 0.93301$. Parameter set: $\gamma_1 = 1/4$, $\gamma_2 = 1 - \gamma_1 = 3/4$, $\alpha = 1/2$, $c = 0$, $\kappa = 5$, $\beta = 6$.

We find that there is exactly one pair of u-values $u_{l/r}$ at which the (un)stable manifolds of the saddle points in the region $D(u) > 0$ can be formally connected by a shock, such that the potential $\Phi(u)$ remains constant; see Fig. 2. In other words, the correct admissible standing shock has been *decided* by the symmetry.

Remark 2.4 *As a consequence of the symmetry, the shock selection values $u = u_{l/r}$ in this case are given by the well-known equal-area rule; see Figure 2(b) and the corresponding formula* (21).

Under continuous variation of the shock selection height and the other model parameters, we expect an entire continuum of such shock connections to persist near to this family of symmetric standing shocks, with continuously varying wavespeeds.[2] Each member of this continuum will constitute a formal shock solution connecting the states u_l and u_r in the desingularised system (5).

2.2 Regularisations of RND models

Following the geometric approach in the previous section, our goal is to describe conditions under which particular shock criteria are *uniquely* selected from within the admissible jump zone. The related issues of nonuniqueness and lack of smoothness suggest that we should consider a 'nearby' system in which locally unique, smoothed shock-fronted solutions are available. Our approach is to add small perturbative high-order regularisation terms.

Regularisation of RND models is typically considered in one of two ways [37, 38]. The first method of regularisation accounts for *viscous relaxation* by adding a small temporal change in the diffusivity:

$$u_t = (\Phi(u) + \varepsilon u_t)_{xx} + f(u), \qquad 0 \leq \varepsilon \ll 1. \tag{7}$$

This is usually referred to as a Sobolev regularisation. The second of these involves adding a small change in the potential to account for interfacial effects, leading to:

$$u_t = (\Phi(u) - \varepsilon^2 u_{xx})_{xx} + f(u), \qquad 0 \leq \varepsilon \ll 1. \tag{8}$$

[2] We revisit this assertion in more detail in Sec. 4.2.

Both regularisation techniques can be viewed as higher order viscous regularisations. They have been widely employed in models of chemical phase-separation, though they have gone relatively unnoticed in biological models until very recently.

Here, we study the possible effects of *both* regularisations in a single RND model, i.e.,

$$u_t = (\Phi(u) + \varepsilon a u_t - \varepsilon^2 u_{xx})_{xx} + f(u), \qquad 0 \leq \varepsilon \ll 1, a \geq 0. \tag{9}$$

Since we only consider small perturbative regularisations $0 < \varepsilon \ll 1$, these models are so-called *singularly perturbed systems* and, as a consequence, the powerful machinery of geometric singular perturbation theory (GSPT) is applicable [7, 21, 45], as we shall explain in this manuscript.

Remark 2.5 *Continuum macroscale models can also be derived from lattice-based microscale models; see [20] for leading order RND models and [2] for (more complicated) regularised RND models.*

3 The GSPT setup for the regularised RND model (9)

We derive conditions based on the specific functions $D(u)$ and $f(u)$ that lead to travelling waves with sharp interfaces (shocks) in one spatial dimension. We introduce a travelling wave coordinate $z = x - ct$ for waves with speed $c \in \mathbb{R}$ and ask for stationary states of the PDE in the co-moving frame. This transforms the regularised RND model (9) into a fourth order ordinary differential equation

$$-cu_z = \Phi(u)_{zz} - \varepsilon a c u_{zzz} - \varepsilon^2 u_{zzzz} + f(u)\,, \tag{10}$$

which we can recast as a *singularly perturbed dynamical system* in standard form

$$\begin{aligned} \varepsilon u_z &= \hat{u} \\ \varepsilon \hat{u}_z &= w + \Phi(u) - \delta \hat{u} \\ v_z &= f(u) \\ w_z &= v + cu\,. \end{aligned} \tag{11}$$

where $(u, \hat{u}) \in \mathbb{R}^2$ are 'fast' variables, $(v, w) \in \mathbb{R}^2$ are 'slow' variables, $\varepsilon \ll 1$ is the singular perturbation parameter, and $\delta := ac$ is an additional lumped system parameter incorporating the wave speed c and the relative contribution of the viscous relaxation a.

Rescaling the 'slow' independent travelling wave variable $dz = \varepsilon dy$ in (11) gives the equivalent fast system

$$\begin{aligned} u_y &= \hat{u} \\ \hat{u}_y &= w + \Phi(u) - \delta \hat{u} \\ v_y &= \varepsilon f(u) \\ w_y &= \varepsilon(v + cu)\,. \end{aligned} \tag{12}$$

with the 'fast' independent travelling wave variable y. These equivalent dynamical systems (11) respectively (12) have a symmetry

$$(\hat{u}, v, c, y) \leftrightarrow (-\hat{u}, -v, -c, -y), \quad \text{respectively} \quad (\hat{u}, v, c, z) \leftrightarrow (-\hat{u}, -v, -c, -z)\,.$$

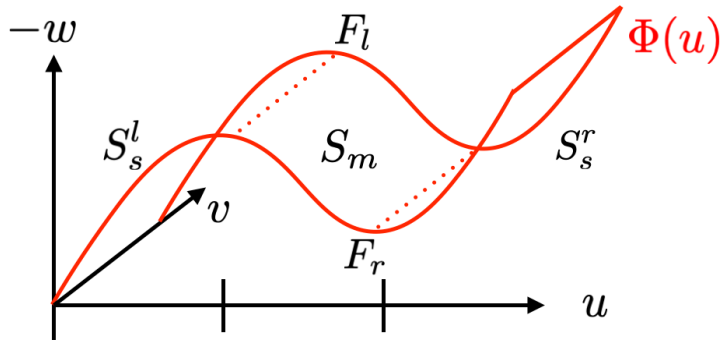


Figure 3: sketch of the two-dimensional critical manifold S projected onto (u, v, w)-space.

The aim is to use methods from GSPT to analyse the travelling wave problem in its 'slow' respectively 'fast' singular limit system, i.e., $\varepsilon \to 0$ in (11) respectively (12), and to infer results on the existence (and stability) of shock-fronted travelling waves in the full regularised RND problem for $\varepsilon \neq 0$.

3.1 The limit on the fast scale - the layer problem

We begin with the 'fast' system (12). Here the limit $\varepsilon \to 0$ gives the *layer problem*

$$\begin{aligned} u_y &= \hat{u} \\ \hat{u}_y &= w + \Phi(u) - \delta\hat{u} \\ v_y = w_y &= 0\,, \end{aligned} \tag{13}$$

i.e., (v, w) are considered parameters. Hence, the flow is along two-dimensional fast fibers $\mathcal{L} := \{(u, \hat{u}, v, w) \in \mathbb{R}^4 : (v, w) = const\}$. The set of equilibria of the layer problem,

$$S := \{(u, \hat{u}, v, w) \in \mathbb{R}^4 : \hat{u} = \hat{u}(u, v) = 0,\ w = w(u, v) = -\Phi(u)\}\,, \tag{14}$$

forms the two-dimensional *critical manifold* of the problem which is a graph over (u, v)-space. In the assumed diffusion-aggregation-diffusion (DAD) setup (3), we have a sign change in the diffusivity along the set

$$F := \{(u, \hat{u}, v, w) \in S : D(u) = 0\}\,, \tag{15}$$

where $F = F_l \cup F_r = \{(u, \hat{u}, v, w) \in S : u = \gamma_1\} \cup \{(u, \hat{u}, v, w) \in S : u = \gamma_2\}$ consists of two disjoint one-dimensional lines. Thus we have a splitting of the critical manifold $S = S_s^l \cup F_l \cup S_m \cup F_r \cup S_s^r$ where

$$\begin{aligned} S_s^l :=& \{(u, \hat{u}, v, w) \in \mathbb{R}^4 : \hat{u} = \hat{u}(u, v) = 0,\ w = w(u, v) = -\Phi(u),\ u < \gamma_1\} \\ S_s^r :=& \{(u, \hat{u}, v, w) \in \mathbb{R}^4 : \hat{u} = \hat{u}(u, v) = 0,\ w = w(u, v) = -\Phi(u),\ u > \gamma_2\} \\ S_m :=& \{(u, \hat{u}, v, w) \in \mathbb{R}^4 : \hat{u} = \hat{u}(u, v) = 0,\ w = w(u, v) = -\Phi(u),\ \gamma_1 < u < \gamma_2\}\,, \end{aligned}$$

see Figure 3. The stability property of this set of equilibria S is determined by the two non-trivial eigenvalues of the layer problem, i.e., the eigenvalues of the Jacobian evaluated along S,

$$J = \begin{pmatrix} 0 & 1 \\ D(u) & -\delta \end{pmatrix}. \tag{16}$$

This matrix has $\mathrm{tr}J = -\delta$ and $\det J = -D(u)$. Hence, for $D(u) > 0$ the outer branches $S_s^{l/r}$ are normally-hyperbolic and of saddle-type (S-type). For $\delta \neq 0$ and $D(u) < 0$ the middle-branch S_m is also normally-hyperbolic, focus/node-type (FN-type), while for $\delta = 0$ and $D(u) < 0$ the middle-branch S_m loses normal-hyperbolicity and is of centre-type (C-type). Loss of normal hyperbolicity happens also along the set $F = F_l \cup F_r$ where $\det J = 0$ independent of δ.

3.2 The limit on the slow scale - the reduced problem

For the slow system (11), the limit $\varepsilon \to 0$ gives the *reduced problem*

$$\begin{aligned} 0 &= \hat{u} \\ 0 &= w + \Phi(u) - \delta\hat{u} \\ v_z &= f(u) \\ w_z &= v + cu\,. \end{aligned} \tag{17}$$

It describes the 'evolution' of the slow variables (v, w) constrained to the 2D critical manifold S (14) which is given as a graph over the (u, v)-coordinate chart. We denote the corresponding embedding $\psi : \mathbb{R}^2 \to \mathbb{R}^4$, i.e., $S = \psi(u, v)$. Therefore, we aim to study the corresponding reduced flow on S in this (u, v)-coordinate chart. By definition, the main requirement on the reduced vector field $R(u, v) \in T\mathbb{R}^2$ is that, when mapped onto the tangent bundle TS via the linear transformation $D\psi$ it has to correspond to the (leading order) slow component of the full four-dimensional vector field constraint to TS, i.e.,

$$D\psi(u, v)R(u, v) = \Pi^S G(\psi(u, v)) = (\frac{v + cu}{-D(u)}, 0, f(u), v + cu)^\top$$

where $\Pi^S G(\psi(u, v))$ is the (oblique) projection of the vector field $G = (0, 0, f(u), v + cu)^\top$ onto the tangent bundle TS of the critical manifold S along fast fibres $\mathcal{L}$ spanned by $\{(1, 0, 0, 0)^\top, (0, 1, 0, 0)^\top\}$. Thus the reduced vector field $R(u, v)$ in the (u, v)-coordinate chart is given by the right-hand side of (4).

Remark 3.1 *The reduced problem is independent of the two choices of regularisation as expected, since it represents the TWP of the original RND model. With the geometric approach, we have the extra information that this reduced flow is constrained to the critical manifold S embedded in the full four-dimensional phase-space.*

We classify all singularities of (4) by analysing the auxiliary system, the desingularised problem (5). The equilibria of the reduced problem (4) respectively desingularised problem (5) and their stability properties are summarised in Table 1. Additionally we know that the two asymptotic end-states $(u_\pm, v_\pm)$ are located on opposite outer branches $S_s^{l/r}$ of the critical manifold while the location of the additional equilibrium state (u_b, v_b) varies under the variation of the system parameters. The Jacobian evaluated at any of these equilibria $(u_{\pm,b}, v_{\pm,b})$ is given by

$$J = \begin{pmatrix} -c & -1 \\ D(u_{\pm,b})f'(u_{\pm,b}) & 0 \end{pmatrix}$$

which has $\mathrm{tr}\, J = -c$ and $\det J = D(u_{\pm,b})f'(u_{\pm,b})$. The types of equilibria are summarized in Table 1. The distinction between Node and Focus (NF) depends on the sign of the discriminant

$\mathcal{D} := c^2 - 4D(u_{\pm,b})f'(u_{\pm,b})$, $\mathcal{D} > 0$ (Node) or $\mathcal{D} < 0$ (Focus). The distinction between stable $c > 0$ and unstable $c < 0$ depends on the sign of the wave speed.[3]

$D(u)$	$f(u)$	(u_-, v_-)	(u_+, v_+)	(u_b, v_b), $\alpha < \gamma_1$	(u_b, v_b), $\gamma_1 < \alpha < \gamma_2$	(u_b, v_b), $\gamma_2 < \alpha$
DAD	bistable	Saddle	Saddle	(un)stable NF	Saddle	(un)stable NF

Table 1: Type of equilibria

Remark 3.2 *The desingularised system* (5) *defines another type of singularities for the original problem through* $D(u) = 0$ *which are known as* folded singularities*; see, e.g., [45]. These will be discussed in more detail in section 3.3.*

3.3 Folded singularities

The desingularised system (5) defines another type of singularity for the reduced problem through $D(u) = 0$ which exists on the fold lines $F_{l/r}$ and are known as *folded singularities*. In our problem, these folded singularities are given by $v_{f_{l/r}} = -cu_{f_{l/r}}$ where $u_{f_{l/r}} = \gamma_{1/2}$, i.e.,

$$(u_{f_{l/r}}, u_{f_{l/r}}) = (\gamma_{1/2}, -c\,\gamma_{1/2})\,.$$

The Jacobian of the desingularised problem evaluated at such a folded singularity is given by

$$J = \begin{pmatrix} -c & -1 \\ D'(u_{f_{l/r}})f(u_{f_{l/r}}) & 0 \end{pmatrix}$$

which has $\operatorname{tr} J = -c$, $\det J = D'(u_{f_{l/r}})f(u_{f_{l/r}})$ and $\mathcal{D} = c^2 - 4D'(u_{f_{l/r}})f(u_{f_{l/r}})$. Hence we are dealing for $\det J < 0$ with a folded saddle (FS), and for $\det J > 0$ with a folded node (FN) or a folded focus (FF) depending on the discriminant $\mathcal{D}$ being positive or negative.

$\phi(u)$	$f(u)$	folded singularity	$\alpha < \gamma_1$	$\gamma_1 < \alpha < \gamma_2$	$\gamma_2 < \alpha$
DAD	bistable	(u_{f_l}, v_{f_l})	FS	'stable' FN/FF	'stable' FN/FF
DAD	bistable	(u_{f_r}, v_{f_r})	'stable' FN/FF	'stable' FN/FF	FS

Table 2: Type of folded singularities on critical manifold S.

Remark 3.3 *The change in type of a folded singularity (FS/FN/FF) under parameter variation coincides with the crossing of the additional equilibrium* (u_b, v_b) *through the corresponding folded singularity (see Table 1). This codimension-one phenomenon is known in the GSPT literature as a folded saddle-node type II (FSN II); see. e.g., [42].*
The term 'stability' indicates only stability properties for the auxiliary system. Folded singularities have no associated stability property since corresponding special solutions known as canards *pass through them in finite time, i.e. these canards represent transient phenomena.*

[3] In case of a standing wave $c = 0$, any NF becomes a Centre.

3.4 Singular heteroclinic orbits

The shock-fronted travelling waves that we seek are found as heteroclinic orbits of the four-dimensional dynamical system (11) connecting the saddle equilibrium end states $(u_\pm, 0, v_\pm, w_\pm) \to (u_\mp, 0, v_\mp, w_\mp)$, i.e., we are seeking evasion fronts $u_- = 0 \to u_+ = 1$ or invasion fronts $u_- = 1 \to u_+ = 0$ of the original travelling front problem. From the point-of-view of GSPT, a key observation is that solutions of the (fast) layer problem (13) and the (slow) reduced problem (4) can be concatenated to form singular heteroclinic orbits,

$$\Gamma_{het}^{\pm} = \Gamma_{l/r} \cup \Gamma_{\pm} \cup \Gamma_{r/l} \tag{18}$$

where Γ_{het}^{+} denotes a singular heteroclinic evasion front connecting $u_- = 0 \to u_+ = 1$, i.e., Γ_l is a slow segment of the unstable manifold of the saddle equilibrium at $u_- = 0$ connecting to u_l, Γ_+ is a fast jump connecting $u_l \to u_r$, and Γ_r is a slow segment of the stable manifold of the saddle equilibrium at $u_+ = 1$ connecting to u_r, while Γ_{het}^{-} denotes a singular heteroclinic invasion front connecting $u_- = 1 \to u_+ = 0$, i.e., Γ_r is a slow segment of the unstable manifold of the saddle equilibrium at $u_- = 1$ connecting to u_r, Γ_- is a fast jump connecting $u_r \to u_l$, and Γ_l, is a segment of the stable manifold of the saddle equilibrium at $u_+ = 0$ connecting to u_l.

Under suitable genericity assumptions, heteroclinic orbits of the regularised system (11) then arise as a codimension-one family of perturbations of the singular connection for $0 < \varepsilon \ll 1$. See [29] for details of this construction in the setting of both 'pure' viscous relaxation and 'pure' Cahn-Hilliard-type regularisation (7) resp. (8). Our objective is to synthesize and greatly extend this analysis in the context of the more general PDE model with composite regularisation (9).

4 Shock selection rules

The asymptotic end states of the 4D composite regularised RND problem (11) are located on saddle branches of the critical manifold (2D layer problem) and the equilibria are also saddles for the 2D reduced problem. This hyperbolic structure persists for the full 4D model, i.e., the fixed asymptotic end-states are saddles with 2 stable and 2 unstable directions.
We extend formally the regularised RND problem (11) by the dummy wavespeed equation $c_y = 0$ and seek 1D intersections of the corresponding 3D centre-stable and the 3D centre-unstable manifold of the two asymptotic end states in 5D phase space, which is a generic.
In the following we construct these shock-fronted travelling waves based on GSPT and identify, in particular, how the composite regularisation parameter a picks a unique shock location and wave speed c.

4.1 Singular fast fronts $\Gamma_\pm$ and generalised shock selection rules

Recall that we introduced the lumped system parameter $\delta = ac$ in the layer problem (13) which takes the composite regularisation parameter a and the wave speed c into account.

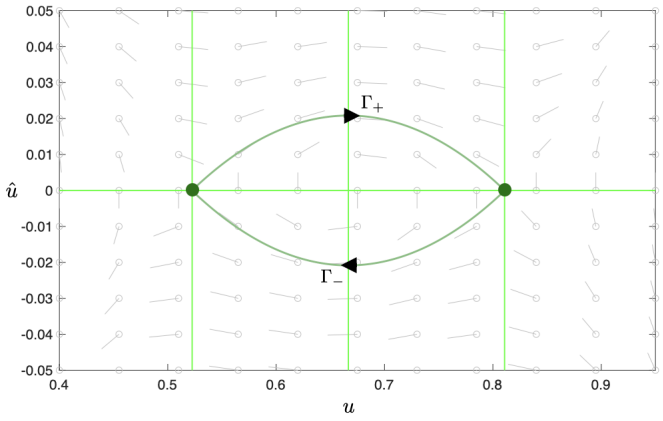


Figure 4: two heteroclinics $\Gamma_+ : p_l \to p_r$ and $\Gamma_- : p_r \to p_l$ for $\delta = 0$ and $w = w_h \approx -0.5648$ in $(u, \hat{u})$-space; other parameter values: $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$.

4.1.1 The $\delta = 0$ case

In this case, the layer problem (13),

$$\begin{aligned} u_y &= \hat{u} \\ \hat{u}_y &= w + \Phi(u) \end{aligned} \tag{19}$$

is Hamiltonian with

$$H(u, \hat{u}) = \frac{\hat{u}^2}{2} - \int (w + \Phi(u)) du \,. \tag{20}$$

Trajectories of this layer problem are confined to level sets of the Hamiltonian (20), i.e., $H(u, \hat{u}) = k$. Possible trajectories that are able to connect equilibrium points on different branches of the critical manifold S are confined to the saddle branches $S_s^{l/r}$ including the boundaries $F_{l/r}$. The corresponding equilibrium points $p_{l/r} = (u_{l/r}, 0, v_{l/r}, -\Phi(u_{l/r})) \in S_s^{l/r} \cup F_{l/r}$ of such connections must fulfill $v_l = v_r$ and $\Phi(u_l) = \Phi(u_r)$ since v and w are constant.

Remark 4.1 *This creates a bound on possible w-values, $w \in [-\Phi(u_{f-}), -\Phi(u_{f+})]$ where $D(u_{f\mp}) = 0$, i.e., confined to the region between the local extrema of Φ; see Figure 1.*

Without loss of generality, set $H(u_l, \hat{u} = 0) = 0$, i.e., $H(u, \hat{u}) = \frac{\hat{u}^2}{2} - \int_{u_l}^{u} (w + \Phi(u)) du$. Then $H(u_r, \hat{u} = 0)$ must be equal zero as well for the existence of a layer connection between these two points. This constraint leads to the well-known *'equal area rule'* (see, e.g. [38]),

$$\boxed{\int_{u_l}^{u_r} (w_h + \Phi(u)) du = 0} \,. \tag{21}$$

This rule allows for $S_s^{l/r}$ to $S_s^{r/l}$ connections, but not to the boundaries $F_{l/r}$ or the centre-type middle branch S_m. Due to the symmetry $(\hat{u}, y) \leftrightarrow (-\hat{u}, -y)$ in (19), there exists automatically a pair of such heteroclinic connections for fixed $w = w_h$, i.e., $\Gamma_+(w_h, 0) : p_l \to p_r$ and $\Gamma_-(w_h, 0) : p_r \to p_l$; see Figure 4.

Remark 4.2 *The equal area rule* (21) *determines the value $w = w_h$ for which this integral vanishes. Since $\delta = ac = 0$ there are two possible cases: for $a = 0$, it is independent of the possible wave speed $c \in \mathbb{R}$. On the other hand, for $c = 0$ it is independent of the viscous relaxation regularisation contribution $a \in \mathbb{R}$. Hence, the only shock-fronted standing waves that our regularised model can produce are those satisfying the equal area rule.*

4.1.2 The $\delta \neq 0$ case

In general, heteroclinic orbits $\Gamma^{\pm}$ of the 2D layer problem (13) connecting $S_s^{l/r}$ to $S_s^{r/l}$ are confined to the upper (Γ_+) or lower (Γ_-) half-plane in $(u, \hat{u})$-space. In these half-planes, the u-dynamics is monotone. Hence, all heteroclinics $\Gamma_\pm$ are graphs over the u-coordinate chart in $(u, \hat{u})$-space, i.e., $\Gamma_\pm : \hat{u}(u) : u \in (u_l, u_r)$. We consider Γ_+ here (the same works for Γ_-). Such a heteroclinic orbit $\hat{u}(u)$ must fulfill

$$\begin{aligned} & \frac{d\hat{u}}{du} = \frac{w + \Phi(u) - \delta\hat{u}}{\hat{u}}\,, \quad \forall u \in (u_l, u_r) \\ \Longrightarrow\ & \frac{d}{du}(\frac{\hat{u}^2}{2}) = \frac{d}{du}\int (w + \Phi(u) - \delta\hat{u})du\,, \quad \forall u \in (u_l, u_r) \\ \Longrightarrow\ & \frac{\hat{u}^2}{2} = \int_{u_l}^{u} (w + \Phi(u) - \delta\hat{u})du\,, \quad \forall u \in (u_l, u_r)\,. \end{aligned}$$

For $u \to u_l$, the last line is fulfilled since $\hat{u}(u_l) = 0$. For $u \to u_r$, where $\hat{u}(u_r) = 0$, we obtain a condition for the existence of a heteroclinic orbit,

$$\boxed{\int_{u_l}^{u_r} (w + \Phi(u))du = \delta \int_{u_l}^{u_r} \hat{u}(u)\, du\,,} \tag{22}$$

which, for $\delta = 0$, gives the equal area rule as established previously. For $\delta \neq 0$ this formula provides a generalised 'equal area rule', i.e., the left hand side must move away from its 'equal area' position given for $w = w_h(0)$ to counteract the right hand side contribution. This gives $w = w_h(\delta)$ for heteroclinic connections defined by the generalised equal-area rule (22). Figure 5 (left) shows an example of a heteroclinic for $\delta \neq 0$.

For sufficiently large $|\delta| \geq \delta_m$, w will necessarily reach its limit w_{sn} where one of the saddle equilibria $p_{l/r}$ goes through a saddle-node bifurcation. Until then, the heteroclinic connection is along the hyperbolic direction, but afterwards it will be along the centre direction which is non-unique and, hence, replaces the codimension-one role of the w-variation. Thus, for fixed $w = w_{sn}$ and for sufficiently large $|\delta| > \delta_m$, there always exists a heteroclinic orbit located at the boundary of the admissible jump zone. Figure 5 (right) shows an example of a heteroclinic orbit for $\delta \approx \delta_m$ connecting S_s^r to F_l.

Remark 4.3 *For $w = w_{sn}$, the left-hand side of the generalised equal-area rule* (22) *is fixed. One concludes that for sufficiently large $|\delta| > \delta_m$, there exists a $\hat{u}(u)$ that fulfills the generalised equal are rule, i.e., $\hat{u}(u)$ fixes the right hand side $\delta \int \hat{u}\, du$ to the correct/desired value.*

Figure 6 summarizes our results on the existence of shocks in the regularised RND model, i.e., the solution branches of $\Delta(w, \delta) = 0$. Under δ-variation, the shock connection varies continuously in the 'jump zone' between the height specified by $w = w_h(0) = -\Phi(u_{infl})$, where u_{infl} denotes

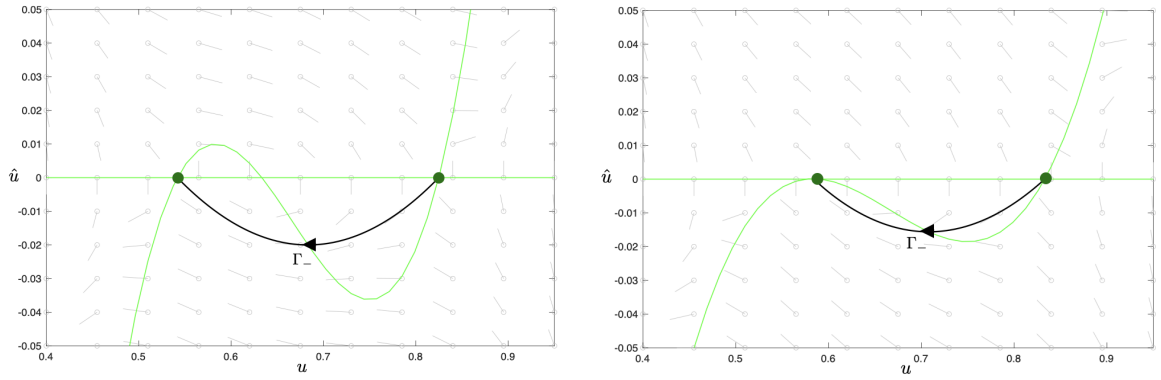


Figure 5: (left) heteroclinic Γ_- for $\delta = 0.1$ and $w = w_h(\delta) \approx -0.5661$, (right) border case heteroclinic Γ_- for $\delta = \delta_m \approx 0.248$ and $w = w_{sn} \approx -0.5671$; other parameter values: $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$.

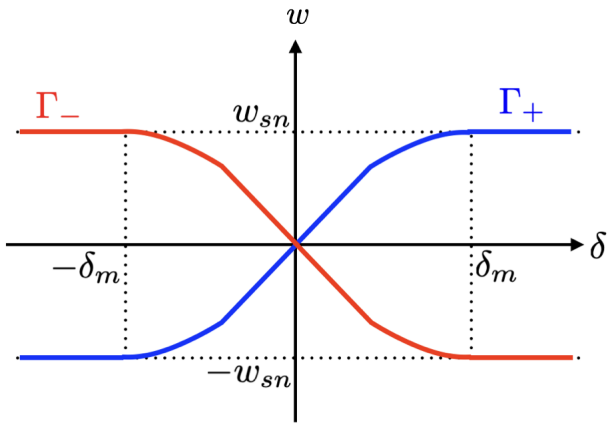


Figure 6: Sketch of complete bifurcation diagram for heteroclinic connections $\Gamma_\pm$ in (δ, w)-space centered at $(w_h(0), 0)$.

the inflection point of the cubic (corresponding to the equal area rule), and a height specified by $w = w_{sn} = -\Phi(F_{l,r})$, where $F_{l,r}$ denote the u-values of the fold points of the critical manifold. One important insight here is that viscous relaxation is the dominant regularising effect for $|\delta| > |\delta_m|$ for shock location selection (in the layer problem).

4.2 Concatenating singular heteroclinic orbits $\Gamma^\pm_{het}$

According to our analysis in the layer problem, we find a curve $w = w_h(\delta)$ in (δ, w) space for which there exist singular fast jumps $\Gamma_\pm$ connecting the outer two branches of the critical manifold; see Figure 6. In the following we aim to concatenate these fast jump segments $\Gamma_\pm$ with slow segments $\Gamma_{l/r}$ of the reduced problem that connect to the given end states $u_\pm$ and, thus, obtain singular heteroclinic orbits $\Gamma^\pm_{het}$ (18) representing singular shock fronted travelling (or standing) waves of our composite regularised RND problem (9).

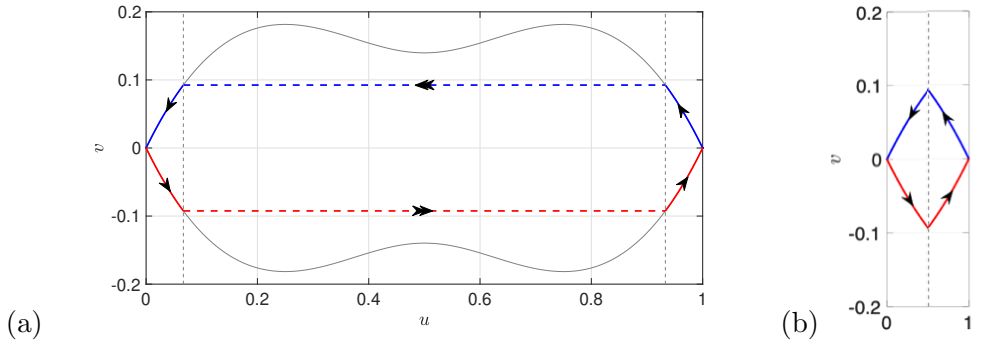


Figure 7: A pair of symmetric standing waves (solid curve segments connected by dashed lines representing shock discontinuities of the solutions, red and blue online) of the desingularised travelling wave equations (5), subject to the equal area shock selection rule; Figure 2a. (b) Corresponding reduced heteroclinic connection problem

We will start with the construction of singular heteroclinic connections in the 'Cahn-Hilliard'-type regularisation limit $a = 0$. Then, we will vary a to produce new families of shock-fronted travelling waves under composite regularisation $a \neq 0$. The system parameters $\beta, \kappa > 0$ are always regarded as fixed for simplicity.

4.2.1 Singular standing and slowly-moving shocks: the $a = 0$ case

As discussed in Section 2.1, the locus of fully symmetric standing waves is specified in $(\gamma_1, \gamma_2, \alpha, c)$-parameter space by the line segment

$$\mathcal{L} = \{(\gamma_1, 1 - \gamma_1, 1/2, 0) : \gamma_1 \in (0, 1/2)\}.$$

We apply a piecewise-smooth variant of the Melnikov method (see Appendix A.2) to show rigorously that $\mathcal{L}$ locally separates (singular) invasion and evasion fronts. We consider a parameter variation in (α, c)-space, where we remind the reader that α determines the location of the middle root of the reaction term $f(u)$. In order to set up the calculation, we first write down the piecewise-defined planar system of the form (49), whose trajectories are topologically conjugate to those of (5) subject to the equal area rule shock condition. Letting $x = (u, v)$, we have

$$\dot{x} = \begin{cases} h_-(x; \alpha, c) := g(x; \alpha, c) & \text{if } u < u_L \\ h_+(x; \alpha, c) := g(x - (u_L - u_R, 0)^T; \alpha, c) & \text{if } u > u_L. \end{cases} \tag{23}$$

We define Σ as a suitable compact segment of $\{u = u_L\}$ intersecting the piecewise-smooth heteroclinic orbit $\gamma(t)$, which is constructed by shifting the right portion of the singular heteroclinic orbit in Fig. 7(a) to the left by $(u_L - u_R)$ and then selecting the phase so that $\gamma(0) \in \Sigma$; see Fig. 7(b). As before, we define a function $\Delta(\alpha, c)$ measuring the distance between the (un)stable manifolds of the corresponding saddle points on Σ, so that $\Delta(1/2, 0) = 0$. We seek to apply the implicit function theorem to obtain a local path

$$c = c(\alpha) = b(\alpha - 1/2) + \mathcal{O}((\alpha - 1/2)^2)$$

so that $\Delta(\alpha, c(\alpha)) = 0$ for $\alpha \in (1/2 - \alpha_0, 1/2 + \alpha_0)$, i.e. it suffices to show that $D_c\Delta(1/2, 0) \neq 0$. The corresponding leading-order coefficient is then given by

$$b = -\frac{D_\alpha \Delta(1/2, 0)}{D_c \Delta(1/2, 0)}. \tag{24}$$

The Melnikov integrals (see (54)) involve the partial derivatives $D_\alpha g(u, v)$ and $D_c g(u, v)$. For $u < u_L$, we have

$$\begin{aligned} D_\alpha g(u, v) &= (0, \kappa u(u-1)D(u)) \\ D_c g(u, v) &= (-u, 0), \end{aligned} \tag{25}$$

and similar expressions can be derived for $u > u_L$ using the piecewise-smooth problem (23). For $(\alpha, c) = (1/2, 0)$, we define a pair of exponentially decaying solutions to the relevant adjoint problem $\psi_\pm$ on either segment $\{u < u_L\}$ and $\{u > u_L\}$, with the orientation at $t = 0$ chosen so that the first components satisfy

$$\begin{aligned} &\psi_{-,1}(s) > 0 \text{ for } s < 0 \text{ and} \\ &\psi_{+,1}(s) < 0 \text{ for } s \geq 0, \end{aligned}$$

i.e. the adjoint solution points to the right on S^r_s and to the left on S^l_s. We highlight that

$$\psi_{-,1}(s) = -\psi_{+,1}(-s)$$

for each $s < 0$; this property is inherited from the vertical reflection symmetry

$$h_{-,2}(u, v; 1/2, 0) = -h_{+,2}(2u_L - u, v; 1/2, 0)$$

satisfied by the vector fields for $u < u_L$.

We now compute the Melnikov integrals defined by (54). Observe that $v_1^- = v_1^+$ at the intersection point $\gamma(0)$, and hence it suffices to compute the 'classical' Melnikov integrals to obtain the leading-order coefficient (24) since the common prefactor $(1/v_1^\pm)$ cancels upon evaluating the ratio.

Let us denote by $\tilde{u} = u - (u_L - u_R)$ the value of u shifted by the shock segment. Dropping the common prefactor, the corresponding Melnikov integral for $D_\alpha\Delta(1/2, 0)$ is

$$\begin{aligned} &\int_{-\infty}^0 \psi_-(s)^\top D_\alpha h_-(\gamma_-(s); 1/2, 0)\, ds + \int_0^\infty \psi_+(s)^\top D_\alpha h_+(\gamma_+(s); 1/2, 0)\, ds \\ &= \int_{-\infty}^0 \psi_{-,2}(s)\kappa u(s)(u(s)-1)D(u(s))\, ds + \int_0^\infty \psi_{+,2}(s)\kappa\tilde{u}(s)(\tilde{u}(s)-1)D(\tilde{u}(s))\, ds < 0 \end{aligned}$$

since $\psi_{\pm,2}(s) > 0$, $\kappa u(s)(u(s)-1)D(u(s)) < 0$, and $\kappa\tilde{u}(s)(\tilde{u}(s)-1)D(\tilde{u}(s)) < 0$ along the heteroclinic connection.
The remaining Melnikov integral for $D_c\Delta(1/2, 0)$ is

$$\begin{aligned} &\int_{-\infty}^0 \psi_-(s)^\top D_c h_-(\gamma_-(s); 1/2, 0)\, ds + \int_0^\infty \psi_+(s)^\top D_c h_+(\gamma_+(s); 1/2, 0)\, ds \\ &= -\left(\int_{-\infty}^0 \psi_{-,1}(s)u(s)\, ds + \int_0^\infty \psi_{+,1}(s)\tilde{u}(s)\, ds \right). \end{aligned}$$

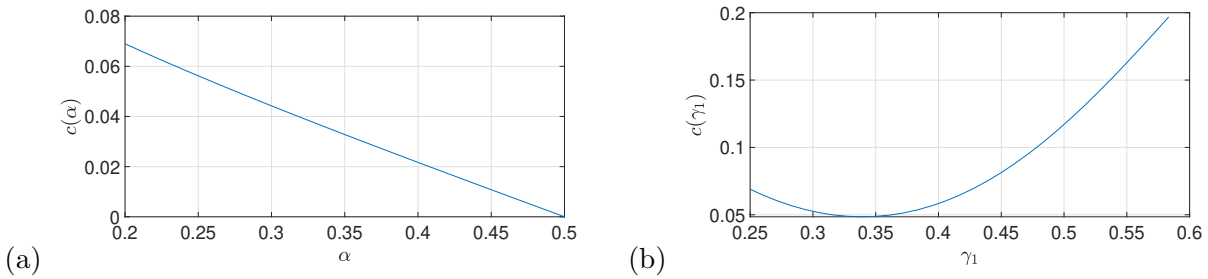


Figure 8: One-parameter continuations of a singular heteroclinic orbit joining the saddle points at $u = 0$ and $u = 1$ for the desingularised slow flow (5), subject to shock selection given by the equal area rule. Initial parameter set: $\gamma_1 = 1/4$, $\gamma_2 = 1 - \gamma_1 = 3/4$, $\alpha = 1/2$, $a = 0$, $c = 0$, $\kappa = 5$, $\beta = 6$. (a) The continuation parameter α is varied from $\alpha = 0.5$ to $\alpha = 0.2$ with c as the free parameter. (b) The continuation parameter γ_1 is then varied from $\gamma_1 = 1/4$ to $\gamma_1 = 7/12 = 0.58\bar{3}$, with c as the free parameter. Note that $c(7/12) = c_* \approx 0.19686$.

In this case, observe that the sign of $\psi_{\pm,1}(s)$ does change across the shock, so we must inspect the integrands more closely. Using the fact that the horizontal components $\psi_{\pm,1}$ have a reflection antisymmetry, while the corresponding measure $\tilde{u}(s)\,ds$ is unevenly weighted toward the heteroclinic segment on $u > u_L$, we conclude that $D_c\Delta(1/2, 0) < 0$ as well. Hence, from (24) we have that $b < 0$.

This result implies that invasion fronts (with $c > 0$) emerge along a locally affine branch as α is varied below $\alpha = 1/2$, resp. evasion fronts ($c < 0$) for $\alpha > 1/2$. By symmetry, a reflected connection from $u = u_-$ to $u = u_+$ coexists for $c = 0$. An analogous Melnikov calculation gives a distinct local branch of bifurcations connecting $u = u_+$ to $u = u_-$. The local bifurcation structure is therefore similar to the one for the shock fronts in Fig. 6. The invasion portion of one of these locally affine branches (i.e. $\alpha < 1/2$, $c > 0$) is depicted in Fig. 8(a) for an example parameter set.

4.2.2 Continuation of singular heteroclinic orbits

This local patch of singular heteroclinic connections serves as a natural starting point for global continuation. For instance, we can demonstrate that the singular heteroclinic connection previously identified in the 'Cahn-Hilliard'-type regularisation setting [29, 30], corresponding to the parameter set

$$\gamma_1 = 7/12,\ \gamma_2 = 3/4,\ \alpha = 1/5,\ a = 0,\ c_* \approx 0.19686,\ \kappa = 5,\ \beta = 6,$$

is 'accessible' from $\mathcal{L}$ via numerical continuation, i.e., they lie on the same path-connected component of the submanifold of heteroclinic connections. Concretely, we begin with the parameter set

$$\gamma_1 = 1/4,\ \gamma_2 = 1 - \gamma_1 = 3/4,\ \alpha = 1/2,\ a = 0,\ c = 0,\ \kappa = 5,\ \beta = 6$$

on $\mathcal{L}$, giving rise to a symmetric standing wave (depicted in Fig. 2). We then perform a sequence of one-parameter continuations: first we vary α from $\alpha = 1/2$ to $\alpha = 1/5$ (Fig. 8(a)), and then we vary γ_1 from $\gamma_1 = 1/4$ to $\gamma_1 = 7/12$ (Fig. 8(b)), both times leaving c free. As expected, the wavespeed $c = c_*$ found after these continuations agrees with the value computed in [29, 30], and the corresponding singular connections are identical.

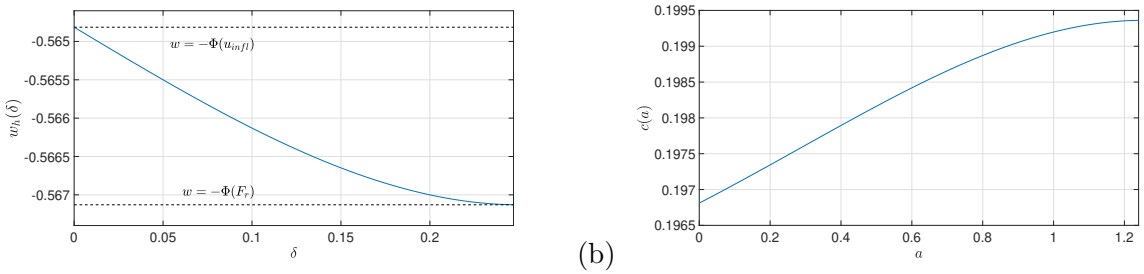


Figure 9: (a) Bifurcation diagram in (δ, w)-space for the shock height selection in the layer problem, corresponding to a segment of the branch Γ_-, as shown in Fig. 6. Here, $\delta_m \approx 0.248$. (b) Bifurcation diagram in (a, c)-space for the singular heteroclinic orbits. Parameter set: $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$, $\kappa = 5$, $\alpha = 1/5$.

Remark 4.4 *We can also fix $c = 0$ and choose some other free parameter (e.g. we can trace families of asymmetric standing waves). Note that the reduced problem* (5) *remains Hamiltonian in this case (see Remark 2.2); but in view of the jump condition, the disjoint slow portions of the singular heteroclinic connections that we seek will typically lie on different energy surfaces. These slow segments happen to lie on the same level set in the symmetric case (e.g. Fig. 2), but this scenario is exceptional.*

Altogether, the line of symmetric standing waves can be continued to a regular codimension-1 submanifold of singular heteroclinic bifurcations, by e.g. letting the first three parameters vary and leaving the wavespeed c as a free parameter.

4.2.3 Singular heteroclinic connections for $a > 0$

Let us now extend the parameter space in the above analysis by the regularisation weighting parameter a. We demonstrate that the heteroclinic orbits in the previous section persist as transversal intersections of (un)stable manifolds of the saddle points in the extended parameter space.

Let $a \geq 0$ be specified; then as c varies, the corresponding extreme values $u = u_L(c)$, $u_R(c)$ selected by the shock also vary continuously according to the height condition $w_h(\delta) = w_h(ac) = -\Phi(u)$ specified by the layer problem; see Fig. 9(a).

Varying c simultaneously varies the (un)stable manifolds of the saddle points in the reduced problem. Suppose that the stable manifold $W^s(u_+, c)$ of the saddle point at $u_+ = 0$ has its first intersection with the cross-section $\{u = u_L(c)\}$ at a point $p_l(c)$, and similarly that the unstable manifold $W^u(u_-, c)$ of the saddle point at $u_- = 1$ has its first intersection the cross-section $\{u = u_R(c)\}$ at $p_r(c)$. Under variation of c we can then locate a locally unique value $c = c(a)$ at which the v-coordinates of $p_l(c)$ and $p_r(c)$ coincide; i.e., so that the shock simultaneously connects two slow trajectories that join the two saddle points.

Altogether, we have defined a bifurcation problem for the singular heteroclinic orbits with respect to the regularisation weighting parameter a; see Figure 9(b) for the resulting bifurcation diagram

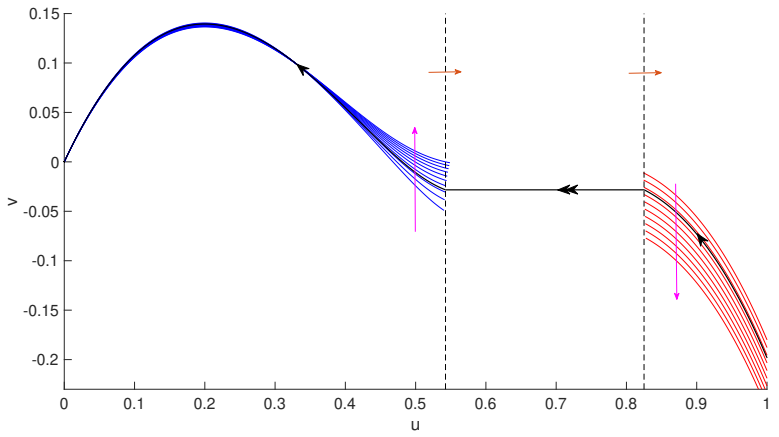


Figure 10: Singular heteroclinic orbit for $(a_*, c_*) \approx (0.5182, 0.19817)$, formed as a transversal intersection of the stable manifold $W^s(u_0, c)$ (blue curves) and unstable manifold $W^u(u_1, c)$ (red curves) as c is increased within the interval $[0.18, 0.25]$. Jump values $u = u_L(c_*)$, $u_R(c_*)$ satisfying $-w_h(\delta_* = a_* c_*) = \Phi(u)$ denoted by dashed black lines. Vertical magenta arrows: direction of variation of (un)stable manifolds; horizontal orange arrows: variation of the endpoints of the shock, as c increases. Parameter set: $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$, $\kappa = 5$, $\alpha = 1/5$.

for our example parameter set. Figure 10 depicts an example of a singular heteroclinic connection formed under variation of the wavespeed parameter c for fixed $a > 0$. We emphasize that each such invasion shock front formed within the parameter interval $a \in [0, a_m]$, with $a_m \approx 1.2465$, satisfies a distinct generalised area rule! For the parameter set in the figure, a fixed wavespeed $c_m \approx 0.1994$ is selected for each $a > a_m$, since this is the wavespeed at which the portions of the singular heteroclinic connection that lie on the critical manifold connect to a viscous shock at the fixed height specified by $w_h(\delta) = w_{sn} = -\Phi(u)$, which is in turn satisfied for each $\delta > \delta_m$ as depicted in Figure 6.

4.3 Persistence of heteroclinics for $a \geq 0$ and $\varepsilon > 0$

After checking standard transversality conditions with respect to slow invariant manifolds and fast fibre bundles arising from Fenichel theory [7], and then applying well-known GSPT estimates (e.g. the Exchange Lemma), this family of singular heteroclinic orbits defined for $\varepsilon = 0$ generically perturbs to a codimension-one manifold of heteroclinic bifurcations of (12) in (a, c, ε) parameter space. An example of such an orbit, which satisfies a generalised area rule for a nontrivial value of $a > 0$, is depicted in Figure 11.

4.4 (Non)monotone wave transition and termination

We now use numerical continuation to explore the eventual fate of the monotone waves we have constructed. We choose α (which specifies the middle root of the reaction term) as the continuation parameter, and we focus on the singular limit $\varepsilon = 0$. Varying α within the interval $0 < \alpha < \gamma_1$, the corresponding fixed point $(\alpha, -c\alpha)$ of the desingularised problem begins to play a significant role in

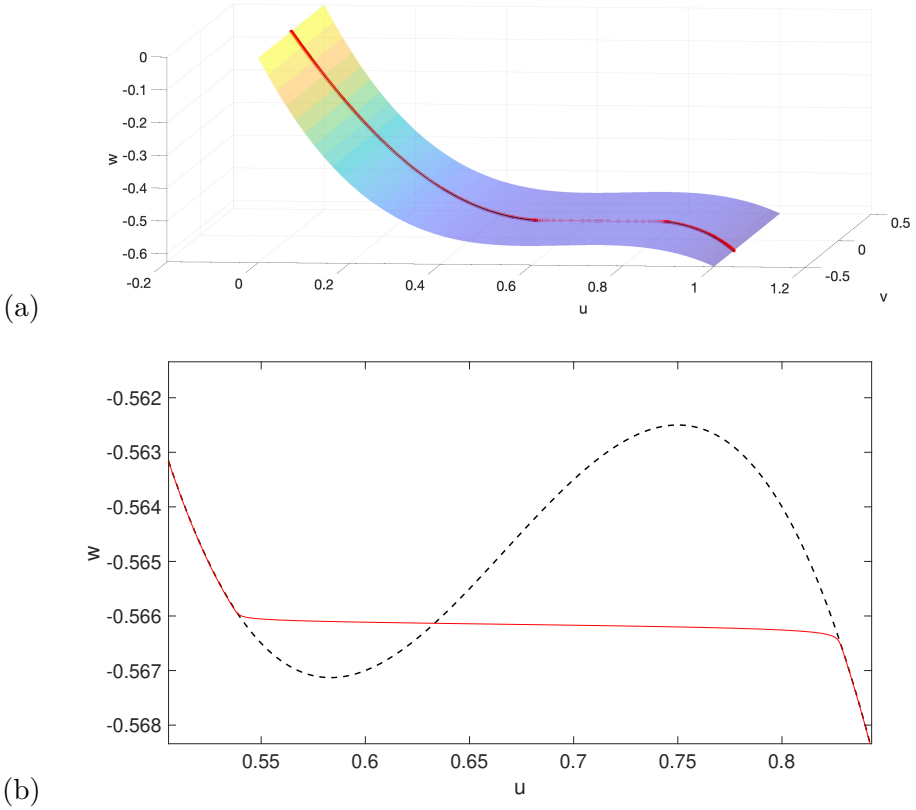
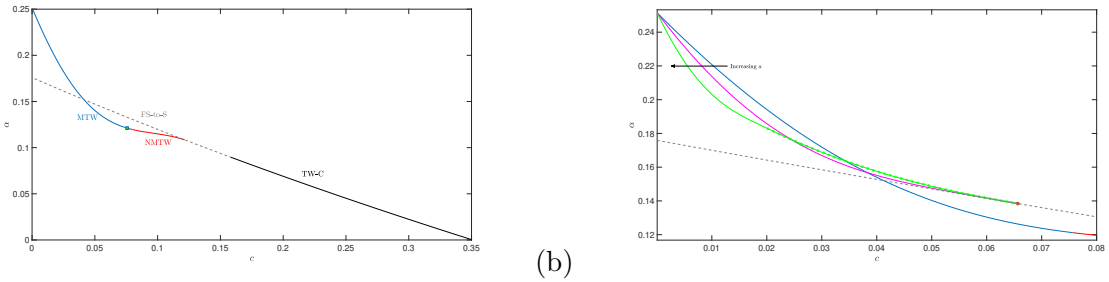


Figure 11: (a) A (u, v, w)-projection of a heteroclinic orbit of (12) for $\varepsilon = 10^{-4}$, $a \approx 0.5182$, and $c = 0.19826$. Black curve segments on the critical manifold (obscured by red curve) denote the slow portions of the corresponding singular heteroclinic orbit depicted in Fig. 10. (b) Side view of the heteroclinic orbit demonstrating that the connection satisfies the perturbed generalized area rule $w_h(\delta) = -\Phi(u)$. Parameter set: $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$, $f_m = 5$, $\alpha = 1/5$. Computations performed with the `bvp4c` boundary value solver in MATLAB 2021a, with a relative error tolerance of 10^{-5}.

the curving of the stable manifold $W^s(0)$ on S_s^l. At the same time, the folded singularity at $u = u_{f_l}$ is now of folded saddle (FS) type (see Table 2), allowing for the existence of canard solutions crossing from the stable middle branch (when $\delta > 0$) of the critical manifold into the saddle-type left branch.

The interplay between these equilibria, together with the relevant shock rule defined by $\delta = ac$, controls the termination of monotone waves, as well as the the emergence of two new types of solutions: *nonmonotone*[4] shock-fronted travelling waves, and shock-fronted travelling wave solutions containing nontrivial slow passage through regions of negative diffusion (i.e. the wave trajectory now has a canard segment, visiting the slow manifold near the middle branch S_m for $O(1)$ periods of 'time' with respect to the travelling wave frame coordinate z). We now show how all of these new phenomena are organised and explained by global singular bifurcations.

[4](Non)monotonicity here refers always to the quantity u.



(a) (b)

Figure 12: (a) Singular bifurcation diagram depicting a monotone (blue, MTW) to nonmonotone (red, NMTW) shock-fronted travelling wave transition, via a codimension-two tangency bifurcation (green square) of the stable manifold of p_0 with the landing curve. The family of nonmonotone waves terminates on the folded-saddle-to-saddle heteroclinic bifurcation (FS-to-S) curve (grey dashed curve). For each fixed $a > 0$, the FS-to-S curve contains an open subset (black solid curve) where shock-fronted travelling waves with canard segments (TW-C) begin to appear. Parameter set: $\beta = 1$, $\gamma_1 = 0.4$, $\gamma_2 = 0.75$, $\kappa = 3$, $a = 0.5$. For this parameter set, we have $\delta_m \approx 0.2121$. (b) (Non)monotone wave bifurcation curves for three different values of a: (i) $a = 0.5$ (blue and red solid), (ii) $a = a_* \approx 3.2304$ (magenta solid), and (iii) $a = 10.605$ (green solid and with bubble markers). Solid red, blue, magenta, and green curves denote interpolated shock rules and the green curve with bubble markers denotes the transition to monotone waves with viscous-type shocks. Orange dot: codimension-three scenario for $(a_*, c_*, \alpha_*) \approx (3.2304, 0.0657, 0.1384)$ in which the monotone waves terminate at an FS-to-S connection precisely as the shock rule changes from interpolated type to viscous type. Remaining parameters as in (a).

Let us first continue the family of shock-fronted travelling waves in the parameter α for small values of $a > 0$. For decreasing α, the stable manifold $W^s(0)$ eventually develops a tangency to the jump curve on S_s^l at the end of the shock; see the transition from the blue to red curves in Fig. 12(a) and the tangency in phase space in Fig. 13(a).This global singular bifurcation heralds the termination of the family of monotone waves and the birth of nonmonotone shock-fronted travelling waves at increasing wavespeeds; see Fig. 13(b). This new family of nonmonotone waves can be continued in α, where they too eventually terminate in a one-parameter family of folded saddle-to-saddle (FS-to-S) heteroclinic connections on S_s^l.

Choosing larger fixed values of $a > 0$ and repeating this continuation, the endpoints of the shock begin to move to the right relatively quickly as the wavespeed c increases, as a result of the dependence of the shock selection rule on the parameter $\delta = ac$. Indeed, for sufficiently large $a > 0$, the endpoints of the shock move too quickly for a tangency with $W^s(0)$ to form, before the shock rule changes from 'interpolated' type to 'viscous' type for $\delta > \delta_m$. The shockwaves then remain monotone until they terminate at the FS-to-S branch; see the green curve, and in particular the transition from interpolated shock rules (green solid) to viscous shock rules (green with bubble markers), in Fig. 12(b).

We numerically identify the existence of an intermediate value $a = a_*$ at which the family of monotone waves terminate on the FS-to-S branch at a point (c_*, α_*) precisely when the shock selection rule changes from 'interpolated' to 'viscous' type; see e.g. the magenta curve terminating

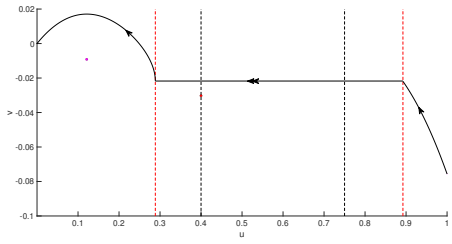
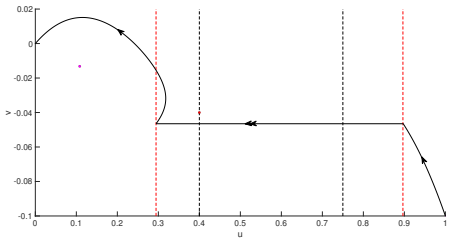
(a)  (b) 

Figure 13: (a) Travelling wave at the moment of tangency of the stable manifold of the left saddle point with the landing curve of the shock. Parameters: $(c, \alpha) \approx (0.0756, 0.1212)$ (corresponding to the location of the green square in Fig. 12(a)). (b) A (singular) nonmonotone travelling wave for $(c, \alpha) \approx (0.1, 0.115)$. Red point: folded saddle at $(u_{fl}, -cu_{fl})$; magenta point: equilibrium at $(\alpha, -c\alpha)$. Remaining parameters: $\beta = 1$, $\gamma_1 = 2/5$, $\gamma_2 = 3/4$, $\kappa = 3$, $a = 1/2$.

at the orange dot on the FS-to-S curve in Fig. 12(b). This boundary case can be interpreted as a codimension-three global singular bifurcation for $(a, c, \alpha) = (a_*, c_*, \alpha_*)$, whose unfolding includes the emergence of nonmonotone traveling waves.

4.4.1 Shock-fronted travelling waves with canard segments

We now turn to the existence of shock-fronted travelling wave solutions containing canard segments passing through the FS point. We can continue the codimension-one family of FS-to-S heteroclinic connections (the grey dashed curve in Fig. 13(a)). Recalling that S_m is attracting when $\delta > 0$, it is then reasonable to search for parameter values along this branch for which the singular stable fast fibre bundle over the unstable manifold of the FS point (with respect to the orientation-reversed desingularised problem) on S_m intersects transversely the unstable fast fibre bundle sitting over $W^u(1, -c)$ (in $\mathbb{R}^4$). Such transverse intersections persist robustly under parameter variation, i.e. the codimension-one FS-to-S manifold will intersect the parameter set where such transverse crossings exist in an open neighborhood. See the black solid curve (labeled TW-C) representing such a subset in Fig. 12(a).

These transverse crossings along the FS-to-S curve give rise to shock-fronted travelling waves containing singular canard segments that connect S_m to S_s^l. See Fig. 14 for an example; here, the relevant transverse intersection of the fast fibre bundles is depicted in the figure by projecting the corresponding segment of $W^u(1, -c)$ onto S_m via the layer flow.

Remark 4.5 *Canards may also arise via slow passage through a folded node (FN) point. However, the stability classification of the critical manifold and the 'stability' of the FN singularities (relative to the desingularised problem) do not appear to be compatible with the existence of an invasion shock front (i.e. with $c > 0$) connecting the unstable manifold of the equilibrium point at $u = 1$ to the stable manifold of the equilibrium point at $u = 0$, and also containing a canard segment passing through a FN point.*

Let us discuss the relevant 'shock selection rule' (i.e. allowable layer connections) in this context; in particular, we highlight the importance of the regularisation weighting parameter a in the formation

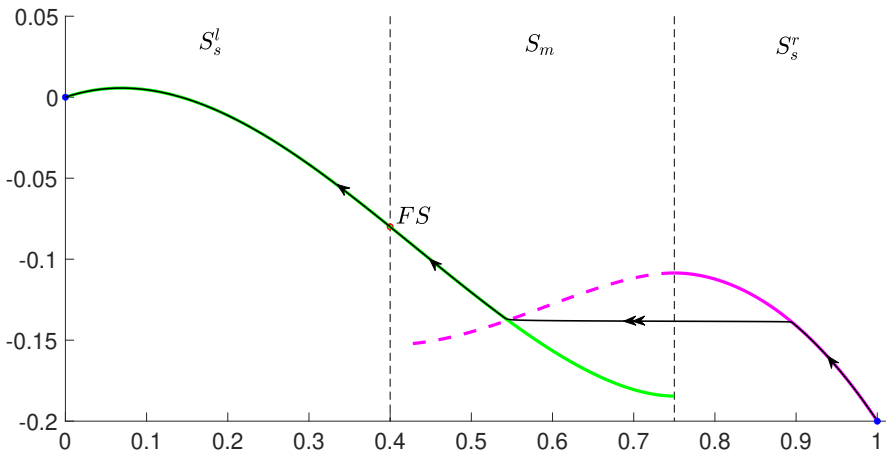


Figure 14: Singular heteroclinic orbit with a singular (vrai) canard segment through a folded saddle (FS) point. Green curve: stable manifold $W^s(0,0)$; magenta solid curve: unstable manifold $W^u(1,-c)$; magenta dashed curve: segment of the projection of $W^u(1,-c)$ onto S_m; black dashed lines: fold curves. Parameter set: $\beta = 1$, $\gamma_1 = 2/5$, $\gamma_2 = 3/4$, $\kappa = 3$, $\alpha \approx 0.068984$, $c = 0.2$.

of such shocks. This parameter does not appear in the desingularised slow problem, but it plays a crucial role in determining the allowable heights w of the corresponding fast fibers that can connect from S_s^r to S_m. In particular, $\delta = ac$ must be *sufficiently large* so that the relevant segment of $W^u(1,-c)$ can be projected onto S_m via the layer flow to transversely intersect the appropriate branch of the unstable manifold of the FS point. This constraint arises from the direction in which the (un)stable manifolds break apart away from the heteroclinic bifurcation curve Γ_- in Fig. 6. In particular, suitable connections from S_s^r to S_m are found in the open region to the *top-right* of Γ_- only.

Remark 4.6 *It is an interesting problem to determine parameter sets such that the (non)monotone travelling wave branches intersect the travelling wave-with-canard branch (e.g. intersections of the red and/or blue curves with the solid black curve in a bifurcation diagram drawn as in Fig. 12(a)). In such a scenario, these intersections would correspond to the simultaneous coexistence of two distinct types of travelling wave solutions in the singular limit, and could be interpreted as codimension-two branch switching bifurcations.*

Remark 4.7 *It would also be interesting to investigate the (non)monotonicity of the travelling waves with canards. These waves are apparently monotone near the critical manifold; see e.g. Fig. 14. However, the fast connection from S_s^r to S_m may introduce winding in the shock layer, because S_m contains subsets where the nontrivial layer eigenvalues become complex, i.e. a shock departing from S_s^r can wind around as it approaches S_m.*

5 Spectral stability of monotone shock-fronted travelling waves

We now assess the stability of the monotone travelling waves constructed in the previous section, with a focus on extending the analysis initiated in [30] and [31]. We adopt the standard approach

of determining the spectral stability of the linearised operator L associated with the PDE (1) near the travelling wave. Our first step is to write down a suitable coordinate representation of L.

Let $\tilde{u}(z,t) = u(z) + \nu e^{\lambda t} p(z) + \mathcal{O}(\nu^2)$ denote a perturbation of a travelling wave $u(z)$ of (1), where λ is the temporal eigenvalue parameter and the variables in the linear term are assumed to separate. Inserting this solution into (1) and collecting terms of linear order in ν, we obtain the equation

$$(f'(u) - \lambda)p = -(cp_z + (D(u)p)_{zz} + \varepsilon a(\lambda p_{zz} - cp_{zzz}) - \varepsilon^2 p_{zzzz}). \tag{26}$$

Defining linearised variables $y := (p, q, r, s)$ corresponding to the nested derivatives on the right (analogously to what is done for the travelling wave system (11)), we arrive at the following nonautonomous linear system:

$$\begin{aligned} \varepsilon \dot{p} &= q \\ \varepsilon \dot{q} &= (D(u) + \varepsilon a\lambda)p + s - \delta q \\ \dot{r} &= (f'(u) - \lambda)p \\ \dot{s} &= r + cp, \end{aligned} \tag{27}$$

or more compactly,

$$\dot{y} = M(z, \lambda, \varepsilon)y, \tag{28}$$

where $M(z, \lambda, \varepsilon)$ is a convenient matrix representation of L. We highlight the terms $\varepsilon a\lambda p$ and $-\delta q$ arising due to the viscous relaxation contribution.

By general theory (see e.g. [22]), the spectrum $\sigma(L)$ of L can be decomposed into its *point* and *essential* spectrum $\sigma(L) = \sigma_p(L) \cup \sigma_c(L)$. Our task is to ensure that the spectrum is bounded within the left half complex plane, except for an eigenvalue at the origin that necessarily exists due to translational invariance. We must also check that the translational eigenvalue is simple.

5.1 Essential spectrum and sectoriality

We first give a brief overview of recent results for the essential spectrum. In [31], it was shown that for each $a \geq 0$, the essential spectrum is bounded well inside the left-half plane. To briefly summarise the approach, the *Fredholm borders* of $\sigma_e(L)$ are computed by tracking changes in the Fredholm index of the asymptotically constant matrices $M_\pm(\lambda, \varepsilon) := \lim_{z \to \pm\infty} M(z, \lambda, \varepsilon)$, which characterise the hyperbolic dynamics near the tails of the wave. We obtain the following parametrisations for the dispersion relations (with $k \in \mathbb{R}$):

$$\lambda_\pm(k) = \frac{f'(u_\pm) - D(u_\pm)k^2 - \varepsilon^2 k^4}{1 + a\varepsilon k^2} + ick. \tag{29}$$

We first verify that the essential spectrum lies entirely within the left half plane. This can be seen by noting that the denominator of the real part of $\lambda_\pm(k)$ in (29) is strictly positive, and then applying Descartes' rule of signs to the numerator, noting that $f'(u_\pm) < 0$, and $D(u_\pm)$ and ε are both positive. Since the numerator is an even polynomial, there are no real roots of the real part of $\lambda_\pm(k)$. Therefore, the essential spectrum is entirely contained in the left half plane.

Let us next recall that a linear operator L is *sectorial* if there exists $M > 0$ and $\eta < \pi/2$ such that $\lambda \notin \sigma(L)$ whenever $|\lambda| > M$ and $|\arg(\lambda)| < \pi - \eta$ (with $-\pi < \arg(\lambda) \leq \pi$) (see Prop 2.2 in [1]; see also Definition 1.3.1 in [16]). Sectoriality provides key resolvent estimates for the linearised operator: in particular, it allows us to conclude *nonlinear* stability from spectral stability, i.e. the existence of a neighborhood of initial conditions of the shock-fronted travelling wave tending to a translate of the wave exponentially quickly as time increases. Simultaneously, this property allows us to deduce the existence of a maximal compact contour K in the complex plane containing all of the point spectrum inside it.

We showed that the essential spectrum is asymptotically vertical (obstructing sectoriality) in the viscous relaxation limit (7) (see [31]), whereas the linearised operator *is* sectorial in the nonlocal limit (8) (see [30]). In this section, we demonstrate that sectoriality persists when $a > 0$. We follow the general approach of the proof of Prop 2.2 in Sec. 5-B of [1]: we identify a suitable rescaling of the linearised variables in (27) in powers of $|\lambda|$, such that the $|\lambda| \to \infty$ limiting system takes an especially simple form. We use this simple form to determine that there is no unstable-to-stable connection made (i.e. there can be no spectrum for sufficiently large values of $|\lambda|$), as long as λ lies inside a suitable sector specified by the constraint on $\arg(\lambda)$.

It turns out that the appropriate choice of rescaling weights depends on whether $a = 0$ or $a > 0$. This dichotomy is not unexpected, in view of the distinct asymptotic behaviour of the dispersion relations (29) in each case: the relations flair to the left with quartic growth when $a = 0$, but with only quadratic growth when $a > 0$.

We first consider the case $a = 0$, repeating the analysis in [30] but with a focus on rescaling the representation (27). Define the rescaled quantities

$$\tilde{p} = p, \quad \tilde{q} = q/|\lambda|, \quad \tilde{r} = r/|\lambda|^3, \quad \tilde{s} = s/|\lambda|^2, \quad \tilde{z} = z|\lambda|.$$

Writing the eigenvalue system (27) in terms of the rescaled variables $(\tilde{p}, \tilde{q}, \tilde{r}, \tilde{s})$ with rescaled 'time' $\tilde{z}$, and then taking the limit $|\lambda| \to \infty$, we arrive at the hyperbolic, constant coefficient linear system

$$\begin{aligned} \varepsilon \dot{\tilde{p}} &= \tilde{q} \\ \varepsilon \dot{\tilde{q}} &= \tilde{s} \\ \dot{\tilde{r}} &= -e^{i\arg\lambda}\tilde{p} \\ \dot{\tilde{s}} &= \tilde{r}. \end{aligned} \tag{30}$$

The (spatial) eigenvalues of (30) are obtained from the roots of the quartic characteristic polynomial

$$e^{i\arg\lambda} + \varepsilon^2 \mu^4 = 0,$$

which can be solved explicitly to obtain the expressions

$$\mu = \frac{e^{i(\arg\lambda + m\pi)/4}}{\sqrt{\varepsilon}}, \quad m = 0,\, 1,\, 2,\, 3.$$

Note that this matches the 'large-scale' eigenvalues in [30] (c.f. eq. (20)), which were calculated using a different choice of rescaling weights. We also highlight that the asymptotic eigenvalue

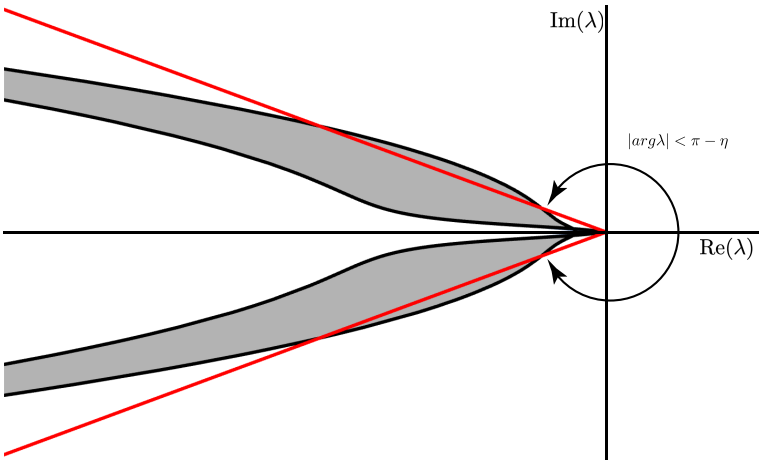


Figure 15: A schematic of the dispersion relations (black) from (29) enclosing the essential spectrum (grey), and an asymptotically bounding sector (red online). A bounding angle η is chosen such that the viscous relaxation contribution causes the dispersion relations to 'flair out' above the sector. The contributions from the Fourier transform of the nonlocal terms eventually take over for sufficiently large values of $|\lambda|$, and the dispersion relations return to being on the left of the bounding sector.

problem is independent of the wavespeed parameter.

Choose any $\eta \in (0, \pi/2)$. Then for each fixed $\varepsilon > 0$, two of these eigenvalues have strictly positive real part and the remaining two have strictly negative real part whenever $|\arg(\lambda)| < \pi - \eta$. Since (30) is autonomous, the corresponding unstable subbundle forms an attractor near to which solutions of the eigenvalue problem remain close by for all y, i.e. for sufficiently large $|\lambda|$ within the sector specified by the choice of η, there can be no contribution to the spectrum of the linearised operator. Furthermore, the eigenvalues remain well separated as $\varepsilon \to 0$.

Now we consider the case $a > 0$. In order to control the additional λ-dependent term in (27) when $|\lambda|$ grows large, we must also include $|\lambda|$ in the rescaling. We introduce an auxiliary rescaling parameter $\sigma > 0$ and write

$$\tilde{p} = p, \quad \tilde{q} = q\sigma, \quad \tilde{r} = r\sigma^3, \quad \tilde{s} = s\sigma^2, \quad \tilde{z} = z/\sigma, \quad |\lambda| = 1/\sigma^2. \tag{31}$$

With respect to this scaling, we have

$$\begin{aligned} \varepsilon\dot{\tilde{p}} &= \tilde{q} \\ \varepsilon\dot{\tilde{q}} &= \left(\sigma^2 D(u) + \varepsilon a e^{i\arg(\lambda)}\right)\tilde{p} + \tilde{s} - \sigma\delta\tilde{q} \\ \dot{\tilde{r}} &= (\sigma^4 f'(u) - \sigma^2 e^{i\arg(\lambda)})\tilde{p} \\ \dot{\tilde{s}} &= \tilde{r} + \sigma^3 c\tilde{p}. \end{aligned} \tag{32}$$

We are concerned with the dynamics near the limit $\sigma \to 0$. As in the purely nonlocal case, the

limiting linear system has constant coefficients, but it is now *nonhyperbolic*:

$$\begin{aligned} \varepsilon\dot{\tilde{p}} &= \tilde{q} \\ \varepsilon\dot{\tilde{q}} &= \varepsilon a e^{i\arg\lambda}\tilde{p} + \tilde{s} \\ \dot{\tilde{r}} &= 0 \\ \dot{\tilde{s}} &= \tilde{r}, \end{aligned} \tag{33}$$

with eigenvalues

$$\mu = 0,\, 0,\, \pm\sqrt{\frac{a}{\varepsilon}}e^{i\arg(\lambda)/2}. \tag{34}$$

The analysis here is more delicate than in the previous case: we do not have access to an attractor (an unstable 2-plane bundle) in the large $|\lambda|$ limit, and the weak (un)stable directions degenerate to a two-dimensional center subspace. We resort to standard perturbation theory. The eigenvalues of (32) can be determined to arbitrary order in σ (i.e. $1/|\lambda|^{1/2}$); we find that the pair of zero eigenvalues perturbs as

$$\mu = \pm\sqrt{\frac{1}{a\varepsilon}}\,\frac{1}{|\lambda|^{1/2}} + \mathcal{O}\left(\frac{1}{|\lambda|}\right). \tag{35}$$

We remind the reader that the system (32) is nonautonomous, but the relevant contributions from $D(u)$ and $f'(u)$ are bounded and do not affect the signs of the two smaller eigenvalues at leading order. An invariant attractor over the unstable subbundle can be constructed explicitly by projectivizing the system and then using the theory of relatively invariant sets for nonautonomous systems (see Sec. B in [11]), but we avoid these technical details here. The argument for sectoriality in terms of bounding angles η then follows as in the purely nonlocal case. See Fig. 15 for a depiction of a bounding sector in relation to the essential spectrum.

Remark 5.1 *The scaling weights n_p, n_q, ... in $\tilde{p} = p/|\lambda|^{n_p}$, $\tilde{q} = q/|\lambda|^{n_q}$, etc. can be chosen such that the limiting system* (30) *is autonomous, and so that the exponent of $|\lambda|$ balances to zero in the slow equation. This rescaling procedure can be interpreted geometrically as an extended Poincaré compactification of the vector field* (27) *'at infinity'; see [44].*

Our choice of weights is also consistent with the scaling derived using the method of dominant balance in the WKB approximation of (26) *(see [3]). Furthermore, the corresponding eigenvalue expansions in* (34)–(35) *match the output of the WKB calculations.*

We have shown that the appropriate cone estimate required for sectoriality holds for each $a \geq 0$, extending the result in [30]. This result has the following interesting implication: it is possible to retain sectoriality (and hence nonlinear stability) for travelling waves containing viscous shocks, when viscous relaxation is counterbalanced by nonlocal regularisation. We interpret this as reflecting the heuristic that 'pure' viscous relaxation is a somewhat degenerate regularisation, whereas nonlocal effects act as a generic perturbation.

5.2 Computation of the point spectrum

Let us fix the parameter set $\beta = 6$, $\gamma_1 = 7/12$, $\gamma_2 = 3/4$, $\kappa = 5$, $\alpha = 1/5$. In both the viscous relaxation limit (7) and the pure nonlocal regularization limit (8), there exist only two eigenvalues in the point spectrum for sufficiently small $\varepsilon > 0$: the simple translational eigenvalue $\lambda_0 = 0$, and another simple real eigenvalue $\lambda_1 \approx -0.8$ that does not affect the stability of the travelling wave (see [30, 31]). In this section we augment these results by sampling the point spectrum of the corresponding linearised operator for $\delta > 0$ and small values of $\varepsilon > 0$.

We now outline the strategy to compute the point spectrum. For each $\lambda \in \mathbb{C}$ to the right of the essential spectrum, we can define an unstable complex 2-plane bundle $\varphi_+(z, \lambda, \varepsilon)$ extending from the unstable subspace of the saddle point at $u = 1$, resp. a stable complex 2-plane bundle $\varphi_-(z, \lambda, \varepsilon)$ extending from the stable subspace of the saddle point at $u = 0$, by using the eigenvalue problem (27). A (spatial) eigenvalue $\lambda \in \sigma_p(L)$ is found whenever φ_- and φ_+ have a nontrivial intersection at some value z; see [1] for details.

In view of this geometric characterization for the spatial eigenvalues, we will use a *Riccati-Evans function* to compute these intersections. The eigenvalue problem (27) induces a nonlinear flow on the Grassmannian $Gr(2, 4)$ of complex 2-planes in $\mathbb{C}^4$. On a suitable coordinate patch of $Gr(2, 4)$, the nonlinear flow is defined using a *matrix Riccati equation* of the form

$$W' = C + DW - WA - WBW, \tag{36}$$

where W is a complex-valued 2×2 matrix variable defined using frame coordinates for the 2-planes, and the 2×2 matrices A, B, C, D are defined via a block decomposition of the linear operator M in (28):

$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right). \tag{37}$$

See [13] and [28] for the derivation of (36), and [30] for the specific construction in the case of 'purely' nonlocal regularisation (8).

In terms of the representation (38) of the projectivised dynamics, φ_+ is equivalent to the unique trajectory of (36) that converges to the unstable subspace of the saddle point at $u = 1$ as $z \to -\infty$; there is also an analogous characterisation of φ_-. We can now formulate a shooting problem defined on a suitable cross section that intersects the travelling wave transversely, say $\Sigma = \{u = 0.7\}$, with the corresponding intersection point $z_0 \in \Sigma$. The unstable bundle $\varphi_+(z, \lambda, \varepsilon)$ is flowed forward from $u = 1$ and the stable bundle $\varphi_-(z, \lambda, \varepsilon)$ is flowed backward from $u = 0$. Suppressing the notation for ε, the Riccati-Evans function $E(z_0, \lambda)$ is defined by

$$E(z_0, \lambda) = \det(\varphi_+(z_0, \lambda) - \varphi_-(z_0, \lambda)). \tag{38}$$

We can then find eigenvalues $\lambda \in \sigma_p(L)$ by locating zeroes of $E(z_0, \lambda)$; see [13]. Using the argument principle, we locate these zeroes by computing the winding number of E along suitably chosen contours in the complex plane and to the right of the essential spectrum.

We use a semicircular contour K with increasingly large radii $R = 10^3, 10^4, 10^5$ opening in the right half complex plane, with a small semicircular detour that avoids the translational eigenvalue at the origin (see Fig. 16(a)). We sampled the interval $[0, a_m]$ with a grid of 100 equally

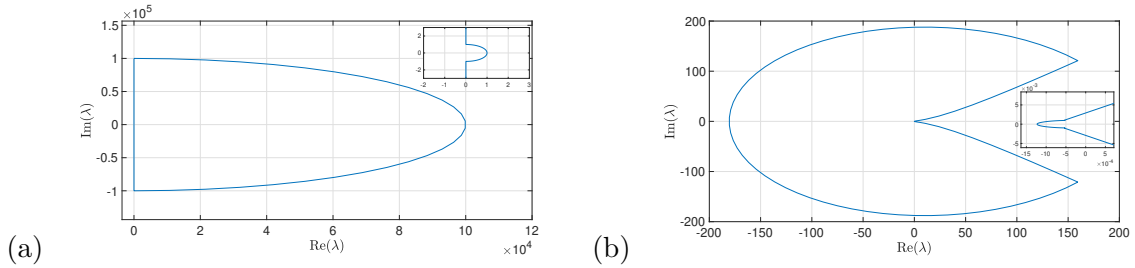


Figure 16: (a) Semicircular contour K of radius 10^5 along which the Riccati-Evans function (38) is evaluated. A small detour avoids the translational eigenvalue at the origin (inset). (b) Image $E(z_0, K)$ in the complex plane. Note that the image does not wind around the origin (inset). Parameter set as in Fig. 11.

spaced points, and we fix $\varepsilon = 10^{-4}$. For each a in this sample, we evaluate $E(z_0, \lambda)$ around K to find a winding number of 0 (see e.g. Fig. 16(b)). Thus we obtain strong numerical evidence that there is no point spectrum within K, and hence that the corresponding family of shock-fronted travelling waves remains nonlinearly stable for each $a \in [0, a_m]$ and for each sufficiently small $\varepsilon > 0$.

We also investigated the zeroes of $E(z_0, \lambda)$ along the real line. For each sampled parameter value δ, we found evidence of only one simple translational eigenvalue $\lambda_0 = 0$ and one other simple real eigenvalue $\lambda_1 \approx -0.8$. We note that this result is entirely consistent with the calculation of the point spectrum in the 'purely nonlocal' and 'purely viscous' regularisation cases [30] and [31]; indeed, these eigenvalues are accounted for by the corresponding reduced *slow eigenvalue problem* defined on the critical manifold. As depicted in Fig. 9(b), singular heteroclinic connections are formed over a small range of wavespeeds $c \in [0.1973, 0.1993]$ as $a \geq 0$ is varied. This in turn slightly perturbs the dynamics of the reduced eigenvalue problem along the singular heteroclinic connection. As a consequence of these small variations, the (singular limit of the) secondary eigenvalue λ_1 moves continuously within the small interval $[-0.81, -0.79]$ on the real line as a is varied. The key point that we would like to emphasize is that a slow eigenvalue problem (defined for $\varepsilon = 0$) continues to approximate the eigenvalues of the 'full' problem (defined for $0 < \varepsilon \ll 1$) when $a > 0$.

Remark 5.2 *It is interesting to ask whether the translational eigenvalue can bifurcate under variation of the regularisation parameter a via e.g. transcritical crossings with the secondary eigenvalue $\lambda_1 = \lambda_1(a)$. In other words, is it possible to destabilize a (regularized) shock-fronted travelling wave of* (9) *just by re-weighting the regularization?*

We argue that such a destabilization scenario is not possible for families of monotone travelling waves. We showed using comparison methods in [31] that if a projectivized solution along a singular heteroclinic orbit of the reduced eigenvalue problem has no winds at $\lambda = \lambda_0 \in \mathbb{R}$, then no further winds are generated for each real $\lambda > \lambda_0$, and furthermore, there are no nontrivially complex eigenvalues—i.e. there are no more eigenvalues to the right of λ_0. Now suppose that the translational eigenvalue had a transcritical bifurcation for some critical value $a = a_c$ such that $\lambda_1(a) > 0$ for $a > a_c$. Then it must be true for nearby values of $a > a_c$ that the variational solution winds around at least once, i.e., the singular heteroclinic of (17) *in (u, v)-space makes a full revolution.*

But this is possible only if the wave is nonmonotone.

Remark 5.3 *It is also interesting to determine the spectral stability of the nonmonotone waves and waves with canards (see Sec. 4.4) that arise in our RND PDE, and in particular to determine whether the corresponding eigenvalue problem can still be investigated using slow-fast splittings as done in [30, 31]. We conjecture that the nonmonotonicity in either the tails or the shock fronts (Remark 4.7) of the waves can play a role in destabilization.*

Acknowledgement

This work has been funded by the Australian Research Council DP200102130 grant.

References

[1] J. Alexander, R. A. Gardner, and C. K. R. T. Jones. A topological invariant arising in the stability analysis of traveling waves. *J. Reine Angew. Math*, 410:167–212, 1990.

[2] K. Anguige and C. Schmeiser. A one-dimensional model of cell diffusion and aggregation, incorporating volume filling and cell-to-cell adhesion. *Journal of Mathematical Biology*, 58(3):395–427, jun 2008.

[3] C.M. Bender and S.A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers I.* Springer New York, 1999.

[4] Alexander P. Browning, Wang Jin, Michael J. Plank, and Matthew J. Simpson. Identifying density-dependent interactions in collective cell behaviour. *Journal of The Royal Society Interface*, 17(165):20200143, 2020.

[5] J. W. Cahn. On spinoidal decomposition. *Acta. Metall.*, 9(795-801), 1961.

[6] D.S. Cohen and J.D Murray. A generalized diffusion model for growth and dispersal in a population. *J. Math. Biology*, 12:237–249, 1981.

[7] N. Fenichel. Geometric singular perturbation theory. *J Differential Equations*, 31:53–98, 1979.

[8] A. E. Fernando, K. A. Landman, and M. J. Simpson. Nonlinear diffusion and exclusion processes with contact interactions. *Physical Review E*, 81, January 2010.

[9] Paul C. Fife. Models for phase separation and their mathematics. 2000.

[10] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:355–369, 1937.

[11] R. Gardner and C.K.R.T. Jones. Stability of Travelling Wave Solutions of Diffusive Predator-Prey Systems. *Transactions of the AMS*, 327:465–524, 1991.

[12] A. Granados, S.J. Hogan, and T.M. Seara. The scattering map in two coupled piecewise-smooth systems, with numerical application to rocking blocks. *Physica D*, (269):1–20, 2014.

[13] K. E. Harley, P. van Heijster, R. Marangell, G. J. Pettet, T. V. Roberts, and M. Wechselberger. (in)stability of travelling waves in a model of haptotaxis. *submitted. 20 pages*, 2019.

[14] K. E. Harley, P. van Heijster, R. Marangell, G. J. Pettet, and M. Wechselberger. Existence of Travelling Wave Solutions for a Model of Tumour Invasion. *SIAM Journal of Applied Dyn Sys*, 13(1):366–396, 2014.

[15] K. E. Harley, P. van Heijster, R. Marangell, G. J. Pettet, and M. Wechselberger. Novel solutions for a model of wound healing angiogenesis. *Nonlinearity*, 27:2975–3003, 2014.

[16] D. Henry. *Geometric Theory of Semilinear Parabolic Equations*. Number 840 in Lecture Notes in Mathematics. Springer–Verlag, New York, 1980.

[17] T. Hillen and K. Painter. A user's guide to pde models for chemotaxis. *Journal of Mathematical Biology*, 58(1–2):183–217, 2009.

[18] J. Hilliard. Spinodal decomposition. *in* phase transformations. *Am. Soc. Metals*, pages 497–560, 1970.

[19] Klaus Höllig. Existence of infinitely many solutions for a forward backward heat equation. *Transactions of the American Mathematical Society*, 278(1):299–316, 1983.

[20] S. T. Johnston, R. E. Baker, D. L. S. McElwain, and M. J. Simpson. Co-operation, competition and crowding: a discrete framework linking allee kinetics, nonlinear diffusion, shocks and sharp-frontec travelling waves. *Scientific Reports*, 7(42134), 2017.

[21] C. K. R. T. Jones. Geometric singular perturbation theory. *in Dynamical Systems, Springer Lecture Notes Math.*, 1609:44–120, 1995.

[22] T. Kapitula and B. Sandstede. Stability of bright solitary-wave solutions to perturbed nonlinear schrödinger equations. *Physica D: Nonlinear Phenomena*, 124(1):58–103, 1998.

[23] J. Keener and J. Sneyd. *Mathematical Physiology*, volume 8 of *Interdisciplinary Applied Mathematics: Mathemarical Biology*. Springer, Hong Kong, 1998.

[24] E. F. Keller and L. A. Segel. Travelling bands of chemotactic bacteria: A theoretical analysis. *J. theor. Biol.*, 30:235–248, 1971.

[25] P. Kukučka. Melnikov method for discontinuous planar systems. *Nonlinear Analysis*, (66):2698–2719, 2006.

[26] K. A. Landman, G. J. Pettet, and D. F. Newgreen. Chemotactic cellular migration: Smooth and discontinuous travelling wave solutions. *SIAM Journal of Applied Mathematics*, 63:1666–1681, 2003.

[27] K. A. Landman, M. J. Simpson, and G. J. Pettet. Tactically-driven nonmonotone travelling waves. *Physica D*, 237:678–691, 2008.

[28] V. Ledoux, S. Malham, and V. Thümmler. Grassmannian spectral shooting. *Mathematics of Computation*, 79:1585–1619, 2010.

[29] Y. Li, P. van Heijster, M. J. Simpson, and M. Wechselberger. Shock-fronted travelling waves in a reaction–diffusion model with nonlinear forward–backward–forward diffusion. *Physica D*, 423(132916), 2021.

[30] I. Lizarraga and R. Marangell. Nonlocal stability of shock-fronted travelling waves under nonlocal regularization. *arXiv:2211.07824 (Preprint)*, 2022.

[31] I. Lizarraga and R. Marangell. Spectral stability of shock-fronted travelling waves under viscous relaxation. *Journal of Nonlinear Science*, 33(82), 2023.

[32] P. K. Maini, L. Malaguti, C. Marcelli, and S. Matucci. Aggregative movement and front propagation for bi-stable population models. *Mathematical Models and Methods in Applied Sciences*, 17(9):1351–1368, 2007.

[33] P. K. Maini, F. Sanchez-Garduno, and J. Perex-Velazques. A nonlinear degenerate equation for direct aggreation and traveling wave dynamics. *Discrete and Continuous Dynamical Systems Series B*, 13(2), March 2010.

[34] B. P. Marchant, J. Norbury, and J. A. Sherratt. Travelling wave solutions to a haptotaxis-dominated model of malignant invasion. *Nonlinearity*, 14:1653–1671, 2001.

[35] J.D. Murray. *Mathematical Biology I: An introduction*. Number 17 in Interdisciplinary Applied Mathematics. Springer, Hong Kong, 3rd edition, 2002.

[36] A. Novick-Cohen and R. L. Pego. Stable patterns in a viscous diffusion equation. *Transactions of the American Mathematical Society*, 324(1), March 1991.

[37] V. Padron. Effect of aggregation on population recovery modelled by a forward-backward pseudoparabolic equation. *Transactions of the American Mathematical Society*, 356(7):2739–2756, 2003.

[38] R. L. Pego. Front migration in the nonlinear Cahn–Hilliard equation. *Proc. R. Soc. Lond. A*, 422:261–278, 1989.

[39] C. J. Penington, B. D. Hughes, and K. A. Landman. Building macroscale models from microscale probablistic models: A general probablistic approach for nonlinear diffusion and multispecies phenomena. *Physical Review E*, 84, October 2011.

[40] Mathieu Poujade, Erwan Grasland-Mongrain, A Hertzog, J Jouanneau, Philippe Chavrier, Benoît Ladoux, Axel Buguin, and Pascal Silberzan. Collective migration of an epithelial monolayer in response to a model wound. *Proceedings of the National Academy of Sciences*, 104(41):15988–15993, 2007.

[41] J. Smoller. *Shock waves and reaction diffusion equations*, volume 258 of *Grundlehren der mathematischen Wissenshaften*. Springer-Verlag, 1982.

[42] P. Szmolyan and M. Wechselberger. Canards in $\mathbb{R}^3$. *J Differential Equations*, 177:419–453, 2001.

[43] A. Vanderbauwhede. Bifurcation of degenerate homoclinics. *Results in Mathematics*, 21(1-2):211–223, mar 1992.

[44] M. Wechselberger. Extending Melnikov theory to invariant manifolds on non-compact domains. *Dynamical Systems*, 17(3):215–233, aug 2002.

[45] M. Wechselberger and G. J. Pettet. Folds, canards and shocks in advection-reaction-diffusion models. *Nonlinearity*, 23:1949–1969, 2010.

[46] G.B. Whitham. *Linear and Nonlinear Waves*. Wiley, New York, 1st edition, 1999.

[47] T.P. Witelski. Shocks in nonlinear diffusion. *Applied Mathematics Letters*, 8(5):27–32, 1995.

[48] T.P. Witelski. The structure of internal layers for unstable nonlinear diffusion equations. *Studies in Applied Mathematics*, 96:277–300, 1996.

A Melnikov theory for autonomous vector fields

One of the main analytical tools that deals with the existence and bifurcation of heteroclinic orbits in dynamical systems is known as *Melnikov theory*. We follow the treatment of Vanderbauwhede [43] for sufficiently smooth autonomous dynamical systems in arbitrary dimensions; see also the extension to heteroclinic problems on non-compact domains in [44]. Here, we provide a succinct summary of this theory for autonomous problems, which we then apply to our layer problem presenting an alternative point-of-view to the heteroclinic orbit analysis presented in section 4.1. We then adapt Melnikov theory to the case of planar piecewise-smooth dynamical systems which is needed to show the existence singular heteroclinic orbits in section 4.2. The Melnikov method has been adapted to the piecewise-smooth setting in different nonsmooth contexts, with an emphasis on time-periodic perturbations of planar vector fields with either Hamiltonian or trace-free structure (see e.g. [12, 25]); however, the following Vanderbauwhede-style presentation for piecewise smooth autonomous vector fields is new, to the best of our knowledge.

A.1 The smooth case

We work with a system of the form

$$x' = h(x, \mu) \tag{39}$$

with $x = x(t) \in \mathbb{R}^n$, $n \geq 2$, h sufficiently smooth, and the parameters are denoted by $\mu \in \mathbb{R}^m$, $m \geq 1$. We assume the existence of a unique heteroclinic connection $\Gamma = \{\gamma(t) \in \mathbb{R}^n : t \in \mathbb{R}\}$ between saddle equilibria p_- and p_+ for some $\mu = \hat{\mu}$, i.e., $\Gamma = W^u_{\hat{\mu}}(p_-) \cap W^s_{\hat{\mu}}(p_+)$ is the one-dimensional intersection of the unstable manifold $W^u_{\hat{\mu}}(p_-)$ of p_- and the stable manifold $W^s_{\hat{\mu}}(p_+)$ of p_+ with $\dim W^u_{\hat{\mu}}(p_-) = l_u$, $1 \leq l_u < n$, and $\dim W^s_{\hat{\mu}}(p_+) = l_s$, $1 \leq l_s < n$.[5] We further assume that $l_u + l_s - 1 < n$. Otherwise, the intersection $W^u_{\hat{\mu}}(p_-) \cap W^s_{\hat{\mu}}(p_+)$ is transverse implying the persistence of this heteroclinic orbit for nearby μ-values.[6]

We define a suitable cross section Σ of the unique heteroclinic orbit Γ. Without loss of generality we assume the intersection of Γ with Σ occurs at $t = 0$. The main task is to measure the distance between the stable and unstable manifolds $W^u_\mu(p_-)$ and $W^s_\mu(p_+)$ in this cross section Σ for nearby μ-values. Melnikov theory defines a corresponding distance function[7] $\Delta = \Delta(\mu)$, noting that $\Delta(\hat{\mu}) = 0$. In the following, we derive computable formulas for $\Delta = \Delta(\mu)$.

Let $x(t) = \gamma(t) + X(t)$, with $X \in \mathbb{R}^n$, which transforms (39) to the non-autonomous problem

$$X' = A(t)X + g(X, t, \mu)$$

with the non-autonomous matrix $A(t) := D_x h(\gamma(t); \mu_0)$ and the nonlinear remainder $g(X, t; \mu) = h(\gamma + X; \mu) - h(\gamma; \mu_0) - A(t)X$. The linear equation

$$X' = A(t)X \tag{40}$$

[5] For the general setup with possible higher dimensional heteroclinic connections we refer to [43, 44].

[6] Note that these stable and unstable manifolds $W^u_{\hat{\mu}}(p_-)$ and $W^s_{\hat{\mu}}(p_+)$ persist for nearby μ-values due to the robustness of the hyperbolic saddle equilibria. With a slight abuse of notation, we will denote these perturbed saddle equilibria by $p_\pm$ independent of μ.

[7] In higher dimensional problems, the distance function is vector-valued.

is the *variational equation* along $\gamma(t)$. The corresponding *adjoint equation* is given by

$$\Psi' + A^\top(t)\Psi = 0\,.$$

Note that solutions of the variational and adjoint equations preserve a constant angle along γ, i.e., $(\partial/\partial t)(\Psi^\top(t)X(t)) = 0, \forall t \in \mathbb{R}$. We can use this fact to define a splitting of the vector space $\mathbb{R}^n$ along $\gamma(t)$. Following Vanderbauwhede in [43], we define a splitting of the vector space $\mathbb{R}^n$ at $t = 0$ in the following way:

$$\mathbb{R}^n = \text{span}\,\{h(\gamma(0), \hat{\mu})\} \oplus Y$$

where Y is a complementary subspace that admits a further sub-splitting into dynamically distinct subspaces. In view of our assumption on Γ, we have a splitting

$$Y = V_s \oplus V_u \oplus W,$$

where $V_{s,u}$ denote subspaces complementary to $T_{\gamma(0)}W^u_{\hat{\mu}}(p_-) \cap T_{\gamma(0)}W^s_{\hat{\mu}}(p_+)$ inside the respective (un)stable tangent spaces $T_{\gamma(0)}W^{s,u}$ and W denotes the orthogonal complement of $T_{\gamma(0)}W^s + T_{\gamma(0)}W^u$.[8]

Based on this setup, we now specify $\Sigma = \gamma(0) + Y$ as a suitable cross-section to perform computations. A key observation of Vanderbauwhede in [43] is that—by means of projection maps $P_-(t)$ (for $t \leq 0$) and $Q_+(t)$ (for $t \geq 0$), defined using properties of exponential dichotomies, together with the variation of parameters formula—the (un)stable manifolds of the saddle points can be characterised analytically as the set of initial conditions whose (forward respectively backward) trajectories have bounded norms in appropriately chosen function spaces. Indeed, $W^s_\mu(p_+)$ and $W^u_\mu(p_-)$ are locally expressible as graphs over the images of the corresponding projection maps near the intersection point $\gamma(0) \in \Sigma$. For a suitable neighborhood Ω containing $\gamma(0)$, we have

$$W^s_\mu(p_+) \cap \Omega = \{\xi + \beta_+(\xi, \mu) : \xi \in \omega_+ \subset \text{Im}(P_+(0))\} \tag{41}$$

and

$$W^u_\mu(p_-) \cap \Omega = \{\eta + \beta_-(\eta, \mu) : \eta \in \omega_- \subset \text{Im}(Q_-(0))\}, \tag{42}$$

where the graphs $h_\pm$ are defined using the transition matrix $\Phi(t, s)$ of (40), the projection operators $P_-(0)$ and $Q_+(0)$, and the corresponding (un)stable manifold segments $\gamma_\pm$ of the saddle equilibria, i.e.

$$\beta_+(\xi, \mu) := -Q_+(0) \int_0^\infty \Phi(0, s) g(s, \gamma_+(\xi, \mu)(s), \mu) ds \tag{43}$$

and

$$\beta_-(\eta, \mu) := P_-(0) \int_{-\infty}^0 \Phi(0, s) g(s, \gamma_-(\eta, \mu)(s), \mu) ds. \tag{44}$$

[8] When $\dim(T_{\gamma(0)}W^u_{\hat{\mu}}(p_-) \cap T_{\gamma(0)}W^s_{\hat{\mu}}(p_+)) > 1$, the splitting for Y generalises to incorporate an additional subspace $Y = U \oplus V_s \oplus V_u \oplus W$. In this case, U can be chosen complementary to span $\{h(\gamma(0), \hat{\mu})\}$ inside this intersection; see [43] for details.

Note that the splitting defines a local coordinate system in the neighborhood Ω. The cross-section is given in local coordinates by $\xi = \eta = 0$. We may therefore define a suitable distance function on Σ by

$$\Delta(\mu) := \beta_-(0,\mu) - \beta_+(0,\mu). \tag{45}$$

We can specify an orthonormal basis $\{\psi_1^0, \cdots, \psi_k^0\}$ of W, $k = n+1-(l_u+l_s)$,[9] so that each solution $\psi_i(t)$ of the adjoint equation with initial conditon $\psi_i(0) = \psi_i^0$ decays exponentially for $t \to \pm\infty$.[10] Melnikov theory then establishes the following formula for the distance function, expressed in components (see e.g. [43, 44]):

$$\begin{aligned}\Delta_i(\mu) &= \int_{-\infty}^{0} \psi_i(s)^\top g(\gamma_-(\mu)(s), s; \mu)\, ds + \int_0^{\infty} \psi_i(s)^\top g(\gamma_+(\mu)(s), s; \mu)\, ds \\ &=: \int_{-\infty}^{\infty} \psi_i(s)^\top g(\gamma(\mu)(s), s; \mu)\, ds, \qquad i = 1, \ldots, k\,,\end{aligned} \tag{46}$$

where γ in the final line should be interpreted as the representative of $\gamma_\pm$ in the relevant domain.

Remark A.1 *This distance function (vector) $\Delta : \mathbb{R}^m \to \mathbb{R}^k$ is well-defined since the vector $\psi(s)$ decays exponentially in forward and backward time and $\gamma_\pm$ is bounded. Solving $\Delta(\mu) = 0$ is well defined for $m \geq k$. For $m = k$ and $rk(D_\mu \Delta) = k$, $\mu = \hat{\mu}$ is an isolated zero. For $m > k$, we expect the solution set to be a submanifold of codimension k.*

Remark A.2 *The distance formula has a particularly simple representation which is independent of the projection operators, because we have some freedom in choosing $P_-(t)$ and $Q_+(t)$; indeed, we can choose them so that their kernels are orthogonal to the adjoint solution $\psi(t)$ at $t = 0$. This is relevant for the derivation of the piecewise-smooth Melnikov function in the next section.*

In general, one cannot solve $\Delta(\mu) = 0$ explicitly, but one can calculate its leading order Taylor series expansion. Taking partial derivatives of (46) with respect to μ_j and referring to the definition of the nonlinear remainder $g(X, t; \mu)$, we obtain the following formulas, defined component-wise:

$$D_{\mu_j}\Delta_i(\hat{\mu}) = \int_{-\infty}^{\infty} \psi_i(s)^\top \frac{\partial h}{\partial \mu_j}(\gamma_0(s), s; \hat{\mu})\, ds, \qquad j = 1, \ldots m\,, \quad i = 1, \ldots k\,. \tag{47}$$

These first-order derivative terms are known as first-order *Melnikov integrals*.

Remark A.3 *In the planar case that we consider, we have $k = 1$ and $Y = W$. A suitable adjoint solution $\psi(t)$ can be chosen so that $\psi(0)$ is orthogonal to $h(\gamma(0), \hat{\mu})$ at $t = 0$.*

A.1.1 Shocks in the layer problem: the small $|\delta|$ case

Here we apply Melnikov Theory to establish heteroclinic connections in the layer problem for sufficiently small $|\delta| > 0$. This complements the analysis in section 4.1.

[9] Based on our assumptions, $\dim W = n + 1 - (l_u + l_s) \geq 1$.

[10] Such solutions of the adjoint quation with exponential decay for $t \to \pm\infty$ exist due to the exponential dichotomy properties induced by the saddle endpoints of this heteroclinic orbit Γ.

Define $x = (u, \hat{u})^\top$ and $h(x; w, \delta) = (\hat{u}, w + \Phi(u) - \delta\hat{u})^\top$ such that the layer problem is given in vector form by

$$x' = h(x; w, \delta)\,, \qquad x \in \mathbb{R}^2\,.$$

This system possesses heteroclinic orbits $\Gamma_\pm(y)$ for $w = w_h$ and $\delta = 0$, i.e., $\Gamma'_\pm = h(\Gamma_\pm; w_h, 0)$. Let $x = \Gamma_\pm + X$, $X \in \mathbb{R}^2$ which transforms the layer problem to the non-autonomous problem

$$X' = A(y)X + g(X, y; w, \delta)$$

with the non-autonomous matrix $A(y) := D_x h(\Gamma_\pm; w_h, 0)$ and the nonlinear remainder $g(X, y; w, \delta) = h(\Gamma_\pm + X; w, \delta) - h(\Gamma_\pm; w_h, 0) - A(y)X$. Without loss of generality, we define the splitting of the vector space at $y = 0$ by

$$\mathbb{R}^2 = \text{span}\,\{h(\Gamma_\pm(0); w_h, 0)\} \oplus W$$

where W is spanned by a solution of the adjoint equation that decay exponentially for $y \to \pm\infty$; here, this space is one-dimensional and we denote the corresponding solution by $\psi(y) = (\psi_1(y), \psi_2(y))^\top$.

We measure the distance $\Delta \in \mathbb{R}$ between the one-dimensional stable and unstable manifolds emanating from the saddle-equilibria $p_{l/r}$ in a suitable cross section $\Sigma = W$. This distance function $\Delta = \Delta(w, \delta)$ depends on the system parameters, and we have $\Delta(w_h, 0) = 0$. Melnikov theory establishes the following distance function formula

$$\begin{aligned}\Delta(w,\delta) &= \int_{-\infty}^{0} (\psi(s)^\top g(X_-(w,\delta)(y), y; w, \delta))ds - \int_{\infty}^{0} (\psi(s)^\top g(X_+(w,\delta)(y), y; w, \delta))ds\,,\\ &= \int_{-\infty}^{\infty} (\psi(s)^\top g(X(w,\delta)(y), y; w, \delta))ds\end{aligned} \tag{48}$$

where $X_\pm(w, \delta)(y)$ denotes the corresponding (un)stable manifold segments of the saddle equilibria $p_{l/r}$ from the corresponding saddle equilibrium to the cross section Σ, and X is the representative of these sets in the relevant domain. If, e.g., $D_w\Delta(w_h, 0) \neq 0$ then $w = w_h(\delta) = w_h + b\delta + O(\delta^2)$ solves $\Delta(w_h(\delta), \delta) = 0$ for $\delta \in (-\delta_0, +\delta_0)$, $\delta_0 > 0$. The leading order expansion parameter b is then given by

$$b = -\frac{D_\delta\Delta(w_h, 0)}{D_w\Delta(w_h, 0)}\,,$$

and the corresponding first-order Melnikov integrals can be calculated as follows:

$$(D_w\Delta(w_h,0), D_\delta\Delta(w_h,0)) = (\int_{-\infty}^{\infty} (\psi(s)^\top D_w h(\Gamma_\pm(s); w_h, 0))ds,\ \int_{-\infty}^{\infty} (\psi(s)^\top D_\delta h(\Gamma_\pm(s); w_h, 0))ds)$$

We have $D_w h(\Gamma_\pm(0); w_h, 0) = (0, 1)^\top$ and, hence,

$$D_w\Delta(w_h,0) = \int_{-\infty}^{\infty} (\psi(s)^\top D_w h(\Gamma_\pm(s); w_h, 0))ds = \int_{-\infty}^{\infty} \psi_2(s)ds \neq 0,$$

based on the observation that the ψ_2-component does not change sign along $\Gamma_\pm$, i.e., it is a monotone function along $\Gamma_\pm$. The integral is well-defined since $\psi_2(y)$ is decaying exponentially for $y \to \pm\infty$. Hence, $w = w_h(\delta) = w_h + b\delta + O(\delta^2)$ solves $\Delta(w(\delta), \delta) = 0$ for $\delta \in (-\delta_0, +\delta_0)$. We also have $D_\delta h(\Gamma_\pm(0); w_h, 0) = (0, -\hat{u}(y))^\top$ and, hence,

$$D_\delta\Delta(w_h,0) = \int_{-\infty}^{\infty} (\psi(s)^\top D_\delta h(\Gamma_\pm(s); w_h, 0))ds = -\int_{-\infty}^{\infty} \hat{u}(s)\psi_2(s)ds \neq 0\,,$$

based on a similar observation as above, i.e., both terms do not change sign under the variation along $\Gamma_\pm$. Hence,

$$b = \frac{D_\delta \Delta(w_h, 0)}{D_w \Delta(w_h, 0)} = -\frac{\int_{-\infty}^{\infty} \hat{u}(s)\psi_2(s)ds}{\int_{-\infty}^{\infty} \psi_2(s)ds} \neq 0$$

and we have a leading order affine solution $w(\delta)$ to $\Delta(w, \delta) = 0$ near $(w_h, 0)$. This confirms the transverse crossing of the heteroclinic branches near $(0, w_h)$ as shown in Figure 6.

A.2 The piecewise-smooth planar case

We now adapt the Melnikov method developed in Appendix A.1 to a subclass of piecewise-smooth planar problems. Fix $\delta > 0$ and $x_\Sigma \in \mathbb{R}$, and consider a pair of planar vector fields $h_\pm(x, \mu) = (h^1_\pm(x, \mu), h^2_\pm(x, \mu))$, such that h_- is smoothly defined on a compact subset $V_- \subset \{x_1 < x_\Sigma + \delta\}$ and h_+ is smoothly defined on a compact subset $V_+ \subset \{x_1 > x_\Sigma - \delta\}$. Both vector fields $h_\pm$ are smoothly dependent on the parameters $\mu \in \mathbb{R}^m$.

We assume that $V_- \cap V_+$ is nonempty and contains a (vertical) cross-section $\Sigma \subset \{x_1 = x_\Sigma\}$ of both vector fields $h_\pm$ for each $\mu \in P \subset \mathbb{R}^m$ for a non-trivial subset P in parameter space. In particular, both vector-fields are assumed to cross (the interior of) Σ unidirectionally, e.g., here from left to right, for each $\mu \in P \subset \mathbb{R}^m$. With this setup, we may define a piecewise smooth vector field on $U_- \cup U_+$, $U_- := V_- \cap \{x_1 < x_\Sigma\}$ and $U_+ := V_+ \cap \{x_1 > x_\Sigma\}$, as follows:

$$\dot{x} = \begin{cases} h_-(x, \mu) & \text{if } x \in U_- \\ h_+(x, \mu) & \text{if } x \in U_+. \end{cases} \tag{49}$$

We assume the existence of a *piecewise-smooth heteroclinic orbit* which connects a saddle equilibrium $p_- \in U_-$ to one at $p_+ \in U_+$ for some $\mu = \hat{\mu}$ in (the interior of) P with the following properties: specify an unstable manifold segment $W^u(p_-) \subset U_- \cup \Sigma$, defined for $t \leq 0$ by a trajectory $\gamma_-(t)$ satisfying $\dot{x} = h_-(x, \hat{\mu})$, and a stable manifold segment $W^s(p_+) \subset U_+ \cup \Sigma$ defined for $t \geq 0$ by a trajectory $\gamma_+(t)$ satisfying $\dot{x} = h_+(x, \hat{\mu})$, such that $\gamma(0) := \gamma_-(0) = \gamma_+(0) \in \Sigma$. The piecewise-smooth heteroclinic orbit γ_0 is then defined by the set

$$\gamma_0 = \left(\bigcup_{t<0} \gamma_-(t) \right) \cup \gamma(0) \cup \left(\bigcup_{t>0} \gamma_+(t) \right).$$

See Figure 17 for a depiction of such a piecewise-defined γ_0. As in the smooth case, we seek to determine the persistence of γ_0 under parameter variation via a Melnikov calculation on Σ.

Much of the setup in Appendix A.1 carries over for the problems $\dot{x} = h_\pm(x, \mu)$ defined separately on the subsets $V_\pm$. On V_-, we can specify a section Σ_- through $\gamma(0)$ which is aligned with a backward exponentially decaying adjoint solution $\psi_-(t)$ at $t = 0$. The associated graph representation of the unstable manifolds in a neighborhood of $\gamma(0)$ is given by

$$\beta_-(\eta, \mu) := P_-(0) \int_{-\infty}^{0} \Phi_-(0, s) g_-(s, \gamma_-(\eta, \mu)(s), \mu) ds, \tag{50}$$

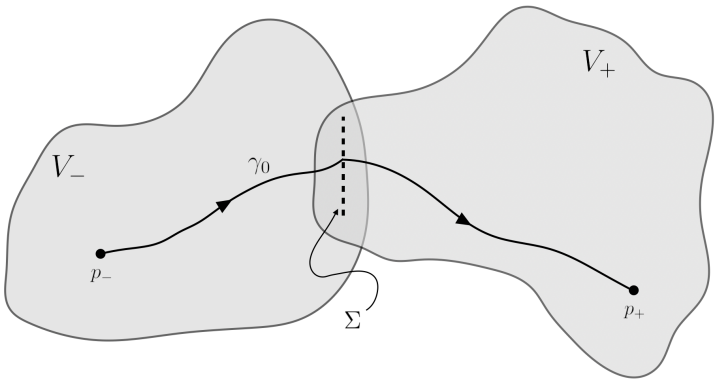


Figure 17: A sketch of a piecewise-smooth heteroclinic connection γ_0.

where we denote the appropriate objects in V_- by the $(-)$ subscript. Similarly, near an appropriately chosen cross-section Σ_+ which is aligned with a forward exponentially decaying adjoint solution $\psi_+(t)$ at $t = 0$ there is a graph representation of the stable manifolds on V_+, given by

$$\beta_+(\xi, \mu) := -Q_+(0) \int_0^\infty \Phi_+(0, s) g_+(s, \gamma_+(\xi, \mu)(s), \mu) ds. \tag{51}$$

The main conceptual difference in deriving the distance function in the piecewise-smooth case is that the local coordinates used to specify the (un)stable manifold segments near Σ are adapted to the sections Σ_- and Σ_+, neither of which is aligned with Σ in general (e.g. see Fig. 18). This introduces extra terms into the separation function when measuring the distance between $W^u(p_-)$ and $W^s(p_+)$ on Σ, as we now show.

We denote by $m_- \in \mathbb{R}$ the slope of Σ relative to the linear subspace specified by $\mathrm{Im}(P_-(0))$ and let $m_+ \in \mathbb{R}$ denote the slope of Σ relative to the linear subspace specified by $\mathrm{Im}(Q_+(0))$. We also denote by $\xi^*(\mu) + \beta_+(\xi^*(\mu), \mu)$ respectively $\eta^*(\mu) + \beta_-(\eta^*(\mu), \mu)$ the intersections of $W^s_\mu(p_+)$ respectively $W^u_\mu(p_-)$ on the section Σ. The quantities $\eta^*(\mu)$ and $\xi^*(\mu)$ are defined using the implicit function theorem as follows. Let u_+ and u_- denote unit vectors in the linear subspaces $\mathrm{Im}(P_+(0))$ respectively $\mathrm{Im}(Q_-(0))$ so that $\eta = \rho_- u_-$ and $\xi = \rho_+ u_+$ for $\rho_\pm \in \mathbb{R}$. Recalling that $\psi_\pm(0)$ are convenient unit basis vectors for the orthogonal complements $\mathrm{Im}(P_-(0))$ and $\mathrm{Im}(Q_+(0))$, we find the following defining equations for $\eta^*(\mu) = \rho^*_-(\mu)u_-$ and $\xi^*(\mu) = \rho^*_+(\mu)u_+$ in terms of the auxiliary functions $\rho^*_\pm(\mu)$:

$$m_\pm \psi_\pm(0)^\top \beta_\pm(\rho^*_\pm(\mu)u_\pm, \mu) = \rho^*_\pm(\mu)$$

The distance function (45) is therefore replaced by

$$\Delta(\mu) = \eta^*(\mu) - \xi^*(\mu) + \beta_-(\eta^*(\mu), \mu) - \beta_+(\xi^*(\mu), \mu). \tag{52}$$

See Figure 18 for a sketch of the local coordinates on the patch V_-. A similar sketch can be drawn for the patch V_+ to measure the location of the stable manifold $W^s_\mu(p_+)$ for μ sufficiently close to μ_0.

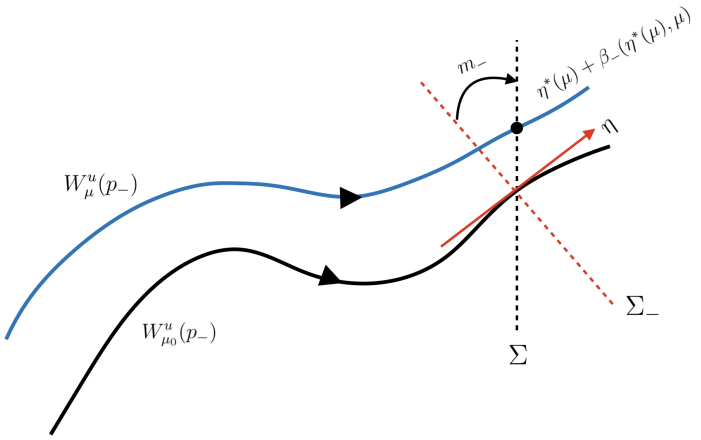


Figure 18: Local coordinates near Σ on V_-.

Remark A.4 *The formula* (52) *also generalises the classical distance formula* (45) *for sections* Σ *that are nonorthogonal to the flow in the smooth case.*

In local coordinates, the section Σ intersects both Σ_- and Σ_+ at the origin, which is also where the unperturbed heteroclinic (defined for $\mu = \hat{\mu}$) intersects Σ. In other words, $\eta^*(\hat{\mu}) = \xi^*(\hat{\mu}) = 0$, and so $\Delta(\hat{\mu}) = 0$ as expected. As before, our task is to derive a computable formula for the Melnikov integrals analogous to (47). Note that $\Delta(\mu)$ is differentiable with image in Σ. Let $e_2 := (0,1)^T$ be a basis vector for Σ, so that $\Delta(\mu) = G(\mu)e_2$ for some real-valued function G. Then

$$\begin{aligned} D_{\mu_i}G(\mu_0) &= e_2^\top D_{\mu_i}\Delta(\hat{\mu}) \\ &= \left(\int_{-\infty}^{0} \psi_-(s)^\top \frac{\partial h_-}{\partial \mu_i}(\gamma_-(s), s; \hat{\mu})\, ds\right) e_2^\top \left(m_- u_- + \psi_-(0)\right) \\ &+ \left(\int_{0}^{\infty} \psi_+(s)^\top \frac{\partial h_+}{\partial \mu_i}(\gamma_+(s), s; \hat{\mu})\, ds\right) e_2^\top \left(m_+ u_+ + \psi_+(0)\right). \end{aligned} \tag{53}$$

Let $u_\pm = (v_{1,\pm}, v_{2,\pm})^T$ denote unit basis vectors obtained from normalising the vector fields $h_\pm(x, \hat{\mu})$ evaluated at $\gamma(0)$. An algebra calculation shows that $m_\pm = v_{2,\pm}/v_{1,\pm}$ and we may choose $\psi_\pm(0) = (-v_{2,\pm}, v_{1,\pm})^T$. Then (53) has the particularly simple form

$$D_{\mu_i}G(\hat{\mu}) = \frac{1}{v_1^-}\int_{-\infty}^{0} \psi_-(s)^\top \frac{\partial h_-}{\partial \mu_i}(\gamma_-(s), s; \hat{\mu})\, ds + \frac{1}{v_1^+}\int_{0}^{\infty} \psi_+(s)^\top \frac{\partial h_+}{\partial \mu_i}(\gamma_+(s), s; \hat{\mu})\, ds. \tag{54}$$

The formula (54) gives the promised generalisation of (47) in the piecewise-smooth planar setting. Note that if $u_- = u_+$ (so that the heteroclinic γ connects smoothly across the discontinuity surface Σ), then (54) is identical to (47) up to a positive prefactor (with the improper integral appropriately interpreted as a sum of two integrals evaluated separately on either side of the discontinuity).

In fact, we also recover the classical Melnikov integral (up to a positive prefactor) if $v_1^- = v_1^+$ only. This observation is useful when the piecewise-smooth problem inherits a reflection symmetry in the vertical component from a smooth problem; see Sec. 4.2.1.