



Universiteit  
Leiden  
The Netherlands

## Geometry of vegetation patterns: understanding patterns in dryland ecosystems and beyond

Jaïbi, O.

### Citation

Jaïbi, O. (2021, November 24). *Geometry of vegetation patterns: understanding patterns in dryland ecosystems and beyond*. Retrieved from <https://hdl.handle.net/1887/3243448>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3243448>

**Note:** To cite this publication please use the final published version (if applicable).

---

# Geometry of Vegetation Patterns

Understanding patterns in dryland ecosystems and beyond



Olfa Jaïbi

---

# Geometry of Vegetation Patterns

*Understanding patterns in dryland ecosystems and  
beyond*

Proefschrift

ter verkrijging van  
de graad van doctor aan de Universiteit Leiden,  
op gezag van rector magnificus prof.dr.ir. H. Bijl,  
volgens besluit van het college voor promoties  
te verdedigen op woensdag 24 november 2021  
klokke 15.00 uur

door

Olfa Jaïbi

**Promotor:**

Prof.dr. Arjen Doelman

Universiteit Leiden

**Copromotor:**

Dr. Martina Chirilus-Bruckner

Universiteit Leiden

**Promotiecommissie:**

Prof.dr. Frank van der Duijn Schouten

Leiden University

Prof.dr. Roeland Merks

Leiden University

Prof.dr. Ehud Meron

Ben-Gurion University

Prof.dr. Vivi Rottschäfer

Leiden University

Prof.dr. Peter van Heijster

Wageningen University

© Olfa Jaïbi, 2021.

Print: Groenprint.

**Front Cover:** Cover design by Bram Bergs.



*To Dikkie and Patat.  
To those who still don't have access to education.*



# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1	Ecological background . . . . .	4
1.1	Ecological context . . . . .	5
1.2	Ecosystem models studied . . . . .	7
1.3	Ecosystem dynamics . . . . .	9
2	Mathematical concepts . . . . .	10
2.1	Reaction-diffusion equations, orbits and patterns . . . . .	10
2.2	Geometric singular perturbation theory . . . . .	17
3	The onset of patterns . . . . .	20
4	Content of thesis . . . . .	24
4.1	Multistability of striped vegetation patterns . . . . .	24
4.2	Existence of (novel) localized vegetation patterns in a generalized ecosystem model . . . . .	25
	Appendices . . . . .	29
1.A	Turing bifurcation on sloped terrains . . . . .	29
<b>2</b>	<b>Multi-stability of model and real dryland ecosystems through spa- tial self-organization</b>	<b>31</b>
1	Introduction . . . . .	32
2	Theory . . . . .	34
2.1	Model description . . . . .	34
2.2	Theoretical outcomes . . . . .	34
2.3	Testable predictions . . . . .	37
3	Data acquisition & processing . . . . .	38
4	Results . . . . .	39
5	Discussion . . . . .	40
	Appendices . . . . .	43
2.A	Dimensional extended-Klausmeier . . . . .	43
2.B	Description of study sites . . . . .	43
2.C	Data sets . . . . .	44
2.C.1	Topographical data . . . . .	44
2.C.2	Biomass measurements . . . . .	45
2.C.3	Optical data . . . . .	45

2.D	Data processing . . . . .	46
2.D.1	Spectral analysis, direction of anisotropy and wavelength . . . . .	46
2.D.2	Pattern classification . . . . .	48
2.D.3	Cross-spectral analysis and migration speed . . . . .	48
2.D.4	Assessment of uncertainty in calculations of slope gra- dient and aspect from topographical data . . . . .	49
2.D.5	Assessment of uncertainty in estimation of pattern fre- quency from optical imagery . . . . .	50
<b>3</b>	<b>The existence of localized vegetation patterns in a systematically re- duced model for dryland vegetation</b>	<b>55</b>
1	Introduction . . . . .	56
2	Set-up of the existence problem . . . . .	62
2.1	The fast reduced problem . . . . .	62
2.2	The slow reduced limit problems . . . . .	64
2.3	Critical points and homogeneous background states . . . . .	68
2.4	The slow flows of the $\varepsilon \neq 0$ system . . . . .	70
3	Localized front patterns . . . . .	75
3.1	$W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$ and its touch down points on $\mathcal{M}_\varepsilon^+$ . . . . .	76
3.2	Traveling 1-front patterns – primary orbits . . . . .	77
3.3	Traveling 1-front patterns by the perturbed integrable flow on $\mathcal{M}_\varepsilon^+$ . . . . .	81
3.4	Stationary 1-front patterns . . . . .	86
3.5	Stationary homoclinic 2-front patterns: vegetation spots and gaps . . . . .	89
3.6	Spatially periodic multi-front patterns . . . . .	96
4	Simulations and discussion . . . . .	100
4.1	Simulations . . . . .	100
4.2	Discussion . . . . .	106
	Appendices . . . . .	111
3.A	Derivation of the model equations in one spatial dimension . . . . .	111
3.B	The derivation of the scaled model . . . . .	112
3.C	Lemma 2.6 and the Bogdanov-Takens bifurcation scenario . . . . .	113
<b>4</b>	<b>Outlook</b>	<b>115</b>
	<b>Bibliography</b>	<b>121</b>
	<b>Samenvatting</b>	<b>135</b>
	<b>Curriculum Vitae</b>	<b>139</b>

# Introduction

---

Understanding nature has been one of mankind's biggest yearnings. We are continuously trying to decipher nature's adjustments as a complex environment of which we, the human species, only represents a fraction, albeit a very impactful one [50]. A universal feature of nature is its ability to adapt to external shifting factors. This adaptation results in catastrophic changes [130, 131, 160]- such as the extinction of animal species [17] - or gradual ones, evolving towards, conceivably temporary, new states [81].

One of the major shifting factors throughout history has been climate change, especially since the Industrial Revolution [110]. Climate change is, arguably, the most severe challenge facing planet Earth during the 21st century [57], also according to the United Nations [3].

Soil degradation - which is mainly caused by human activities - has a direct impact on climate change [149, 156]. This directly results in desertification, a form of land degradation in arid and semi-arid areas. Arid and semi-arid areas are land surfaces receiving less than 400mm of annual precipitation [167]. They cover more than 40% of the global land surface [37]. This development has led to escalating concerns about the impact of climate change on desertification and biodiversity loss [5, 47], thereby propelling the search for indicators of imminent ecosystem shifts and the understanding thereof – from a theoretical perspective – in the last two decades. In addition to that, aerial photographs and, later on, satellite images have opened up a research field area that goes beyond laboratory dimensions, trying to understand the mysteries of these landscapes such as spatial plant distributions [58, 78, 102, 112, 152].

One emergent feature that came to light while studying (semi-)arid landscapes with the aid of aerial photographs, is nature's adaptability to 'self-organize' in order to sustain itself under changing environmental conditions [32, 102, 106].

Self-organization is the emergence of patterns and order in a system by internal processes, rather than external constraints or forces [68]. As a result, local interactions between individuals may often spawn global features. Self-organization and the formation of patterns encapsulates all fields and scales, from the smallest molecular DNA structure [85] to the Copernican solar system model [170] and beyond, making it a universal feature of nature. Ecological examples of self-organization range from the structure of animal social groups to spatial patterns associated with plant distributions [68]. A few visual examples are given in Figure 1.1.



(a) The cheetah has a spotted body with a stripy tail [4].



(b) Google Earth satellite image of a fractal pattern in the Egyptian desert ( $27^{\circ}28'N$ ;  $32^{\circ}14'E$ ).

**Figure 1.1:** Different patterns at different scales in nature.

Vegetation patterns form a characteristic feature of semi-arid regions through the formation of distinct spatial plant distributions [137]. Water scarcity hampers the ecosystem into spatial separation of plants; patches of dense biomass and areas of bare soil appear, induced by a positive feedback between local vegetation growth and water redistribution towards areas of high biomass [126]. This manifests itself in different shapes and on different spatial scales as can be seen in Figure 1.2.

Lots of research within the ecological community has already been done about this specific type of land degradation [58, 61, 66, 87, 88, 100, 138, 141, 145, 161, 172, 176]. The proposed ecological models studied in these papers have already shown to numerically reproduce the vegetation patterns observed [14, 90, 124].

Given the magnitude and tragic consequences of the process of land degradation and desertification, a few questions naturally present themselves

- How does the study of vegetation patches in any way relate to mathematics where - stereotypically - research relies on pencil, paper and nowadays some computational power?



(a) Fairy circles in the Namibian desert [1]. (b) Striped pattern of woody vegetation (Acacia) that grows parallel to dry slopes in Australia [2].

**Figure 1.2:** Different vegetation patterns located in different parts of the world.

- How can the abstractness and clarity of mathematics help with the randomness and complexity of the real ecological world?
- Can ecology and mathematics fruitfully join efforts in order to get a better understanding and grasp of this pressing phenomenon?

The answers to these questions are not that simple nor straightforward. On the one hand, aerial observations do not form proof of a universal feature. Adding to that, solely relying on numerical simulations has its restrictions as the latter rely on a whole set of specific conditions, making it harder to formulate generalized statements about an ecosystem. On the other hand, the complexity of vegetation ecosystems limits rigorous mathematical analysis. Fortunately, pattern formation models can play an important role in validating (or refuting) the results acquired from topographical data and numerics.

The challenge and aim of this thesis is bridging the gap between ecology and mathematics, real world observations and theorems and finding a common ground to further build new research on. It transcends the study of spatial vegetation structures to enclose a whole variety of patterns that fit into the generalized mathematical framework of dynamical systems. The motivation behind this piece of work is threefold:

- To assess the extent to which conceptual models' predictions and results apply to real world observations as has been done in Chapter 2.
- To extend the ecological insights acquired from an ecosystem model and apply rigorous mathematical analysis to prove the existence of various (already observed) spatial patterns as has been done in Chapter 3.
- To exploit the rich nonlinear structure of an ecosystem model and to investigate the existence of a broad class of novel spatial patterns, both from an

ecological and a mathematical perspective as has been done in Chapter 3.

So how does mathematics lend a hand to ecology and vice versa? A few of the models proposed within the ecological community rely on reaction-(advection)-diffusion equations [14, 90, 124], which have played a key role in understanding the spatio-temporal evolution of complex dynamical systems. Two of these models are studied within this thesis, and have been meticulously selected to achieve the aims listed above: the most conceptual model at hand, in order to maximize data and model prediction comparison and in contrast, one of the most elaborated ecosystem models that had not been analyzed in such a mathematical way before. Both these models are, deducibly, or reaction-(advection)-diffusion type.

A reaction-diffusion equation, as the term suggests, comprises a reaction term and a diffusion term, describing the evolution in time and space of a certain concentration or density of a component of a system. The addition of a diffusion term to a reactive system, according to Turing [150], would disrupt a system in a way it may change from a homogeneous spatial state to a heterogeneous patterned state, giving rise to all sorts of spatial structures [150].

Pattern formation in this type of equations is of high interest both in ecology and mathematics. Understanding the cross-fertilization of ecology and mathematics along the specific problem at hand can be hard. The complexity of the domain of pattern formation compels the boundaries of different areas of research to fuse. Trying to bridge the gaps and breaches between different fields of science can be a tough quest, starting with the linguistic barrier of mere definitions and concepts. Therefore, a thorough and detailed presentation of the basics of each side of the spectrum is required, in order to set a firm and clear working ground to advance with.

First, an ecological background is given, exploring the methodologies and mechanisms preceding the models discussed. Second, an overview of the mathematical concepts and tools is given, as a basic foundation of the work within this thesis. Last but not least, a brief outline of the studies carried out in this work is presented.

## 1 Ecological background

*The pattern of variation shown by the distribution of species among quadrants of the earth's surface chosen at random hovers in a tantalizing manner between the continuous and the discontinuous [164].*

## 1.1 Ecological context

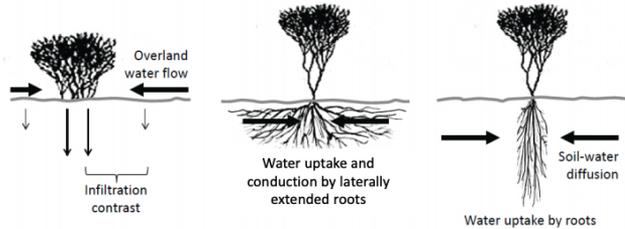
The complexity and diversity of factors involved in the study of ecological systems, in particular vegetation ecology, demands for a reflection on the scientific methodology adopted and for a diligent modeling approach. This has been cause for divisiveness within the ecological community since the early twentieth century [154]. However, the concept of patterns in vegetation had already become standard since Watt's 1947 paper on *pattern and process in the plant community* [163].

It is to be expected that the most consistent and clearest ecological patterns will almost invariably have several explanations, driven by various mechanisms all pulling in the same direction. According to Whittaker [166], two main principles are at play when it comes to spatial vegetation dynamics, forming the skeleton of the different modeling approaches adopted so far:

- The principle of species individuality – each species has its unique environmental requirements and distribution, according to its own genetic structure, physiological characteristics and population dynamics. The individualistic nature of the species reflects itself in the individualistic nature of plant community distributions.
- The principle of community continuity, i.e., continuum principle – communities which occur along continuous environmental gradients usually intergrade continuously, with gradual changes in population levels of species along the gradient.

These two principles resulted in two main modeling approaches in use in studies of plant population dynamics [106]; agent-based models, i.e., stochastic computational algorithms, based on the individuality principle and often capture the characteristics of a plant in great detail. The continuum principle was incorporated in partial differential equations (PDE's) which focus on deterministic processes at small spatial scales rather than addressing individual plants. The plant population is then described by a continuous biomass areal density [106]. The second approach has been adopted throughout this thesis, as it fits into the well-developed theory of PDE's. Within these ecosystem models, different types of environmental variables can be recognized, e.g. of abiotic or biotic type. Abiotic variables such as precipitation and soil nitrogen content directly determine plant growth and success. Biotic variables such as the competition from other plants, pathogens, herbivores and fungi (mycorrhizae), may be beneficial or detrimental to plant growth or have complex effects contingent on abiotic variables [154]. The main abiotic variable is the amount of water a plant can acquire in order to survive. Therefore, in order to understand the basic dynamics of plant population, most models have stripped these complex ecosystems to their "bare essentials": biomass and water, and the interaction thereof.

Water transport has two main effects; providing an extra source of water to vegetation patches as well as inhibiting vegetation growth in the patch surroundings. As shown and explained in Figure 1.3, there are different ways water transport contributes to biomass and root system growth.



**Figure 1.3:** Schematic illustration of three different modes of water transport capable of inducing pattern-forming feedback in water-limited vegetation.

*Left: Infiltration feedback:* typically presents itself in soil areas covered in soil crust that change the rate of surface-water infiltration inducing overland-water flow towards densely vegetated areas.

*Middle: Root-augmentation feedback:* laterally extended root systems allow water uptake and conduction from a larger domain.

*Right: Soil-water diffusion feedback:* confined root systems create a soil-water gradient due to high water uptake resulting in lateral water diffusion [106].

The most simple reaction-(advection)-diffusion system of partial differential equation that would heuristically describe the dynamics of a plant population in a water-scarce environment takes the form

$$\begin{cases} \text{change in biomass} & = \text{growth} - \text{mortality} + \text{spread through seed dispersal}, \\ \text{change in water} & = \text{precipitation} - \text{evaporation} - \text{uptake} + \text{spread of water}. \end{cases} \quad (1.1)$$

In the above, the abiotic variable has been "limited" to precipitation and evaporation, i.e., water availability, and the biotic variable is modeled through seed dispersal and plant growth/mortality.

The above heuristic concept has been translated into several PDE models for dryland vegetation [62, 63, 90, 97, 124, 161]. The simplest two-component model that takes into account both biomass and water was proposed by Klausmeier [90]. The extended version of this model (system (1.2)) is studied in Chapter 2. More detailed models include two water variables representing soil-water content and overland-water flow [62, 79, 142].

Finding a fruitful balance in the realm of adequately modeling vegetation patterns in semi-arid ecosystems is a double-edged blade. Models that capture

the plant population dynamics in more detail are more advantageous in that they provide better defined and measurable parameters, though being less prone to mathematical analysis. In contrast, considering specific contexts for vegetation patterns such as steady rainfall, a single plant species, high soil-water diffusivity, allows for model simplifications that facilitate mathematical assessment of the model [107]. The latter method has been adopted for the analysis of the ecosystems studied, which we will present next.

## 1.2 Ecosystem models studied

*“Dynamic models bridge over two other modeling concepts in ecology; descriptive models, which presents relationships between measured variables, either graphically or algebraically and conceptual models, which seek to unravel basic mechanisms that underlie observed behaviors, but remain at the level of qualitative propositions [106].”*

Two dynamical ecosystem models have been studied in detail in this thesis, both two-component reaction–diffusion models describing the interaction between plant biomass and water. They exhibit spatial vegetation patterns that have been observed in semi-arid regions [61, 102]. Both models have been analyzed from different perspectives and with different purposes. The mathematical processes and patterns enclosed within these two models are of the highest interest, from ecological and analytical perspectives. The proposed models do not only grant deeper understanding of the process of desertification but also dig into an even richer realm of patterns, with novel connections and mathematical frameworks that have yet to be explored in more detail.

The first model is the extended-Klausmeier model [155], based on the model proposed by Klausmeier [90]:

$$\begin{cases} \frac{\partial w}{\partial t} = e \frac{\partial^2 w}{\partial x^2} + \frac{\partial(vw)}{\partial x} + a - w - wn^2, \\ \frac{\partial n}{\partial t} = \frac{\partial^2 n}{\partial x^2} - mn + wn^2. \end{cases} \quad (1.2)$$

where  $w(x, t) \geq 0$  and  $n(x, t) \geq 0$  represent areal densities of soil water and biomass, respectively, and  $x \in \mathbb{R}$ ,  $t \in \mathbb{R}^+$  are the space and time coordinates. The reaction terms model the change in water as a combined effect of rainfall (+ $a$ ), evaporation ( $-w$ ) and uptake by plants ( $-wn^2$ ). Note that the coupling to plants is nonlinear, as more plants enable a better uptake of water through a bigger root system. The change of plant biomass comes from mortality ( $-mn$ ) and plant growth ( $+wn^2$ ). Note here again that the plant growth scales nonlinearly, as

more present biomass encourages water uptake and thus growth. Dispersion by plants is modeled as diffusion and the movement of water as a combined effect of diffusion (+ $e$ ) and advection. The latter is due to gradients in the terrain, which are proportional to the slope parameter  $v$ .

The second model is a reduced version of the three-component model proposed by Gilad et al. for flat terrains [63]:

$$\begin{cases} \frac{\partial \tilde{B}}{\partial T} &= \Lambda \tilde{W} \tilde{B} (1 - \tilde{B}/K)(1 + E\tilde{B}) - M\tilde{B} + D_B \frac{\partial^2 \tilde{B}}{\partial \tilde{X}^2}, \\ \frac{\partial \tilde{W}}{\partial T} &= P - N(1 - R\tilde{B}/K)\tilde{W} - \Gamma \tilde{W} \tilde{B} (1 + E\tilde{B}) + D_W \frac{\partial^2 \tilde{W}}{\partial \tilde{X}^2}, \end{cases} \quad (1.3)$$

where  $\tilde{B}(\tilde{X}, T) \geq 0$  and  $\tilde{W}(\tilde{X}, T) \geq 0$  represent areal densities of biomass and soil water, respectively, and  $\tilde{X} \in \mathbb{R}$ ,  $T \in \mathbb{R}^+$  are the space and time coordinates. In the biomass ( $\tilde{B}$ ) equation,  $\Lambda$  represents the biomass growth rate coefficient,  $K$  the maximal standing biomass,  $E$  is a measure for the root-to-shoot ratio,  $M$  the plant mortality rate and  $D_B$  the seed-dispersal or clonal growth rate, while in the water ( $\tilde{W}$ ) equation,  $P$  represents the precipitation rate,  $N$  the evaporation rate,  $R$  the reduction of the evaporation rate due to shading,  $\Gamma$  the water-uptake rate coefficient and  $D_W$  the effective soil water diffusivity. Notice that the power of the factor  $(1 + E\tilde{B})$  in both equations is unity, whereas in the reduced model in [175] the power is two. This difference stems from the consideration in the study of one space dimension rather than two.

System (1.3) represents a similar set-up to system (1.2). The equation lacks a first derivative, that is, an advection term, as it is constrained to flat terrains. Note that the original three component model [62] does account for topography. A simple glimpse at both equations discloses that system (1.3) takes more parameters into account and exhibits a less trivial mathematical structure than system (1.2). This allows for a more realistic modeling of real ecosystems. In addition to reproducing banded vegetation, the model has also reproduced the famous fairy circles [14, 61].

A closer look at both models reveals more contrast. In the extended-Klausmeier model, plant growth is unbounded ( $+wn^2$ ), while in the Gilad et al. model, it is bounded ( $(1 - \tilde{B}/K)$ ). Once the biomass reaches a certain height, it "suffers" from self-shading, which prohibits the process of photosynthesis and therefore further growth. In addition, the water and biomass growth/loss in the extended-Klausmeier model are coupled nonlinearly in the same manner in both equations while in the Gilad et al. model, the coupling between the two variables is more elaborated.

From an ecological perspective, system (1.3), which represents the master equation of this thesis, thus accounts for a vaster parameter space and a more realistic representation of dryland ecosystems. Unfortunately, this also makes it harder to analyze and numerically simulate, as numerics require specific initial conditions, parameter values and boundary conditions. To be blunt, it boils down to trying to grasp the behavior of an infinite dimensional continuous space in a finite dimensional discrete manner, which has its obvious limitations. Rigorous mathematical analysis of ecosystems like (1.2) and (1.3) has formed a bridge between ecologists and mathematicians within the pattern formation community. Setting the basis for the methods used in this type of analysis naturally leads us to the mathematical part of this introduction.

### 1.3 Ecosystem dynamics

Consider the compactly reformulated PDE of system (1.3), as done in Appendix B of Chapter 3

$$\begin{cases} B_t &= (aW - 1)B + WB^2 - WB^3 + B_{xx}, \\ W_t &= \Psi - [\Phi + \Omega B + \Theta B^2]W + \frac{1}{\varepsilon^2}W_{xx}, \end{cases} \quad (1.4)$$

where the emphasis has to be made on the introduction of the small parameter  $\varepsilon^2 = \frac{D_B}{D_W} \ll 1$ , that is,  $\varepsilon$  is asymptotically small compared to 1, which will play a crucial role in the tools presented in this introduction and in the proofs of Chapter 3.

The classical approach to studying systems like (1.4) is to search for stationary and/or uniformly traveling solutions, that is, solutions that are stationary with respect to a co-moving frame with constant speed  $c$ . Introduce a new coordinate  $\xi = x - ct$ , with speed  $c \in \mathbb{R}$ , an a priori free  $\mathcal{O}(1)$  parameter (w.r.t. the asymptotically small parameter  $\varepsilon$ ). By setting

$$(B(x, t), W(x, t)) = (b(\xi), w(\xi)) \quad (1.5)$$

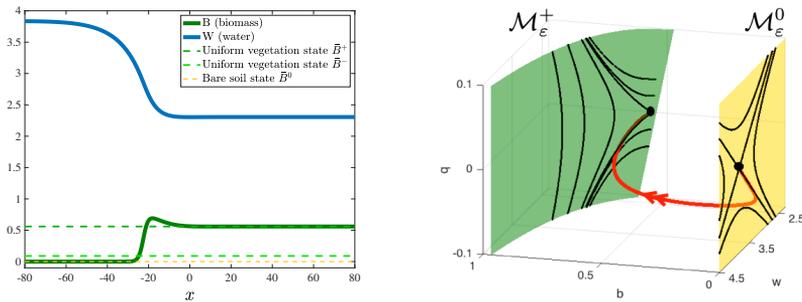
and introducing

$$p = b_\xi, \quad q = \frac{1}{\varepsilon}w_\xi, \quad (1.6)$$

PDE (1.4) reduces to the four-dimensional ‘spatial’ dynamical system

$$\begin{cases} b_\xi &= p, \\ p_\xi &= wb^3 - wb^2 + (1 - aw)b - cp, \\ w_\xi &= \varepsilon q, \\ q_\xi &= \varepsilon(-\Psi + [\Phi + \Omega b + \Theta b^2]w) - \varepsilon^2 cq. \end{cases} \quad (1.7)$$

The fully vegetated state and the bare soil state are equilibria of this system, while solutions connecting the two - so called heteroclinic orbits - represent interfaces of vegetation patterns and, hence, the fundamental entities to study in order to understand pattern formation.



(a) Spatial profiles of the  $B$ - and  $W$ -components of a stationary front solution of the original, unscaled model (1.3) – corresponding to a stationary 1-front.

(b) The profile from (a) as a projection in  $(b, w, q)$ -space. The  $B$ - and  $W$ -coordinates were extracted from the simulation of (a) and  $q$  was computed by numerical differentiation

**Figure 1.4:** (a) Profile of a stationary 1-front connection of the original, unscaled model (1.3) connecting the bare soil state to the vegetated state in real space (a) and in phase space (b).

Although (1.7) represents a 4-dimensional system of ordinary differential equations (ODE's), it is still rather difficult to analyze in full. A "simple" connection between the bare soil state and the fully vegetated state as seen in the simulation from Figure 1.4a takes the complex geometrical form of Figure 1.4b when studied in the context of (1.7). The rigorous mathematical construction of such heteroclinic solutions in 4-D is a challenging task making use of various theories and techniques, the basics of which we review next.

## 2 Mathematical concepts

### 2.1 Reaction-diffusion equations, orbits and patterns

Similar to the complexity of ecosystems and ecology, only a limited amount of statements can be made about most nonlinear dynamical systems and their behavior [72]. Nonetheless, investigating their "somewhat" simpler parts, such as stationary or uniformly traveling solutions, yields a bigger understanding of the dynamics of the system than one might intend to think. In order to properly

address the questions that the dryland models in Chapter 2 and Chapter 3 present, some mathematical concepts and theorems are of indispensable importance.

The following can be seen as a stripped down, hands-on, guide to the basic understanding of the tools and techniques used in the following chapters, accompanied with a simple example to set the scene.

We are looking for stationary solutions of the reaction-diffusion equation

$$\frac{\partial u}{\partial t} = D\Delta u + f(u). \quad (1.8)$$

where  $u(x, t) \in \mathbb{R}^m$  is a vector of state variables describing the densities of a population or a concentration at position  $x \in \mathbb{R}^n \subset \Omega$  with  $\Omega$  an open set, at time  $t \in \mathbb{R}^+$ . On the right side of the equation,  $D$  represents a constant diagonal "diffusion" matrix,  $\Delta$  represents the Laplacian operator working on  $u$  and  $f(u)$  represents a smooth (nonlinear) vector field  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , the "reaction" term of the equation.

Introducing the traveling wave coordinate  $\xi = x - ct$ , with speed  $c \in \mathbb{R}^n$ , above solution is referred to as a traveling wave  $\Phi(\xi) = \Phi(x - ct)$  if it satisfies

$$D\Phi'' + c\Phi' + f(\Phi) = 0. \quad (1.9)$$

Solutions to (1.9) can take different forms, from single fixed points to whole orbits/trajectories. This is quantified within the notion of invariant sets. An invariant set is a set such that a solution of (1.8) through each point of the set lies in the specific set for all time. We can further specify forward and backward invariant sets, respectively. A forward invariant set  $\Lambda^+$  is a set for which  $\Phi(\xi) \in \Lambda^+$  for all  $\xi \geq \xi_+$ . A backward invariant set  $\Lambda^-$  is a set for which  $\Phi(\xi) \in \Lambda^-$  for all  $\xi \leq \xi_-$ . A rough sketch of invariant sets is given in Figure 1.5.

The notion of invariant sets can be cast in the more general notion of invariant manifolds. For sufficiently smooth systems, the invariant set represents an invariant manifold. For a detailed definition, we refer the reader to classical books on differential dynamical systems [104, 169]. Similar to forward and backward invariant sets, we can specify unstable and stable manifolds  $W^s$  and  $W^u$  for a set  $\Lambda$ , respectively:

$$W^s(\Lambda) = \{\xi_0 \in \mathbb{R}^n : \Phi(\xi, \xi_0) \rightarrow \Lambda \text{ for } \xi \rightarrow \infty\}, \quad (1.10)$$

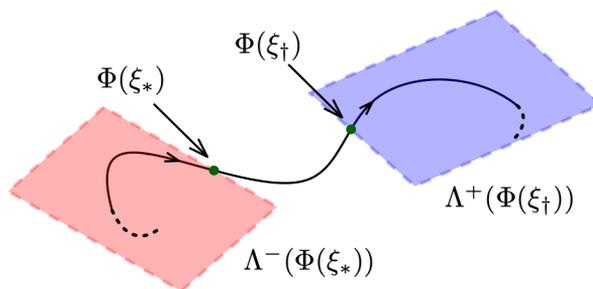
$$W^u(\Lambda) = \{\xi_0 \in \mathbb{R}^n : \Phi(\xi, \xi_0) \rightarrow \Lambda \text{ for } \xi \rightarrow -\infty\}. \quad (1.11)$$

Again, for sufficiently smooth systems,  $W^{s,u}(\Lambda)$  are differentiable manifolds.

Once we focus our attention on fixed points (equilibria) and trajectories connected to them, we can classify the different orbits (that are stationary in the

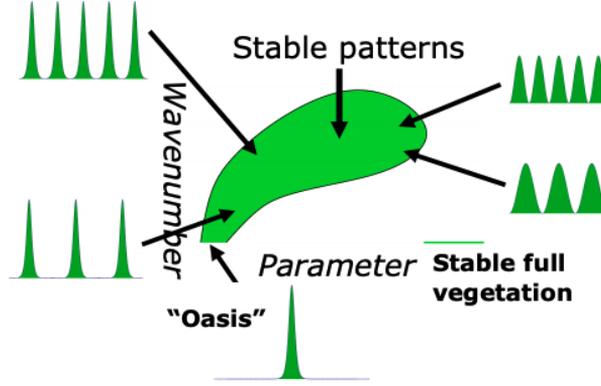
co-moving frame) we are interested in, depending on their behavior within the invariant sets:

- A **homoclinic orbit**  $\gamma(p)$  is a closed orbit where the stable manifold  $W^s(p)$  of a fixed point  $p$  intersects with the unstable manifold  $W^u(p)$  of the same fixed point  $p$ , that is,  $\gamma(p) \subset W^s(p) \cap W^u(p)$  and all points tend to the same fixed point  $p$  for  $\xi \rightarrow \pm\infty$ .
- A **heteroclinic orbit** A heteroclinic orbit  $\gamma(p_1, p_2)$  is an orbit where for two different fixed points  $p_1$  and  $p_2$  the unstable manifold  $W^u(p_1)$  of  $p_1$  intersects with the stable manifold  $W^s(p_2)$  of  $p_2$ , that is,  $\gamma(p_1, p_2) \subset W^u(p_1) \cap W^s(p_2)$ , and for which all points tend to  $p_1$  and  $p_2$  for  $\xi \rightarrow -\infty$  and  $\xi \rightarrow +\infty$ , respectively.



**Figure 1.5:** Sketch of unstable (blue) and stable (red) invariant sets  $\Lambda^+$  and  $\Lambda^-$ , respectively, of a trajectory  $\Phi$ . For every  $\xi \geq \xi_+$ , it holds that  $\Phi(\xi) \in \Lambda^+(\Phi(\xi_+))$  and for every  $\xi \leq \xi_*$ , it holds that  $\Phi(\xi) \in \Lambda^-(\Phi(\xi_*))$ .

Besides these two orbits, frequently studied solutions are **periodic orbits**, to which Chapter 2 is basically dedicated. Periodic orbits have the characteristic that they contain no fixed points and that there exists a specific  $0 < L < \infty$  with the property that  $\Phi(\xi + L) = \Phi(\xi)$  for all  $\xi \in \mathbb{R}^n$ . This is equivalent to having a distinct wavenumber. A system of equations such as (1.8) generally displays a whole continuous families of periodic orbits [39]. From the point of view of the underlying PDE (1.8), these orbits are spatially periodic (traveling) ‘wave trains’. Of course, one is especially interested in those wave trains that are stable and thus observable as solutions of the PDE. The family of stable “observable” periodic orbits within (parameter, wavenumber)-space is referred to as the “Busse balloon” [18]. This concept is extensively studied within Chapter 2, where theory and real world data are put side by side. A schematic representation of a section of a Busse balloon is given in Figure 1.6.



**Figure 1.6:** A schematic representation of a section of a Busse balloon, where the stable patterns are depicted in (parameter, wavenumber)-space. The green area encloses all stable periodic orbits. Outside this area the periodic patterns are unstable. As the parameter is declined, the patterns become narrower and vice versa. Per parameter value, there exists a whole set of observable stable patterns with different wavenumbers.

In order to get a visual understanding of the potential different patterns presented above, consider the following one-component toy model in one spatial dimension:

$$u_t = u_{xx} - u^3 - u^2 + (1 - a)u \quad (1.12)$$

where  $a \in \mathbb{R}$  is a free parameter we can vary. Note that this is the equivalent of PDE (1.8) in this simplified setting, with diffusion term  $D\Delta u = u_{xx}$  and nonlinear reaction term  $f(u) = u^3 - u^2 + (1 - a)u$ . Introducing  $v = u_x$  and looking at stationary solutions, that is setting  $u_t = 0$ , we get the following two dimensional system of ODE's

$$\begin{cases} u_x = v, \\ v_x = u^3 - u^2 + (1 - a)u, \end{cases} \quad (1.13)$$

System (1.13) represents a stripped down version of equation (1.9) where we have set for simplicity  $c = 0$  (i.e., we only consider stationary patterns). It exhibits up to three fixed points, given by:

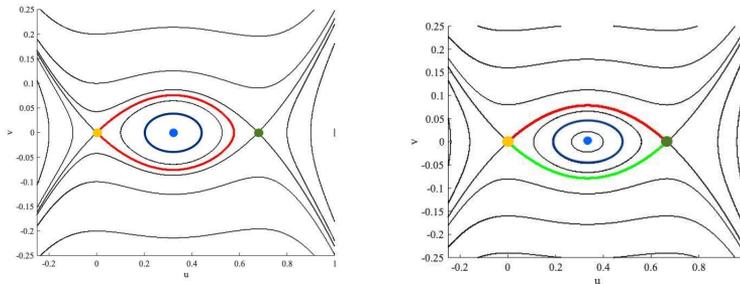
$$\begin{cases} (u_0, v_0) = (0, 0) \\ (u_{\pm}, v_0) = \left( \frac{1 \pm \sqrt{4a - 3}}{2}, 0 \right). \end{cases} \quad (1.14)$$

Clearly, for  $a < \frac{3}{4}$ , system (1.13) possesses only one fixed point  $(u_0, v_0) = (0, 0)$ .

For  $a = \frac{3}{4}$ , it possesses two fixed points  $(u_0, v_0) = (0, 0)$  and  $(u_{\pm}, v_{\pm}) = (\frac{1}{2}, 0)$ . For  $a = 1$ , it again possesses two fixed points  $(u_0, v_0) = (u_-, v_0) = (0, 0)$  and  $(u_+, v_0) = (1, 0)$ . For  $a \in (\frac{3}{4}, \infty) \setminus \{1\}$  we get three distinct fixed points as given in (1.14). Different values of  $a$  thus generate different possible connections - orbits - between the different fixed points. To illustrate the possible orbits we focus on the case where  $a > \frac{3}{4}$ . System (1.13) represents a Hamiltonian system. The solutions to the set of equations lay on the level sets of the Hamiltonian  $\mathcal{H} = b$  with  $b \in \mathbb{R}$  where

$$\mathcal{H}(u, v) = \frac{v^2}{2} - \frac{u^4}{4} + \frac{u^3}{3} - \frac{(1-a)u^2}{2}. \quad (1.15)$$

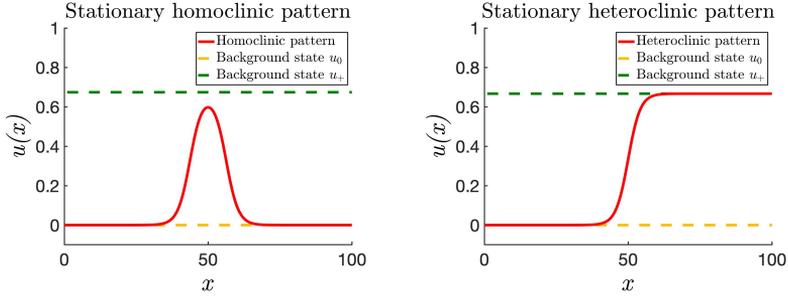
This yields *phase portraits*, that can look different, depending on the value of  $a$ , as can be seen in Figure 1.7.



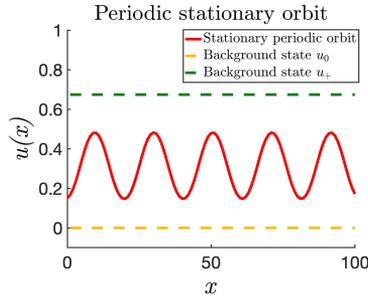
- (a) Homoclinic orbit connecting the fixed point  $(u_0, v_0)$  to itself for  $a = \frac{7}{9} + \frac{1}{250}$ . (b) Heteroclinic orbits connecting the fixed point  $(u_0, v_0)$  to the fixed point  $(u_+, v_0)$  with  $a = \frac{7}{9}$ . Note that due to the symmetry of the system we have two heteroclinic orbits (red and green).

**Figure 1.7:** Phase portraits of solutions of system (1.13).

Each orbit in a phase portrait corresponds to a distinct spatial pattern in real space associated to (1.13). A homoclinic orbit in the ODE (1.13) system corresponds to what is commonly referred to as a pulse in the PDE (1.12) as shown in Figure 1.8a, a heteroclinic orbit similarly corresponds to a 1-front connection (Figure 1.8b) and a periodic orbit simply corresponds to a periodic pattern (Figure 1.8c).



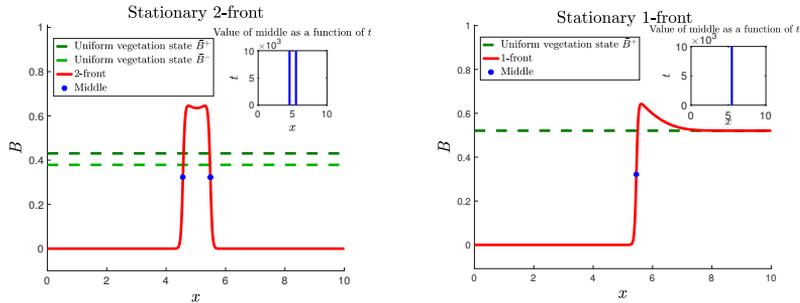
- (a) The stationary homoclinic orbit corresponding to the phase portrait of Figure 1.7a and  $a = \frac{7}{9} + \frac{1}{250}$ .
- (b) The unique stationary heteroclinic orbit connecting  $(u_0, v_0)$  to  $(u_+, v_0)$  in the phase portrait of Figure 1.7b for  $a = \frac{7}{9}$ , where  $v \geq 0 \quad \forall x$ .



- (c) A stationary periodic orbit for  $a = \frac{7}{9} + \frac{1}{200}$ .

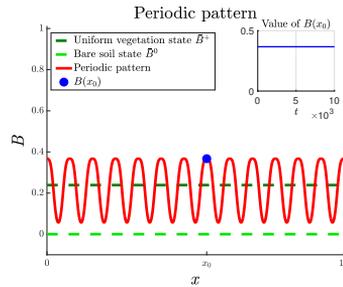
**Figure 1.8:** Spatial profile of the different stationary patterns exhibited by system (1.12).

The analysis and spatial patterns extracted from toy model (1.12) are pretty simple and straightforward. The patterns obtained from toy model (1.12) are similar to those observed in real ecosystems, such as the ones described in Chapter 3, where the ecosystem model involved (1.3) is more realistic yet more mathematically involved. This yields the following "real ecosystem patterns counterpart" of the patterns in Figure 1.8.



(a) A homoclinic stationary 2-front spot pattern of system (1.4).

(b) The unique heteroclinic stationary 1-front pattern of system (1.4).



(c) A small amplitude stationary spatially periodic solution generated by a Turing bifurcation in system (1.4).

**Figure 1.9:** Spatial profile of the different patterns exhibited by system (1.4). For the exact parameter settings, we refer to section 4 of Chapter 3.

The spatial profile of the homoclinic and heteroclinic patterns of Figure 1.9 is non-monotonous, due to the geometry of the invariant slow manifolds of model (1.3). Due to the complexity of the model at hand, basic ODE solving tools often fail to deliver. Luckily, when it comes to connections, especially homoclinic and heteroclinic, a mathematical theory comes in handy: geometrical singular perturbation theory.

## 2.2 Geometric singular perturbation theory

Geometrical singular perturbation theory is a geometric approach revolving around scale separation within a dynamical system, as is the case in the models presented in Section 1.2. The geometric approach endeavors to identify the central dynamical structures such as invariant sets and invariant manifolds, present in the phase space of the system. Often within reaction-diffusion equations, the time evolution of different variables differs by several orders of magnitude. Similarly, their spatial evolution can take place on various length scales. A ‘magnification’ of (one of) these scale separations can be very insightful. Assumptions with respect to processes that have different time or length scales might simplify the problem we want to understand, yielding a first insight into the structures and dynamics involved. Assuming that very fast processes regulate instantly or that slow processes stand still are examples of these possible magnifications. They describe the system in some *limiting case*, which is usually easier to analyze mathematically. The foundation of this approach was set by Fenichel [51–54]. Since then, these mathematical methods have evolved and found their way towards applications, including ecological models. For a more mathematical introduction, including proofs of the theorems, we refer to the surveys of Jones [83] and Kaper [86].

Consider the system of singularly perturbed ordinary differential equations in general form:

$$\begin{cases} u_x = f(u, v, \varepsilon), \\ v_x = \varepsilon g(u, v, \varepsilon). \end{cases} \quad (1.16)$$

In the equation above,  $u \in \mathbb{R}^m$ ,  $v \in \mathbb{R}^n$ . The constant real parameter  $\varepsilon$  is small, i.e.,  $0 < \varepsilon \ll 1$ . The functions  $f$  and  $g$  are assumed to be sufficiently smooth. ‘Sufficiently smooth’ here means at least  $C^1$  in  $u, v$  and  $\varepsilon$ .

A reformulation of system (1.16) in terms of the rescaled variable  $y = \varepsilon x$  yields

$$\begin{cases} \varepsilon u_y = f(u, v, \varepsilon), \\ v_y = g(u, v, \varepsilon). \end{cases} \quad (1.17)$$

Systems (1.16) and (1.17) are equivalent for  $\varepsilon \neq 0$ . System (1.16) is referred to as the *fast system* and system (1.17) is referred to as the *slow system*. For  $\varepsilon = 0$ , the fast system (1.16) reduces to an  $m$ -dimensional *reduced fast system* on a manifold  $\{f(u, v_0) = 0\}$  in  $(m + n)$ -dimensional space with the  $n$ -dimensional variable  $v$  as a constant parameter

$$\begin{cases} u_x = f(u, v_0, 0), \\ v = v_0. \end{cases} \quad (1.18)$$

For  $\varepsilon = 0$ , the slow system (1.17) reduces to an  $n$ -dimensional *reduced slow system* with

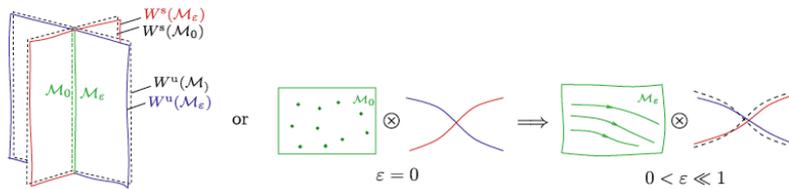
$$\begin{cases} 0 = f(u, v, 0), \\ v_y = g(u, v, 0). \end{cases} \quad (1.19)$$

Under (1.18) the solutions are defined in  $\mathbb{R}^{m+n}$ , but are in fact an  $n$ -parameter family of  $m$ -dimensional systems. Moreover, the solutions under (1.18) on the  $n$ -dimensional set  $f(u, v, 0) = 0$  are trivial. On the other hand, (1.19) does prescribe a nontrivial solution on  $f(u, v, 0) = 0$ , but at the same time it is limited to only this set.

Geometric singular perturbation theory seeks to exploit the properties of these fast-slow decompositions and the intersections of various manifolds in order to establish the existence of different desired orbits (periodic, heteroclinic, homoclinic). The two reduced systems (1.18) and (1.19) offer different insights into the behavior of orbits and geometric structures in the fast and slow regimes, respectively, but in either formulation the  $\varepsilon > 0$  system (1.16) can't be described in full. The goal of geometric singular perturbation theory is to analyze the dynamics of system (1.16) with  $\varepsilon$  nonzero but small by suitably combining the dynamics of the two limits of the reduced fast and slow systems. Certain geometric structures in the phase space of (1.16) with  $0 < \varepsilon \ll 1$  can be found close to counterparts from (1.18). These structures are usually easier to locate. In addition, certain other geometric constructs of (1.16) with  $0 < \varepsilon \ll 1$  lie close to objects in the phase space of (1.19). By exploiting this decomposition into fast and slow, the geometric approach reduces the full singularly perturbed system to separate lower-dimensional regular perturbation problems in the fast and slow regimes, respectively.

The basic ingredient is to combining these 'puzzle pieces' is illustrated in Figure 1.10 and is as follows. Suppose we are given an  $n$ -dimensional manifold  $\mathcal{M}_0$  which is contained in the set  $\{f(u, v, 0) = 0\}$ . Note that this implies that  $\mathcal{M}_0$  consists of critical points of the reduced fast system (1.18). Of all the systems like (1.16), we focus on the one for which the manifold  $\mathcal{M}_0$  consists of only hyperbolic points of the reduced fast system (1.18), that is, critical points whose

eigenvalues  $\lambda$  of the Jacobian  $\frac{\partial f}{\partial u}(u, v, 0)|_{\text{hyperbolic point}}$  are uniformly bounded away from the imaginary axis. Then the so-called critical manifold  $\mathcal{M}_0$ , is said to be normally hyperbolic, and persists as a *locally invariant slow manifold*  $\mathcal{M}_\varepsilon$  of the full problem (1.16) that is  $\mathcal{O}(\varepsilon)$  close to  $\mathcal{M}_0$ . The restriction of the flow (1.17) to  $\mathcal{M}_\varepsilon$  is a small perturbation of the trajectories of the limiting problem (1.19). Moreover, the stable and unstable manifolds  $\mathcal{W}^s(\mathcal{M}_0)$  and  $\mathcal{W}^u(\mathcal{M}_0)$  of  $\mathcal{M}_0$  persist as manifolds  $\mathcal{W}^s(\mathcal{M}_\varepsilon)$  and  $\mathcal{W}^u(\mathcal{M}_\varepsilon)$  too. They lie within  $\mathcal{O}(\varepsilon)$  distance of, and are diffeomorphic to,  $\mathcal{W}^s(\mathcal{M}_0)$  and  $\mathcal{W}^u(\mathcal{M}_0)$  respectively [77]. These persistence theorems, proved by Fenichel [51–54], form the basis for the construction of global singular orbits. By ‘gluing’ together fast and slow pieces obtained in the fast and slow regimes, respectively, and verifying the persistence of these global structures for small  $\varepsilon \neq 0$ , the desired orbits can be obtained for the full system.



**Figure 1.10:** Unperturbed critical manifolds  $\mathcal{M}_0$  consisting of fixed points ( $\bullet$ ) and their local stable and unstable manifolds  $\mathcal{W}^{s,u}(\mathcal{M}_0)$ . The manifolds persist for  $0 < \varepsilon \ll 1$  as perturbed manifolds  $\mathcal{M}_\varepsilon$  with trajectories on them [77].

### 3 The onset of patterns

The study of the formation and nonlinear evolution of ecosystems canonically emerges from combining the pre-requisites of Section 1 and Section 2 of this introduction. This section is focused on providing the mathematical mechanisms by which spatio-temporal heterogeneities may occur in a reaction-diffusion system, and their interpretation in terms of the underlying ecological mechanisms. In doing so, we lay out the basics of the localized structures considered in Chapter 2 and Chapter 3 of this thesis.

Consider equation (1.8), in the simplest case that still generates rich pattern dynamics, that is, a 2-component reaction-diffusion equation in one spatial dimension, the unbounded domain  $\mathbb{R}$ :

$$\begin{cases} u_t = u_{xx} + f(u, v; \mu), \\ v_t = dv_{xx} + g(u, v; \mu), \end{cases} \quad (1.20)$$

where  $u(x, t), v(x, t) \in \mathbb{R}$  with  $x \in \mathbb{R}$  and  $t \in \mathbb{R}^+$ . The coefficient  $d$  is a constant diffusion coefficient and the vector fields  $f(u, v; \mu), g(u, v; \mu) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  represent the reaction terms of the equations where we specified their dependence on (a number of) constant parameters  $\mu \in \mathbb{R}^m$  of the system.

In the ecological setting of this thesis, system (1.20) models the spatial interactions between biomass density and water in vegetation ecosystems on a flat terrain. In reality, ecosystems are subject to another spatial property of their environment: its topography. This is characterized by the (local) slope of the terrain, which can have a significant effect on the vegetation dynamics. It has been observed that vegetation aligns in resilient striped patterns perpendicular to the slope of the terrain due to an oriented flow of water downhill [6, 9, 33, 101, 152]. Therefore, we will study the effect of the presence of slope on the classical ‘flat terrain’ Turing destabilization.

Including slope into system (1.20), the displacement of water is mathematically modeled as a combined effect of diffusion ( $dv_{xx}$ ) and advection ( $sv_x$ ), where  $s \in \mathbb{R}$  is a measure for the slope of the terrain [90]. The latter is a topographical effect, which in general depends on the spatial variable  $x$ , that is, takes the form  $h(x)$ . This would yield addition terms  $h_x v_x + h_{xx} v$  in (1.20) [11]. Thus by modeling the topographical effect as  $sv_x$  with  $s$  a constant, we consider the case of a constant slope,  $h(x) \equiv sx$ .

In the general mathematical setting of the system (1.20), we thus extend the reaction-diffusion equation into a (2-component) reaction-advection-diffusion

equation (in one spatial dimension) of the form

$$\begin{cases} u_t = u_{xx} + f(u, v; \mu), \\ v_t = dv_{xx} + sv_x + g(u, v; \mu). \end{cases} \quad (1.21)$$

In the ecological setting,  $v$  represents the water equation, where the diffusion coefficient  $d \gg 1$  as water diffuses on a much faster timescale than vegetation disperses, the latter being described by the  $u$ -equation. Since  $v_t = sv_x$  yields  $v = v(x + st)$ , i.e., an advective transport with speed  $-s$ , this means that increasing  $x$  corresponds to the uphill direction. Moreover, for the case  $s = 0$ , system (1.21) is equivalent to system (1.20), which has been studied extensively in [39] and the references therein.

Denote the trivial background states of (1.20) and (1.21) as  $(U, V)$ , for which we know that  $f(U, V) = g(U, V) = 0$ . Note that the stationary solutions are the same for both systems, as the additions of the  $sv_x$  term does not alter the steady state solutions.

The spectral – or linearized – stability of  $(U, V)$  of system (1.21) against (bounded) perturbations can be determined by plugging the decomposition,

$$\begin{pmatrix} u(x, t) \\ v(x, t) \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} + \begin{pmatrix} u_p \\ v_p \end{pmatrix} \quad (1.22)$$

with

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{ikx + \lambda t} + c.c. \quad (1.23)$$

into (1.21), with  $k \in \mathbb{R}$ ,  $\lambda \in \mathbb{C}$  and  $(\alpha, \beta) \in \mathbb{C}^2$ . The linearized  $2 \times 2$  eigenvalue problem is then determined by

$$\mathcal{A}(k; \mu) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} f_u - k^2 & f_v \\ g_u & g_v - dk^2 + isk \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \lambda \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (1.24)$$

with

$$f_u = \frac{\partial f}{\partial u}(U(\mu), V(\mu); \mu), f_v(\mu) = \frac{\partial f}{\partial v}(U(\mu), V(\mu); \mu), \text{ etc.} \quad (1.25)$$

Since  $k \in \mathbb{R}$ , the associated characteristic polynomial,

$$\lambda^2 - \lambda[(f_u + g_v - (1+d)k^2 + isk)] + [dk^4 - (g_v + df_u)k^2 + (f_u g_v - g_u f_v)] + isk(f_u - k^2) = 0 \quad (1.26)$$

defines 2 functions  $\lambda_{1,2} : \mathbb{R} \rightarrow \mathbb{C}$  which we assume to be ordered;  $\text{Re}(\lambda_2(k)) \leq \text{Re}(\lambda_1(k))$ . Note that  $\lambda_j$  is symmetric in  $k$  for  $s = 0$  while  $\lambda(-k) = \bar{\lambda}(k)$  for

$s \neq 0$ .

A trivial background state  $(U, V)$  of (1.21) is spectrally stable for those values of the parameter  $\mu$  for which  $\text{Re}(\lambda_{1,2}(k; \mu)) < 0$  for all  $k \in \mathbb{R}$ . Pattern formation sets in (from  $(U, V)$ ) as  $\mu$  crosses through a critical value  $\mu_c$  beyond which there are values of  $k$  for which  $\text{Re}(\lambda_1(k; \mu)) > 0$ . Here we focus on the Turing mechanism, i.e., we assume that  $(U, V)$  is marginally stable at  $\mu = \mu_c$  and that there is a  $k_c \neq 0$  such that  $\text{Re}(\lambda_1(\pm k_c; \mu_c)) = 0$  while  $\text{Re}(\lambda_1(k; \mu_c)) < 0$  for all  $k \in \mathbb{R}$  with  $k \neq \pm k_c$ . Note that this implicitly implies that  $\text{Re}(\lambda_{1,2})(0, \mu_c) < 0$ , i.e., that  $(U, V)$  is stable against spatially homogeneous perturbations [113, 150]. It follows from the smoothness of  $\lambda_{1,2}(k; \mu)$ , that  $k = k_c$  must be a local maximum of  $\text{Re}(\lambda_1(k; \mu))$ :  $\text{Re}(\frac{d\lambda_1}{dk}(\pm k_c; \mu_c)) = 0$  for all  $k$  in neighborhoods of  $\pm k_c$ . Analyzing this for the classical case  $s = 0$  yields

$$k_c = k_c(\mu) = \pm \sqrt{\frac{df_u + g_v}{2d}} \quad \text{with} \quad df_u + g_v > 0, \quad (1.27)$$

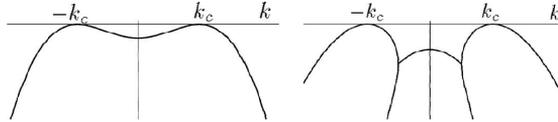
(see for instance [39, 113]). The critical value  $\mu_c$  can be computed by plugging (1.27) back into (1.26). Further manipulations, for which we refer the reader to [39], yield the classical conditions for the destabilization by the Turing mechanism

$$f_u > 0, g_v < 0, \quad (1.28)$$

$$f_u + g_v < 0, \quad (1.29)$$

$$f_u g_v - f_v g_u > 0, \quad (1.30)$$

where we have used that  $d > 1$ . Note that the two components  $u(x, t)$  and  $v(x, t)$  are thus diffusing with different speeds, thereby fulfilling the roles of an *activator*  $u$  and an *inhibitor*  $v$  within the equation, that is,  $f_u(U(\mu_c), V(\mu_c); \mu_c) > 0$  and  $g_v(U(\mu_c), V(\mu_c); \mu_c) < 0$  Figure 1.11, depicts two different critical eigenvalue curves  $\lambda_{1,2}(k)$  for the distinct possible cases in the classical scenario  $s = 0$ .



**Figure 1.11:** Two critical eigenvalue curves in the classical case  $s = 0$ . (a,b): The real parts of the solutions  $\lambda_{1,2}(k)$  of (1.26) for a Turing destabilization, i.e.,  $k_c \neq 0$  with  $\lambda_{1,2}(k) \in \mathbb{R}$  for all  $k \in \mathbb{R}$  in (a), and parameter combinations such that  $\lambda_{1,2}(k) \notin \mathbb{R}$  for certain  $k$  in (b) citeReview.

To investigate the impact of the slope on the classical Turing bifurcation and to minimize the technicalities of the stability analysis, we assume that  $s$  is small, i.e., we introduce  $0 < \delta \ll 1$  and set  $s = \delta\tilde{s}$ .

We know from the above classical Turing analysis that  $\lambda_1(k; \mu)$ , that is, the critical  $\lambda_j$ , is real for  $(k, \mu)$  close to  $(k_c, \mu_c)$ . Thus we may conclude from (1.26) that

$$\lambda_1(k; \mu) = \lambda_r(k; \mu) + \lambda_i(k; \mu) = \lambda_r + i\delta\tilde{\lambda}_i(k; \mu) \quad (1.31)$$

A straightforward perturbation analysis yields that

$$k_c(s) = k_c(0) + \mathcal{O}(\delta^2) \quad \text{and} \quad \mu_c = \mu_c(0) + \mathcal{O}(\delta^2), \quad (1.32)$$

with  $k_c(0), \mu_c(0)$  as determined by the classical Turing analysis (i.e., in (1.20)/(1.21) with  $s = 0$  – see Appendix 1.A). That same analysis yields

$$\lambda_i(k_c) = \delta\tilde{\lambda}_i(k_c) = -\frac{\tilde{s}}{d-1}k_c\delta + \mathcal{O}(\delta^3). \quad (1.33)$$

Thus we find by (1.22) that the most critical perturbation, i.e., the perturbation that is the first to start growing at the Turing destabilization is given by

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} \sim \begin{pmatrix} \alpha_c \\ \beta_c \end{pmatrix} e^{ik_c x + \lambda(k_c, \mu_c)t} + c.c. = \begin{pmatrix} \alpha_c \\ \beta_c \end{pmatrix} e^{ik_c \left( x - \frac{\delta\tilde{s}}{d-1}t \right) (1 + \mathcal{O}(\delta^2))} + c.c. \quad (1.34)$$

and travels with speed

$$c = \frac{\delta\tilde{s}}{d-1} + \mathcal{O}(\delta^3). \quad (1.35)$$

The interpretation in terms of (1.21) as ecosystem model in biomass  $u$  and water  $v$  confirms the ecological intuition that vegetation patterns should move uphill: as was already argued, the term  $sv_x$  in (1.21) determines the downhill direction by its associated ‘advective speed’ =  $-s$ . Since water diffuses much faster than biomass, i.e.,  $d \gg 1$ , in (1.35) Turing patterns travel in the opposite direction: biomass travels towards the down flowing water.

While the conditions for the onset of a Turing destabilization are relatively straightforward, the determination of the nature of the pattern that is exhibited is a more difficult problem since beyond the bifurcation point, a finite band of wavenumbers is growing exponentially. Therefore, it is *a priori* unclear whether the perturbations also evolve into stable (spatially periodic) Turing patterns. This is actually only the case if the associated Turing bifurcation is supercritical [39]. The nature of the Turing bifurcation can be determined by a weakly nonlinear amplitude equation approach – we refrain from going in the details here.

## 4 Content of thesis

The nature of the research done in this thesis is diverse. Two different "paths" have been taken in executing interdisciplinary research in ecology and mathematics through the lens of pattern formation as will become clear in Chapter 2 and Chapter 3. Nonetheless, the foundation of the work at hand is concordant: an ecological phenomenon is modeled with explicit two-component ecosystem models, allowing for rigorous mathematical analysis and numerical computations, that reproduce the patterns observed in nature. The mixture of analyzing the underlying mathematical structure of the ecosystem model and exploring the ecological data and parameters at hand allows for a deeper understanding of the mechanisms at play, thereby leading to qualitative conclusions about the ecosystem, the patterns it exhibits and the knowledge thereof. This represents a step towards a better comprehension and theoretical predictions of the behavior of the process of desertification as well as pattern formation in ecosystems. The work done in this thesis even goes beyond the latter, by unraveling new insights in both the ecological and mathematical community.

### 4.1 Multistability of striped vegetation patterns

Chapter 2, titled "*Multistability of model and real dryland ecosystems through spatial self-organization*"<sup>1</sup>, represents a novel study providing a meticulous comparison between theoretical model predictions and empirical data of spatial vegetation patterns in dryland ecosystems.

Model predictions by the extended-Klausmeier model, which exhibits different patterns (such as bands and gaps) are put side by side with state of the art aerial and topographical observations of banded vegetation patterns obtained in two regions in Somalia and characteristics thereof (biomass density, spatial spread of bands). Since it is considered the most "conceptual" reaction-diffusion model at hand for these types of ecosystems, the extended-Klausmeier model constitutes a perfect candidate to assess basic yet fundamental correlations between theory and data.

Regularly spaced banded vegetation patterns have long been considered to be leading indicators for the proximity of an ecosystem to desertification and changes thereof. They are characterized by their wavenumber/period as well as their size and biomass density. In this study, we establish multi-stability of banded vegetation patterns, which implies the co-existence of a whole (continu-

---

<sup>1</sup>Joint work with Robbin Bastiaansen, Vincent Deblauwe, Maarten B. Eppinga, Koen Siteur, Stéphane Mermoz, Alexandre Bouvet, Arjen Doelman and Max Rietkerk; has appeared as publication in *PNAS* [13].

ous) range of wavenumbers of significant spread, for the same set of parameters of the ecosystem model. That is, an ecosystem can display a whole family of stable (i.e., observable) banded patterns for a given set of parameters. This is theoretically contained in the notion of Busse balloon, the set of stable spatially-periodic solutions in (parameter, wavenumber)-space to the reaction-diffusion model. The concept of Busse balloon introduces the possibility of an ecosystem to increase its resilience - a mechanism that cannot take place in the classical mono-stable point of view and which thus offers a way to avoid collapse into the desert state. A disruption of stable banded patterns in a relatively large region was thought to result in a catastrophic shift towards desertification, with the disappearance of all vegetation bands. This study indicates that once a stable banded pattern is destabilized due to environmental changes, it may adapt its wavenumber (as well as its biomass density) to a new value which might still be contained within the Busse balloon, therefore still being stable (thus observable), making the shift in ecosystem change less drastic.

For this study, we did not consider changes in yearly rainfall, as the data over longer time periods was insufficient. Therefore, the main parameter worked with in the Busse balloon was the local slope of the terrain, using topographical and biomass data to gather precise information about the two regions studied. In addition to the notion of multi-stability, this study shows a vast spread in the wavenumber, biomass density and migration speed of banded vegetation patterns of the site, further corroborating model predictions thereof.

## 4.2 Existence of (novel) localized vegetation patterns in a generalized ecosystem model

Chapter 3, titled "*The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation*"<sup>2</sup>, focuses on the two-component reaction-diffusion model for vegetation biomass and soil water content, as introduced in (1.3), which was obtained in Appendix 3.B by a systematic reduction of the three-component Gilad et al. model for dryland ecosystem dynamics [62]. The starting motivation was to gain a fundamental understanding of the underlying mathematical mechanism behind the infamous 'fairy circles' of the Namibian Desert presented in [175], which are localized extended gaps patterns (with an excess of vegetation at the edge). For this, the local two-component model obtained in Appendix 3.A is fully adequate, given the environmental characteristics of the present vegetation and soil (such as confined root systems and sandy soil). Exploring the model in this context resulted in a change of course and a broadening of the study. The goal of this study then became two-fold: first, to analytically prove the existence of a multitude of vegetation patterns that have

---

<sup>2</sup>Joint work Arjen Doelman, Martina Chirilus-Bruckner and Ehud Meron; has appeared as publication in *Physica D* [82].

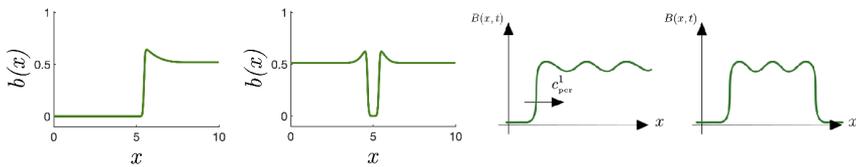
been observed in nature and can be found in numerical simulations of the model [175]. Second, to go beyond the existing observations and exploit the nonlinear richness of the model at hand in order to successfully construct and prove the existence of novel, both from an ecological and mathematical perspective, multi-front patterns.

In comparison to more conceptual models, such as the much studied extended-Klausmeier model, the Gilad et al. model possesses a more involved nonlinear structure. Although this yields a more complicated mathematical analysis, it has a strong advantage in that it can be more directly linked to ecological mechanisms and observations. That being said, no basic mathematical analysis on the patterns generated by the Gilad et al. model had been done before, until this study. Therefore, an in depth analysis and set-up of the basic mathematical characteristics of the model is required. This has been done in detail in the first part of the study, where the slow/fast dynamics of the 4-dimensional spatial ODE as presented in (1.7) are analyzed.

Building on this, a geometric singular perturbation analysis is applied to the rescaled version of the model introduced in (1.3). Geometric singular perturbation theory addresses the disparate length scales associated with biomass and water, and focuses on the strongly nonlinear ‘far-from-equilibrium’ regime, where desertification transitions typically take place and vegetation patterns exist.

The study proves the existence of ‘basic’ 1-front invasion patterns and 2-front spot/gap patterns which have a direct ecological interpretation and appear to be stable in simulations of the model. In fact, the existence of novel countably many distinct traveling 1-front patterns connecting the bare soil to the vegetated state is established as well as the existence of (traveling) 2-front spot and gap patterns for a whole open set of parameters of the model. A glimpse of the various patterns simulated and constructed can be seen in Figure 1.12.

The basic 1- and 2-front patterns, in addition, form the building blocks of the novel multi-front patterns. Based on the fast/slow dynamics of the 4-dimensional ODE, a whole set of localized vegetation spots embedded in bare soil with a varying number of spatial oscillations is constructed, as well as periodic versions of the latter as presented in Figure 1.12. In fact, we argue that these orbits/patterns are not specific for the model considered here, but will also occur in a much more general (singularly perturbed reaction-diffusion) setting as well as in an ecological setting. The study calls for further analysis of the findings in terms of (spectral) stability of the constructed (multi-)front patterns and their application in terms of ecology.



**Figure 1.12:** *Left half:* Two basic patterns exhibited by numerical simulations of model (1.3): a traveling (heteroclinic) invasion front and a stationary homoclinic, 2-front vegetation gap, i.e., a fairy circle. *Right half:* Two sketches of ‘higher order’ localized patterns constructed in Chapter 3: a 1-front connection between the bare soil state and a spatially periodic vegetation state and a representation of a ‘higher order’ localized (stationary, homoclinic 2-front) spot pattern with a countable number of ‘spatial oscillations’.



# Appendices

---

## 1.A Turing bifurcation on sloped terrains

By setting

$$s = \delta\tilde{s}, \quad \lambda = \lambda_r + i\lambda_i = \lambda_r + i\delta\tilde{\lambda}_i, \quad 0 < \delta \ll 1 \quad (1.A.36)$$

equation (1.26) expands to

$$\begin{aligned} \lambda_r^2 &+ 2i\delta\lambda_r\tilde{\lambda}_i + \delta^2\tilde{\lambda}_i^2 - \lambda_r[(f_u + g_v) - (1+d)k^2] \\ &- i\delta\tilde{\lambda}_i[(f_u + g_v) - (1+d)k^2] - i\lambda_r\tilde{s}\delta k + \tilde{\lambda}_i\tilde{s}\delta^2 k \\ &+ [dk^4 - (g_v + df_u)k^2 + (f_u g_v - g_u f_v)] + i\delta\tilde{s}k(f_u - k^2) = 0. \end{aligned} \quad (1.A.37)$$

Isolating both real and complex parts of (1.A.37) gives:

**Re:**

$$\lambda_r^2 + \delta^2\tilde{\lambda}_i^2 - \lambda_r[(f_u + g_v) - (1+d)k^2] + \tilde{\lambda}_i\tilde{s}\delta^2 k + [dk^4 - (g_v + df_u)k^2 + (f_u g_v - g_u f_v)] = 0. \quad (1.A.38)$$

**Im:**

$$2\delta\lambda_r\tilde{\lambda}_i - \delta\tilde{\lambda}_i[(f_u + g_v) - (1+d)k^2] - \lambda_r\tilde{s}\delta k + \delta\tilde{s}k(f_u - k^2) = 0. \quad (1.A.39)$$

At  $k = k_c$ : it holds that  $\frac{\partial\lambda_r}{\partial k} = \lambda_r = 0$ . This gives

$$\begin{cases} 4dk^3 - 2(g_v + df_u)k + \delta^2(-2\tilde{\lambda}_i\tilde{\lambda}_{i,k} + \tilde{s}\tilde{\lambda}_i + \tilde{s}k\tilde{\lambda}_{i,k}) = 0, \\ -\tilde{\lambda}_{i,k}[(f_u + g_v) - (1+d)k^2] + 2\tilde{\lambda}_i(1+d)k + \tilde{s}(f_u - 3k^2) = 0. \end{cases} \quad (1.A.40)$$

At  $\lambda_r = 0$  we get

$$\begin{cases} dk^4 - (g_v + df_u)k^2 + (f_u g_v - g_u f_v) + \delta^2\tilde{\lambda}_i(\tilde{s}k - \tilde{\lambda}_i) = 0, \\ -\tilde{\lambda}_i[(f_u + g_v) - (1+d)k^2] + \tilde{s}k(f_u - k^2) = 0. \end{cases} \quad (1.A.41)$$

Expanding  $k_c$  as

$$k_c = k_{c,0} + \delta^2 k_{c,2} + \dots \quad (1.A.42)$$

yields

$$4dk_{c,0}^3 - 2(g_v + df_u)k_{c,0} = 0. \quad (1.A.43)$$

which confirms that here that  $k_{c,0}$  is the critical wave number of the standard Turing bifurcation (i.e., for  $s = 0$ ). Substituting this in equation (1.A.41) for  $\tilde{\lambda}_i$  yields:

$$-\tilde{\lambda}_i[(f_u + g_v) - (1 + d)k_{c,0}^2] + \tilde{s}k_{c,0}(f_u - k_{c,0}^2) = \mathcal{O}(\delta^2) \quad (1.A.44)$$

and thus

$$\tilde{\lambda}_i(k_c) = \frac{f_u - k_{c,0}^2}{(f_u + g_v) - (1 + d)k_{c,0}^2} \cdot \tilde{s}k_{c,0} + \mathcal{O}(\delta^2) \quad (1.A.45)$$

$$= \frac{f_u - \frac{g_v + df_u}{2d}}{(f_u + g_v) - (1 + d)\frac{g_v + df_u}{2d}} \cdot \tilde{s} \sqrt{\frac{(g_v + df_u)}{2d}} + \mathcal{O}(\delta^2) \quad (1.A.46)$$

$$= -\frac{\tilde{s}}{d-1} \cdot k_c + \mathcal{O}(\delta^2). \quad (1.A.47)$$

This confirms (1.32) and (1.33),

$$\lambda_i(k_c) = -\frac{\delta\tilde{s}}{d-1} \cdot k_c + \mathcal{O}(\delta^3), \quad \mu_c = \mu_{c,0} + \mathcal{O}(\delta^2). \quad (1.A.48)$$

# Multi-stability of model and real dryland ecosystems through spatial self-organization

---

Spatial self-organization of dryland vegetation constitutes one of the most promising indicators for an ecosystem's proximity to desertification. This insight is based on studies of reaction-diffusion models that reproduce visual characteristics of vegetation patterns observed on aerial photographs. However, until now, the development of reliable early warning systems has been hampered by the lack of more in-depth comparisons between model predictions and real ecosystem patterns. In this chapter, we combined topographical data, (remotely sensed) optical data and in-situ biomass measurements from two sites in Somalia to generate a multi-level description of dryland vegetation patterns. We performed an in-depth comparison between these observed vegetation pattern characteristics and predictions made by the extended-Klausmeier model for dryland vegetation patterning. Consistent with model predictions, we found that for a given topography, there is multi-stability of ecosystem states with different pattern wavenumbers. Furthermore, observations corroborated model predictions regarding the relationships between pattern wavenumber, total biomass and maximum biomass. In contrast, model predictions regarding the role of slope angles were not corroborated by the empirical data, suggesting that inclusion of small-scale topographical heterogeneity is a promising avenue for future model development. Our findings suggest that patterned dryland ecosystems may be more resilient to environmental change than previously anticipated, but this enhanced resilience crucially depends on the adaptive capacity of vegetation patterns.

## 1 Introduction

A key aim of ecological modeling is to generate an understanding of the mechanisms driving observed patterns [99]. A significant challenge in this pursuit, however, is that multiple alternative processes may generate the same emergent outcome [48, 70, 99, 153], a phenomenon also referred to as equifinality [93, 94]. As a result, modeling efforts may reveal that a particular ecological pattern can be explained by a suite of alternative driver mechanisms. Therefore, a match between a pattern simulated with a mechanistic model and a pattern observed in a real ecosystem may only constitute limited support for the modeled mechanism being its true driver [70, 93, 94].

Pattern-oriented modeling [69, 70] aims to address the challenge of equifinality of alternative model formulations. In this approach, model assessment is based on the degree to which the output corresponds to observed patterns. A distinction is made between strong and weak patterns. Strong patterns are the dominant emergent features a model should reproduce, such as the cycles within predator and prey population sizes, or a spatial distribution of vegetation patches [69, 94]. Weak patterns are typically qualitative relationships, such as the existence of a population over a specific timespan, or a positive association between one state variable and another [69, 94]. Rather than comparing model output to a single strong pattern, additional comparisons to multiple weak patterns, at different scales or levels of organization, provide more power to model validation and selection procedures [69, 70, 94].

A specific type of ecological patterns that has received considerable attention is regular spatial patterning of sessile biota [126]. On flat terrain, the reported patterns are gaps, labyrinths, and spots [124, 161]. On sloping grounds banded patterns form, their regular spacing enabling a description of the characteristic band-inter-band period and wavenumber. Evidence is accumulating that these patterns are self-organized, meaning that the larger-scale patterning is driven by internal ecosystem processes operating at smaller scales [125, 126]. The crucial component in this self-organization process is a long-range negative effect of biota on itself, either directly or through modulation of resource availability. In cases where this long-range negative feedback is coupled to a locally positive feedback, the mechanism creating pattern formation may be linked to the existence of alternative stable states, as well as the possibility of so-called catastrophic shifts between these states [125]. This phenomenon has been most prominently studied in (semi-)arid ecosystems, where decreases in resource availability or increases in grazing pressure may trigger catastrophic shifts from vegetated states to desert states without vegetation [103, 115, 127]. In this context, the formation of regular spatial vegetation patterns may indicate proximity to a threshold of catastrophic change [125].

There is a long tradition in the scientific literature of explaining regular spatial patterning with reaction-diffusion models [29, 114, 150]. In line with this work, a variety of reaction-diffusion models has been applied to investigate self-organization in (semi-)arid ecosystems [62, 90, 124, 161]. Despite the broad support for the findings obtained with these models and their implications for (semi-)arid ecosystem functioning, comparisons of model results with empirical data have mainly been limited to comparison of a single strong pattern, namely the spatial distribution of vegetation patches. Until now, the few studies considering additional weak patterns have shown that reaction-diffusion model simulations successfully reproduce associations between pattern shape and aridity, and associations between pattern shape and slope of the terrain [34]. In addition, models that account for sloped terrain also seem to capture the observed migration of the location of banded patterns in uphill direction [33]. Despite these promising agreements between model results and empirical data, a more systematic comparison between model results and data, based on multiple patterns at different levels of organization [69, 70], was still lacking.

Advanced model analyses that have recently been applied to ecological models have yielded a number of findings which, when confronted with high quality remote sensing products, makes a more systematic comparison possible. More specifically, recent theoretical studies have shown that for a given environmental condition (i.e., a given parameter combination), not a single ecosystem state, but multiple ecosystem states with patterns spanning a range of wavenumbers may be stable, hence observable [143, 146, 155]. The range of observable patterns, across a range of environmental conditions forms a bounded region in (parameter,wavenumber)-space. This region is referred to as the Busse balloon, after F.H. Busse, who studied similar phenomena in the field of fluid dynamics [18]. Although the patterned ecosystem states in the Busse balloon are defined by their wavenumber, other properties, like migration speed and spatially averaged biomass, have also been studied [138] and are suggested to depend on the position of a system within the Busse balloon. These theoretical findings provide multiple additional weak patterns that can be compared to empirical data, providing opportunities for more powerful tests of the validity of the developed reaction-diffusion models to describe dryland ecosystems.

The aim of this study was to confront theoretical findings regarding pattern wavenumber, biomass and migration speed with the same pattern properties derived from aerial imagery and remote sensing products of banded vegetation patterns in the Horn of Africa, a location with prominent undisturbed presence of vegetation pattern formation. Hence, a multi-level comparison between theory and data in line with the pattern-oriented modeling approach was conducted [69, 70, 94].

## 2 Theory

### 2.1 Model description

Multiple reaction-diffusion models of dryland vegetation dynamics include a mechanism in which vegetation acts as an ecosystem engineer, locally increasing the influx of available water [62, 90, 124, 161]. Despite the different nuances between these models, a number of predictions can be robustly derived from these frameworks. One of the simplest of these ecosystem models – and the archetype considered in this article – is an extended version of the dryland ecosystem model by Klausmeier [90, 155], which we will refer to as the extended-Klausmeier model. This model describes the interaction between water,  $w$ , and plant biomass,  $n$ . A non-dimensional version of this model is used for the purposes of this article. A dimensional version of the model and the physical meaning of its parameters can be found in Appendix 2.A. The model is given by the following equations

$$\begin{cases} \frac{\partial w}{\partial t} = e \frac{\partial^2 w}{\partial x^2} + \frac{\partial(vw)}{\partial x} + a - w - wn^2, \\ \frac{\partial n}{\partial t} = \frac{\partial^2 n}{\partial x^2} - mn + wn^2. \end{cases} \quad (2.1)$$

The reaction terms model the change in water as a combined effect of rainfall ( $+a$ ), evaporation ( $-w$ ) and uptake by plants ( $-wn^2$ ). The change of plant biomass comes from mortality ( $-mn$ ) and plant growth ( $+wn^2$ ). Dispersion by plants is modeled as diffusion and the movement of water as a combined effect of diffusion and advection. The latter is due to gradients in the terrain, which are proportional to the slope parameter  $v$ .

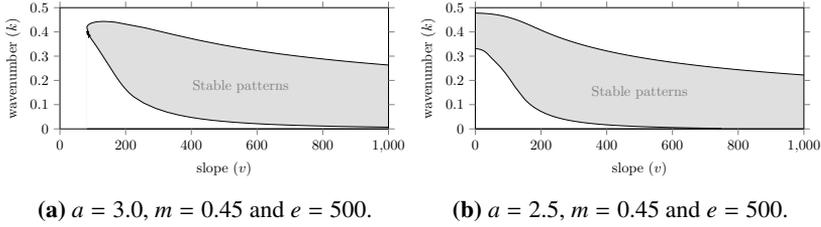
### 2.2 Theoretical outcomes

#### 1 Multi-stability of patterned states

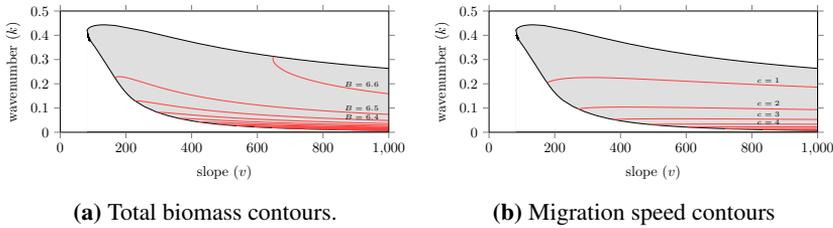
Reaction-advection-diffusion equations in general – and the extended-Klausmeier model in particular – exhibit a vast variety of spatial patterns [96, 117]. However, not all feasible patterns are stable solutions of these models. Which patterned states are stable (hence, observable) depends on the combination of the model parameters. For regular patterns, the concept of the Busse balloon can help to illustrate this dependency [18]. A Busse balloon is a model dependent shape in the (parameter, wavenumber)-space that indicates all combinations of parameter and wavenumber that represent stable solutions of the model. If, for a given set of model parameters, a wavenumber  $k$  lies within the Busse balloon, then regular patterns with wavenumber  $k$  are observable. So, in measurements, all (non-transient) patterns are expected to be present in the Busse balloon.

Typically, the Busse balloon is a high-dimensional structure due to the number of parameters in a system. Therefore, usually, only one parameter is varied when a Busse balloon is visualized. This produces a 2D-slice of the full Busse balloon. In the context of desertification research, the straightforward choice would be to vary the rainfall [146]. However, mean annual rainfall was relatively constant in our study sites during the observation period considered. Instead, topography (i.e., the slope gradient) comprised the main source of environmental variation within our study areas. Thus, relevant theoretical predictions for our study sites can be generated by varying the slope parameter  $\nu$  (while keeping rainfall constant). Here, we present two of such 2D-Busse balloon slices for the extended-Klausmeier model (Figure 2.1), which were constructed by tracking the boundary of the Busse balloon using numerical continuation methods [122, 141, 143, 146]. The shaded region in these figures indicates the combinations of pattern wavenumber  $k$  and slope  $\nu$  for which stable solutions exist. Thus, the model shows multi-stability; a given slope  $\nu$  can sustain a continuous range of wavenumbers  $k$ . That is, knowing all current parameter values of a system is not enough to predict the pattern, but only gives a range of possible wavenumbers – as indicated by the Busse balloon. For patterns with wavenumbers above this range, there are too few resources to sustain all bands; below this range, there is an abundance of resources that leads to the formation of additional vegetation bands. It is in general not possible to predict which of these wavenumbers is selected at a specific location; small changes in the (entire) history of environmental changes can have large impacts on the wavenumber that is currently selected [138, 140]. To understand these hysteretic dynamics, it is vital to acknowledge that model patterns do not change their wavenumber unless they have to [141, 146]: if an environmental change forces the system outside of the Busse balloon, the current pattern has become unstable, and will need to adapt into a new pattern that is again stable – thus part of the Busse balloon. During this (fast) adaption, only part of the vegetation bands are lost, while the remaining bands increase in volume; these adaptations thus have limited effect on the total biomass in the system [146]. Hence, multiple wavenumber adaptations are expected to occur after each other that will, gradually, lead to a complete desertification of the system [146]. Both the moment of a destabilization and the then occurring wavenumber adaption can be vastly different depending on (historical) environmental conditions [11, 138, 140]. Thus, indeed, precisely which wavenumber  $k$  gets selected at each of these destabilizations is difficult to predict.

Numerical simulations help to get an insight in the kind of wavenumber distribution one ought to expect in observations. To illustrate the typical spread in wavenumber, a total of 200 simulations on a flat terrain ( $\nu = 0$ ) were run, where the rainfall parameter was slowly decreased from  $a = 3$  to  $a = 0.5$ . The initial configurations for these runs were chosen randomly, but close to the equilibrium state of uniform biomass before the onset of patterns (between 90% and 110%



**Figure 2.1:** *(slope, wavenumber)*-Busse balloon slices for the extended-Klausmeier model for two different values of the rainfall parameter  $a$ . A banded pattern solution to the extended-Klausmeier model with slope  $v$  and wavenumber  $k$  is stable if the  $(v, k)$ -combination lies inside the Busse balloon. This indicates that a wide spread of  $(v, k)$ -combinations yields stable banded patterns. The latter are therefore expected for a broad range of wavenumbers – and not for specific  $(v, k)$ -choices only. The shape of a Busse balloon can change between models and between parameter values. This is illustrated in the figures which were computed for different  $a$ -values.

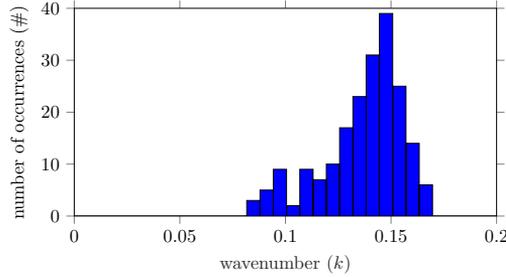


**Figure 2.2:** *(slope, wavenumber)*-Busse balloon slices for the extended Klausmeier model that include contours for the total biomass (per area)  $B$  (a) and the migration speed  $c$  (b). Biomass (per area) is positively correlated with both wavenumber  $k$  and slope  $v$ ; the migration speed is negatively correlated with the wavenumber  $k$ . Model parameters used:  $a = 3, m = 0.45, e = 500$ .

of the uniform vegetated equilibrium state). At the end of each simulation – after several pattern selections – the wavenumber of the remaining pattern was measured. This gives a snapshot of the wavenumber distribution, similar to the snapshots acquired from observations. Note that a similar experiment was done before, albeit on a much smaller scale [141]. The histogram of the resulting wavenumbers is shown in Figure 2.3. It shows a substantial spread, which goes from a wavenumber of 0.08 to 0.16 (a difference of 100%).

### 2 Biomass & migration speed

Besides a wavenumber, each ecosystem state also has a specific biomass and a specific pattern migration speed. The biomass of regular patterned states has been studied using numerical simulations [146] and more general formulas have



**Figure 2.3:** Histogram demonstrating a spread in wavenumber ( $k$ ) at the end of 200 simulations of the extended-Klausmeier model on a flat terrain ( $\nu = 0$ ) with model parameters  $e = 500$  and  $m = 0.45$ . These simulations had a random initial configuration close to a stable fully vegetated state. A climate change was simulated by decreasing the rainfall parameter  $a$  linearly from 3 to 0.5 over the course of  $10^5$  time unit, causing several pattern selections and corresponding changes in wavenumber.

been derived for patterns with small wavenumber [11]. Both indicate that the biomass (per unit area) is positively correlated with both the wavenumber  $k$  of the pattern and the slope parameter  $\nu$  [146]; see also Figure 2.2a. This has a physical interpretation: both steeper slopes and higher wavenumbers (lower wavelengths) reduce the time it takes for water to reach vegetation bands, and thereby reduce water losses during the transportation process. As a result, the vegetation will be able to harvest water from the uphill inter-bands more effectively. The biomass per wavelength is also of interest. The same studies indicate that the band biomass (per wavelength) is increased when the wavenumber  $k$  is decreased and when the slope  $\nu$  is increased. Hence, vegetation bands are expected to have more biomass when other vegetation is farther away, because of the larger (upslope) inter-band area water can be collected from.

The theoretical predictions for migration speed (of a pattern's location) are a bit more subtle. For terrains with a constant slope, numerical simulations have been done [136, 139] and general formulas have been determined for patterns with small wavenumber [11, 133]. In these idealized limit cases, migration speed is negatively correlated with wavenumber  $k$  and positively correlated with slope  $\nu$ . However, beyond these idealizations, numerical computations show the contour lines are slightly humped, see Figure 2.2b. This indicates a (slightly) negative correlation between speed and slope  $\nu$  for large slopes.

### 2.3 Testable predictions

The theoretical findings in this section lead to predictions that can be confronted with the field data. First of all, the model possesses a Busse balloon, which should lead to a wide spread in observable pattern wavenumbers (Figures 2.1 and 2.3). Moreover, biomass and migration speed are affected by pattern

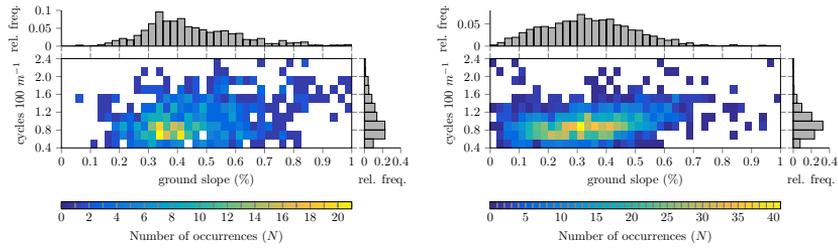
wavenumber. The biomass (per unit area) is expected to be positively correlated with both the wavenumber and the slope of the terrain (Figure 2.2a). Migration speed is expected to decrease as a function of pattern wavenumber; the effect of slope on the migration speed is context-specific, as it can be either positive or negative depending on the specific topographical and environmental conditions (Figure 2.2b).

### **3 Data acquisition & processing**

For this comparison study, two sites were selected in Somalia. The first one (8°0'14" to 8°15'11"N; 47°11'54" to 47°31'4"E) is located in the Haud pastoral region, which will be referred to as the 'Haud' site. The other site (9°18'49" to 9°34'34"N; 48°8'15" to 48°43'15"E) is located in the Sool-Plateau pastoral area and will be called the 'Sool' site. Both sites mainly exhibit banded vegetation and have ground slopes ranging from 0% to 1%. Vegetation mainly constitutes of perennial grasses, which typically have an average lifetime of 1-7 years [19, 95, 171]. A more detailed description of these sites can be found in Appendix 2.B; a map with the location of these sites along with the mean annual rainfall in these areas is shown in Figure 2.D.1.

To study the pattern properties in these study areas, each site was divided into square windows (of size 750m × 750m for the Haud site and of size 1010m × 1010m for the Sool site). As has been done in previous studies, the type of pattern (e.g. bare soil, banded vegetation), along with its wavenumber, was determined using spectral analysis [7, 27, 34, 118]. Only those windows were kept that exhibited banded vegetation with a wavenumber that could be determined with enough certainty (i.e., between 0.4 and 2.5 cycles per 100m). Moreover, windows with a too large curvature were ignored, because the theoretical predictions only apply to terrains with a constant slope. To obtain data on the migration speed of the banded vegetation, a cross-spectral analysis was performed, along the lines of previous studies [8, 33, 66]. A more in-depth explanation of the processing steps can be found in Appendix 2.D.

The topographical data used in this article were derived from the Advanced Land Observation Satellite (ALOS) World 3D (AW3D) digital raster elevation model; biomass data for the Haud site have been retrieved from a recently made map on (above-ground) biomass of African savannahs and woodlands [15] (no reliable data for the Sool site was available). Finally, optical data was acquired from various sources: three multi-spectral WorldView-2 images were mosaicked and used as reference for the Haud site; a panchromatic Ikonos "Geo" Imagery was acquired for the same site. For the Sool site, six WorldView-2 images were used and a panchromatic Satellite Pour l'Observation de la Terre (SPOT) 4 image preprocessed to level 2A was used as reference layer (©Cnes 2004 – Spot Image



(a) Frequency distribution for the Haud site (b) Frequency distribution for the Sool site

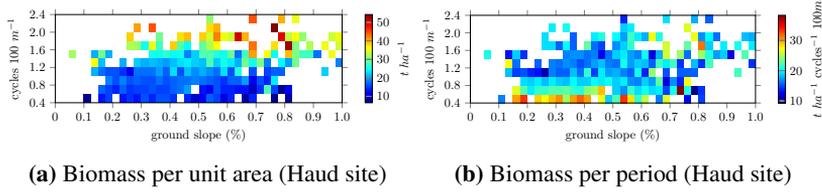
**Figure 2.4:** Frequency distribution of banded patterns as function of ground slope and wavenumber (number of cycles per 100m) for the Haud site (a) and Sool site (b). The distribution on the right indicates the relative frequency of banded vegetation with corresponding wavenumber. The color gradient indicates the amount of windows ( $N$ ).

distribution). Moreover, two  $7\mu\text{m}$  digitized panchromatic declassified Corona spy satellite image, national intelligence reconnaissance system, available from the US Geological Survey, were obtained for the Haud and the Sool sites. More information about these data sets can be found in Appendix 2.C.

## 4 Results

### Empirical Busse balloon

The most prominent pattern property studied in this article is the pattern wavenumber, which was derived from aerial imagery. The resulting distribution of wavenumbers is reported in Figure 2.4 (a map with the spatial distribution of wavenumbers over the study sites is shown in Figure 2.D.2). These figures show the number of windows that have a particular slope-wavenumber combination. Also given is the relative frequency that indicates the spread of wavenumbers across all windows. The data display banded vegetation with wavenumbers roughly between 0.4 and 2.0 cycles per 100m. Importantly, this large spread is present for all of the ground slope values which had a representative sample size and could not be explained by present heterogeneities in elevation or rainfall. This shows that for a given environmental condition not a single wavenumber pattern, but rather multiple patterns spanning a sizable range of wavenumbers are observable. Additionally, measurements used to determine the migration speed show barely any changes in wavenumber over the scope of 39 years (consistent with [66]), indicating that these patterns are in fact quite stable. Therefore, the observations are in agreement with the existence of a Busse balloon in the real ecosystem.



**Figure 2.5:** Biomass distribution per area (a) and per period (b) as a function of ground slope and wavenumber (cycles per 100m) for the Haud site. The color gradient indicates the amount of biomass measured for a particular (*slope, wavenumber*)-combination.

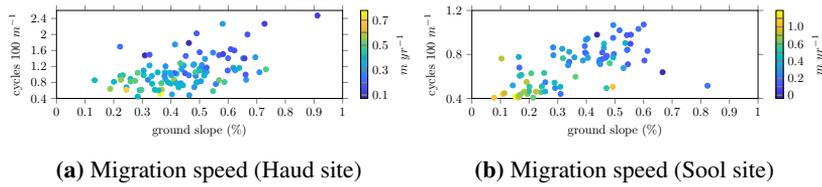
### Biomass and migration speed

The processed biomass data for the Haud site is shown in Figure 2.5. In Figure 2.5a the relation between biomass per area (in  $t\ ha^{-1}$ ) is plotted against the ground slope and the wavenumber. From the same data the biomass per period is computed – which is biomass per area divided by the window’s wavenumber. The resulting plot is given in Figure 2.5b. The measurements of biomass show agreement with theoretical predictions of model studies; in both, the total biomass increases (all slopes:  $r^2 = 0.64$ ,  $n = 714$ ,  $P < 0.001$ ; linear regression) and the biomass per period decreases when the wavenumber increases (all slopes:  $r^2 = 0.09$ ,  $n = 714$ ,  $P < 0.001$ ; linear regression). However, a more in-depth inspection reveals disagreements. For one, the effect of ground slope is not strongly present in the data, though its effect is clear in the extended-Klausmeier model (Figure 2.2a). Additionally, the more refined details of wavenumber dependence also differ (it is concave in the theoretical model and convex in the real-life data).

The migration speed is plotted in Figure 2.6 for both the Haud and the Sool sites. These measurements show an increase in speed when the wavenumber decreases (Haud:  $r^2 = 0.43$ ,  $n = 104$ ,  $P < 0.001$ ; Sool:  $r^2 = 0.45$ ,  $n = 79$ ,  $P < 0.001$ ; linear regression), corroborating theoretical predictions (see Figure 2.2b).

## 5 Discussion

Leading ecological frameworks emphasize the potential role of regular spatial vegetation patterns as indicators for proximity to catastrophic ecosystem shifts [125, 129]. In these frameworks, however, mono-stability of patterns is implicitly assumed, suggesting that for a given environmental condition there is only one stable vegetated state, i.e., a single pattern with a specific wavelength [125, 129]. Subsequent theoretical insights have challenged this view, highlighting the possibility of multi-stability of patterns, bounded by the so-called Busse balloon. In this study, we provide empirical evidence corroborating the existence of a Busse balloon for stable vegetation patterns in dryland ecosystems.



**Figure 2.6:** Observed (average) migration speed of vegetation bands in the Haud (a) and the Sool (b) sites over the course of 39 years as a function of ground slope and wavenumber (cycles per 100m). The color gradient indicates the migration speed for a particular (slope, wavenumber)-combination. The sign indicates the direction of migration relative to the slope, with positive and negative values indicating upslope and downslope migration respectively.

Specifically, our two study sites in Somalia revealed the sustained (i.e., over a 39 year period) co-occurrence of banded vegetation with wavenumbers varying over a substantial range. Our findings have major implications for the way in which vegetation patterns indicate ecosystem resilience and mediate ecosystem responses to environmental change.

Specifically, the existence of a Busse balloon implies that an ecosystem’s resilience can no longer merely be defined by the magnitude of environmental change it can cope with [80]. In these systems there is not one tipping point, but a cascade of destabilizations – indicated by the boundary of the Busse balloon. When environmental changes push a patterned ecosystem beyond the boundary of the Busse balloon, a wavelength adaptation occurs, and typically part of the vegetation patches are lost, while the remaining patches grow in size. The extent of these adaptations depends on the rate of environmental change [138, 144–146]. Moreover, human activities or natural variations can cause local disturbances, diminishing the regularity of ecosystem patterns. The recovery process from such disturbances may involve a rearrangement of patches in the landscape [11, 146]. Again, the extent to which such recovery is possible depends on the rate of environmental change that the ecosystem is exposed to [145]. Hence, the existence of a Busse balloon of stable dryland vegetation patterns suggests that adaptability of patches to changing environmental conditions provides a more comprehensive indicator for the ecosystem’s resilience than the shape of the pattern itself, as suggested in current leading frameworks [125, 129]. To fully comprehend the consequences of this, it is necessary to provide a more thorough understanding of what determines the spatial rearrangement of vegetation patches resulting from disturbances, environmental changes, and spatial heterogeneities in the landscape.

The pattern-oriented modeling approach was mainly developed to aid model development and design, but the approach can also be used to evaluate the suc-

cess of existing models to explain multiple strong and weak patterns observed [69]. This so-called ‘reverse pattern-oriented modeling’ approach [69] was used in the current study. Such systematic comparisons between model predictions and empirical data can be part of an iterative process toward further model improvement [93, 94]. In this context, it is interesting to note the discrepancy that we observed between model predictions and field measurements of the influence of the ground slope on pattern migration speeds. Because topography critically changes the distribution of water within ecosystems, it also alters the migration speed of patterns. Therefore, it is of interest to determine the effects of more complex topographies for dryland ecosystem dynamics.

Moreover, the available empirical data aligns with theoretical predictions on both strong and weak patterns. However, environmental conditions were characterized by differences in slope gradient only. Although, indeed, the topography comprised the main source of environmental variation, other less pronounced heterogeneities are present and can cause spreads in wavenumber. The observed spread could not be attributed to variation in rainfall or elevation, but the role of other heterogeneities (e.g. soil composition and grazing activity) could not be fully determined for lack of precise and accurate data sets. When these become more readily available, further research might infer to which extent the observed wavenumber spread is explained by these environmental drivers.

Since their appearance on aerial photographs in the 1950s [102], the origin of regular vegetation patterns in dryland ecosystems has been a topic of fascination within the scientific community. The study of these patterns through reaction-diffusion modeling subsequently highlighted the importance of these patterns for the functioning of dryland ecosystems, and their response to environmental change. The recent increase in the availability of optical and topographical data provides unprecedented opportunities to confront model predictions with empirical data [94, 140]. In this study, we combined these data sources with in-situ measurements of biomass, enabling the comparison of multiple pattern characteristics of Somalia drylands with predictions derived from reaction-diffusion modeling. The empirical evidence corroborates theories of multi-stability of patterned vegetation states, improving our understanding of these systems’ resilience to environmental changes. In addition, our results call for more detailed investigations of the role of small-scale topographic variability in pattern formation and migration.

# Appendices

---

## 2.A Dimensional extended-Klausmeier

The dimensional extended-Klausmeier model is given by (2.A.2). The model used throughout the chapter, equation (2.1), can be obtained from the dimensional version by the right set of scaling. Following [146, Appendix A], the required scaling is given in (2.A.3) for the variables and in (2.A.4) for the parameters of the model.

$$\begin{cases} \frac{\partial W}{\partial T} = E \frac{\partial^2 W}{\partial X^2} + \frac{\partial(VW)}{\partial X} + A - LW - RWN^2, \\ \frac{\partial N}{\partial T} = D \frac{\partial^2 N}{\partial X^2} + RJWN^2 - MN \end{cases} \quad (2.A.2)$$

$$w = \frac{\sqrt{R}J}{\sqrt{L}}W \quad n = \frac{\sqrt{R}}{\sqrt{L}}N \quad x = \frac{\sqrt{L}}{\sqrt{D}}X \quad t = LT \quad (2.A.3)$$

$$a = \frac{\sqrt{R}J}{L\sqrt{L}}A \quad m = \frac{1}{L}M \quad v = \frac{1}{\sqrt{LD}}V \quad e = \frac{E}{D} \quad (2.A.4)$$

In these equations, water is supplied to the system at a rate  $+A$ , modeling uniform rainfall. Because of evaporation, water is lost at a rate  $-LW$ ; water is also lost through uptake by plants, at rate  $-RWN^2$ . The parameter  $J$  models the increase of biomass per unit of water consumed, which results in the reproduction of plants at rate  $+RJWN^2$ . Plant mortality is modeled as  $-MN$ . The parameter  $V$  is the speed at which water flows downhill; this is proportional to the slope gradient. Finally,  $E$  is the diffusion coefficient of water;  $D$  is the diffusion coefficient of vegetation, modeling the dispersal of biomass. See also [90].

## 2.B Description of study sites

For this study, two sites in Somalia were selected that exhibit mostly banded vegetation. The Haud site is a  $35\text{km}$  by  $28\text{km}$  study area ( $8^\circ 0' 14''$  to  $8^\circ 15' 11''\text{N}$ ;  $47^\circ 11' 54''$  to  $47^\circ 31' 4''\text{E}$ ) at  $650\text{-}750\text{m}$  elevation in the Haud pastoral region –

see also Figure 2.D.1. Here, banded vegetation dominates the landscape with some minor occurrences of gapped vegetation on flat ground on the summits of rolling hills. Bands display a broad range of wavelengths (from 60m to 200m). Ground slope ranges from 0 to 1%. Mean annual precipitation, ranging from 210mm to 270mm, is distributed in two rainy seasons around spring (April-May) and fall (September-November) separated by two dry seasons. Rainfall data was extracted from Climate Hazards Group InfraRed Precipitation with Stations [59]. Estimates were provided by [35].

In the north-eastern corner of this area near Kalabaydh city, the soils of the bands and inter-bands are very similar [78]. Moreover, the large perennial tussock grass *Andropogon kelleri* dominates the core of the band along with some scattered small trees and bushes [78]. Characteristically, plants are sparsely distributed on the downslope side of the bands. Along this edge and below it, in the bare inter-band, dead trees of all of the species found within the bands were present. Along the bands' upslope side, some initial colonization by two perennial grass species, tussock-forming *Chrysopogon aucheri* var. *quinqueplumis* and stoloniferous *Dactyloctenium scindicum*, was observed [78]. Although the lifespan of perennial grasses is highly variable, ranging from less than a year to multiple decades [95], the average lifespan of perennial grasses in arid and semi-arid environments is typically 1-7 years [19, 95, 171]. Upon inspection of satellite imageries taken 39 years apart, an upslope migration speed of  $0.3m\ yr^{-1}$  was observed [33].

The Sool site is an approximately 77km by 29km study site, located 190km to the NE of the Haud site ( $9^{\circ}18'49''$  to  $9^{\circ}34'34''N$ ;  $48^{\circ}8'15''$  to  $48^{\circ}43'15''E$ ); it is located in the Sool-Plateau pastoral area, which has more arid conditions (100mm-140mm) and higher elevations (900m-1000m) – see also Figure 2.D.1. Here, the ground slope ranges from 0 to 1%, and ground is either bare or covered with banded vegetation which sometimes displays a dashed physiognomy. To the authors knowledge, there is no published record of the composition of these vegetated bands and associated soils. Remote sensing analysis of vegetation dynamic in this area over the last decades have shown a continuous upslope migration of the bands as well as a change in band width. However, no change in wavelength has been observed [33, 66].

## 2.C Data sets

### 2.C.1 Topographical data

For both sites, topographical data was retrieved from the ALOS World 3D 30m (AW3D30, v. 2.1) digital raster elevation model. This model describes the height above sea level (in m, rounded to the nearest integer), at a ground resolution of

approximately  $30m$  at the equator. The elevation data was preprocessed for the removal of artifacts by applying a global soft-thresholding on its dual tree complex wavelet transform. Specifically, we set a threshold of 0.9 on the first five dual-tree complex wavelet transform levels. From the preprocessed data, we calculated the slope gradient (in %) and slope aspect (in degrees). We first extracted square DEM windows of 33 by 33 cells (i.e., approximately  $990m \times 990m$ ), centered on the image windows. We then applied a least squares fitting procedure of an unconstrained quadratic surface on the unweighted elevation values. From the first derivatives of this fitted surface, evaluated at the focal cell, we could then calculate slope gradient and aspect analytically, following [146]. Complex topographic features were discarded from subsequent analysis by ignoring windows (see below) with quadratic fit RMSE above  $1m$  or a total curvature (as defined by [132]) above  $10^{-10}$  radians per  $m^2$ .

### 2.C.2 Biomass measurements

Recently, a map has been made with data on (above-ground) biomass of African savannahs and woodlands at a resolution of  $25m$  [15], which provides the biomass data of the patterns studied in this article. This map is built from 2010 L-band PALSAR mosaic produced by JAXA following a method adapted from [105], while the perturbing sources that affect the SAR data have been minimized: the environmental effects (soil and vegetation moisture) were reduced by stratifying the African continent into wet/dry season areas, and the speckle noise inherent to SAR data acquisitions was decreased by applying a multi-image filter developed by [16] that preserves the spatial resolution of the images. Then, the sensitivity of the radar backscatter to AGB was analyzed to develop a direct model relating the PALSAR backscatter to AGB, calibrated with the help of in-situ and ancillary data. The in-situ data were composed of 144 selected field plots, located in 8 countries (Cameroon, Burkina Faso, Malawi, Mali, Ghana, Mozambique, Botswana and South Africa), with plot size larger than  $0.25ha$  and a mean plot size of  $0.89ha$ .

### 2.C.3 Optical data

Three multi-spectral WorldView-2 images, acquired on December 25th 2011, January 21st 2012 and July 21st 2012, were mosaicked and used as reference orthoimage for the Haud site. For the diachronic study, a panchromatic Ikonos “Geo” imagery, with a  $1m$  nominal ground resolution, was used as the reference layer. It was acquired on January 7th 2006. Orthorectification was performed using a rational polynomial coefficient (RPC) camera model block adjustment without ground control points [71].

A mosaic of six WorldView-2 images, acquired between February 3rd 2011 and September 12th 2013, was used for the Sool sites. For the diachronic study,

a panchromatic SPOT4 image preprocessed to level 2A, with a 10m nominal ground resolution, was used as reference layer (©Cnes 2004 – Spot Image distribution). It was acquired on February 18th 2004.

Two 7 $\mu$ m digitized panchromatic declassified Corona spy satellite image, national intelligence reconnaissance system, available from the USGS, were acquired on February 28th 1967 (KH-4A, mission 1039, AFT camera) and December 12th 1967 (KH-4B, mission 1102, FWD camera), respectively for the Haud and the Sool. The images were co-registered with the orthorectified reference imagery. Co-registration was performed using a third-order polynomial adjustment using landmarks such as geological features, crossroads, isolated trees, or large termite nests. We obtained an RMS adjustment error below the KH-4A ground resolution, which is 3m for this area. The resolution of the imagery was then lowered through pixel averaging to match the coarsest image pair.

The analysis of pattern wavelength was performed over the full area of the study sites. However, for the diachronic study, a subset of each of the sites covered by the historic and reference image was selected. Projection and datum for all data sets were WGS 1984, UTM Zone 38N and 39N respectively for the Haud and the Sool sites.

## **2.D Data processing**

### **2.D.1 Spectral analysis, direction of anisotropy and wavelength**

On visible light digital images over drylands, bright pixels correspond to bare soil, intermediate gray-scale levels to closed grass cover, and darker pixels to woody vegetation. As a first approximation, gray-scale levels can thus be considered as a monotonically decreasing function of the above-ground biomass [28]. This approximation allows us to analyze the spatio-temporal dynamics of biomass organization through image analysis techniques.

We used a Fourier windowing technique equivalent to short time Fourier transforms to obtain spatial maps of dominant pattern wavelength  $\lambda$  and orientation  $\theta$  from the satellite images as previously used for banded vegetation systems [7, 27, 34, 118]. We applied a two-dimensional (2D) Fourier transform to obtain the power spectrum within square, non-overlapping moving windows. In order to maintain resolution and signal-to-noise ratio a boxcar windowing function to signal was applied. This choice is, in this case, reasonable as only one periodic component is expected to be present in the vegetation. The technique provides information about the local wave-vector  $k = k_x\hat{x} + k_y\hat{y}$ . The two-dimensional (2D) fast Fourier transform  $\tilde{f}(k_x, k_y)$  of the pattern of biomass

$f(x, y)$  was obtained for each window  $\tilde{f}(k)$  of size  $L \times L$ . As  $L$  increases, the spatial resolution, i.e., localization in space of frequency or orientation change, is reduced. Conversely, as  $L$  decreases, the frequency resolution is decreased, i.e., the likelihood of separating frequency components close together in Fourier space. To optimize both,  $L$  was chosen to be at least  $3\lambda$ , i.e.,  $750m$  and  $1010m$  respectively for the Haud and the Sool sites.

To separate the characteristics of the signal that are meaningful for this study, each  $k$ , of frequency  $2\pi/\lambda$  (wavenumber), was decomposed into its orientation  $\theta$  and its magnitude. For each window, the power spectrum  $S(k) = |\tilde{f}(k)|^2$  was computed. The power spectrum measures how the variation, or power, of the pattern is distributed over the wavevectors  $k$ , of different frequencies and spatial directions. To identify the dominant  $k$  in each window,  $S(k)$  was binned into annular rings of unit width [123]. The resulting radial spectrum thus quantifies the contribution of successive ranges of spatial frequencies to the image variance across all orientations.

To deconvolve the natural  $1/k$  scaling of the power spectrum, the total power within each annular ring,  $S(k)$ , was computed instead of the mean power. The location peak of this total power is used to define the most energetic wavenumber,  $k_1$ . To compensate for the discrete  $k$ -resolution in Fourier space, the location of the weighted average  $k_1 := (\sum_k kS(k)) / (\sum_k S(k))$  was computed over all rings that formed part of the peak and contained more than 70% of the peak power.

The patterns were characterized in terms of level and orientation of anisotropy (i.e., direction orthogonal to the long axis of the bands) following [34]. The average pattern orientation was studied using the circular mean direction weighted by the power spectrum values,  $\bar{\theta} := \frac{1}{2} \arctan 2(S, C)$ , where

$$S := \left( \sum_k 2k_{x,y} \sin \theta_{x,y} \right) / \left( \sum_k k \right)$$

and

$$C := \left( \sum_k 2k_{x,y} \cos \theta_{x,y} \right) / \left( \sum_k k \right).$$

The norm of the resultant vector,  $\bar{R} := \sqrt{S^2 + C^2} / (\sum_k k)$ , was used as an index of pattern anisotropy. The division by the sum of periodogram amplitudes ensures bounding between zero (perfect isotropy) and one (all variance concentrated in one direction, i.e., perfect bands). Pattern orientation features were extracted from the power spectrum, within the frequency ring characterizing periodic vegetation patterns, i.e., between 0.4 and 2.5 cycles  $100m^{-1}$  for both the Haud and the Sool site, to exclude anisotropy sources resulting from large scale gradients or small scale (anthropogenic) features.

### **2.D.2 Pattern classification**

The vegetation cover of each window was quantified by converting the gray-scale intensity image to a binary image using the Otsu thresholding method [116]. Windows with less than 15% vegetation cover were considered as bare soil and discarded. Windows with dominant patterns within the acceptable range (i.e., between 0.4 and 2.5 cycles  $100m^{-1}$  for both sites) and with anisotropy index above 0.2 were considered as banded patterns.

### **2.D.3 Cross-spectral analysis and migration speed**

Scale specific comparisons between pairs of periodic 2D signals – in this case, images taken at different dates – can be performed through 2D Fourier cross-spectral analysis. In principle, this means identifying the frequencies and orientations of patterns dominating in any two images as well as possible shifts among them. Correction of radiometric variability between dates is not required since Fourier coefficients are invariant to linear rescaling of gray-scale levels. A detailed mathematical development of the analysis can be found in [8]. The procedure can be summarized as follows [33, 66].

To assess band migration distance for each temporal pair of image windows, a coherency spectrum and a phase spectrum were computed. The coherency spectrum expresses the correlation between the frequency components of the Fourier spectra of the pair of windows. For each spatial frequency, the coherency value is interpreted in a similar way to the classical Pearson's coefficient but in absolute values, because the sign of the correlation is expressed by the phase spectrum. For each window pair, the maximum value of coherency and its associated frequency were recorded along the direction of maximal anisotropy computed for the first acquisition date. Window pairs with a maximum coherency below 0.9 were rejected from the analysis, because this indicates that pattern characteristics (wavelength and orientation) changed between the dates. Rejected windows often corresponded to man-made perturbations or ephemeral patterns, which are not the subject of this study. The obtained frequency value therefore corresponds to a pattern of constant scale and orientations dominating at both acquisition dates. The corresponding phase-spectrum value provides the phase difference, i.e., the angular distance, between the selected frequency components at both dates in the maximal anisotropy direction. This value is defined between  $-180$  and  $180$  degrees, with the sign representing the forward (+) vs. backward (–) displacement, with the direction of reference as the upslope direction. Angular distances were then converted into meters by multiplying the phase difference by the wavelength, which in turn were converted to an average migration speed for the time period. This conversion allows for inter-site comparisons independent of varying time intervals. An inherent limitation to this procedure is that only migration distances not exceeding half the wavelength

will be correctly estimated (phase saturation). This condition was verified by visual inspection prior to image analysis.

#### 2.D.4 Assessment of uncertainty in calculations of slope gradient and aspect from topographical data

As the digital raster elevation model contains errors, these will propagate into derived estimates of slope gradient and aspect. Ideally, one would use ground observations (e.g. using differential global navigation satellite systems) to assess the magnitude of these errors (e.g. [119]). Since such observations are not available for our study areas, we used a simulation method to evaluate the propagation of error from the elevation data to the estimates of slope gradient and aspect. Specifically, we created artificial elevation grids with a fixed (from here referred to as ‘true’) slope and aspect, and added to these grids (normally distributed) random errors with a similar standard deviation as observed in the AW3D30 dataset. The resulting elevation values were rounded to the nearest integers, as this is also done in the AW3D30 dataset. Then, utilizing the same procedures as described above (see section 2.C2.C.1), we derived slope gradients and aspects from these simulated grids. Comparison of this ‘observed’ slope and aspect and the ‘true’ slope and aspect of the grid yielded insight in the propagation of errors from the elevation model to the calculated metrics. Seven fixed slope levels were considered: 0; 0.025; 0.05; 0.1; 0.2; 0.3 and 0.4%. For each level, we simulated 10,000 replicate grids of errors that were added to the fixed slope level. For each replicate, the aspect was a randomly assigned value between 0 and 360 degrees.

Following the above procedure, we found that the distributions of errors in the calculated slope gradients and aspects were relatively small, for all slope levels considered (Figure 2.D.3). Because the estimated slope is bounded between zero and positive infinity, a small positive bias was observed for slopes less than 0.1%. For slopes of 0.025% and higher, the RMSE is 0.010% and 95% of the observed errors for slopes of 0.1% or higher are within  $\pm 0.016\%$  (5<sup>th</sup> and 95<sup>th</sup> percentiles). For aspect, the magnitude of errors was inversely proportional to the magnitude of the slopes (abscissa; Figure 2.D.3). For slopes of 0.2% and higher, the RMSE is 2.9 degrees or lower, and 95% of the observed errors are within  $\pm 4.8$  degrees (5<sup>th</sup> and 95<sup>th</sup> percentiles). These results show that the errors in calculated slopes and aspects were relatively small compared to the observed range in the dataset. Hence, it is unlikely that correlations between pattern metrics and slope gradients, as observed in the main text, are strongly affected by the errors originating from the underlying topographical database.

Moreover, it should be noted that the above procedure may even be overestimating the errors associated with the AW3D30 dataset (from here referred

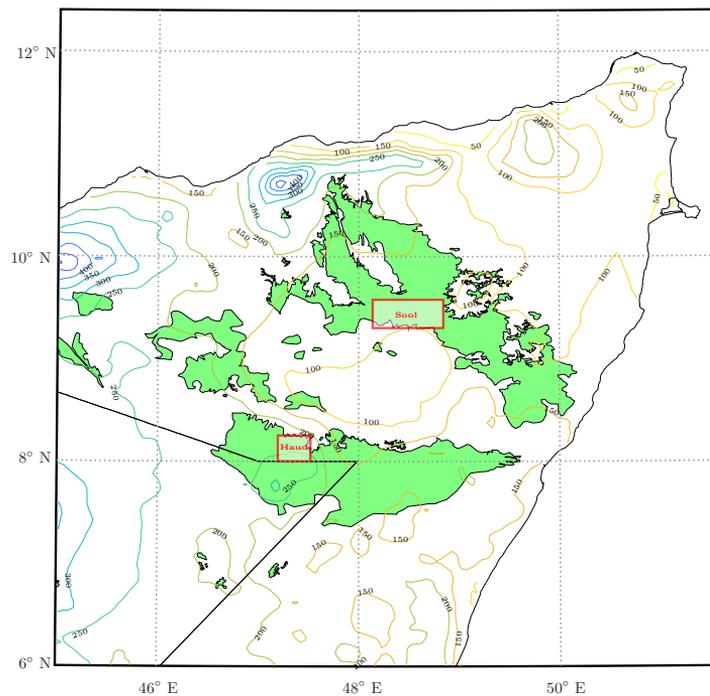
to as the 30m elevation dataset). To generate the simulated errors, we used the global average standard deviation of the difference between the original AW3D 5m elevation dataset (from here referred to as the 5m dataset), from which the 30m dataset has been derived, and a reference LiDAR dataset. This standard deviation is 1.73m for gently sloping terrain (below 17.6%) [148]. However, as the 30m elevation dataset was produced by averaging non-overlapping windows of 7 by 7 pixels of the 5m elevation dataset, the resulting standard deviation will be lower [147]. Additionally, the ground slope in our study areas is at the lower end of the 0-17.6% range, namely below 1.5%, and therefore likely to suffer from smaller errors than reported for the whole range. Finally, the Somalian area we are studying displays relatively small errors in elevation measurements compared to other areas of the world [148].

### **2.D.5 Assessment of uncertainty in estimation of pattern frequency from optical imagery**

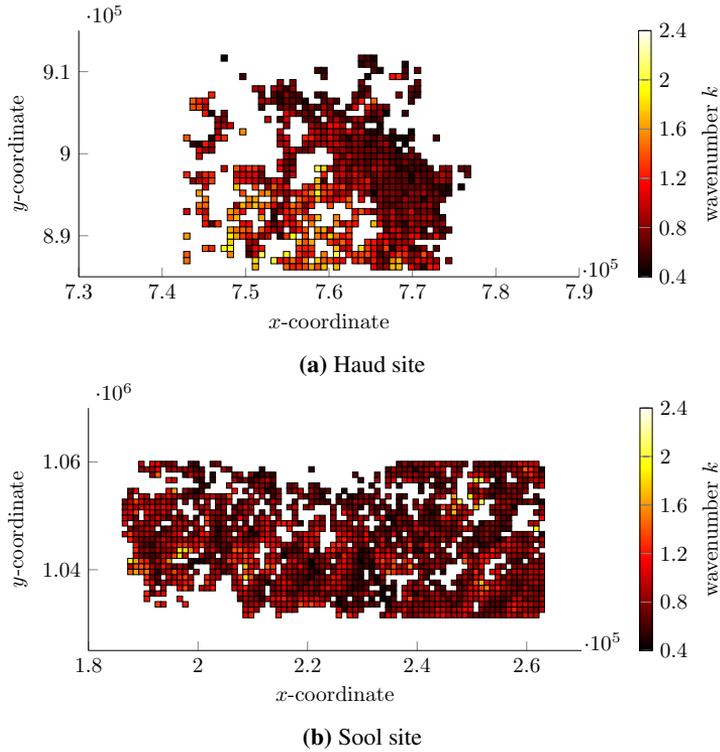
The estimation of the dominant pattern frequency using a Fourier windowing technique introduces an unknown uncertainty in these estimations. This uncertainty stems from the decomposition of the signal into a discrete set of frequencies and from the noise in the analyzed (non-stationary) signal. To assess the model's uncertainty in the wavenumber estimations, we have used a simulation method. Specifically, we simulated 200 synthetic images, representing a two-dimensional sinusoid of each frequency class between 0.4 and 2.4 cycles  $100m^{-1}$ , with step size of 0.2. The directions of the sinusoidal waves were selected randomly and the signal was standardized to have zero mean and standard deviation of one. To mimic real images of vegetation patterns, we have added red noise with zero mean and standard deviation of 0.5 to each simulated signal. Red noise is a self-similar, or fractal, random spatial structure; this is a desirable property here because these are common in nature and especially in natural landscapes [89]. The noise was created using the Fourier synthesis technique with an energy spectrum exponent of 0.5 [36]. Finally, in order to account for the fact that reflectance values are constant over the width of both vegetated and bare bands, the signal was converted to binomial values; that is, values between 0 and 1 were rounded to the nearest integer value. Several examples of simulated bands are presented in Figure 2.D.5.

Using the characteristics of the optical image windows of the Haud site (cell-size of 2.36m and windows 317 by 317 cells), the root-mean-square error of the estimated frequency was 0.082 cycles  $100m^{-1}$ . For the Sool (cell-size of 2.36m and windows 425 by 425 cells), the root-mean-square error was 0.044 cycles  $100m^{-1}$ . The magnitude of this error is significantly less than the observed variability in frequency in both sites (for every slope bin of Figure 3 in the main text, Levene's test,  $P < 0.001$ ), showing significance of the observed

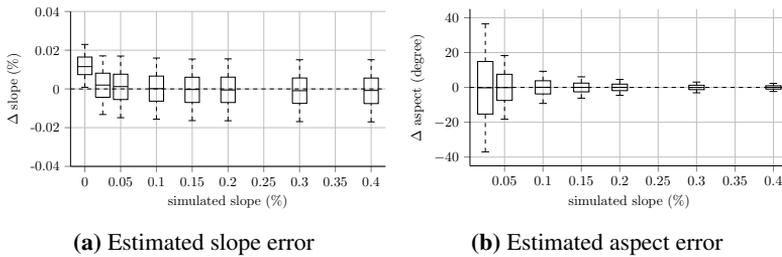
wavenumber spread in both study sites.



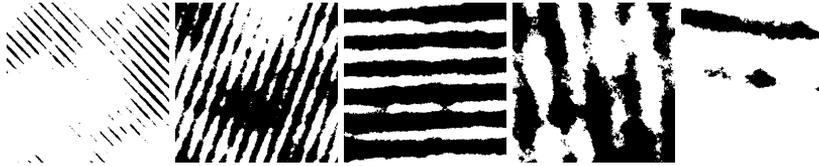
**Figure 2.D.1:** Locations of the study sites and rainfall gradient in the Horn of Africa. The ‘Haud’ site (8°0′14″ to 8°15′11″N; 47°11′54″ to 47°31′4″E) has a mean annual rainfall of 210–270mm yr<sup>-1</sup> whilst the ‘Sool’ site (9°18′49″ to 9°34′34″N; 48°8′15″ to 48°43′15″E) has a mean annual rainfall of 100–140mm yr<sup>-1</sup>. The distribution of periodic vegetation pattern shown in green is adapted from [32]. Precipitation data was extracted from Climate Hazards Group InfraRed Precipitation with Stations [59] and is averaged over the years 1981–2013.



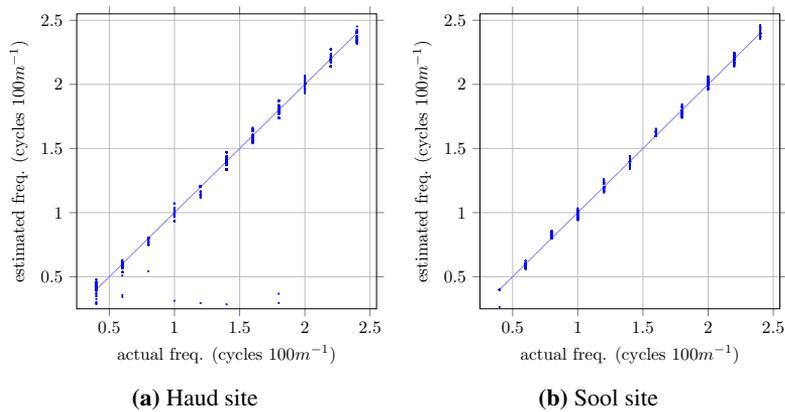
**Figure 2.D.2:** The distribution of the measured banded pattern's wavenumber over the Haud site (a) and the Sool site (b). Here, darker red indicates a lower wavenumber and lighter yellow a higher wavenumber. On the  $x$ - and  $y$ -axes the UTM coordinates of the locations are given.



**Figure 2.D.3:** Slope (a) and aspect (b) estimation error from simulated topographical surfaces. Median errors are shown as horizontal bars with 25<sup>th</sup>-75<sup>th</sup> percentile ranges (boxes) and 5<sup>th</sup> and 95<sup>th</sup> percentile outlier cutoffs (whiskers). Note that aspect error could range from  $-180$  to  $+180$  degrees but has been cropped to largest measured error for visual purpose.



**Figure 2.D.4:** Examples of simulated vegetation patterns with frequency decreasing from left ( $2.4 \text{ cycles } 100\text{m}^{-1}$ ) to right ( $0.4 \text{ cycles } 100\text{m}^{-1}$ ).



**Figure 2.D.5:** Pattern frequency estimation error for the Haud (a) and the Sool (b) sites. Actual frequency of the simulated patterns and the corresponding estimation of these frequencies is shown by the blue dots. The straight line represents the perfect estimation line.



# The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation

---

In this chapter we consider the 2-component reaction-diffusion model that was recently obtained by a systematic reduction of the 3-component Gilad et al. model for dryland ecosystem dynamics [62]. The nonlinear structure of this model is more involved than other more conceptual models, such as the extended-Klausmeier model, and the analysis a priori is more complicated. However, the present model has a strong advantage over these more conceptual models in that it can be more directly linked to ecological mechanisms and observations. Moreover, we find that the model exhibits a richness of analytically tractable patterns that exceeds that of Klausmeier-type models. Our study focuses on the 4-dimensional dynamical system associated with the reaction-diffusion model by considering traveling waves in 1 spatial dimension. We use the methods of geometric singular perturbation theory to establish the existence of a multitude of heteroclinic/homoclinic/periodic orbits that ‘jump’ between (normally hyperbolic) slow manifolds, representing various kinds of localized vegetation patterns. The basic 1-front invasion patterns and 2-front spot/gap patterns that form the starting point of our analysis have a direct ecological interpretation and appear naturally in simulations of the model. By exploiting the rich nonlinear structure of the model, we construct many multi-front patterns that are novel, both from the ecological and the mathematical point of view. In fact, we argue that these orbits/patterns are not specific for the model considered here, but will also occur in a much more general (singularly perturbed reaction-diffusion) setting. We conclude with a discussion of the ecological and mathematical implications of our findings.

## **1 Introduction**

Ecosystems consist of organisms that interact among themselves and with their environment. These interactions involve various kinds of feedback processes that may combine to form positive feedback loops and instabilities when environmental conditions change [107, 108]. In many ecosystems – drylands, peatlands, savannas, mussel beds, coral reefs, and ribbon forests – the leading feedback processes have different spatial scales: a short-range facilitation by local modification of the environment versus a long-range competition for resources [126]. Like the well-established activator-inhibitor principle in bio-chemical systems [113], the combination of these scale-dependent feedback mechanisms can induce instabilities that result in large-scale spatial patterns, which are similar to a wide variety of vegetation patterns observed in drylands, peatlands, savannas and undersea [13, 32, 61, 66, 125, 128, 134]. Varying climatic conditions and human disturbances may continue to propel ecosystem dynamics. Ecosystem response to decreasing rainfall, for example, may take the form of abrupt collapse to a nonproductive ‘desert state’ [125, 130, 161], or involve gradual desertification, consisting of a cascade of state transitions to sparser vegetation [12, 146], or gradual vegetation retreat by front propagation [14, 174]. Understanding the dynamics of spatially extended ecosystems has become an active field of research in the last two decades – within communities of ecologists, environmental scientists, mathematicians and physicists. Apart from its obvious environmental and societal relevance, the phenomena exhibited pose fundamental challenges to the research field of pattern formation.

Several models of increasing complexity have been proposed in the past two decades. Of these, the models that have received most attention are the one-component model by Lefever and Lejeune [97], the two-component models by Klausmeier [90] and von Hardenberg et al. [161], and the three-component models by Rietkerk et al. [124] and Gilad et al. [62]. A basic difference between these models is the manner by which they describe water dynamics. The Lefever-Lejeune model does not describe water dynamics at all, the Klausmeier model does describe water dynamics but does not make a clear distinction between soil water and surface water [155], while the von Hardenberg et al. model only takes soil water into account. The Rietkerk model and the Gilad et al. model describe both soil water and surface water dynamics and, therefore, capture more aspects of real dryland ecosystems. A major difference between these two models is the inclusion of water conduction by laterally spread roots, as an additional water-transport mechanism, in the Gilad et al. model.

Despite these differences, all models appear to share a similar bifurcation structure, as analytical and numerical-continuation studies reveal [38, 67, 98, 173], except the Klausmeier model. This structure includes, in particular, a stationary

uniform instability (i.e., involving the monotonic growth of spatially uniform perturbations) of the bare soil (zero biomass) state as the precipitation rate exceeds a threshold value. The Klausmeier model fails to capture that instability, leaving the bare soil state stable at all precipitation values. This behavior limits the applicability of the Klausmeier model to ecological contexts where the bare soil state is stabilized at relatively high precipitation rates, e.g. by high evaporation rates. Nevertheless, of all models, the Klausmeier model and its extension to include water diffusion have been studied to a greater extent [21, 135, 139, 155], partly because the extended form coincides with the much studied Gray-Scott model for autocatalytic chemical reactions – see [11, 24, 133] and the references therein.

All models have been analyzed mathematically to various extents. Two main analytical approaches can be distinguished in these studies (see however Goto et al. [64]); linear stability and weakly nonlinear analysis near instability points [30, 56, 65, 67, 98, 155], and singular perturbation analysis, based on the disparate length scales associated with biomass (short) and water (long) [11, 21, 24, 133]. Studies of the first category are strictly valid only near instability points, although they do capture essential parts of the bifurcation structure even far from these points and are quite insightful in this respect. By contrast, studies of the second category apply to the strongly nonlinear ‘far-from-equilibrium’ regime, where desertification transitions take place, and are, potentially, of higher ecological interest. So far, however, these studies have been limited to the simpler and less realistic Klausmeier model.

In this chapter we apply a geometric singular perturbation analysis to a reduced version of the Gilad et al. model in order to study the existence of various forms of localized patterns. Singular perturbation theory has already been applied to three-component models – see for instance [46, 159] – and could be applied, in principle, to the non-local three-component Gilad et al. model. Here we choose to consider ecological contexts that allow to reduce that model to a local two-component model for the vegetation biomass and the soil water content. Specifically, we assume soil types characterized by high infiltration rates of surface water into the soil, such as sandy soil, and plant species with laterally confined root zones (see 3.A for more details). These conditions are met, for example, by Namibian grasslands showing localized and extended gap patterns (‘fairy circles’) [175]. We further simplify the problem by assuming one space dimension. The reduced model reads:

$$\begin{cases} \frac{\partial \tilde{B}}{\partial T} = \Lambda \tilde{W} \tilde{B} (1 - \tilde{B}/K)(1 + E\tilde{B}) - M\tilde{B} + D_B \frac{\partial^2 \tilde{B}}{\partial \tilde{X}^2}, \\ \frac{\partial \tilde{W}}{\partial T} = P - N(1 - R\tilde{B}/K)\tilde{W} - \Gamma \tilde{W} \tilde{B} (1 + E\tilde{B}) + D_W \frac{\partial^2 \tilde{W}}{\partial \tilde{X}^2}, \end{cases} \quad (3.1)$$

where  $\tilde{B}(\tilde{X}, T) \geq 0$  and  $\tilde{W}(\tilde{X}, T) \geq 0$  represent areal densities of biomass and soil water, respectively, and  $\tilde{X} \in \mathbb{R}$ ,  $T \in \mathbb{R}^+$  are the space and time coordinates. In the biomass ( $\tilde{B}$ ) equation,  $\Lambda$  represents the biomass growth rate coefficient,  $K$  the maximal standing biomass,  $E$  is a measure for the root-to-shoot ratio,  $M$  the plant mortality rate and  $D_B$  the seed-dispersal or clonal growth rate, while in the water ( $\tilde{W}$ ) equation,  $P$  represents the precipitation rate,  $N$  the evaporation rate,  $R$  the reduction of the evaporation rate due to shading,  $\Gamma$  the water-uptake rate coefficient and  $D_W$  the effective soil water diffusivity. Notice that the power of the factor  $(1 + E\tilde{B})$  in both equations is unity, whereas in the reduced model in [175] the power is two. This difference stems from the consideration in this study of one space dimension rather than two (see 3.A).

From the ecological point of view, the advantage in studying model (3.1) over the much analyzed Klausmeier model lies in the fact that it has been systematically derived from a more extended model that better captures relevant ecological processes, such as water uptake by plant roots (controlled by  $E$ ), reduced evaporation by shading (controlled by  $R$ ), and late-growth constraints, such as self-shading (controlled by  $K$ ) – see [62, 63, 106, 134]. As a consequence, (mathematical) insights in (3.1) can be linked to ecological observations and mechanisms in a direct fashion. Naturally, there also is a disadvantage to analyzing a model that incorporates concrete ecological mechanisms: the more involved – algebraically more complex – nonlinear structure of (3.1) a priori makes it less suitable for an analytical study than the Klausmeier model (or other more conceptual models). However, that apparent disadvantage turned around into an advantage: we will find that the reduced model transcends by far the Klausmeier model in terms of richness of analytically tractable pattern solutions.

The model equations (3.1) represent a singularly perturbed system, because of the low seed-dispersal rate as compared with soil water diffusion, that is,  $D_B \ll D_W$  [63, 155, 175]. To make this explicit and to simplify (3.1) as much as possible, we introduce the following scalings,

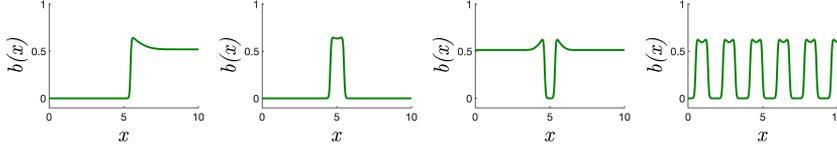
$$B = \frac{\tilde{B}}{\alpha}, \quad W = \frac{\tilde{W}}{\beta}, \quad t = \delta T, \quad x = \gamma \tilde{X}, \quad (3.2)$$

and set,

$$\alpha = K - \frac{1}{E}, \quad \beta = \frac{MK}{\alpha^2 \Lambda E}, \quad \gamma = \sqrt{\frac{\alpha^2 \beta \Lambda E}{KD_B}}, \quad \delta = \frac{\alpha^2 \beta \Lambda E}{K}. \quad (3.3)$$

By also introducing our main parameters,

$$a = \frac{KE}{(KE - 1)^2}, \quad \varepsilon^2 = \frac{D_B}{D_W} \ll 1, \quad (3.4)$$



**Figure 3.1:** The 4 basic patterns exhibited by numerical simulations of model (3.5): a traveling (heteroclinic) invasion front (Theorem 3.4), a stationary, homoclinic, 2-front vegetation spot (Theorem 3.11), a stationary homoclinic, 2-front vegetation gap (Theorem 3.13), and a stationary, spatially periodic multi-front spot/gap pattern (Theorem 3.15) – see Remark 4.1 for the precise parameter values.

we arrive at,

$$\begin{cases} B_t = (aW - 1)B + WB^2 - WB^3 + B_{xx}, \\ W_t = \Psi - [\Phi + \Omega B + \Theta B^2]W + \frac{1}{\varepsilon^2}W_{xx}, \end{cases} \quad (3.5)$$

in which,

$$\Psi = \frac{\alpha^2 P \Lambda E}{M^2 K}, \quad \Phi = \frac{N}{M}, \quad \Omega = \frac{\alpha}{M} \left( \Gamma - \frac{R}{K} \right), \quad \Theta = \frac{\alpha^2 \Gamma E}{M}. \quad (3.6)$$

A more detailed derivation of the scaled equations (3.5) from (3.1) is given in 3.B. Since the signs of the parameters in (3.5) will turn out to be crucial in the upcoming analysis, we note explicitly that  $a, \Psi, \Phi, \Theta \geq 0$  while  $\Omega \in \mathbb{R}$ , i.e.,  $\Omega$  may be negative.

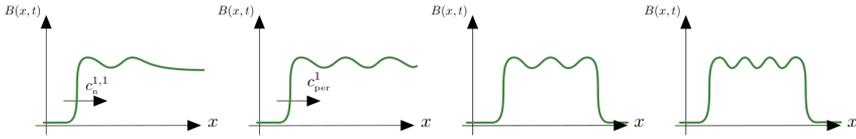
We study pattern formation in (3.1) by analyzing (3.5) using the methods of (geometric) singular perturbation theory [83, 86] and thus ‘exploit’ the fact that  $\varepsilon \ll 1$  (3.4). In fact – apart from some observations in section 2.3 and the discussion section 4.2 – we focus completely on the ‘spatial’ 4-dimensional dynamical system that is obtained from (3.5) by considering ‘simple’ solutions that are stationary in a co-moving frame traveling with constant speed  $c$ . More specifically, in this chapter we study the existence of traveling (and stationary) solutions – in particular localized (multi-)front solutions connecting a (uniform) bare soil state to a uniform vegetation state, or a bare soil state to itself (with a ‘passage’ along a vegetated state), etc. – by taking the classical approach of introducing a (uniformly) traveling coordinate  $\xi = x - ct$ , with speed  $c \in \mathbb{R}$  an a priori free  $\mathcal{O}(1)$  parameter (w.r.t. the asymptotically small parameter  $\varepsilon$ ). By setting  $(B(x, t), W(x, t)) = (b(\xi), w(\xi))$  and introducing  $p = b_\xi$  and  $q = \frac{1}{\varepsilon}w_\xi$ ,

PDE (3.5) reduces to

$$\begin{cases} b_\xi = p, \\ p_\xi = wb^3 - wb^2 + (1 - aw)b - cp, \\ w_\xi = \varepsilon q, \\ q_\xi = \varepsilon \left( -\Psi + \left[ \Phi + \Omega b + \Theta b^2 \right] w \right) - \varepsilon^2 cq. \end{cases} \quad (3.7)$$

Fig. 3.1 shows four basic patterns that naturally appear in simulations of (3.5) and have identifiable ecological counterparts: vegetation fronts (ecotones), isolated vegetation spots and gaps, and periodic patterns [49, 55, 61, 109]. These patterns are rigorously constructed by the methods of singular perturbation theory in section 3. From the geometrical point of view, these constructions are natural and thus relatively straightforward: all patterns in Fig. 3.1 ‘jump’ between two well-defined slow manifolds (of (3.7)) – see Theorems 3.4, 3.11, 3.13, and 3.15. Therefore, the main work in establishing these results lies in resolving technical issues (which can be done by the preparations of section 2). However, the preparations of section 2 also form the origin of the construction of a surprisingly rich ‘space’ of traveling and/or stationary patterns that goes way beyond those exhibited in Fig. 3.1 – see for instance the sketches of Fig. 3.2. These are novel patterns, at least from the point of view of explicit rigorous mathematical constructions in multi-component reaction-diffusion equations. However, similar patterns have been analyzed as (perturbations of) heteroclinic networks in a more abstract framework – see [120, 121] and the references therein. Moreover, patterns similar to those of Fig. 3.2 have been observed in simulations of the Klausmeier-Gray-Scott model [172], although with parameter settings beyond that for which the mathematical singular-perturbation approach can be applied.

Here, our motivation to study these patterns is primarily ecological; however, we claim that patterns like these must also occur generically in the setting of a completely general class of singularly perturbed 2-component reaction-diffusion systems – as we will motivate in more detail in section 4.2. Thus, our explicit analysis of model (3.5) provides novel mathematical insights beyond that of the present ecological setting. The driving mechanism behind these patterns originates from the perturbed integrable flow on the slow manifolds associated with (3.7) – see sections 2.2 and 2.4. The perturbation terms are generically introduced by the  $\mathcal{O}(\varepsilon)$  differences between the slow manifold and its  $\varepsilon \rightarrow 0$  limit, and they transform the (Hamiltonian) integrable reduced slow flow to a (planar) ‘nonlinear oscillator with nonlinear friction’ that can be studied by explicit Melnikov methods. Typically, one for instance expects (and finds: Theorem 2.4) persistent periodic solutions on the slow manifold. Associated with these persisting periodic solutions, one can subsequently construct heteroclinic 1-front connections between a critical point – representing the uniform bare soil state in the ecological setting – and such a periodic pattern (Theorems 3.5 and 3.9 and



**Figure 3.2:** Four sketches of ‘higher order’ localized patterns constructed in this chapter. (a) A secondary traveling 1-front, the second one in a countable family of traveling 1-fronts between the bare soil state and a uniform vegetated state – all traveling with different speeds – that starts with the primary 1-front of Fig. 3.1(a) (Theorem 3.6). (b) The limiting orbit of the family sketched in (a): a 1-front connection between the bare soil state and a spatially periodic vegetation state (Theorem 3.5). (c,d) The first 2 representations of a (countable) ‘higher order’ family of localized (stationary, homoclinic 2-front) spot patterns with an increasing number of ‘spatial oscillations’ (Theorem 3.12).

Fig. 3.2b) and a countable family of ‘higher order’ heteroclinic 1-fronts between critical points that limits on these orbits (Theorem 3.6 and Fig. 3.2a – where we note that Fig. 3.1a represents the very first – primary – member of this family). In the case of (stationary) localized spot patterns, one can construct a countable family of connections that follow the periodic orbit for arbitrarily many ‘spatial oscillations’ (Theorem 3.12 and Fig. 3.2c, 3.2d). Combining these insights with the ideas of [44], one may even construct many increasingly complex families of spatially periodic and aperiodic multi-spot/gap patterns (Corollary 3.16 and section 3.6). Moreover, we can explicitly study the associated bifurcation scenarios: in section 3.3 we present a scenario of cascading saddle-node bifurcations of heteroclinic 1-front connections starting from no such orbits to countably many – all traveling with different speed (Theorem 3.6 and Figs. 3.1a, 3.2a and 3.2b) – back to 1 unique 1-front pattern (of the type presented in Fig. 3.1a) – see Fig. 3.9 in section 3.3.

Finally, we illustrate our analytical findings by several numerical simulations of PDE model (3.1)/(3.7) – see also Fig. 3.1. We did not systematically investigate the question whether all heteroclinic/homoclinic/periodic (multi-)front orbits of (3.7) constructed here indeed may be (numerically) observed as stable patterns in (3.7), either for general parameter combinations in (3.5) or for the more restricted class of ecologically relevant parameter combinations. This will be the subject of future work, as will be the analytical question about the spectral stability of the constructed patterns. These issues will be discussed more extensively in section 4.2, where we will also discuss further implications of our findings – both from the mathematical as well as from the ecological point of view.

The set-up of this chapter is as follows. Section 2 is a preparatory section: in section 2.1 and 2.2 we consider the fast and slow reduced problems associated with (3.7), followed by a brief section – section 2.3 – in which we discuss the

nature (and stability) of the critical points of (3.7) as uniform vegetated states in (3.5); in section 2.4 we study the full, perturbed, slow flow on the slow manifolds (leading to Theorem 2.4). All localized patterns are constructed in section 3, which begins with (another) preparatory section – section 3.1 – in which we set up the geometry of orbits ‘jumping’ between slow manifolds. The primary traveling 1-front patterns of Fig. 3.1a are constructed in section 3.2, the associated higher order 1-fronts of Figs. 3.2a and 3.2b in section 3.3. Stationary patterns are considered in 3.4 – on 1-fronts – and 3.4 – on 2-fronts of spot and gap type as shown in Figs. 3.1b, 3.1c and Fig. 3.2(c,d); various families of spatially periodic multi-front patterns – including the basic ones of Fig. 3.1d – are constructed in section 3.6. Section 4 begins with section 4.1 in which we show various numerically obtained patterns – some of them beyond the analysis of the present chapter – and ends with discussion section 4.2.

**Remark 1.1.** *While the original model (3.1) has 8 parameters –  $(\Lambda, \Gamma, R, K, E, M, N, P)$  – (neglecting  $D_B, D_W$  which are represented by  $\varepsilon$ ), rescaled model (3.5) has only 5 parameters –  $(a, \Psi, \Phi, \Omega, \Theta)$ . We will formulate our results by stipulating conditions on  $(a, \Psi, \Phi, \Omega, \Theta)$  and refrain from giving a corresponding range for the original parameters. Moreover, we notice that we have implicitly assumed that  $\alpha > 0$ , i.e., that  $EK > 1$  (3.3). This is a technical assumption (and not unrealistic from ecological point of view), the case  $0 < EK < 1$  can be treated in a completely analogous way – see 3.B.*

## 2 Set-up of the existence problem

We first notice that (3.7) is the ‘fast’ description of the ‘spatial ODE’ associated with (3.5). By introducing  $X = \varepsilon\xi$  ( $= \varepsilon(x - ct)$ ) we obtain its equivalent slow form,

$$\begin{cases} \varepsilon b_X = p, \\ \varepsilon p_X = wb^3 - wb^2 + (1 - aw)b - cp, \\ w_X = q, \\ q_X = -\Psi + [\Phi + \Omega b + \Theta b^2]w - \varepsilon cq. \end{cases} \quad (3.8)$$

Note that these systems possess a  $c \rightarrow -c$  symmetry that reduces to a reversibility symmetry for  $c = 0$ ,

$$(c, \xi, p, q) \rightarrow (-c, -\xi, -p, -q) \text{ or } (c, X, p, q) \rightarrow (-c, -X, -p, -q). \quad (3.9)$$

### 2.1 The fast reduced problem

The fast reduced limit problem associated to (3.7) is a two-parameter family of planar systems that is obtained by taking the limit  $\varepsilon \rightarrow 0$  in (3.7),

$$b_{\xi\xi} = w_0 b^3 - w_0 b^2 + (1 - aw_0)b - cb_{\xi}, \quad (w, q) \equiv (w_0, q_0) \in \mathbb{R}^2. \quad (3.10)$$

These planar systems can have up to 3 (families of) critical points (parameterized by  $(w_0, q_0)$ ) given by,

$$(b_0, p_0) = (0, 0), \quad (b_{\pm}, p_{\pm}) = (b_{\pm}(w_0), 0) = \left( \frac{1}{2} \pm \sqrt{a + \frac{1}{4} - \frac{1}{w_0}}, 0 \right). \quad (3.11)$$

Clearly,  $(b_0, w_0)$  represents the (homogeneous) bare soil state  $B(x, t) \equiv 0$ , the other two solutions correspond to uniform vegetation states and only exist for  $w_0 > 4/(1 + 4a)$ . The critical points also determine 3 two-dimensional invariant (slow) manifolds,  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^{\pm}$ ,

$$\begin{aligned} \mathcal{M}_0^0 &= \{(b, p, w, q) \in \mathbb{R}^4 : b = 0, p = 0\}, \\ \mathcal{M}_0^{\pm} &= \left\{ (b, p, w, q) \in \mathbb{R}^4 : b = b_{\pm}(w) = \frac{1}{2} \pm \sqrt{a + \frac{1}{4} - \frac{1}{w}}, p = 0 \right\}. \end{aligned} \quad (3.12)$$

A straightforward analysis yields that the critical points  $(b_+, 0)$  are saddles for all  $c \in \mathbb{R}$  and that the points  $(b_0, p_0) = (0, 0)$  are saddles for all  $c$  as long as  $w_0 < 1/a$ . Therefore, we consider in this chapter  $w_0$  such that,

$$w_0 \in \mathcal{U}_a = \left\{ w_0 \in \mathbb{R} \mid \frac{4}{1 + 4a} < w_0 < \frac{1}{a} \right\}, \quad (3.13)$$

so that (parts of) the manifolds  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  are normally hyperbolic for all  $w_0$  that satisfy (3.13) (and thus persist as  $\varepsilon$  becomes nonzero [83, 86]); moreover, all stable and unstable manifolds  $W^{s,u}(\mathcal{M}_0^0)$  and  $W^{s,u}(\mathcal{M}_0^+)$  are 3-dimensional. (In this chapter, we do not consider the manifold  $\mathcal{M}_0^-$  for several reasons: (i) it is not normally hyperbolic in the crucial case of stationary patterns (i.e., for  $c = 0$ , under the – natural – assumption that the water concentration  $w_0$  does not become negative), (ii) critical points for the full system (3.7) that limit on  $\mathcal{M}_0^-$  as  $\varepsilon \rightarrow 0$  cannot correspond to stable homogeneous states of PDE (3.5) – see section 2.3.)

The manifolds  $W^{s,u}(\mathcal{M}_0^0)$  and  $W^{s,u}(\mathcal{M}_0^+)$  are determined by the stable and unstable manifolds of  $(0, 0)$  and  $(b_+, 0)$ . By the (relatively) simple cubic nature of (3.10) we do have explicit control over these manifolds in the relevant cases that they collide, i.e., that there is a heteroclinic connection between  $(0, 0)$  and  $(b_+, 0)$ . Although this is a classical procedure – see [113] – we provide a brief sketch here.

We assume that a heteroclinic solution of (3.10) between  $(0, 0)$  and  $(b_+, 0)$  can also be written as a solution of the first order equation

$$b_{\xi} = nb(b_+(w_0) - b), \quad (3.14)$$

where  $n$  is a free pre-factor (and we know that this assumption provides all possible heteroclinic connections). Taking the derivative (w.r.t.  $\xi$ ) yields an

equation for  $b_{\xi\xi}$  that must equal (3.10) – that we write as  $b_{\xi\xi} = w_0 b(b - b_-)(b - b_+) - cb_{\xi}$ . Working out the details yield explicit expressions for  $n$  and  $c$ ,

$$\begin{aligned} n &= n^{\pm}(w_0) = \pm \sqrt{\frac{1}{2}w_0}, \\ c &= c^{\pm}(w_0) = \pm \sqrt{\frac{1}{2}w_0} \left( 3 \sqrt{a + \frac{1}{4} - \frac{1}{w_0}} - \frac{1}{2} \right). \end{aligned} \quad (3.15)$$

Thus, for a given  $c$ , there is a heteroclinic connection between  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  at the ‘level’  $w_0 = w_h^{\pm}(c)$  if  $w_0$  solves (3.15). A direct calculation yields that  $c^{\pm}(w_0)$  are strictly monotonic function of  $w_0$  with inverse

$$w_h^{\pm}(c) = \frac{4(9 + 2c^2)^2}{\left( 3\sqrt{2c^2(1 + 4a)} + 4(2 + 9a) \mp \sqrt{2}c \right)^2}. \quad (3.16)$$

We conclude that for a given  $c$ , there may be ‘parabolic’ – by the relation between  $b$  and  $p$  (3.14) – two-dimensional intersections  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  and  $W^s(\mathcal{M}_0^0) \cap W^u(\mathcal{M}_0^+)$  explicitly given by,

$$\begin{aligned} W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+) &= \left\{ 0 < b < b_+(w_h^+), p = n^+(w_h^+)b(b_+(w_h^+) - b), w = w_h^+ \right\}, \\ W^s(\mathcal{M}_0^0) \cap W^u(\mathcal{M}_0^+) &= \left\{ 0 < b < b_+(w_h^-), p = n^-(w_h^-)b(b_+(w_h^-) - b), w = w_h^- \right\} \end{aligned} \quad (3.17)$$

(where we recall that  $q = q_0 \in \mathbb{R}$  is still a free parameter). See Lemma 3.2 for a further discussion and analysis (for instance on the allowed  $c$ -intervals for which the heteroclinic connections exist:  $w_h^{\pm}(c)$  must satisfy (3.13)).

In the case of stationary patterns ( $c = 0$ ), fast reduced limit problem (3.10) is integrable, with Hamiltonian  $\mathcal{H}_f$  given by,

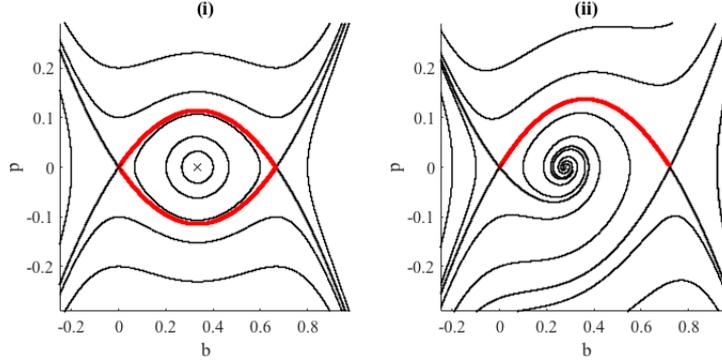
$$\mathcal{H}_f(b, p; w_0) = \frac{1}{2}p^2 - \frac{1}{2}(1 - aw_0)b^2 + \frac{1}{3}w_0b^3 - \frac{1}{4}w_0b^4, \quad (3.18)$$

which is gauged such that  $\mathcal{H}_f(0, 0; w_0) = 0$ . This system has a heteroclinic connection between  $(0, 0)$  and  $(b_+^0, 0)$  for  $w_0 = w_h^{\pm}(0)$  such that  $\mathcal{H}_f(b_+(w_0), 0; w_0) = \mathcal{H}_f(0, 0; w_0) = 0$ . It follows by (3.11) and (3.18) that  $w_h^+(0) = w_h^-(0) = 9/(2 + 9a)$  (which agrees with (3.16)) – see Fig. 3.3.

## 2.2 The slow reduced limit problems

The slow reduced limit problem is obtained by taking the limit  $\varepsilon \rightarrow 0$  in (3.8). It is a planar problem in  $(w, q)$ ,

$$w_{XX} = -\Psi + \left[ \Phi + \Omega b + \Theta b^2 \right] w. \quad (3.19)$$



**Figure 3.3:** Numerical simulations of dynamics of the fast reduced system (3.7) for  $a = \frac{1}{4}$  and two choices of  $w_0 \in \mathcal{U}_a$  (3.13), both featuring a heteroclinic orbit between the saddle points  $(0, 0)$  and  $(b_+(w_0), 0)$ : **(i)**  $w_0 = 9/(2 + 9a)$ ,  $c = c_{\pm}(w_0) = 0$ ; **(ii)**  $w_0 = 9/(2 + 9a) + 0.1$ ,  $c = c_+(w_0) \approx 0.17$ .

restricted to  $(p, b)$  such that,

$$p = 0, \quad wb^3 - wb^2 + (1 - aw)b = 0$$

i.e., (3.19) describes the (slow) flow on the (slow) manifolds  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^{\pm}$  (3.12). The flow on  $\mathcal{M}_0^0$  is linear,

$$w_{XX} = -\Psi + \Phi w, \quad (3.20)$$

with critical point  $P_0^0 = (0, 0, \Psi/\Phi, 0) \in \mathcal{M}_0^0$  of saddle type – that corresponds to the uniform bare soil state  $(B(x, t), W(x, t)) \equiv (0, \Psi/\Phi)$  of (3.5) – that has the stable and unstable manifolds (on  $\mathcal{M}_0^0$ ) given by

$$\begin{aligned} W^s(P_0^0)|_{\mathcal{M}_0^0} &:= \ell_0^s = \left\{ (b, p, w, q) \in \mathcal{M}_0^0 : q = -\sqrt{\Phi} \left( w - \frac{\Psi}{\Phi} \right) \right\}, \\ W^u(P_0^0)|_{\mathcal{M}_0^0} &:= \ell_0^u = \left\{ (b, p, w, q) \in \mathcal{M}_0^0 : q = \sqrt{\Phi} \left( w - \frac{\Psi}{\Phi} \right) \right\} \end{aligned} \quad (3.21)$$

Since we focus on orbits – patterns – that ‘jump’ between  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  (in the limit  $\varepsilon \rightarrow 0$ ), we do not consider the flow on  $\mathcal{M}_0^-$  but focus on (the flow on)  $\mathcal{M}_0^+$ ,

$$w_{XX} = -\mathcal{A} + (\mathcal{B} + a\Theta)w + \mathcal{C}w \sqrt{a + \frac{1}{4} - \frac{1}{w}}, \quad (3.22)$$

where

$$\mathcal{A} = \Psi + \Theta \geq 0, \quad \mathcal{B} = \Phi + \frac{1}{2}\Omega + \frac{1}{2}\Theta \in \mathbb{R}, \quad \mathcal{C} = \Omega + \Theta \in \mathbb{R}, \quad (3.23)$$

and we notice explicitly that  $\mathcal{B}$  and  $\mathcal{C}$  may be negative (since  $\Omega$  may be negative). For  $w$  satisfying (3.13), we define,

$$\mathcal{W} = \sqrt{a + \frac{1}{4} - \frac{1}{w}} \geq 0, \quad \mathcal{D} = \mathcal{B} + a\Theta - \left(a + \frac{1}{4}\right)\mathcal{A} \in \mathbb{R}, \quad (3.24)$$

and conclude that the critical points  $P_0^{+,j} = (b_+(w_0^{+,j}), 0, w_0^{+,j}, 0) \in \mathcal{M}_0^+$  are determined as solutions of the quadratic equation,

$$\mathcal{A}\mathcal{W}^2 + \mathcal{C}\mathcal{W} + \mathcal{D} = 0. \quad (3.25)$$

Thus, the points  $P_0^{+,j}$  exist for parameter combinations such that  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0$ . There are 2 critical points if additionally  $\mathcal{C} < 0$  and  $\mathcal{D} > 0$  and only 1 if  $\mathcal{D} < 0$ .

Clearly, the flow (3.22) is integrable, with Hamiltonian given by

$$\mathcal{H}_0^+(w, q) = \frac{1}{2}q^2 + \mathcal{A}w - \frac{1}{2}(\mathcal{B} + a\Theta)w^2 - \mathcal{C}\mathcal{J}_0^+(w), \quad (3.26)$$

with, for  $\tilde{a} = a + \frac{1}{4}$ ,

$$\mathcal{J}_0^+(w) = \frac{1}{4\tilde{a}}(2\tilde{a}w - 1) \sqrt{\tilde{a}w^2 - w} - \frac{1}{8\tilde{a}\sqrt{\tilde{a}}} \ln \left| \frac{1}{2}(2\tilde{a}w - 1) + \sqrt{\tilde{a}} \sqrt{\tilde{a}w^2 - w} \right|. \quad (3.27)$$

Hence, if non-degenerate, the critical points  $P_0^{+,j}$  are either centers –  $P_0^{+,c}$  – or saddles –  $P_0^{+,s}$ . Notice that, except the uniform bare soil state  $(0, \Psi/\Phi)$ , all critical points correspond to uniform vegetation states  $(B(x, t), W(x, t)) \equiv (\bar{B}, \bar{W})$  in (3.5) – see section 2.3. In the case that there is only 1 critical point  $P_0^+ \in \mathcal{M}_0^+$ , it can either be of saddle or center type:  $P_0^+$  is a saddle if,

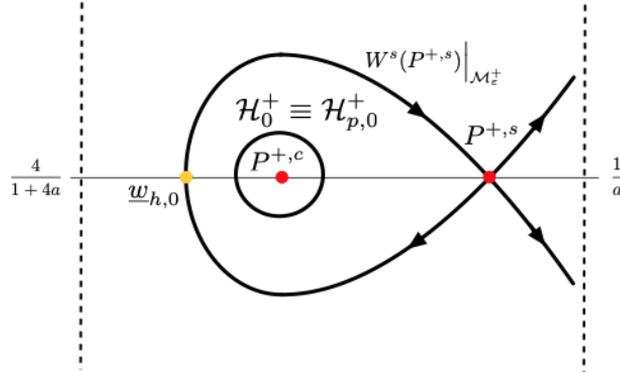
$$\mathcal{E} = \mathcal{B} + a\Theta + \frac{1}{2}\mathcal{C} \left( \mathcal{W} + \frac{a + \frac{1}{4}}{\mathcal{W}} \right) > 0 \quad (3.28)$$

where  $\mathcal{W} > 0$  is the solution of (3.25). We notice that the stable and unstable manifolds (restricted to  $\mathcal{M}_0^+$ ) of the saddle point  $P_0^{+,s} \in \mathcal{M}_0^+$  are represented by,

$$W^s(P_0^{+,s}) \cup W^u(P_0^{+,s})|_{\mathcal{M}_0^+} = \left\{ (b, p, w, q) \in \mathcal{M}_0^+ : \mathcal{H}_0^+(w, q) \equiv \mathcal{H}_0^{+,s} := \mathcal{H}_0^+(w_0^{+,s}, 0) \right\}. \quad (3.29)$$

In the upcoming analysis, we will be especially interested in the case of 2 critical points  $P_0^{+,s}, P_0^{+,c} \in \mathcal{M}_0^+$ , therefore we investigate this situation on some more detail. First, we introduce  $\mathcal{D}^{SN}$  and  $\sigma \geq 0$  by setting,

$$\mathcal{D}(\sigma^2) = \mathcal{D}^{SN} - \mathcal{A}\sigma^2 = \frac{\mathcal{C}^2}{4\mathcal{A}} - \mathcal{A}\sigma^2 > 0 : \sigma = \sqrt{\frac{\mathcal{D} - \mathcal{D}^{SN}}{\mathcal{A}}}, \quad (3.30)$$



**Figure 3.4:** Phase portrait of the unperturbed flow (3.22) on  $\mathcal{M}_0^+$  for parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  such that (3.40) holds.

so that the solutions of (3.25) are given by  $\mathcal{W} = \mathcal{W}^{SN} \pm \sigma = -\frac{\mathcal{C}}{2\mathcal{A}} \pm \sigma$ . We rewrite (3.22) in terms of  $(a, \mathcal{A}, \mathcal{C}, \mathcal{D})$

$$w_{XX} = -\mathcal{A} + \left( \mathcal{D} + \left( a + \frac{1}{4} \right) \mathcal{A} \right) w + \mathcal{C} w \sqrt{a + \frac{1}{4} - \frac{1}{w}}. \quad (3.31)$$

Clearly,  $\sigma = 0$  corresponds to the degenerate saddle-node case in which  $P_0^{+,s}$  and  $P_0^{+,c}$  merge,

$$P_0^{SN} = \left( b_+(w_0^{SN}), 0, w_0^{SN}, 0 \right) \quad (3.32)$$

with

$$w_0^{SN} = \frac{4\mathcal{A}^2}{(1+4a)\mathcal{A}^2 - \mathcal{C}^2}, \quad (1+4a)\mathcal{A}^2 - \mathcal{C}^2 \neq 0, \quad (3.33)$$

where we note that  $w_0^{SN}$  satisfies (3.13) for  $0 < \mathcal{C}^2 < \mathcal{A}^2$  (independent of  $a$ ). In fact, we can consider the unfolding of the saddle-node bifurcation by the additional assumption that  $0 < \sigma \ll 1$ ,

$$w_0^{+,j} = w_0^{SN} \pm w_1^{SN} \sigma + \mathcal{O}(\sigma^2) = w_0^{SN} \pm 2\mathcal{W}^{SN} (w_0^{SN})^2 \sigma + \mathcal{O}(\sigma^2), \quad (3.34)$$

( $j = 1, 2$ ), where the “+”-case represents the saddle  $P_0^{+,s}$  and the “-”-case the center  $P_0^{+,c}$ :  $w_0^{+,c} < w_0^{+,s}$  – see Fig. 3.4. In this parameter region, the slow reduced system (3.22) features a homoclinic orbit  $(w_{\text{hom}}, q_{\text{hom}})$  to  $P_0^{+,s}$  and a family of periodic solutions around the center point  $P_0^{+,c}$  (Fig. 3.4).

**Remark 2.1.** We conclude from (3.31) that the reduced slow flow on  $\mathcal{M}_0^+$  is fully determined by the values of  $(a, \mathcal{A}, \mathcal{B}, \mathcal{D})$ . Clearly, the (linear) mapping  $(\Psi, \Phi, \Omega, \Theta) \mapsto (\mathcal{A}, \mathcal{B}, \mathcal{D})$  has a kernel: we can vary one of the parameters – for instance  $\Phi$  – and determine  $(\Psi, \Omega, \Theta)$  such that this does not have an effect on the

reduced flow (3.31) on  $\mathcal{M}_0^+$  (by choosing  $(\Psi(\Phi), \Theta(\Phi), \Omega(\Phi))$  such that  $(\mathcal{A}, \mathcal{B}, \mathcal{D})$  are kept at a chosen value). We will make use of this possibility extensively in section 3.

### 2.3 Critical points and homogeneous background states

Since the critical points  $P^j = (b_j, p^j, w^j, q^j)$  of the full  $\varepsilon \neq 0$  system (3.7) must have  $p^j = q^j = 0$ , their  $(b, w)$  coordinates are determined by the intersections of the  $b$ - and  $w$ -nullclines,

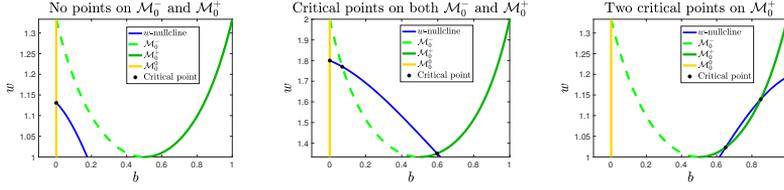
$$wb^3 - wb^2 + (1 - aw)b = 0, \quad -\Psi + [\Phi + \Omega b + \Theta b^2]w = 0, \quad (3.35)$$

where we recall that the  $b$ -nullcline determines the slow manifolds  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^\pm$  – see Fig. 3.5. Hence, all critical points  $P^j$  must correspond to critical points of the slow reduced flows on either one of the (unperturbed) slow manifolds  $\mathcal{M}_0^0$ ,  $\mathcal{M}_0^+$  or  $\mathcal{M}_0^-$ . This immediately implies that  $P^1 = P_0^0 = (0, 0, \Psi/\Phi, 0) \in \mathcal{M}_0^0$ . The (potential) critical points on  $\mathcal{M}_0^-$  can be determined completely analogously to  $P_0^{+,j} \in \mathcal{M}_0^+$  in section 2.2 – the only difference is that the term  $+C\mathcal{W}$  in (3.25) must be replaced by  $-C\mathcal{W}$ . Thus, we conclude that there are two additional critical points  $P^2$  and  $P^3$  if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} \geq 0$  (and that  $P^1 = P_0^0$  is the unique critical point if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} \leq 0$ ). Moreover, if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} \geq 0$  then,

- if  $\mathcal{D} < 0$ , then  $P^2 = P_0^- \in \mathcal{M}_0^-$  and  $P^3 = P_0^+ \in \mathcal{M}_0^+$ ;
- if  $\mathcal{D} > 0$  and  $\mathcal{C} > 0$ , then  $P^2 = P^{-,1}$ ,  $P^3 = P^{-,2}$  and both  $P^{-,j} \in \mathcal{M}_0^-$ ;
- if  $\mathcal{D} > 0$  and  $\mathcal{C} < 0$ , then  $P^2 = P^{+,1}$ ,  $P^3 = P^{+,2}$  and both  $P^{+,j} \in \mathcal{M}_0^+$ .

Naturally, the critical points  $P^j$  correspond to homogeneous background states  $(B(x, t), W(x, t)) \equiv (\bar{B}^j, \bar{W}^j)$  of the full PDE (3.5). In this chapter, we focus on the *existence* of patterns in (3.5) and do not consider the stability of these patterns (which is the subject of work in progress). However, there is a strong relation between the local character of critical points  $P^j$  in the spatial system (3.7) and their (in)stability as homogeneous background pattern in (3.5) – see for instance [39]. Therefore, we may immediately conclude,

- the bare soil state  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  is stable as solution of (3.5) for  $\Psi/\Phi < 1/a$ , i.e., as long as  $(0, \Psi/\Phi)$  corresponds to a critical point on the normally hyperbolic part of  $\mathcal{M}_0^0$  (3.13);
- background states  $(\bar{B}, \bar{W})$  that correspond to critical points on  $\mathcal{M}_0^-$  are unstable;
- a background state  $(\bar{B}, \bar{W})$  that corresponds to a center point on  $\mathcal{M}_0^+$  is unstable;
- a background state  $(\bar{B}, \bar{W})$  that corresponds to a saddle point on  $\mathcal{M}_0^+$  is stable as solution of (3.5) if one additional (technical) condition on the parameters of (3.5) is satisfied.



(a) No intersections of the  $w$ -nullcline with either  $\mathcal{M}_0^-$  or  $\mathcal{M}_0^+$  ( $a = 0.75, \Psi = 0.1131, \Omega = 0.0369, \Phi = 0.1, \Theta = 0.2131$ ).

(b) A unique intersection of the  $w$ -nullcline with both  $\mathcal{M}_0^-$  and  $\mathcal{M}_0^+$  ( $a = 0.75, \Psi = 1.9, \Omega = 0.1, \Phi = 0.1, \Theta = 0.3, \Theta = 0.5$ ).

(c) Two intersections of the  $w$ -nullcline with  $\mathcal{M}_0^+$  and none with  $\mathcal{M}_0^-$  ( $a = 0.75, \Psi = 0.2983, \Omega = -0.4517, \Phi = 0.5, \Theta = 0.2017$ ).

**Figure 3.5:** Various relative configurations of the nullclines (3.35) and the associated critical points for  $w \in \mathcal{U}_a$  (3.13).

Of course this motivates our choice to study homoclinic and heteroclinic connections between the saddle points on  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  in this chapter.

**Remark 2.2.** *The singular perturbation point of view also immediately provides insight in the possible occurrence of a Turing bifurcation in (3.5). In the setting of (3.7) – with  $c = 0$  – a Turing bifurcation corresponds to a reversible 1 : 1 resonance Hopf bifurcation [76], i.e., the case of a critical point with 2 colliding pairs of purely imaginary eigenvalues. By the slow/fast nature of the flow of (3.7), such a critical point cannot lay inside one of the 3 possible reduced slow manifolds  $\mathcal{M}_0^0, \mathcal{M}_0^-$  or  $\mathcal{M}_0^+$  (critical points not asymptotically close to the boundaries must have 2  $\mathcal{O}(\varepsilon)$  and 2  $\mathcal{O}(1)$  eigenvalues). Thus, critical points that may undergo a Turing/reversible 1 : 1 Hopf bifurcation have to be asymptotically close to the edge of  $\mathcal{M}_\varepsilon^+$  where it approaches  $\mathcal{M}_\varepsilon^-$  (where we note that we a priori do not claim that  $\mathcal{M}_\varepsilon^-$  persists). Indeed, the bifurcation appears in that region – although we refrain from going into the details. See Fig. 3.18a for a thus found spatially periodic Turing pattern in (3.5).*

**Remark 2.3.** *By directly focusing on (3.35) – and thus by not following the path indicated by the singularly perturbed structure of (3.7) – the uniform vegetation background states can also be computed in a more straightforward way: assuming*

*$b \neq 0$ , yields  $w = -\frac{1}{b^2 - b - a}$ , which implies that  $(\Theta + \Psi)b^2 + (\Omega - \Psi)b + (\Phi - a\Psi) = 0$ . Hence it follows (for  $(\Omega - \Psi)^2 - 4(\Theta + \Psi)(\Phi - a\Psi) \geq 0$ ) that*

$$(b_{1,2}, w_{1,2}) = \left( \frac{-(\Omega - \Psi) \pm \sqrt{(\Omega - \Psi)^2 - 4(\Theta + \Psi)(\Phi - a\Psi)}}{2(\Theta + \Psi)}, -\frac{1}{b_{1,2}^2 - b_{1,2} - a} \right).$$

### 2.4 The slow flows of the $\varepsilon \neq 0$ system

Condition (3.13) was chosen such that the points  $(0, 0, w, q) \in \mathcal{M}_0^0$  and  $(b_+(w), 0, w, q) \in \mathcal{M}_0^+$  are saddles for the fast reduced limit problem (3.10) (so that the associated background states may be stable as trivial, homogeneous, patterns of (3.5) – section 2.3). Thus, where (3.13) holds,  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  are normally hyperbolic and they thus persist as  $\mathcal{M}_\varepsilon^0$  and  $\mathcal{M}_\varepsilon^+$  for  $\varepsilon \neq 0$  [83, 86]. Clearly,  $\mathcal{M}_0^0$  is also invariant under the flow of the full system (3.7):  $\mathcal{M}_\varepsilon^0 = \mathcal{M}_0^0$ . Moreover, the flow on  $\mathcal{M}_\varepsilon^0$  is only a slight  $-\mathcal{O}(\varepsilon)$  – (linear) perturbation of the unperturbed flow (3.20) on  $\mathcal{M}_0^0$  – due to the (asymmetric)  $-\varepsilon cq$  term. As a consequence, only the orientation of the (un)stable manifolds  $W^{s,u}(P_0^0)|_{\mathcal{M}_\varepsilon^0} = \ell_\varepsilon^{s,u}$  undergoes an  $\mathcal{O}(\varepsilon)$  change w.r.t.  $\ell_0^{s,u}$  (3.21).

The situation is very different for  $\mathcal{M}_\varepsilon^+$ . A direct perturbation analysis yields,

$$\mathcal{M}_\varepsilon^+ = \left\{ (b, p, w, q) \in \mathbb{R}^4 : b = b_+(w) + \varepsilon cq b_1(w) + \mathcal{O}(\varepsilon^2), p = \varepsilon q p_1(w) + \mathcal{O}(\varepsilon^2) \right\}, \quad (3.36)$$

with

$$p_1(w) = \frac{1}{2w^2 \sqrt{a + \frac{1}{4} - \frac{1}{w}}}, \quad b_1(w) = \frac{p_1(w)}{2wb_+(w) \sqrt{a + \frac{1}{4} - \frac{1}{w}}} \quad (3.37)$$

and  $b_+(w)$  as defined in (3.11). Since we only consider situations in which there are critical points (of the full flow)  $P^{+,j}$  on  $\mathcal{M}_0^+$ , and thus on  $\mathcal{M}_\varepsilon^+$ , we know (and use) that  $\mathcal{M}_\varepsilon^+$  is determined uniquely. The slow flow on  $\mathcal{M}_\varepsilon^+$  is given by

$$w_{XX} = -\mathcal{A} + (\mathcal{B} + a\Theta)w + \mathcal{C}w \sqrt{a + \frac{1}{4} - \frac{1}{w}} + \varepsilon cq p_1(w) + \mathcal{O}(\varepsilon^2), \quad (3.38)$$

(cf. (3.22)), with

$$\rho_1(w) = (\Omega + 2b_+(w)\Theta)wb_1(w) - 1 = \left( \mathcal{C} + 2\Theta \sqrt{a + \frac{1}{4} - \frac{1}{w}} \right)wb_1(w) - 1. \quad (3.39)$$

Thus, for  $c \neq 0$  the flow on  $\mathcal{M}_\varepsilon^+$  is a perturbed integrable planar system with ‘nonlinear friction term’  $\varepsilon cq p_1(w)$ . In the case that there is only one critical point  $P^{+,s}$  of saddle type on  $\mathcal{M}_0^+$  – and thus on  $\mathcal{M}_\varepsilon^+$  – the impact of this term is asymptotically small. The situation is comparable to that of the flow on  $\mathcal{M}_\varepsilon^0$  w.r.t. the flow on  $\mathcal{M}_0^0$ . The stable and unstable manifolds of  $P^{+,s}$  restricted to the slow manifolds remain close:  $W^{u,s}(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  is  $\mathcal{O}(\varepsilon)$  close to  $W^{u,s}(P^{+,s})|_{\mathcal{M}_0^+}$  (for  $\mathcal{O}(1)$  values of  $(w, q)$ ) and the span  $W^{u,s}(P^{+,s}) \cup W^{u,s}(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  has become slightly asymmetric – cf. (3.29). This is drastically different in the case that there are 2 critical points  $P^{+,c}$  – the center – and  $P^{+,s}$  – the saddle – on  $\mathcal{M}_\varepsilon^+$ . We deduce

by classical dynamical system techniques – such as the Melnikov method (see for instance [72]) – the following (bifurcation) properties of (3.38), and thus of (3.10).

**Theorem 2.4.** *Let parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  of (3.7) be such that there is a center  $P^{+,c} = (b_+(w^{+,c}), 0, w^{+,c}, 0)$  and a saddle  $P^{+,s} = (b_+(w^{+,s}), 0, w^{+,s}, 0)$  on  $\mathcal{M}_\varepsilon^+$  and assume that the unperturbed homoclinic orbit  $(w_{\text{hom},0}(X), q_{\text{hom},0}(X))$  to  $P^{+,s}$  of (3.19) on  $\mathcal{M}_0^+$  lies entirely in the  $w$ -region in which both  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  are normally hyperbolic. More explicitly, assume that,*

$$\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0, \mathcal{C} < 0, \mathcal{D} > 0 \quad \text{and} \quad \frac{4}{1+4a} < \underline{w}_{h,0} < w^{+,c} < w^{+,s} < \frac{1}{a} \quad (3.40)$$

(3.23), (3.24), (3.13), where  $(\underline{w}_{h,0}, 0)$  is the intersection of  $(w_{\text{hom},0}(X), q_{\text{hom},0}(X))$  with the  $w$ -axis – see Fig. 3.4. Then, for all  $c \neq 0$  (but  $\mathcal{O}(1)$  w.r.t.  $\varepsilon$ ) and  $\varepsilon$  sufficiently small,

- there is a co-dimension 1 manifold  $\mathcal{R}_{\text{Hopf}} = \mathcal{R}_{\text{Hopf}}(a, \Psi, \Phi, \Omega, \Theta)$  such that a periodic solution (dis)appears in (3.38) – and thus in (3.7) – for parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  that cross through  $\mathcal{R}_{\text{Hopf}}$ ; moreover,  $\mathcal{R}_{\text{Hopf}}$  is at leading order (in  $\varepsilon$ ) determined by  $\rho_1(w^{+,c}) = 0$  (3.39);
- there is a co-dimension 1 manifold  $\mathcal{R}_{\text{hom}} = \mathcal{R}_{\text{hom}}(a, \Psi, \Phi, \Omega, \Theta)$  such that for  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{R}_{\text{hom}}$ , the unperturbed homoclinic solution  $(w_{\text{hom},0}(X), q_{\text{hom},0}(X))$  on  $\mathcal{M}_0^+$  persists as homoclinic solution to  $P^{+,s}$  of (3.38)/(3.7); moreover,  $\mathcal{R}_{\text{hom}}$  is at leading order determined by,

$$\Delta H_{\text{hom}} = c \int_{\underline{w}_{h,0}}^{w^{+,s}} \rho_1(w) \sqrt{2\mathcal{H}_0^{+,s} - 2Aw + (\mathcal{B} + a\Theta)w^2 + 2\mathcal{C}\mathcal{J}_0^+(w)} dw = 0. \quad (3.41)$$

with  $\mathcal{H}_0^{+,s}$ ,  $\mathcal{J}_0^+(w)$  as defined in (3.29), (3.27).

- there is an open region  $\mathcal{S}_{\text{per}}$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space – with  $\mathcal{R}_{\text{Hopf}} \cup \mathcal{R}_{\text{hom}} \subset \partial\mathcal{S}_{\text{per}}$  – such that for all  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{\text{per}}$ , one of the (restricted) periodic solutions  $(w_{p,0}(X), q_{p,0}(X))$  of the integrable flow (3.22) on  $\mathcal{M}_0^+$  persists as a periodic solution  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  of (3.38)/(3.7) on  $\mathcal{M}_\varepsilon^+$ ; the stability of the periodic orbit on  $\mathcal{M}_\varepsilon^+$  is determined by (the sign of)  $c$ .

The flow on  $\mathcal{M}_\varepsilon^+$  is reversible for  $c = 0$ : there always is a one-parameter family of periodic solutions on  $\mathcal{M}_\varepsilon^+$  enclosed by a homoclinic loop if (3.40) holds, i.e., the phase portrait remains as in the  $\varepsilon = 0$  case of Fig. 3.4, it is not necessary to restrict parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  to  $\mathcal{S}_{\text{per}}$  or to  $\mathcal{R}_{\text{hom}}$ , for  $c = 0$ .

**Proof.** A periodic solution  $(w_{p,0}(X), q_{p,0}(X))$  of the unperturbed flow (3.22) on  $\mathcal{M}_0^+$  is described by the value  $\mathcal{H}_{p,0}^+$  of the Hamiltonian  $\mathcal{H}_0^+(w, q)$  (3.26), where necessarily  $\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$  – with  $\mathcal{H}_0^{+,c} < \mathcal{H}_0^{+,s}$  the values of  $\mathcal{H}_0^+(w, q)$  at the center  $P_0^{+,c}$ , resp. saddle  $P_0^{+,s}$  (cf. (3.29)). We define  $L_{p,0} = L_{p,0}(\mathcal{H}_{p,0}^+)$  as the period – or wavelength – of  $(w_{p,0}(X), q_{p,0}(X))$  and  $\underline{w}_{p,0} = \underline{w}_{p,0}(\mathcal{H}_{p,0}^+)$  and  $\bar{w}_{p,0} = \bar{w}_{p,0}(\mathcal{H}_{p,0}^+)$  as the minimal and maximal values of  $w_{p,0}(X)$ , i.e.  $\underline{w}_{p,0} \leq$

$w_{p,0}(X) \leq \bar{w}_{p,0}$  – see Fig. 3.4.

Hamiltonian  $\mathcal{H}_0^+(w, q)$  (3.26) becomes a slowly varying function in the perturbed system (3.38),

$$\frac{d\mathcal{H}_0^+}{dX}(w, q) = \varepsilon c q^2 \rho_1(w) + \mathcal{O}(\varepsilon^2).$$

Thus, unperturbed periodic solution  $(w_{p,0}(X), q_{p,0}(X))$  on  $\mathcal{M}_0^+$  persists as periodic solution  $(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  of (3.38) on  $\mathcal{M}_\varepsilon^+$  – with  $|L_{p,\varepsilon} - L_{p,0}|, |\bar{w}_{p,\varepsilon} - \bar{w}_{p,0}| = \mathcal{O}(\varepsilon)$  and, by definition,  $\underline{w}_{p,\varepsilon} = \underline{w}_{p,0}$  – if,

$$\int_0^{L_{p,\varepsilon}} \frac{d\mathcal{H}_0^+}{dX}(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) dX = \varepsilon c \int_0^{L_{p,\varepsilon}} (q_{p,\varepsilon}(X))^2 \rho_1(w_{p,\varepsilon}(X)) dX + \mathcal{O}(\varepsilon^2) = 0. \quad (3.42)$$

The approximation of  $(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  by  $(w_{p,0}(X), q_{p,0}(X))$  yields, together with (3.26),

$$\begin{aligned} & \int_0^{L_{p,\varepsilon}} (q_{p,\varepsilon}(X))^2 \rho_1(w_{p,\varepsilon}(X)) dX \\ &= \int_0^{L_{p,0}} q_{p,0}(X)^2 \rho_1(w_{p,0}(X)) dX + \mathcal{O}(\varepsilon) \\ &= 2 \int_{\underline{w}_{p,0}}^{\bar{w}_{p,0}} \rho_1(w) \sqrt{2\mathcal{H}_{p,0}^+ - 2Aw + (B + a\Theta)w^2 + 2C\mathcal{J}_0^+(w)} dw + \mathcal{O}(\varepsilon). \end{aligned} \quad (3.43)$$

Thus, unperturbed periodic solution/pattern  $(w_{p,0}(X), q_{p,0}(X))$  persists as periodic solution on  $\mathcal{M}_\varepsilon^+$  for parameter combinations such that,

$$\begin{aligned} \Delta H(\mathcal{H}_{p,0}^+) &= c \int_{\underline{w}_{p,0}(\mathcal{H}_{p,0}^+)}^{\bar{w}_{p,0}(\mathcal{H}_{p,0}^+)} \rho_1(w) \sqrt{2\mathcal{H}_{p,0}^+ - 2Aw + (B + a\Theta)w^2 + 2C\mathcal{J}_0^+(w)} dw \\ &= 0. \end{aligned} \quad (3.44)$$

Note that this expression does not depend on the speed  $c$  – see however Remark 2.8 – but that (the sign of)  $c$  indeed determines the stability of  $(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  on  $\mathcal{M}_\varepsilon^+$ . For given  $\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$ , condition (3.44) determines a codimension 1 manifold  $\mathcal{R}_{\text{per}}(\mathcal{H}_{p,0}^+)$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space for which a periodic orbit  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  on  $\mathcal{M}_\varepsilon^+$  exists. Clearly

$$\mathcal{S}_{\text{per}} \subset \cup_{\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})} \mathcal{R}_{\text{per}}(\mathcal{H}_{p,0}^+).$$

Moreover,  $\underline{w}_{p,0}(\mathcal{H}_{p,0}^+) \uparrow w^{+,c}$  and  $\bar{w}_{p,0}(\mathcal{H}_{p,0}^+) \downarrow w^{+,c}$  as  $\mathcal{H}_{p,0}^+ \downarrow \mathcal{H}_0^{+,c}$ , so that (3.44) indeed reduces to  $\rho_1(w^{+,c}) = 0$  as  $\mathcal{H}_{p,0}^+ \downarrow \mathcal{H}_0^{+,c}$ :  $\mathcal{R}_{\text{Hopf}} = \mathcal{R}_{\text{per}}(\mathcal{H}_0^{+,c})$ . Note that  $\rho_1(w) \rightarrow -\infty$  as  $w \downarrow 4/(1 + 4a)$  – recall that  $\mathcal{C} < 0$  – and that

$$\rho_1\left(\frac{1}{a}\right) = a^2(\Omega + 2\Theta) - 1 = -(1 - a^2\mathcal{C}) + a^2\Theta$$

can be made positive by choosing  $\Theta$  sufficiently large:  $\rho_1(w)$  must change sign for  $\Theta$  not too small (in fact, it can be shown by straightforward analysis of (3.39) that  $\rho_1(w)$  may change sign twice (at most)). It thus follows that  $\mathcal{R}_{\text{Hopf}} \neq \emptyset$  and consequentially that  $\mathcal{S}_{\text{per}}$  is nonempty. Since  $\underline{w}_{p,0}(\mathcal{H}_{p,0}^+) \downarrow \underline{w}_{h,0}$  and  $\overline{w}_{p,0}(\mathcal{H}_{p,0}^+) \uparrow w^{+,s}$  as  $\mathcal{H}_{p,0}^+ \uparrow \mathcal{H}_0^{+,s}$ , it follows that  $\Delta H(\mathcal{H}_{p,0}^+) \rightarrow \Delta H_{\text{hom}}$  and thus that  $\mathcal{R}_{\text{hom}} = \mathcal{R}_{\text{per}}(\mathcal{H}_0^{+,s})$ , which also can be shown to be non-empty – see Lemma 2.6.  $\square$

Of course, Theorem 2.4 has a direct interpretation in terms of traveling waves in the full PDE (3.5),

**Corollary 2.5.** *Let the conditions formulated in Theorem 2.4 hold, then for all  $c \in \mathbb{R}$   $\mathcal{O}(1)$  w.r.t.  $\varepsilon$ ,*

- *there is a traveling spatially periodic wave (train) solution  $(B_{p,\varepsilon}(\varepsilon(x-ct)), W_{p,\varepsilon}(\varepsilon(x-ct)))$  of (3.5) for  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{\text{per}}$ ;*
- *there is a traveling pulse  $(B_{\text{hom},\varepsilon}(\varepsilon(x-ct)), W_{\text{hom},\varepsilon}(\varepsilon(x-ct)))$  in (3.5) – homoclinic to the background state  $(\bar{B}^{+,s}, \bar{W}^{+,s}) = (b_+(w^{+,s}), w^{+,s})$  – for  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{R}_{\text{hom}}$ .*

It is possible to (locally) get full analytical control over the set  $\mathcal{S}_{\text{per}}$  and its boundary manifolds  $\mathcal{R}_{\text{Hopf}}$  and  $\mathcal{R}_{\text{hom}}$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space by considering the unfolding of the saddle-node bifurcation on  $\mathcal{M}_\varepsilon^+$  as in section 2.2 (cf. (3.34)).

**Lemma 2.6.** *Let the conditions formulated in Theorem 2.4 hold, introduce  $\sigma > 0$  as in (3.30) and consider  $\sigma$  sufficiently small (but still  $\mathcal{O}(1)$  w.r.t.  $\varepsilon$ ). Then, system (3.38)/(3.7) has a periodic solution*

*$(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  on  $\mathcal{M}_\varepsilon^+$  for all  $(a, \Psi, \Phi, \Omega, \Theta)$  such that,*

$$\frac{5}{7}\sigma w_1^{SN} \rho'_1(w_0^{SN}) + \mathcal{O}(\sigma^2) < \rho_1(w_0^{SN}) < \sigma w_1^{SN} \rho'_1(w_0^{SN}) + \mathcal{O}(\sigma^2), \quad (3.45)$$

where  $\rho_1(w)$ ,  $\sigma$ ,  $w_0^{SN}$  and  $w_1^{SN}$  are explicitly given in terms of the parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  in (3.39), (3.30), (3.34) (with (3.23), (3.24)):  $\mathcal{S}_{\text{per}}$  is given by (3.45) and its boundaries  $\mathcal{R}_{\text{Hopf}}$  and  $\mathcal{R}_{\text{hom}}$  by the upper, respectively lower, boundary of (3.45).

**Proof.** For  $\mathcal{D}$   $\mathcal{O}(\sigma^2)$  close to  $\mathcal{D}^{SN}$  (3.30), the unperturbed flow (3.22) on  $\mathcal{M}_0^+$  can be given locally, i.e., in an  $\mathcal{O}(\sigma)$  neighborhood of the critical points  $P^{+,c} = (b_+(w_0^{+,c}), 0, w_0^{+,c}, 0)$  and  $P^{+,s} = (b_+(w_0^{+,s}), 0, w_0^{+,s}, 0)$  with  $w_0^{+,c} = w_0^{+,1} < w_0^{+,2} = w_0^{+,s}$  (3.34), be given by its quadratic approximation,

$$w_{XX} = \tilde{\alpha}(w - w_0^{+,c})(w - w_0^{+,s}) + \mathcal{O}(\sigma^3) = \tilde{\alpha} \left( (w - w_0^{SN})^2 - \sigma^2 (w_1^{SN})^2 \right) + \mathcal{O}(\sigma^3), \quad (3.46)$$

(3.34), where  $\tilde{\alpha} > 0$  is the second derivative of the right-hand side of (3.22) evaluated at  $w_0^{SN}$ . Thus, the integral  $\mathcal{H}_0^+$  (3.26) can locally be given by,

$$\mathcal{H}_0^+(w, q) = \frac{1}{2}q^2 - \tilde{\alpha} \left( \frac{1}{3}(w - w_0^{SN})^3 - \sigma^2 (w_1^{SN})^2 w \right) + \mathcal{O}(\sigma^4). \quad (3.47)$$

Direct evaluation yields that the stable/unstable manifolds of  $P^{+,s}$  (restricted to  $\mathcal{M}_0^+$ ) are given by,

$$\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s} = \tilde{\alpha}\sigma^2(w_1^{SN})^2 \left( w_0^{SN} + \frac{2}{3}\sigma w_1^{SN} \right) + \mathcal{O}(\sigma^4) \quad (3.48)$$

(cf. (3.29)), which implies that the (second) intersection with the  $w$ -axis of the homoclinic orbit connected to  $P^{+,s}$  (in  $\mathcal{M}_0^+$ ) is given by,

$$\underline{w}_{h,0} = w_0^{SN} - 2\sigma w_1^{SN} + \mathcal{O}(\sigma^2) \left( < w_0^{+,c} = w_0^{SN} - \sigma w_1^{SN} + \mathcal{O}(\sigma^2) \right) \quad (3.49)$$

(cf. Theorem 2.4)). Now, we consider parameter combinations such that  $\rho_1(w)$  has a zero  $\mathcal{O}(\sigma)$  close to  $w_0^{SN}$ , i.e., we set

$$\rho_1(w) = \tilde{\beta} \left( w - (w_0^{SN} + \sigma\mu) \right) + \mathcal{O}(\sigma^2),$$

where  $\sigma\mu$  represents the position of the zero and  $\tilde{\beta} = \rho_1'(w_0^{SN})$ . Hence, the condition  $\Delta H_{\text{hom}} = 0$  (3.41) – that determines the manifold  $\mathcal{R}_{\text{hom}}$  – is at leading order (in  $\sigma$ ) given by,

$$c\tilde{\beta} \int_{\underline{w}_{h,0}}^{w^{+,s}} \left( w - (w_0^{SN} + \sigma\mu) \right) \sqrt{2\mathcal{H}_0^{+,s} + 2\tilde{\alpha} \left( \frac{1}{3}(w - w_0^{SN})^3 - \sigma^2(w_1^{SN})^2 w \right)} dw = 0 \quad (3.50)$$

(3.48). Introducing  $\omega$  by  $w = w_0^{SN} + \sigma\omega$  and using (3.34), (3.49), we reduce (3.50) to,

$$\tilde{\beta}\sigma^3 \sqrt{\frac{2}{3}\tilde{\alpha}\sigma} \sqrt{\sigma} \int_{-2w_1^{SN}}^{w_1^{SN}} (\omega - \mu) \sqrt{\omega^3 - 3(w_1^{SN})^2\omega + 2(w_1^{SN})^3} d\omega = 0.$$

Thus, the homoclinic orbit to  $P^{+,s}$  (in  $\mathcal{M}_0^+$ ) persists for  $\mu$  such that,

$$\int_{-2w_1^{SN}}^{w_1^{SN}} (\omega - \mu)(\omega - w_1^{SN}) \sqrt{\omega + 2w_1^{SN}} d\omega = 0$$

(at leading order in  $\sigma$  (and in  $\varepsilon$ )). Straightforward integration yields that  $\mu = \mu_{\text{hom}} = -\frac{5}{7}\omega_1^{SN} + \mathcal{O}(\sigma)$ , i.e., that on  $\mathcal{R}_{\text{hom}}$ , the zero of  $\rho_1(w)$  must be at  $w_0^{SN} - \frac{5}{7}\sigma w_1^{SN} + \mathcal{O}(\sigma^2) > w_0^{+,c} = w_0^{SN} - \sigma w_1^{SN} + \mathcal{O}(\sigma^2)$ .

We conclude that for  $\sigma$  (and  $\varepsilon$ ) sufficiently small, the boundaries  $\mathcal{R}_{\text{Hopf}}$  and  $\mathcal{R}_{\text{hom}}$  of the domain  $\mathcal{S}_{\text{per}}$  are given by  $\rho_1(w_0^{SN} - \sigma w_1^{SN} + \mathcal{O}(\sigma^2)) = 0$  (first bullet of Theorem 2.4), respectively  $\rho_1(w_0^{SN} - \frac{5}{7}\sigma w_1^{SN} + \mathcal{O}(\sigma^2)) = 0$  – which is equivalent to the boundaries of (3.45) by Taylor expansion (in  $\sigma$ ). Finally, we

notice that for parameter values between  $\mathcal{R}_{\text{Hopf}}$  and  $\mathcal{R}_{\text{hom}}$ , i.e., for which (3.45) holds, one of the periodic orbits between the center point and the homoclinic loop must persist – in other words, for parameter combinations that satisfy (3.45),  $\Delta H(\mathcal{H}_{p,0}^+) = 0$  (3.44) for certain  $\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$ .  $\square$

**Remark 2.7.** *Lemma 2.6 ‘rediscovers’ the periodic solutions associated to a Bogdanov-Takens bifurcation. In 3.C we present a brief embedding of our result into the normal form approach to the Bogdanov-Takens bifurcation scenario.*

**Remark 2.8.** *A higher order perturbation analysis yields that the  $\mathcal{O}(\varepsilon)$  corrections to  $\mathcal{R}_{\text{Hopf}}$  and  $\mathcal{R}_{\text{hom}}$  – and thus to  $\mathcal{S}_{\text{per}}$  – explicitly depend on  $c$ .*

**Remark 2.9.** *Of course one can also establish the persistence of periodic orbits of the slow reduced flow – as in Theorem 2.4 – under the assumption that there is only one critical point  $P^{+,c}$  of center type on  $\mathcal{M}_0^+$ , instead of focusing on the present case in which the reduced slow flow (3.19) has a homoclinic orbit on  $\mathcal{M}_0^+$  (Theorem 2.4). Since we decided to focus on situations in which there is a saddle point on  $\mathcal{M}_0^+$  – that is potentially stable as homogeneous background state in (3.5) (section 2.3) – we do not consider this possibility here. Note however that the analysis of this case is essentially the same as presented here. See also Remark 3.7.*

### 3 Localized front patterns

In this section we use the slow-fast geometry of the phase space associated to (3.7) to establish a remarkably rich variety of localized vegetation patterns (potentially) exhibited by model (3.5). First, we consider various kinds of traveling and stationary ‘invasion fronts’ that connect the bare soil state to a uniform or an ‘oscillating’ vegetation state and their associated bifurcation structures (sections 3.2, 3.3 and 3.4), next we study stationary homoclinic 2-front spot and gap patterns (section 3.5) and finally spatially periodic multi-front (spot/gap) patterns (section 3.6). As starting point, we need to control the intersection of  $W^u(P^0)$  and  $W^s(\mathcal{M}_\varepsilon^+)$ .

**Remark 3.1.** *We start by considering localized patterns that correspond to orbits in  $W^u(P^0)$ , i.e., patterns that approach the bare soil state  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  of (3.5) as  $x \rightarrow -\infty$ . In fact, the upcoming results on 1-fronts are all on orbits in (3.7) that connect  $P^0 \in \mathcal{M}_\varepsilon^0$  either to a critical point or to a persisting periodic orbit in  $\mathcal{M}_\varepsilon^+$  (Theorem 2.4): all constructed 1-fronts originate from the uniform bare soil state. The existence of 1-front patterns that approach  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  as  $t \rightarrow +\infty$  is embedded in these results through the application of the symmetry (3.9).*

### 3.1 $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$ and its touch down points on $\mathcal{M}_\varepsilon^+$

A (traveling) front pattern between the bare soil state  $(0, \Psi/\Phi)$  and a (potentially stable) uniform vegetation state  $(\bar{B}, \bar{W})$  of (3.5) corresponds to a heteroclinic solution  $\gamma_h(\xi) = (w_h(\xi), p_h(\xi), b_h(\xi), q_h(\xi))$  of (3.7) between the critical points  $P^0 = P_0^0 = (0, 0, \Psi/\Phi, 0) \in \mathcal{M}_\varepsilon^0$  and  $P^{+,s} = (b_+(w^{+,s}), 0, w^{+,s}, 0) \in \mathcal{M}_\varepsilon^+$  – see section 2.3. We know by Fenichel’s second Theorem that, by the normal hyperbolicity  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$ , their stable and unstable manifolds  $W^{s,u}(\mathcal{M}_0^0)$  and  $W^{s,u}(\mathcal{M}_0^+)$  persist as  $W^{s,u}(\mathcal{M}_\varepsilon^0)$  and  $W^{s,u}(\mathcal{M}_\varepsilon^+)$  for  $\varepsilon \neq 0$  as  $w \in (1/(a + 1/4), 1/a)$  (3.13), [83, 86]. Thus,  $\gamma_h(\xi) \subset W^u(P^0) \cap W^s(P^{+,s}) \subset W^u(\mathcal{M}_\varepsilon^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  – where we note that the manifolds  $W^u(P^0)$  and  $W^s(P^{+,s})$  are 2-dimensional, while  $W^u(\mathcal{M}_\varepsilon^0)$  and  $W^s(\mathcal{M}_\varepsilon^+)$  are 3-dimensional (and that the intersections take place in a 4-dimensional space).

We know by (3.17) that  $W^u(\mathcal{M}_0^0)$  and  $W^s(\mathcal{M}_0^+)$  intersect transversely – and thus that  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  is 2-dimensional. Since  $W^u(\mathcal{M}_\varepsilon^0)$  and  $W^s(\mathcal{M}_\varepsilon^+)$  are  $C^1$ - $\mathcal{O}(\varepsilon)$  close to  $W^u(\mathcal{M}_0^0)$  and  $W^s(\mathcal{M}_0^+)$ , it immediately follows that  $W^u(\mathcal{M}_\varepsilon^0)$  and  $W^s(\mathcal{M}_\varepsilon^+)$  also intersect transversely, that  $W^u(\mathcal{M}_\varepsilon^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  is 2-dimensional and at leading order (in  $\varepsilon$ ) given by (3.17). Since  $W^u(P^0) \subset W^u(\mathcal{M}_\varepsilon^0)$ ,  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  is a 1-dimensional subset of  $W^u(\mathcal{M}_\varepsilon^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  – i.e., an orbit – that follows  $W^u(P^0)|_{\mathcal{M}_\varepsilon^0} = \ell_\varepsilon^u$  (3.21) exponentially close until its  $w$ -component reaches  $w_h^+(c)$  (3.16) at which it ‘takes off’ from  $\mathcal{M}_\varepsilon^0$  to follow the fast flow along the ‘parabolic’ manifold given by (3.17), all at leading order in  $\varepsilon$  – see sections 2.1, 2.4. Since  $w, q$  only vary slowly (3.7), the  $(w, q)$ -components of the orbit  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  remain constant at leading order during its fast jump: it ‘touches down’ on  $\mathcal{M}_\varepsilon^+$  with (at leading order) the same  $(w, q)$ -coordinates (Remark 3.3). Therefore, we define the touch down curve  $\mathcal{T}_{\text{down}}(c) \subset \mathcal{M}_\varepsilon^+$  as the set of touch down points of the orbits  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  that take off from  $\mathcal{M}_\varepsilon^0$  exponentially close to the intersection  $\ell_\varepsilon^u \cap \{w = w_h^+(c)\}$  (3.16), parameterized by  $c$ ; it is at leading order (in  $\varepsilon$ ) given by,

$$\mathcal{T}_{\text{down}}(c) = \left\{ \left( b_+(w_h^+(c)), 0, w_h^+(c), \sqrt{\Phi} \left( w_h^+(c) - \frac{\Psi}{\Phi} \right) \right) \right\} \quad (3.51)$$

In terms of the projected  $(w, q)$ -coordinates by which the dynamics on  $\mathcal{M}_\varepsilon^+$  are described (3.38),  $\mathcal{T}_{\text{down}}(c)$  describes a smooth 1-dimensional manifold  $\mathcal{I}_{\text{down}} = \{(w_{\text{down}}(c), q_{\text{down}}(c))\}$  parameterized by  $c$  with boundaries (its endpoints): the family of base points of the Fenichel fibers of  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  on  $\mathcal{M}_\varepsilon^+$  – Remark 3.3; at leading order in  $\varepsilon$ ,  $\mathcal{I}_{\text{down}}$  is a straight interval with endpoints determined by the bounds (3.13) on  $w = w_h^+(c)$ .

**Lemma 3.2.** *At leading order in  $\varepsilon$ ,*

$$\mathcal{I}_{\text{down}} = \left\{ \left( w_h^+(c), \sqrt{\Phi} \left( w_h^+(c) - \frac{\Psi}{\Phi} \right) \right), c \in \left[ -\frac{1}{\sqrt{2(1+4a)}}, \frac{1}{\sqrt{2a}} \right] \right\}.$$

The map  $[-\frac{1}{\sqrt{2(1+4a)}}, \frac{1}{\sqrt{2a}}] \rightarrow \mathcal{I}_{\text{down}}$  is bijective and

$$w_h^+ \left( -\frac{1}{\sqrt{2(1+4a)}} \right) = \frac{4}{1+4a}, \quad w_h^+(0) = \frac{9}{2+9a}, \quad w_h^+ \left( \frac{1}{\sqrt{2a}} \right) = \frac{1}{a}.$$

Expression (3.16) a priori does not exclude the possibility that  $w_h^+$  has several extremums as function of  $c$ , in fact  $\frac{d}{dc} w_h^+ \left( -\frac{1}{\sqrt{2(1+4a)}} \right) = 0$ . The proof – derivation – of this lemma thus requires some careful, but straightforward, analysis. We refrain from going into the details here.

We conclude this section by noticing that heteroclinic connections  $\gamma_h(\xi)$  between  $P^0 \in \mathcal{M}_\varepsilon^0$  and  $P^{+,s} \in \mathcal{M}_\varepsilon^+$  directly correspond to intersections  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  (Remark 3.3). However, the coordinates of this intersection determine  $c$  (through  $\mathcal{I}_{\text{down}}$ ), while  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  also varies as function of  $c$ . Moreover, by the perturbed integrable nature of the flow on  $\mathcal{M}_\varepsilon^+$  (3.38), there can a priori be (countably) many intersections  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$ . Thus, the analysis is more subtle and richer than (perhaps) expected – as we shall see in the upcoming sections.

**Remark 3.3.** *We (for instance) refer to [41] for a more careful treatment of ‘take off’ and ‘touch down’ points/manifolds. In fact, these points/manifolds correspond to base points of Fenichel fibers (that persist under perturbation by Fenichel’s third Theorem [83, 86]). By construction/definition, an orbit that touches down at a certain (touch down) point on a slow manifold is asymptotic to the orbit of the slow flow that has this point as initial condition. Therefore, if an orbit touches down on a stable manifold of a critical point on the slow manifold, it necessarily is asymptotic to this critical point.*

### 3.2 Traveling 1-front patterns – primary orbits

Our first result – on the existence of *primary* heteroclinic orbits – can be described in terms of the slow reduced flow on  $\mathcal{M}_0^+$ , or more precise, on intersections of the touch down manifold  $\mathcal{I}_{\text{down}}$  and the restricted stable manifold  $W^s(P^{+,s})|_{\mathcal{M}_0^+} \subset \{\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s}\}$  (3.29) of the reduced slow flow (3.22) on  $\mathcal{M}_0^+$ . However, it is a priori unclear whether such intersections may exist and how many of such intersections may occur: the many parameters of system (3.7) have a ‘nontrivial’ effect on  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  and thus on their relative positions. To obtain a better insight in this, we ‘freeze’ the flow of (3.22) by fixing  $a, \mathcal{A}, \mathcal{C}, \mathcal{D}$  at certain values. Since  $\mathcal{B} + a\Theta = \mathcal{D} + (a + \frac{1}{4})\mathcal{A}$  (3.24), this indeed fixes all coefficients of the reduced slow flow (3.22) on  $\mathcal{M}_0^+$ . At the same time, this leaves a 1-parameter freedom in the parameters  $\Phi, \Psi, \Omega, \Theta$ . Defining,

$$\chi = \frac{1}{a} \left( \frac{1}{4}\mathcal{A} - \frac{1}{2}\mathcal{C} + \mathcal{D} \right), \quad (3.52)$$

we see that for all  $\Phi$ , the choices

$$\Psi = \frac{1}{a}\Phi - \chi, \quad \Theta = \mathcal{A} - \frac{1}{a}\Phi + \chi, \quad \Omega = \mathcal{C} - \mathcal{A} + \frac{1}{a}\Phi - \chi \quad (3.53)$$

yield identical slow reduced flows (3.22). On the other hand, the (leading order) interval  $\mathcal{I}_{\text{down}}$  clearly varies as function of  $\Phi$ ,

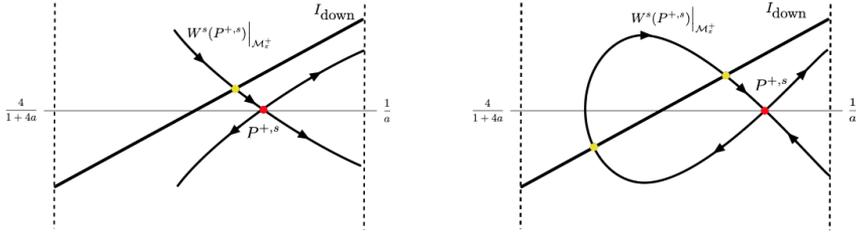
$$\mathcal{I}_{\text{down}}(\Phi) = \left\{ q = \sqrt{\Phi} \left( w - \left( \frac{1}{a} - \frac{\chi}{\Phi} \right) \right), \quad w \in \left( \frac{4}{1+4a}, \frac{1}{a} \right) \right\}. \quad (3.54)$$

Note that for  $\chi > 0$ , the intersection of  $\mathcal{I}_{\text{down}}$  with the  $w$ -axis can be varied between the critical  $w$  values  $4/(1+4a)$  and  $1/a$  by increasing  $\Phi$  from  $a(1+4a)\chi$  to  $\infty$ . In fact,  $\chi > 0$  necessarily holds in case there are 2 critical points on  $\mathcal{M}_\varepsilon^+$  (since in that case  $\mathcal{A}, \mathcal{D} > 0, \mathcal{C} < 0$ ), while  $\chi$  can also be chosen to be positive in the case that there is only 1 critical point on  $\mathcal{M}_\varepsilon^+$ . Thus, by choosing  $\Psi, \Omega, \Theta$  as in (3.53) and varying  $\Phi$  we can control  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$ .

**Theorem 3.4.** *Let  $P^{+,s} = (b_+(w^{+,s}), 0, w^{+,s}, 0) \in \mathcal{M}_0^+$  be a critical point of (3.7) that is a saddle point for the slow reduced flow (3.22) on  $\mathcal{M}_0^+$ , and consider the touch down manifold  $\mathcal{I}_{\text{down}}$  at leading order given in Lemma 3.2 and the restricted stable manifold  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  of the reduced slow flow (3.22). If there is a non-degenerate intersection point  $(\bar{w}_{\text{prim},0}, \bar{q}_{\text{prim},0}) \in \mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$ , then, for  $\varepsilon$  sufficiently small, there exists for  $c = c_{\text{prim}}$  a primary heteroclinic orbit*

$\gamma_{\text{prim}}(\xi) = (w_{\text{prim}}(\xi), p_{\text{prim}}(\xi), b_{\text{prim}}(\xi), q_{\text{prim}}(\xi)) \subset W^u(P^0) \cap W^s(P^{+,s})$  of (3.7) connecting  $P^0 \in \mathcal{M}_\varepsilon^0$  to  $P^{+,s} \in \mathcal{M}_\varepsilon^+$  – where  $c_{\text{prim}} = c_{\text{prim},0} + \mathcal{O}(\varepsilon)$  and  $c_{\text{prim},0}$  is the unique solution of  $w_h^+(c) = \bar{w}_{\text{prim},0}$  (3.16). Departing from  $P^0$  (and at leading order in  $\varepsilon$ ),  $\gamma_{\text{prim}}(\xi)$  first follows  $\ell_0^u \subset \mathcal{M}_0^0$  (3.21) until it reaches the take off point  $(0, 0, \bar{w}_{\text{prim}}, \bar{q}_{\text{prim}})$  from which it jumps off from  $\mathcal{M}_0^0$  and follows the fast flow along  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  (3.17) to touch down on  $\mathcal{M}_0^+$  at  $(b_+(\bar{w}_{\text{prim}}), 0, \bar{w}_{\text{prim}}, \bar{q}_{\text{prim}}) \in W^s(P^{+,s})|_{\mathcal{M}_0^+}$ ; from there, it follows  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  towards  $P^{+,s}$ . Moreover;

- if  $P^{+,s}$  is the only critical point on  $\mathcal{M}_\varepsilon^+$ , i.e., if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0, \mathcal{D} < 0, \mathcal{E} > 0$  (3.28), there is an open region  $\mathcal{S}_{\text{s-prim}}^1$  in  $(a, \Psi, \Phi, \Omega, \Theta)$  parameter space for which  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  intersect transversely; however, there is at most one intersection  $(\bar{w}_{\text{prim}}, \bar{q}_{\text{prim}}) \in \mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$  and thus at most one primary heteroclinic orbit  $\gamma_{\text{prim}}(\xi)$ ; in fact, this is the only possible heteroclinic orbit between  $P^0$  and  $P^{+,s}$ ;
- if there are two critical points on  $\mathcal{M}_\varepsilon^+$ , the center  $P^{+,c}$  and saddle  $P^{+,s}$ , i.e., if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0, \mathcal{C} < 0, \mathcal{D} > 0$ , then there are open regions  $\mathcal{S}_{\text{cs-prim}}^1$ , respectively  $\mathcal{S}_{\text{cs-prim}}^2$ , in  $(a, \Psi, \Phi, \Omega, \Theta)$  parameter space for which  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  have 1, resp. 2, (transversal) intersections, so that there can be (up to) 2 distinct primary heteroclinic orbit  $\gamma_{\text{prim}}^j(\xi)$  that travel with different speeds, i.e.,  $c_{\text{prim}}^2 < c_{\text{prim}}^1$ .



**Figure 3.6:** Sketches of the intersections of  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  in  $\mathcal{M}_0^+$ , i.e., the leading order configurations as described by the (integrable) slow reduced flow (3.22), in the 2 cases considered in Theorem 3.4: there is one critical point  $P^{+,s}$  of saddle type on  $\mathcal{M}_0^+$  or there is a center  $P^{+,c}$  and a saddle  $P^{+,s}$  on  $\mathcal{M}_0^+$ .

A primary heteroclinic orbit  $\gamma_{\text{prim}}(\xi) = (w_{\text{prim}}(\xi), p_{\text{prim}}(\xi), b_{\text{prim}}(\xi), q_{\text{prim}}(\xi))$  corresponds to a (localized, traveling, invasion) 1-front pattern  $(B(x, t), W(x, t)) = (b_{\text{prim}}(x - c_{\text{prim}}t), w_{\text{prim}}(x - c_{\text{prim}}t))$  in PDE (3.5) that connects the bare soil state  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  to the uniform vegetation state  $(\bar{B}, \bar{W}) = (b_+(w^{+,s}), w^{+,s})$ .

In the case of 2 critical points on  $\mathcal{M}_\varepsilon^+$ , we shall see that the primary orbits may only be the first of many ‘higher order’ heteroclinic orbits – see section 3.3. We refer to Fig. 3.6 for sketches of the constructions in  $\mathcal{M}_\varepsilon^0$  that yield the primary heteroclinic orbits  $\gamma_{\text{prim}}(\xi)$  and to Figs. 3.1a, 3.13 and 3.14b for the associated – numerically obtained – primary 1-front patterns in (3.5)– see especially Fig. 3.13b in which the the slow-fast-slow structure of a (numerically obtained) heteroclinic front solutions of (3.5) is exhibited by its projection in the 3-dimensional  $(b, w, q)$ -subspace of the 4-dimensional phase space associated to (3.7).

**Proof.** The existence of the heteroclinic orbit  $\gamma_{\text{prim}}(\xi)$  follows by construction – Remark 3.3 – from an intersection of  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$ . Thus, we first need to show that a (non-degenerate) intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$  implies an intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$ . More precise, since  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  varies with  $c$ , i.e., since  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+} = W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c)$ , we need to determine  $c^*$  such that  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c^*)$  intersects  $\mathcal{I}_{\text{down}} = \{(\bar{w}_{\text{down}}(c), \bar{q}_{\text{down}}(c))\}$  exactly at  $(\bar{w}_{\text{down}}(c^*), \bar{q}_{\text{down}}(c^*))$ .

By the assumption that  $(\bar{w}_{\text{prim},0}, \bar{q}_{\text{prim},0}) \in \mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$  is a non-degenerate intersection point, we know that the intersection is transversal, and thus that  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(\tilde{c})$  – i.e.,  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  for (3.38) with  $c = \tilde{c}$  – also intersects  $\mathcal{I}_{\text{down}}$  transversally as  $\tilde{c}$  is varied around  $c_{\text{prim},0}$  in an  $\mathcal{O}(1)$  fashion. Thus, for  $\tilde{c}$  sufficiently (but  $\mathcal{O}(1)$ ) close to  $c_{\text{prim},0}$ ,  $\mathcal{I}_{\text{down}} \cap W^s(P^0)|_{\mathcal{M}_\varepsilon^+}(\tilde{c}) = (\bar{w}_{\text{down}}(c_i), \bar{q}_{\text{down}}(c_i))$  determines a curve  $c_i = c_i(\tilde{c})$  by  $c_i = \bar{w}_{\text{down}}(c_i)$ . Since the flows of (3.22) and (3.38) are  $\mathcal{O}(\varepsilon)$  close, we know that

$\|(\bar{w}_{\text{prim},0}, \bar{q}_{\text{prim},0}) - (\bar{w}_{\text{down}}(c_i), \bar{q}_{\text{down}}(c_i))\| = \mathcal{O}(\varepsilon)$ , which implies that  $c_i(\tilde{c}) = c_{\text{prim},0} + \mathcal{O}(\varepsilon)$ . Hence, the  $\mathcal{O}(1)$  variation of  $\tilde{c}$  through  $c_{\text{prim},0}$  yields at leading order (in  $\varepsilon$ ) a horizontal line  $c_i(\tilde{c}) \equiv c_{\text{prim},0}$ : there must be a unique intersection  $c_i(\tilde{c}^*) = \tilde{c}^*$ , and thus, by construction,  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(\tilde{c}^*) = (\bar{w}_{\text{down}}(\tilde{c}^*), \bar{q}_{\text{down}}(\tilde{c}^*))$ :  $\tilde{c}^* = c_{\text{prim}}$ .

If  $P^{+,s}$  is the only critical point on  $\mathcal{M}_\varepsilon^+$  – i.e., if  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0$ ,  $\mathcal{D} < 0$ ,  $\mathcal{E} > 0$  – we freeze the flow of (3.22) with  $\mathcal{A}, \mathcal{C}, \mathcal{D}$  such that  $\chi > 0$  (3.52) and define  $\Phi = \Phi^{+,s}$  such that  $\Psi(\Phi)/\Phi = 1/a - \chi/\Phi = w^{+,s}$ , the  $w$ -coordinate of the saddle  $P^{+,s}$  on  $\mathcal{M}_0^+$  – see (3.53), (3.54). Since  $q$  is an increasing function of  $w$  on  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  is decreasing near  $P^{+,s}$  – see Fig 3.6a – it follows that there must be a transversal intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  for values of  $\Phi$  in an (open) interval around  $\Phi^{+,s}$ . Transversality implies that the intersection persists under varying  $\mathcal{A}, \mathcal{C}, \mathcal{D}$  around their initially frozen values, which establishes the existence of the open region  $\mathcal{S}_{s\text{-prim}}^1$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space for which  $\mathcal{I}_{\text{down}}$  and  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  intersect. Moreover, the manifold  $W^s(P^{+,s})|_{\mathcal{M}_0^+} \subset \{\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s}\}$  (3.29) is given by a (strictly) decreasing function  $q^{+,s}|_{\mathcal{M}_0^+}(w)$  for all  $w \in (4/(1+4a), 1/a)$  since it cannot have extremums: zeroes of  $\frac{d}{dw}q^{+,s}|_{\mathcal{M}_0^+}(w)$  correspond to zeroes of  $\frac{\partial}{\partial w}\mathcal{H}_0^+(w, q)$  (3.26) and thus to critical points of (3.22). By assumption, there are no critical points besides  $P^{+,s}$ , which yields that there indeed can be maximally one intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$ .

To control the case with a center  $P^{+,c}$  and saddle  $P^{+,s}$  on  $\mathcal{M}_0^+$ , we again consider the unfolded saddle-node case of Lemma 2.6 and define  $\Phi = \Phi^{+,c}$  such that  $\Psi(\Phi)/\Phi = 1/a - \chi/\Phi = w^{+,c}$ , the  $w$ -coordinate of the center  $P^{+,c}$ . The level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s}\}$  forms a small (w.r.t. the unfolding parameter  $\sigma$ ) homoclinic loop around  $P^{+,c}$  that intersects  $\mathcal{I}_{\text{down}}$  (transversally) in two points  $(\bar{w}_{\text{prim}}^{j,0}, \bar{q}_{\text{prim}}^{j,0})$ ,  $j = 1, 2$  – see Fig. 3.6(b). By varying  $\Phi$  around  $\Phi = \Phi^{+,c}$  and  $\mathcal{A}, \mathcal{C}, \mathcal{D}$  around their initially frozen values, we find the open region  $\mathcal{S}_{\text{cs-prim}}^2$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space for which both elements of the intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  persist: for  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{\text{cs-prim}}^2$ , (3.7) has 2 (distinct) primary heteroclinic orbits  $\gamma_{\text{prim}}^j(\xi)$ ,  $j = 1, 2$ , that correspond to 1-front patterns traveling with speeds  $c_{\text{prim}}^1 \neq c_{\text{prim}}^2$  – where  $c_{\text{prim}}^{j,0}$  is the unique solution of  $w_h^+(c) = \bar{w}_{\text{prim}}^{j,0}$ . Finally, we note that the existence of the open set  $\mathcal{S}_{\text{cs-prim}}^1$  follows by considering  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  for values of  $\Phi > \Phi^{+,s}$  (as defined above).  $\square$

### 3.3 Traveling 1-front patterns by the perturbed integrable flow on $\mathcal{M}_\varepsilon^+$

As in Theorem 2.4, we assume throughout this section that there is a center  $P^{+,c} = (b_+(w^{+,c}), 0, w^{+,c}, 0)$  and a saddle  $P^{+,s} = (b_+(w^{+,s}), 0, w^{+,s}, 0)$  on  $\mathcal{M}_\varepsilon^+$  and – for simplicity – that the unperturbed homoclinic orbit  $(w_{\text{hom},0}(X), q_{\text{hom},0}(X))$  to  $P^{+,s}$  of (3.22) on  $\mathcal{M}_0^+$  – that is a subset of  $W^s(P^{+,s})|_{\mathcal{M}_0^+} \subset \{\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s}\}$  – lies entirely in the  $w$ -region in which both  $\mathcal{M}_0^0$  and  $\mathcal{M}_0^+$  are normally hyperbolic, i.e., we assume that (3.40) holds.

The homoclinic orbit  $(w_{\text{hom},0}(X), q_{\text{hom},0}(X))$  of (3.22) typically breaks open under the perturbed flow of (3.38), and  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  either spirals inwards in backwards ‘time’, i.e., as  $\xi \rightarrow -\infty$ , or not. In the former case, there will be (typically many) further intersections  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$  – see Fig. 3.7. Of course, this is determined by the sign of  $\Delta H_{\text{hom}}$  (3.41): if

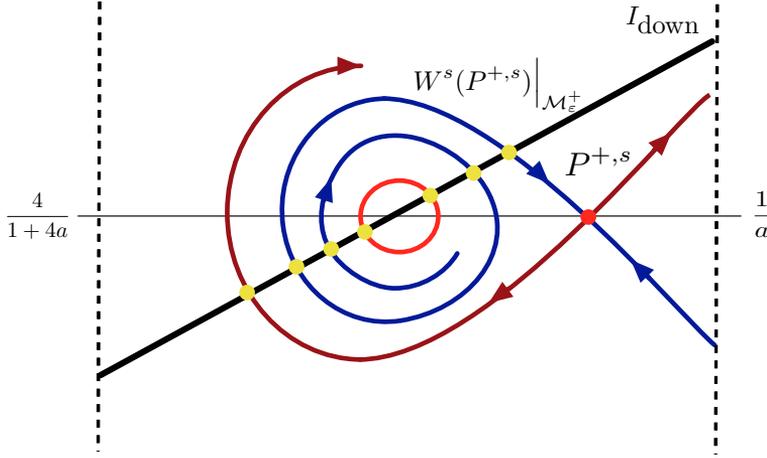
$$\Delta H_{\text{hom}} = c \int_{w_{h,0}}^{w^{+,s}} \rho_1(w) \sqrt{2\mathcal{H}_0^{+,s} - 2\mathcal{A}w + (\mathcal{B} + a\Theta)w^2 + 2\mathcal{C}\mathcal{J}_0^+(w)} dw > 0 \quad (3.55)$$

(at leading order in  $\varepsilon$ ), we may expect further heteroclinic connections  $\gamma_{h,j}$  in (3.7) connecting  $P^0 \in \mathcal{M}_\varepsilon^0$  to  $P^{+,s} \in \mathcal{M}_\varepsilon^+$  beyond the primary orbits  $\gamma_{\text{prim}}^1(\xi)$  established in Theorem 3.4. In fact, it follows directly that  $\gamma_{\text{prim}}^1(\xi)$  and  $\gamma_{\text{prim}}^2(\xi)$  are the only heteroclinic orbits between  $P^0$  and  $P^{+,s}$  if (3.55) does not hold. If (3.55) does hold, the (spiraling part of)  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  clearly must limit – for  $\xi \rightarrow -\infty$  – on either the center  $P^{+,c}$  or, if  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{\text{per}}$ , on the persistent periodic solution  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$  (Theorem 2.4). Therefore, we first formulate a result on the existence of heteroclinic connections between  $P^0 \in \mathcal{M}_\varepsilon^0$  and  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$ . Like in Theorem 3.4, this can be done in terms of the unperturbed flow in  $\mathcal{M}_0^+$ .

**Theorem 3.5.** *Assume that (3.40) holds and that  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{\text{per}}$ . Let  $(w_{p,0}(X), q_{p,0}(X)) \subset \{\mathcal{H}_0^+(w, q) = \mathcal{H}_{p,0}^+\}$  with  $\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$  (3.26) be the periodic solution of (3.22) that persists (on  $\mathcal{M}_\varepsilon^+$ ) as periodic solution  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  on  $\mathcal{M}_\varepsilon^+$  of (3.7). Then there is an open set  $\mathcal{S}_{\text{h-p}} \subset \mathcal{S}_{\text{per}} \cap \mathcal{S}_{\text{cs-prim}}^2$  – with  $\mathcal{S}_{\text{cs-prim}}^2$  defined in Theorem 3.4 – such that there are 2 (non-degenerate) intersection points  $(\bar{w}_{\text{h-p}}^j, \bar{q}_{\text{h-p}}^j) \in \mathcal{I}_{\text{down}} \cap \{\mathcal{H}_0^+(w, q) = \mathcal{H}_{p,0}^+\}$ ,  $j = 1, 2$ , that correspond – for  $\varepsilon$  sufficiently small – to 2 distinct heteroclinic orbits  $\gamma_{\text{h-p}}^j(\xi) = (b_{\text{h-p}}^j(\xi), p_{\text{h-p}}^j(\xi), w_{\text{h-p}}^j(\xi), q_{\text{h-p}}^j(\xi))$  of (3.7) – in which  $c = c_{\text{h-p}}^j$  – between the critical point  $P^0 \in \mathcal{M}_\varepsilon^0$  and the periodic orbit  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$ ; at leading order in  $\varepsilon$ ,  $c_{\text{h-p}}^j$  is determined by  $w_h^+(c) = \bar{w}_{\text{h-p}}^j$ , with  $c_{\text{prim}}^2 < c_{\text{h-p}}^2 < c_{\text{h-p}}^1 < c_{\text{prim}}^1$  (Theorem 3.4).*

*The orbits  $\gamma_{\text{h-p}}^j(\xi)$  correspond traveling 1-front patterns*

*$(B(x, t), W(x, t)) = (b_{\text{h-p}}^j(x - c_{\text{h-p}}^j t), w_{\text{h-p}}^j(x - c_{\text{h-p}}^j t))$  in PDE (3.5) that connect*



**Figure 3.7:** A sketch of the flow (3.38) on  $\mathcal{M}_\varepsilon^+$  for  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{h-p}$  (Theorem 3.5) and  $c = c_{\text{prim}}^1$  (Theorem 3.4) in the case that (3.55). Since  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  ‘wraps around’ the persistent periodic solution  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  (Theorem 3.5) – in backwards time – and since  $\mathcal{I}_{\text{down}}$  intersects this orbit in 2 points (by assumption), there are two countable sets of intersections  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$ .

the bare soil state  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  to the traveling wave train  $(B_{p,\varepsilon}(\varepsilon(x - c_{h-p}^j t)), W_{p,\varepsilon}(\varepsilon(x - c_{h-p}^j t)))$  of Corollary 2.5.

Notice that this result is independent of condition (3.55), i.e., Theorem 3.5 holds independent of the sign of  $\Delta H_{\text{hom}}$ . Moreover, we could formulate similar limiting result concerning heteroclinic 1-front connections between  $P^0 \in \mathcal{M}_\varepsilon^0$  and  $P^{+,c} \in \mathcal{M}_\varepsilon^+$  for  $(a, \Psi, \Phi, \Omega, \Theta)$  on a certain co-dimension 1 manifold. Since the background state associated to  $P^{+,c}$  cannot be stable – section 2.3 – we refrain from going into the details.

**Proof.** The proof goes exactly along the lines of that of Theorem 3.4. □

Theorem 3.5 provides the foundation for a result on the existence of multiple – in fact countably many – distinct traveling 1-front connections between  $P^0 \in \mathcal{M}_\varepsilon^0$  and  $P^{+,s} \in \mathcal{M}_\varepsilon^+$  for an open set in parameter space – see also the sketches in Figs. 3.2a and 3.2b.

**Theorem 3.6.** *Assume that the conditions of Theorem 3.5 hold and let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{h-p}$ . If  $c_{h-p}^j$  and  $c_{\text{prim}}^j$  have the same sign (for either  $j = 1$  or  $2$ ) and if (3.55) holds for  $c$  of this sign, then – for  $\varepsilon$  sufficiently small – there are countably many distinct heteroclinic orbits  $\gamma_h^{j,k}(\xi) = (b_h^{j,k}(\xi), p_h^{j,k}(\xi), w_h^{j,k}(\xi), q_h^{j,k}(\xi))$ ,  $k \geq 0$ , of (3.7) with  $c = c_h^{j,k}$  connecting  $P^0 \in \mathcal{M}_\varepsilon^0$  to  $P^{+,s} \in \mathcal{M}_\varepsilon^+$ . Moreover,*

$\gamma_h^{j,0}(\xi) = \gamma_{\text{prim}}^j(\xi)$ ,  $|c_h^{1,k+1} - c_h^{1,k}| = \mathcal{O}(\varepsilon)$ , and,

$$\begin{aligned} j = 1 : & \quad c_{h-p}^1 < \dots < c_h^{1,k} < \dots < c_h^{1,1} < c_h^{1,0} = c_{\text{prim}}^1, & c_h^{1,k} \downarrow c_{h-p}^1 & \text{ for } k \rightarrow \infty, \\ j = 2 : & \quad c_{\text{prim}}^2 = c_h^{2,0} < c_h^{2,1} < \dots < c_h^{2,k} < \dots < c_{h-p}^2, & c_h^{2,k} \uparrow c_{h-p}^2 & \text{ for } k \rightarrow \infty. \end{aligned} \quad (3.56)$$

Each orbit  $\gamma_h^{j,k}(\xi)$  corresponds to a (localized, traveling, invasion) 1-front pattern  $(B(x, t), W(x, t)) = (b_h^{k,j}(x - c_h^{k,j}t), w_h^{k,j}(x - c_h^{k,j}t))$  in PDE (3.5) that connects the bare soil state  $(\bar{B}, \bar{W}) = (0, \Psi/\Phi)$  to the uniform vegetation state  $(\bar{B}, \bar{W}) = (b_+(w^{+,s}), w^{+,s})$ .

As in the proofs of Theorems 2.4 and 3.4, we can verify that there indeed are open regions in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space for which  $c_{h-p}^j$  and  $c_{\text{prim}}^j$  have the same sign (for either  $j = 1, 2$  or for both) and such that (3.55) holds, by considering the unfolded saddle-node case of Lemma 2.6. In fact, we know from Lemma 3.2 that  $c$  changes sign as the  $w$ -coordinate of the intersection point on  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_0^+}$  passes through  $9/(2 + 9a)$ . Thus (and for instance), all 4 values  $c_{h-p}^j$  and  $c_{\text{prim}}^j$ ,  $j = 1, 2$ , must have the same sign as the entire homoclinic orbit spanned by  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  either is to the left or to the right of  $w = 9/(2 + 9a)$  – more precise, if either  $(\underline{w}_{h,0}, w^{+,s}) \subset (4/(1 + 4a), 9/(2 + 9a))$  or  $(\underline{w}_{h,0}, w^{+,s}) \subset (9/(2 + 9a), 1/a)$  (cf. (3.40)). Note that it follows from (3.32) that  $w_0^{s,N} = 9/(2 + 9a)$  implies that  $\mathcal{C}^2 = \mathcal{A}^2/9$  (independent of  $a$ ), so that we can indeed move the homoclinic loop associated to the unfolded saddle-node – i.e.,  $\sigma \ll 1$  as in (3.34) – through  $w = 9/(2 + 9a)$  by increasing  $\mathcal{C}^2 \in (0, \mathcal{A}^2)$  through  $\mathcal{A}^2/9$ . On the other hand, it is certainly also possible that  $c_{h-p}^j$  and  $c_{\text{prim}}^j$  do not have the same sign. Hence, apart from the PDE point of view – from which it is natural to consider stationary patterns – this gives us an additional motivation to study the sign-changing stationary case  $c = 0$  in more detail, as we will briefly do in Remark 3.10 in section 3.4.

**Proof.** We only consider the case  $j = 1$ , i.e., we assume that  $c_{h-p}^1$  and  $c_{\text{prim}}^1$  have the same sign and that (3.55) holds for  $c = c_{h-p}^1, c_{\text{prim}}^1$ . The proof for  $j = 2$  goes exactly along the same lines.

For  $c = c_{\text{prim}}^1$ ,  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  by assumption spirals inwards in backwards ‘time’ and ‘wraps around’ the (perturbed) periodic orbit  $(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  on (the projection of)  $\mathcal{M}_\varepsilon^+$  – see Fig. 3.7. Since  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{h-p}$ ,  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  must intersect  $\mathcal{I}_{\text{down}}$  countably many times. We define  $(\bar{w}_i^{1,1}, \bar{q}_i^{1,1})$  as the next intersection of  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+(c_{\text{prim}}^1)}$  with  $\mathcal{I}_{\text{down}}$  beyond the 2 primary intersection points: it is the first non-primary intersection point and has  $\bar{q}_i^{1,1} > 0$ . As before,  $(\bar{w}_i^{1,1}, \bar{q}_i^{1,1}) \in \mathcal{I}_{\text{down}} = \{(\bar{w}_{\text{down}}(c), \bar{q}_{\text{down}}(c))\}$  determines the value  $c_i^{1,1}$  through  $\bar{w}_{\text{down}}(c) = \bar{w}_i^{1,1}$  – where we know that  $c_i^{1,1} < c_{\text{prim}}^1$  since the  $w$ -component of  $\mathcal{I}_{\text{down}}$  is a monotonically increasing function of  $q$  (Lemma 3.2). Since the perturbation term in (3.38) is  $\mathcal{O}(\varepsilon)$ , it follows that  $\|(\bar{w}_{\text{prim}}^1, \bar{q}_{\text{prim}}^1) - (\bar{w}_i^{1,1}, \bar{q}_i^{1,1})\| = \mathcal{O}(\varepsilon)$

and thus that  $c_{\text{prim}}^1 - c_i^{1,1} = \mathcal{O}(\varepsilon)$ . An  $\mathcal{O}(\varepsilon)$  change in  $c$  yields an  $\mathcal{O}(\varepsilon^2)$  change in the flow of (3.38), hence for all  $\tilde{c}$   $\mathcal{O}(\varepsilon)$  close to  $c_{\text{prim}}^1$ , the first non-primary intersection of  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(\tilde{c})$  and  $\mathcal{I}_{\text{down}}$  – denoted by  $(\bar{w}_i^{1,1}(\tilde{c}), \bar{q}_i^{1,1}(\tilde{c}))$  – must be  $\mathcal{O}(\varepsilon^2)$  close to  $(\bar{w}_i^{1,1}, \bar{q}_i^{1,1}) \in \mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c_{\text{prim}}^1)$ . Thus, the speed  $c_i^{1,1}(\tilde{c})$  associated to this intersection – by  $\bar{w}_{\text{down}}(c) = \bar{w}_i^{1,1}(\tilde{c})$  – must also be  $\mathcal{O}(\varepsilon^2)$  close to  $c_i^{1,1}$ . The situation is therefore similar to that in the proof of Theorem 3.4: an  $\mathcal{O}(\varepsilon)$  variation of  $\tilde{c}$  around  $c_i^{1,1}$  in (3.38) yields only an  $\mathcal{O}(\varepsilon^2)$  change in the  $c$ -coordinate associated the first non-primary intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(\tilde{c})$  so that there must be a unique  $\tilde{c} = c_h^{1,1}$  such that  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c_h^{1,1}) = (\bar{w}_{\text{down}}(c_h^{1,1}), \bar{q}_{\text{down}}(c_h^{1,1}))$ . This establishes the existence of the first non-primary heteroclinic 1-front orbit  $\gamma_h^{1,1}(\xi)$  for  $c = c_h^{1,1}$  in (3.7).

We can now iteratively consider the first intersection in backwards ‘time’ – denoted by  $(\bar{w}_i^{1,2}, \bar{q}_i^{1,2})$  – of  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c_h^{1,1})$  with  $\mathcal{I}_{\text{down}}$  beyond  $(\bar{w}_{\text{down}}(c_h^{1,1}), \bar{q}_{\text{down}}(c_h^{1,1}))$  with  $\bar{q}_i^{1,2} > 0$  – so that the speed  $c_i^{1,2}$  associated to this intersection is  $\mathcal{O}(\varepsilon)$  close to  $c_h^{1,1}$ . Completely analogous to the above arguments, we deduce the existence of a unique  $c = c_h^{1,2}$  such that  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c_h^{1,2}) = (\bar{w}_{\text{down}}(c_h^{1,2}), \bar{q}_{\text{down}}(c_h^{1,2}))$ , which establishes the existence of the next non-primary 1-front orbit  $\gamma_h^{1,2}(\xi)$  of (3.7) with  $0 < c_h^{1,1} - c_h^{1,2} = \mathcal{O}(\varepsilon)$ . Next, we construct  $\gamma_h^{1,3}(\xi)$  in (3.7) with  $0 < c_h^{1,2} - c_h^{1,3} = \mathcal{O}(\varepsilon)$  through the intersection  $\mathcal{I}_{\text{down}} \cap W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c_h^{1,3}) = (\bar{w}_{\text{down}}(c_h^{1,3}), \bar{q}_{\text{down}}(c_h^{1,3}))$ , etc.

Theorem 2.4 holds independent of  $c$ , which implies that  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c)$  wraps around the periodic orbit  $(w_{p,\varepsilon}(X), q_{p,\varepsilon}(X))$  of (3.38) (in backwards ‘time’) for all  $c$  with the same sign as  $c_{\text{h-p}}^1$  and  $c_{\text{prim}}^1$  (cf. Fig. 3.7). Thus, there must be countably many heteroclinic orbits  $\gamma_h^{1,k}(\xi)$  – every  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}(c)$  intersects  $\mathcal{I}_{\text{down}}$  countably many times – and the associated speeds  $c_h^{1,k}$  must all be between  $c_{\text{h-p}}^1$  and  $c_{\text{prim}}^1$ . Moreover, the decreasing sequence  $\{c_h^{1,k}\}_{k=1}^\infty$  must have a limit that cannot differ from  $c_{\text{h-p}}^1$ :  $c_h^{1,k} \downarrow c_{\text{h-p}}^1$  as  $k \rightarrow \infty$ .  $\square$

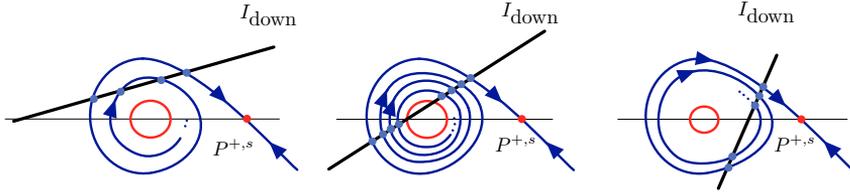
By establishing the existence of countably many distinct heteroclinic connections between  $P^0$  and  $P^{+,s}$ , Theorem 3.6 in a sense considers (one of) the most complex case(s), which is quite far removed from situations in which there are no such connections. To obtain insight in the bifurcations that occur ‘in between’, we can again freeze the reduced slow flow and vary  $\mathcal{I}_{\text{down}} = \mathcal{I}_{\text{down}}(\Phi)$  by increasing  $\Phi$  (from  $a(1+4a)\chi$  to  $\infty$  (3.52), (3.53), (3.53)). We consider the most simple case and assume that the homoclinic orbit of the frozen flow lies entirely in the  $w$ -region  $(4/(1+4a), 9/(2+9a))$  – so that all  $c_h^{j,k}$ s of Theorem 3.6 are positive – and that  $\mathcal{I}_{\text{down}}(\Phi) \cap W^s(P^{+,s})|_{\mathcal{M}_0^+} = \emptyset$  at  $\Phi = a(1+4a)\chi$  (this can easily be achieved by the unfolded saddle-node approach). As  $\Phi$  increases,  $\mathcal{I}_{\text{down}}(\Phi)$

becomes steeper and the intersection  $\mathcal{I}_{\text{down}}(\Phi) \cap \{q = 0\}$  moves over the entire interval determined by (3.13), i.e., from  $4/(1+4a)$  to  $1/a$ . Thus,  $\mathcal{I}_{\text{down}}(\Phi)$  moves through the homoclinic loop spanned by  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  and through the enclosed persistent periodic orbit established by Theorem 2.4. We tune the parameters such that during the passage of the latter, (3.55) holds and  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{h-p}$ , i.e., that Theorem 3.6 can be applied – which is also possible. It should be noticed that although the reduced flow (3.22) is frozen, this is not the case for the perturbed flow (3.38), since  $\rho_1(w)$  varies with  $\Theta$  (3.39) and  $\Theta = \Theta(\Phi)$  (3.53), thus the persistent periodic orbit established in Theorem 2.4 is not frozen, but also varies with  $\Phi$  – this is represented in the sketches of Fig. 3.8 by the decreasing size of the limiting periodic orbit on  $\mathcal{M}_\varepsilon^+$ .

Fig. 3.8 exhibits sketches of 3 configurations of  $\mathcal{I}_{\text{down}}(\Phi)$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  for increasing  $\Phi$ , the associated bifurcation scenario is sketched in Fig. 3.9. In Fig. 3.8a,  $\Phi$  has already passed through the first bifurcation value  $\Phi_{\text{prim}}$  at which the first 2 primary heteroclinic orbits  $\gamma_{\text{prim}}^j(\xi)$ ,  $j = 1, 2$  of Theorem 3.4 are created, and through a second one,  $\Phi_{SN,1}^b - \mathcal{O}(\varepsilon)$  close  $\Phi_{\text{prim}}$  – at which the first 2 secondary orbits appear. This bifurcation is followed by countably subsequent saddle-node bifurcations until  $\Phi$  reaches  $\Phi_{\text{per}}^b$  at which the 2 limiting heteroclinic orbits between  $P^0$  and the persistent periodic solution of Theorem 3.5 appear and we enter into the realm of Theorem 3.6. These orbits next disappear at  $\Phi_{\text{per}}^e$ , Fig. 3.8(b) is similar to Fig. 3.7 and represents the 2 countable families of heteroclinic orbits that exist for  $\Phi \in (\Phi_{\text{per}}^b, \Phi_{\text{per}}^e)$  (Theorem 3.6). All these orbits step-by-step disappear in pairs as  $\Phi$  is increased further: Fig. 3.8c shows the situation with only 5 left – 4 of these will disappear just before  $\Phi$  reaches  $\Phi^{+,s}$  at which  $\mathcal{I}_{\text{down}}(\Phi)$  passes through  $P^{+,s}$ .

We refrain from giving all rigorous details on which the above sketched scenario is based – this is in essence a matter of following the lines set out in the proofs of the preceding results. Moreover, we also refrain from working out all possible alternative bifurcation scenarios that may occur – there are many (sub)cases to consider, some more simple, others more complex than that of Fig. 3.9. Nevertheless, we do briefly come back to this in the upcoming section – where we consider stationary, sign-changing, case  $c = 0$  case.

**Remark 3.7.** *As in Remark 2.9, we note that a result like Theorem 3.5 on the existence of heteroclinic connections between  $P^0$  and a periodic orbit on  $\mathcal{M}_\varepsilon^+$  can also be established under the assumption that there is only one critical point  $P^{+,c}$  of center type on  $\mathcal{M}_0^+$ . Similar remarks can be made about the upcoming Theorems 3.9, 3.12 and Corollary 3.16. We note – also as in Remark 2.9 – that the analysis of these additional cases is essentially the same as already presented.*



**Figure 3.8:** Sketches of 3 relative configurations of  $\mathcal{I}_{\text{down}}(\Phi)$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  for increasing  $\Phi$  that represent 3 distinct stages in the bifurcation scenario of Fig. 3.9.

### 3.4 Stationary 1-front patterns

In this section, we construct stationary heteroclinic 1-front patterns that are similar to those constructed in Theorems 3.4 and 3.5. We immediately note that if the reduced flow on  $\mathcal{M}_0^+$  has an unperturbed homoclinic loop  $W^s(P^{+,s})|_{\mathcal{M}_0^+} \cap W^u(P^0)|_{\mathcal{M}_0^+}$  – as in Fig. 3.4 – that it persists as homoclinic solution of (3.7) on  $\mathcal{M}_\varepsilon^+$  for  $\varepsilon \neq 0$  – since (3.7) with  $c = 0$  is a reversible system (see also Theorem 2.4). Thus, we a priori deduce that there cannot be any further non-primary heteroclinic 1-front connections between  $P^0 \in \mathcal{M}_\varepsilon^0$  and  $P^{+,s} \in \mathcal{M}_\varepsilon^+$  as those of Theorem 3.6 for  $c \neq 0$  (see however also Remark 3.10 for a result similar to Theorem 3.6). In the subsequent sections, we will proceed to construct homoclinic and periodic multi-front patterns – i.e., solutions of (3.7) that jump up and down between  $\mathcal{M}_\varepsilon^0$  and  $\mathcal{M}_\varepsilon^+$  – and show that there is a richness in these kinds of patterns similar to that of Theorem 3.6.

As in the previous sections, we approach the bifurcation analysis by freezing the flow on  $\mathcal{M}_\varepsilon^+$ . Thus, we choose  $\Psi, \Omega, \Theta$  as in (3.53) and vary  $\Phi$ . We know by Lemma 3.2 and (3.54) that for  $c = 0$ , the touch down point of  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  is represented by a vertical line/half line  $\mathcal{J}_{s-d}$  in the  $(w, q)$ -plane

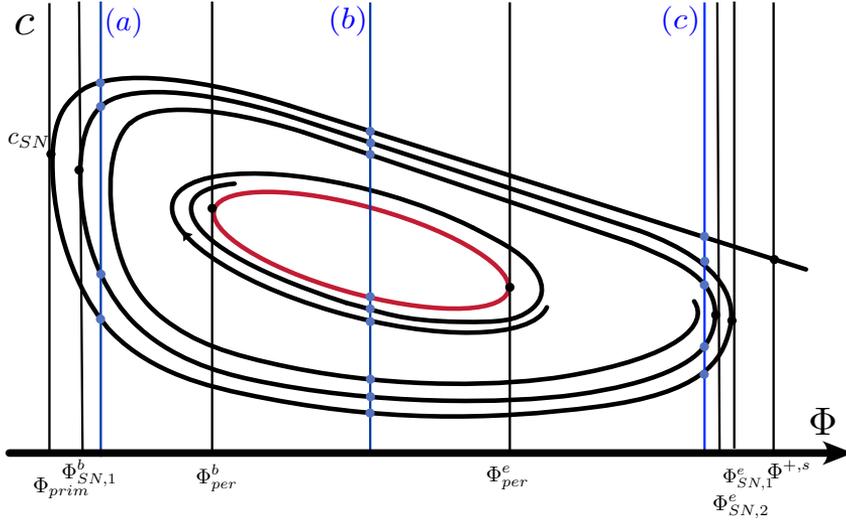
$$\mathcal{J}_{s-d} = \left\{ J_{s-d}(\Phi) = (w_{s-d}(\Phi), q_{s-d}(\Phi)) = \left( \frac{9}{2+9a}, \sqrt{\Phi} \left( \frac{\chi}{\Phi} - \frac{2}{a(2+9a)} \right) \right), \Phi > 0 \right\} \quad (3.57)$$

(at leading order in  $\varepsilon$ ). Clearly, a 1-front connection between  $P^0$  and  $P^{+,s}$  corresponds to those values of  $\Phi$  for which  $J_{s-d}(\Phi) \in W^s(P^{+,s})|_{\mathcal{M}_0^+}$ .

**Theorem 3.8.** *Let  $c = 0$  and let  $\varepsilon$  be sufficiently small. Then, there is a codimension 1 set  $\mathcal{R}_{s-1f}$  in  $(a, \Phi, \Psi, \Omega, \Theta)$ -space for which (stationary) 1-front heteroclinic orbits  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  exists in (3.7). More precise, let  $\Psi, \Omega, \Theta$  as in (3.53), then:*

(A) *If  $P^{+,s} = (w^{+,s}, 0)$  is the only critical point on (the projection of)  $\mathcal{M}_\varepsilon^+$ , i.e., if  $C^2 - 4AD > 0$ ,  $\mathcal{D} < 0$ ,  $\mathcal{E} > 0$  (3.28), then,*

- *if  $\chi > 0$ , then there is a unique value  $\Phi_{s-1f}$  such that there is a 1-front heteroclinic orbit  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  in (3.7);*
- *if  $\chi < 0$ ,  $w^{+,s} < w_{s-d} = 9/(2+9a)$  (3.57) and there are  $\Phi$  such that*



**Figure 3.9:** A sketch of the bifurcation scenario as function of  $\Phi$  representing the appearance in a saddle-node bifurcation of the 2 primary heteroclinic 1-front orbits  $\gamma_{\text{prim}}^j(\xi)$ ,  $j = 1, 2$ , of Theorem 3.4, followed by further saddle-node bifurcations leading to the situation governed by Theorem 3.6 in which countably many 1-front orbits exist; these orbits subsequently disappear in another cascade of saddle-node bifurcations eventually leaving only one (primary) 1-front orbit behind (Theorem 3.4). The relative configurations of  $\mathcal{I}_{\text{down}}(\Phi)$  and  $W^s(P^{+,s})|_{\mathcal{M}_0^+}$  sketched in Fig. 3.8 occur at the  $\Phi$ -values indicated by the vertical (a), (b) and (c) lines.

$\mathcal{H}_0^+(w_{s-d}(\Phi), q_{s-d}(\Phi)) < \mathcal{H}_0^{+,s}$  (3.18), (3.29), then there are 2 values  $\Phi_{s-1f}^j$ ,  $j = 1, 2$  for which 1-front heteroclinic orbits  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  exist in (3.7);

• if  $\chi < 0$  and either one of the above additional conditions does not hold, then there is no such stationary 1-front orbit.

**(B)** If there are two critical points on  $\mathcal{M}_\varepsilon^+$ , the center  $P^{+,c}$  and saddle  $P^{+,s}$ , i.e., if  $C^2 - 4AD > 0$ ,  $C < 0$ ,  $\mathcal{D} > 0$ , then,

• if  $\underline{w}_{h,0} < 9/(2 + 9a) < w^{+,s}$  – with  $\underline{w}_{h,0}$  as defined in Theorem 2.4 – then there are 2 values  $\Phi_{s-1f}^1 < \Phi_{s-1f}^2$  for which 1-front heteroclinic orbits  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  exist in (3.7);

• if  $9/(2 + 9a) > w^{+,s}$  there is a unique value  $\Phi_{s-1f}$  such that a 1-front heteroclinic orbit  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  exists in (3.7);

Every heteroclinic orbit  $\gamma_{s-1f}(\xi) = (w_{s-1f}(\xi), p_{s-1f}(\xi), b_{s-1f}(\xi), q_{s-1f}(\xi))$  corresponds to a stationary 1-front pattern

$(B(x, t), W(x, t)) = (b_{s-1f}(x), w_{s-1f}(x))$  in PDE (3.5) that connects the bare soil state  $(0, \Psi/\Phi)$  to the uniform vegetation state  $(b_+(w^{+,s}), w^{+,s})$ .

We refer to Fig. 3.14a for an example of a numerical simulation of (3.5)

exhibiting a stationary 1-front pattern. Moreover, we notice that – by symmetry (3.9) of (3.7) with  $c = 0$  – the heteroclinic orbit  $\gamma_{s-1f}(\xi) \subset W^u(P^0) \cap W^s(P^{+,s})$  has a counterpart  $\subset W^u(P^{+,s}) \cap W^s(P^0)$ , i.e., an orbit from  $P^{+,s}$  to  $P^0$ . Together, these orbits form a heteroclinic cycle between the saddles  $P^{+,s}$  to  $P^0$ .

**Proof.** The result follows directly by studying the possible intersections of  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  and the (vertical) line  $\mathcal{I}_{s-d}$  in combination with the observation that the range of  $q_{s-d}(\Phi)$  is  $\mathbb{R}$  for  $\chi > 0$ , while it's bounded from above by a negative number for  $\chi < 0$  (3.57). See Fig. 3.10.

Since all periodic orbits on  $\mathcal{M}_\varepsilon^+$  persist for  $c = 0$  (Theorem 2.4), we also ‘automatically’ obtain a result similar to Theorem 3.5 on the existence of heteroclinic connections  $\gamma_{s-p}^j(\xi) = (b_{s-p}^j(\xi), p_{s-p}^j(\xi), w_{s-p}^j(\xi), q_{s-p}^j(\xi))$  of (3.7) between the critical point  $P^0 \in \mathcal{M}_\varepsilon^0$  and one of the periodic orbits  $\gamma_{p,\varepsilon}(X) = (b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$  determined by  $\mathcal{H}_{p,0}^+ \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$  (3.26) – note that this orbit is  $\mathcal{O}(\varepsilon^2)$  close to the level set  $\mathcal{H}_0^+(w, q) = \mathcal{H}_{p,0}^+$  (cf. (3.38) with  $c = 0$ ).

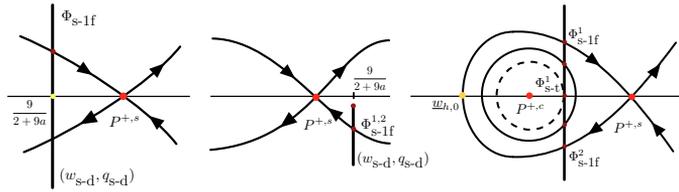
**Theorem 3.9.** *Let  $c = 0$ ,  $\Psi, \Omega, \Theta$  as in (3.53) such that (3.40) holds, let  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$  be a periodic solution of (3.7) determined by  $\mathcal{H} \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,s})$  and let  $\varepsilon$  be sufficiently small. Assume that  $\underline{w}_{h,0} < 9/(2 + 9a) < w^{+,s}$  and define*

$$\Phi_{s-t} = \frac{1}{2}a(2 + 9a)\chi \in (\Phi_{s-1f}^1, \Phi_{s-1f}^2), \quad (3.58)$$

$$\mathcal{H}_{s-t}^+ = \mathcal{H}_0^+(w_{s-d}(\Phi_{s-t}), q_{s-d}(\Phi_{s-t})) = \mathcal{H}_0^+\left(\frac{9}{2 + 9a}, 0\right) \quad (3.59)$$

with  $\Phi_{s-1f}^{1,2}$  as defined in Theorem 3.8,  $(w_{s-d}(\Phi), q_{s-d}(\Phi))$  as in (3.57) and  $\mathcal{H}_{s-t}^+ > \mathcal{H}_0^{+,c}$  (unless  $P^{+,c} = \left(\frac{9}{2 + 9a}, 0\right)$  (restricted to  $\mathcal{M}_0^+$ ) – see Fig. 3.10c). For all  $\mathcal{H} \in (\mathcal{H}_{s-t}^+, \mathcal{H}_0^{+,s})$ , there are 2 values  $\Phi_{p-1f}^{1,2} = \Phi_{p-1f}^{1,2}(\mathcal{H})$  – with  $\Phi_{s-1f}^1 < \Phi_{p-1f}^1 < \Phi_{s-t} < \Phi_{p-1f}^2 < \Phi_{s-1f}^2$  – that determine 2 distinct heteroclinic orbits  $\gamma_{s-p}^j(\xi; \mathcal{H})$  of (3.7) between the critical point  $P^0 \in \mathcal{M}_\varepsilon^0$  and the periodic orbit  $(b_{p,\varepsilon}(X), p_{p,\varepsilon}(X), w_{p,\varepsilon}(X), q_{p,\varepsilon}(X)) \subset \mathcal{M}_\varepsilon^+$  (and  $\gamma_{s-p}^j(\xi; \mathcal{H}) = (b_{s-p}^j(\xi; \mathcal{H}), p_{s-p}^j(\xi; \mathcal{H}), w_{s-p}^j(\xi; \mathcal{H}), q_{s-p}^j(\xi; \mathcal{H}))$ ). The orbits  $\gamma_{s-p}^j(\xi; \mathcal{H})$  ( $j = 1, 2$ ) correspond to stationary 1-front patterns  $(B(x), W(x)) = (b_{s-p}^j(x; \mathcal{H}), w_{s-p}^j(x; \mathcal{H}))$  in PDE (3.5) that connect the bare soil state  $(0, \Psi/\Phi)$  to the spatially periodic pattern  $(b_{p,\varepsilon}(X), w_{p,\varepsilon}(X))$ .

**Remark 3.10.** *Together, Theorems 3.8 and 3.9 provide the possibility to establish a result similar to that of Theorem 3.6 in the case that  $c_{h-p}^j$  and  $c_{prim}^j$  do not*

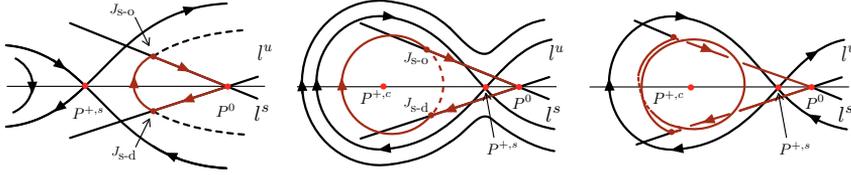


**Figure 3.10:** Sketches of 3 configurations of the line  $\mathcal{J}_{s-d}$  (3.57) and  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  as consider in Theorem 3.8: (a)  $P^{+,s}$  is the only critical point on  $\mathcal{M}_\varepsilon^+$  and  $\chi > 0$ ; (b)  $P^{+,s}$  is the only critical point on  $\mathcal{M}_\varepsilon^+$  and  $\chi < 0$ ; (c) a center and a saddle on  $\mathcal{M}_\varepsilon^+$  with  $\underline{w}_{h,0} < 9/(2+9a) < w^{+,s}$ .

have the same signs. Assume we have – for a certain parameter combination  $(a, \Psi, \Phi, \Omega, \Theta)$  such that (3.40) holds – that  $c_{h-p}^1 < 0 < c_{\text{prim}}^1$ . This implies that the point  $J_{s-d}(\Phi)$  on  $\mathcal{T}_{\text{down}}$  – Lemma 3.2 – must lie between the intersections of  $\mathcal{T}_{\text{down}}$  with the unperturbed homoclinic orbit that determines  $c_{\text{prim}}^1 > 0$  (Theorem 3.8) and the persisting periodic orbit that determines  $c_{h-p}^1 < 0$  (Theorem 3.9). Thus,  $J_{s-d}(\Phi)$  determines a level set  $\mathcal{H}_0^+(w, q) = \mathcal{H}_0^+(w_{s-d}(\Phi), q_{s-d}(\Phi)) = \tilde{\mathcal{H}} \in (\mathcal{H}_0^{+,c}, \mathcal{H}_0^{+,c})$  and we know by Theorem 3.9 that  $\Phi = \Phi_{p-2f}^2(\tilde{\mathcal{H}})$ :  $(w_{s-d}(\Phi), q_{s-d}(\Phi))$  is the touchdown point of a heteroclinic orbit between the bare soil state and the persisting (stationary) periodic orbit determined by the level set  $\mathcal{H}_0^+(w, q) = \tilde{\mathcal{H}}$ . It then follows by arguments similar to those in the proof of Theorem 3.6 that there are countably many  $c$ -values  $0 < \dots < c_h^{1,k} < \dots < c_h^{1,1} < c_h^{1,0} = c_{\text{prim}}^1$  with  $c_h^{1,k} \downarrow 0$  for  $k \rightarrow \infty$  for which non-primary heteroclinic connections between  $P^0$  and  $P^{+,s}$  exist – as in Theorem 3.6. The main difference with Theorem 3.6 is that  $c = 0$  determines a stationary orbit and not an attracting one: for  $c$  slightly above  $c = 0$ , the unstable manifold  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+}$  only spirals inwards very weakly (in backward time). As a consequence, the number of intersections  $W^s(P^{+,s})|_{\mathcal{M}_\varepsilon^+} \cap \mathcal{T}_{\text{down}}$  with  $\mathcal{H}_0^+(w, q) > \tilde{\mathcal{H}}$  increases (without bound) as  $c \downarrow 0$ .

### 3.5 Stationary homoclinic 2-front patterns: vegetation spots and gaps

In this section we construct stationary 2-front patterns that correspond to vegetation spots or vegetation gaps – the latter sometimes also interpreted as fairy circles. These patterns are observed in nature and appear as stable patterns in simulations of (3.1)/(3.5) – see [175] and Figs. 3.14c and 3.15a. The patterns/orbits to be constructed are symmetric with respect to the reversibility symmetry in (3.5) that persists as (3.9) into (3.7) – with  $c = 0$ . As a consequence, we may expect that these patterns are generic, in the sense that they exist in open regions within parameter space – see for instance [39]. Notice that this is unlike the stationary – but non-symmetric – 1-front patterns of the previous section, that only exist for  $(a, \Phi, \Psi, \Omega, \Theta) \in \mathcal{R}_{s-1f}$ , an explicitly determined co-dimension 1



**Figure 3.11:** Sketches of 3 (projected) ‘skeleton structures’ of stationary 2-fronts homoclinic to  $P^0$ : (a) (b) Two examples of the 2 skeleton structures  $\Gamma_{s-2f}$  as considered in Theorem 3.11; (c) The extended skeleton  $\Gamma_{s-2f}^{ext}$  of Theorem 3.12 and an associated higher order 2-front homoclinic with an additional full extra ‘spatial oscillation’ during its passage along  $\mathcal{M}_\varepsilon^+$ .

manifold (Theorem 3.8).

We first consider the (localized) spots: localized vegetated regions embedded within bare soil. Thus, these spots correspond to solutions of (3.7) that are homoclinic to the bare soil state  $P^0$ . Singularly perturbed models of the type (3.7), can have homoclinic (pulse) solutions of various types. The localized vegetation (spot) patterns constructed in [11, 133] in the context of the extended-Klausmeier model – also called generalized Klausmeier-Gray-Scott model [155] – make a fast excursion away from the slow manifold that contains the critical point associated to the bare soil state following a homoclinic solution of the fast reduced system. As a consequence these spots are ‘narrow’, their size scales with  $\varepsilon$ . Although such pulses are also exhibited by the present model – see Fig. 3.15b and Remark 3.14 – we focus here on 2-front patterns, i.e., orbits homoclinic to  $P^0$  that ‘jump’ from  $\mathcal{M}_\varepsilon^0$  to  $\mathcal{M}_\varepsilon^+$ , follow the slow flow on  $\mathcal{M}_\varepsilon^+$  over an  $\mathcal{O}(1)$  distance and jump back again – by its second fast reduced heteroclinic front – to  $\mathcal{M}_\varepsilon^0$ . Since the extended-Klausmeier model only has one slow manifold [11, 133], such orbits cannot exist in that model. Moreover, these patterns have a well-defined width that does not decrease to 0 as  $\varepsilon \downarrow 0$ , a property that appears to be natural in observed ecosystems [175].

The construction of the most simple – primary, cf. section 3.2 – singular ‘skeleton structure’  $\Gamma_{s-2f} \subset \mathbb{R}^4$  – the phase space of (3.7) – of a stationary homoclinic 2-front orbit  $\gamma_{s-2f}(\xi)$  to  $P^0$  of (3.7) is relatively straightforward (but somewhat involved/technical). Since the homoclinic orbit  $\gamma_{s-2f}(\xi) \subset W^u(P^0)$ , it follows  $\ell^u$  along  $\mathcal{M}_\varepsilon^0$ , takes off from  $\mathcal{M}_\varepsilon^0$  by following the fast reduced flow and touches down on  $\mathcal{M}_\varepsilon^+$  near an element of  $\mathcal{I}_{down}$  – see section 3.1. In fact, since we consider stationary spots, the touch down point is near a point  $J_{s-d} \in \mathcal{J}_{s-d}$  (3.57) – where we note that  $\gamma_{s-2f}(\xi)$  cannot exactly touch down on  $\mathcal{I}_{down}/\mathcal{J}_{s-d}$  since  $\gamma_{s-2f}(\xi) \notin \mathcal{W}^s(\mathcal{M}_\varepsilon^+)$  – see the proof of (upcoming) Theorem 3.11. The touch down point  $J_{s-d}$  determines (at leading order) a level set  $\mathcal{H}_{s-d}$  of the Hamiltonian  $\mathcal{H}_0^+(w, q)$  (3.26) of the slow reduced flow (3.22) on  $\mathcal{M}_\varepsilon^+$ :

$\mathcal{H}_{s-d} = \mathcal{H}_{s-d}(\Phi) = \mathcal{H}_0^+(w_{s-d}(\Phi), q_{s-d}(\Phi))$ . As long as it remains (exponentially) close to  $\mathcal{M}_\varepsilon^+$ , the homoclinic orbit-to-be-constructed remains asymptotically close to the level set  $\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}$ . This construction provides the first half of skeleton  $\Gamma_{s-2f}$ , the second part follows by the symmetry (3.9) – with  $c = 0$ . Completely analogous to  $\mathcal{I}_{\text{down}}$ , one can define  $\mathcal{I}_{\text{off}}$  as the points on  $\mathcal{M}_\varepsilon^+$  that determine the evolution of orbits in  $W^s(P^0) \cap W^u(\mathcal{M}_\varepsilon^+)$  after their jump from  $\ell^s \subset \mathcal{M}_\varepsilon^0$  through the fast field *in backwards time*. In fact, it follows by the symmetry (3.9) that  $\mathcal{I}_{\text{off}}$  and its stationary counterpart  $\mathcal{J}_{s-0}$  correspond exactly to the reflections of  $\mathcal{I}_{\text{down}}$  and  $\mathcal{J}_{s-d}$  with respect to the  $q$ -axis – where we consider  $\mathcal{I}_{\text{down}}/\mathcal{J}_{s-d}$  and  $\mathcal{I}_{\text{down}}/\mathcal{J}_{s-d}$  within the (projected) 2-dimensional representation of  $\mathcal{M}_\varepsilon^+$  as in Lemma 3.2 and in (3.57). Thus, for a given  $\Phi$ , the take off point  $J_{s-d}$  is given by  $J_{s-0}(\Phi) = (w_{s-d}(\Phi), -q_{s-d}(\Phi))$ ; this point also lies on the level set  $\mathcal{H}_{s-d}$  – since  $\mathcal{H}_0^+(w, q)$  (of course) also is symmetric in  $q \rightarrow -q$ .

We define the region  $\mathcal{S}_{s-2f}$  in  $(a, \Psi, \Phi, \Omega, \Theta)$ -space for which the point  $J_{s-d}$  (and thus  $J_{s-0}$ ) can be constructed (as above) *and* there is a solution of the slow reduced flow (3.22) on  $\mathcal{M}_\varepsilon^+$  that connects  $J_{s-d}$  to  $J_{s-0}$  – so that  $\Gamma_{s-2f}$  indeed exists as closed singular ‘loop’. Obviously,  $\mathcal{S}_{s-2f} \neq \emptyset$  – see also Fig. 3.11 – however, the fact that both  $J_{s-d}, J_{s-0} \in \{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$  does not necessarily imply that  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f}$  for all values for which these points exist on  $\mathcal{M}_\varepsilon^+$ . For instance, in the case that the saddle  $P^{+,s}$  is the only critical point on  $\mathcal{M}_\varepsilon^+$ ,  $J_{s-d}$  and  $J_{s-0}$  are not connected by a solution of (3.22) if  $\mathcal{H}_{s-d} > \mathcal{H}_0^{+,s}$  – the value of  $\mathcal{H}_0^+(w, q)$  at  $P^{+,s}$  – see Fig. 3.11a. Moreover, if  $\mathcal{H}_{s-d} < \mathcal{H}_0^{+,s}$  there is an additional condition on  $\Phi$  that is determined by the relative positions of  $w^{+,s}$  (the  $w$ -coordinate of  $P^{+,s}$ ),  $9/(2+9a)$  (the  $w$ -coordinate of  $J_{s-d}/J_{s-0}$ ) and  $\Psi/\Phi$  (the  $w$ -coordinate of the bare soil state associated to  $P^0$ ). Here, we refrain from working out the full ‘bookkeeping’ details by which (the boundary of)  $\mathcal{S}_{s-2f}$  is determined – see also a further brief discussion following Theorem 3.11. We refer to Fig. 3.11a for a case  $w^{+,s} < 9/(2+9a) < \Psi/\Phi$  (and implicitly  $\chi > 0$ ) for which  $J_{s-d}$  can only be connected to  $J_{s-0}$  if  $\Phi < \Phi_{s-t} = \frac{1}{2}a(2+9a)\chi$  (3.58) – since we need that  $q_{s-d}(\Phi) < 0$ . Notice that the sketch in Fig. 3.11a in principle also covers a (sub)case of the situation with two critical points  $P^{+,c}$  and  $P^{+,s}$  on  $\mathcal{M}_\varepsilon^+$  and that Fig. 3.11b considers the case  $9/(2+9a) < w^{+,s} < \Psi/\Phi$  for this situation. Clearly, there are no further restrictions on  $\Phi$  if  $\mathcal{H}_{s-d} < \mathcal{H}_0^{+,s}$  if there are two critical points  $P^{+,c}$  and  $P^{+,s}$  on  $\mathcal{M}_\varepsilon^+$ , since the orbits on the level set associated to  $\mathcal{H}_{s-d}(\Phi)$  are periodic, while one again has to impose  $\Phi < \Phi_{s-t}$  to have a connection between  $J_{s-d}$  and  $J_{s-0}$  for level sets outside the homoclinic loop, i.e., for  $\mathcal{H}_{s-d} > \mathcal{H}_0^{+,s}$ . Finally, we note that the skeleton structure  $\Gamma_{s-2f}$  can in principle also be constructed for  $(a, \Psi, \Phi, \Omega, \Theta)$  such that there is no critical point on  $\mathcal{M}_\varepsilon^+$ , or only one critical point that is not a saddle but a center (in the limit  $\varepsilon \rightarrow 0$ ).

Summarizing, the (open) region  $\mathcal{S}_{s-2f}$  is defined such that for parameter combi-

nations  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f}$ , a singular skeleton  $\Gamma_{s-2f} \subset \mathbb{R}^4$  can be constructed as above. In the limit  $\varepsilon \rightarrow 0$ ,  $\Gamma_{s-2f}$  is spanned by a piece of  $\ell^u \subset \mathcal{M}_0^0$  from  $P^0$  up to the ( $\varepsilon \rightarrow 0$  limit of the) take off point from  $\mathcal{M}_0^0$  (that has the same  $(w, q)$ -coordinates as  $J_{s-d}$  in the limit  $\varepsilon \rightarrow 0$ ), the jump through the fast field along (a piece of)  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  (3.17) towards the  $\varepsilon \rightarrow 0$  limit of the (projected) touch down point  $J_{s-d} \in \mathcal{M}_0^+$ , the connection along  $\mathcal{M}_0^+$  to  $J_{s-d}$  by the slow reduced flow (on the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$ ) up to (the  $\varepsilon \rightarrow 0$  limit of) the take off point  $J_{s-o}$ , followed by a fast jump backwards along (a piece of)  $W^s(\mathcal{M}_0^0) \cap W^u(\mathcal{M}_0^+)$  to  $\mathcal{M}_0^0$  and a final piece of  $\ell^s$  (up to  $P^0$ ) – see Figs. 3.11a and 3.11b for 2 sketches of projections of  $\Gamma_{s-2f}$  on  $\mathcal{M}_0^+$  that skip both jumps through the fast field. The proof of the persistence of  $\Gamma_{s-2f}$  for  $\varepsilon \neq 0$  relies heavily on the reversibility symmetry of (3.7) with  $c = 0$  (3.9).

**Theorem 3.11.** *Let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f}$  and  $\Gamma_{s-2f} \subset \mathbb{R}^4$  be the singular skeleton constructed above. Then, there is for  $\varepsilon > 0$  sufficiently small a symmetric homoclinic 2-front orbit*

$\gamma_{s-2f}(\xi) = (b_{s-2f}(\xi), p_{s-2f}(\xi), w_{s-2f}(\xi), q_{s-2f}(\xi)) \subset W^u(P^0) \cap W^s(P^0)$  of (3.7) with  $c = 0$  that merges with  $\Gamma_{s-2f}$  as  $\varepsilon \downarrow 0$ . The associated stationary pattern  $(B(x, t), W(x, t)) = (b_{s-2f}(x), w_{s-2f}(x))$  in (3.5) represents a stationary localized vegetation spot embedded in bare soil.

We refer to Figs. 3.1b and 3.14c for numerical observations of these 2-front spot patterns. In Fig. 3.16, the projection of the 2-front orbit on the  $(w, b)$ -plane is given; it clearly shows the slow-fast-slow-fast-slow nature of the pattern: it first follows  $\mathcal{M}_\varepsilon^0$  (slowly), jumps to  $\mathcal{M}_\varepsilon^+$ , follows the slow flow on  $\mathcal{M}_\varepsilon^+$ , jumps back to  $\mathcal{M}_\varepsilon^0$  and slowly returns to  $P^0$ .

To get some insight in the boundaries of  $\mathcal{S}_{s-2f}$  – and thus is the bifurcations of  $\gamma_{s-2f}(\xi)$  – we can (as usual) ‘freeze’ the flow on  $\mathcal{M}_0^+$  by choosing  $\Psi, \Omega, \Theta$  as in (3.53) and vary  $\Phi$ . We need to be aware though that  $\ell^{u,s}$  and  $P^0$  do vary with  $\Phi$  (i.e., they are not frozen). In the situation sketched in Fig. 3.11a – thus with  $w^{+,s} < 9/(2 + 9a) < \Psi/\Phi$  and  $\chi > 0$  – we see that the distance between the 2 fronts of  $\gamma_{s-2f}(\xi)$  approaches  $\infty$  as  $\Phi \downarrow \Phi_{s-1f} < \Phi_{s-t}$  as defined in Theorems 3.8 and 3.9:  $\gamma_{s-2f}(\xi)$  obtains the character of the superposition of the 1-front heteroclinic orbit  $\gamma_{s-1f}(\xi)$  of Theorem 3.8 between  $P^0$  and  $P^{+,s}$  and its symmetrical counterpart – (3.9) – connecting  $P^{+,s}$  back to  $P^0$ . The other boundary corresponds to  $\Phi \uparrow \Phi_{s-t}$ : the distance traveled along  $\mathcal{M}_\varepsilon^+$  decreases to 0 and  $\gamma_{s-2f}(\xi)$  detaches from  $\mathcal{M}_\varepsilon^+$ . However, since the angle between  $\ell^u$  and  $\ell^s$  is determined by  $\sqrt{\Phi}$  (3.21), this can only happen as also  $\Psi/\Phi \downarrow 9/(2 + 9a)$ , i.e., as the ‘projected triangle’ of Fig. 3.11a that represents the skeleton  $\Gamma_{s-2f}$  entirely contracts to a point – see also Remark 3.14. The bifurcational structure associated to the situation sketched in Fig. 3.11b is quite different: as  $\Phi$  decreases towards  $\Phi_{s-1f}^1 < \Phi_{s-t}$ ,  $\gamma_{s-2f}(\xi)$  does not merge with a (superposition of 2) 1-front(s)  $\gamma_{s-1f}(\xi)$  of Theorem 3.8. In fact,  $\gamma_{s-2f}(\xi)$  does not bifurcate at

all, the distance between the 2 fronts of  $\gamma_{s-2f}(\xi)$  remains bounded as  $\Phi$  passes through  $\Phi_{s-1f}^1$  (the main difference between the cases  $\Phi > \Phi_{s-1f}^1$  and  $\Phi < \Phi_{s-1f}^1$  is the sign of  $\mathcal{H}_{s-d} - \mathcal{H}_0^{+,s}$ :  $\gamma_{s-2f}(\xi)$  follows an orbit of (3.22) outside the homoclinic loop for  $\Phi < \Phi_{s-1f}^1$ ). Moreover, as  $\Phi$  increases towards  $\Phi_{s-t}$ , the distance  $\gamma_{s-1f}(\xi)$  travels along  $\mathcal{M}_\varepsilon^+$  also does not go to 0, in fact  $\gamma_{s-1f}(\xi)$  (almost) follows the entire periodic orbit of the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}(\Phi_{s-t}) = \mathcal{H}_{s-t}\}$ , the critical/limiting orbit of Theorem 3.9. Again, this can only happen if also the projection of  $P^0$  on  $\mathcal{M}_\varepsilon^+$  merges with this periodic orbit. We refrain from going any further into the details of these – and other – bifurcations of  $\gamma_{s-2f}(\xi)$ .

**Proof of Theorem 3.11.** The proof follows the geometrical approach developed in [41, 42], equivalently the more analytical approach of [84] could be employed. The construction of  $\gamma_{s-2f}(\xi)$  is based on the ‘intermediate’ orbit  $\gamma_{i-1f}(\xi) \subset W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$ , the heteroclinic connection between  $P^0$  and  $\mathcal{M}_\varepsilon^+$  that touches down on  $J_{s-d} \in \mathcal{M}_\varepsilon^+$ ;  $\gamma_{i-1f}(\xi)$  follows the slow flow along  $\mathcal{M}_\varepsilon^+$  and is thus asymptotically close to the skeleton  $\Gamma_{s-2f}$  up to the take off point  $J_{s-0}$  ( $\gamma_{i-1f}(\xi) \subset W^s(\mathcal{M}_\varepsilon^+)$  and thus cannot take off from  $\mathcal{M}_\varepsilon^+$ ). The homoclinic orbit  $\gamma_{s-2f}(\xi)$  is constructed as a symmetric orbit – i.e., an orbit that passes through the plane  $\{p = q = 0\}$  at its ‘midpoint’ – that is exponentially close to  $\gamma_{i-1f}(\xi)$  up to the point it takes off from  $\mathcal{M}_\varepsilon^+$ .

Since  $b_-(w) < 1/2 < b_+(w)$  (3.11),  $\gamma_{i-1f}(\xi)$  intersects the hyperplane  $\{b = 1/2\}$  transversally in the point  $P_{i-1f}$  – by definition. We define for some sufficiently small  $\sigma$  (independent of  $\varepsilon$ ), the (bounded) 1-dimensional curve  $C_{i-1f}^\sigma \subset \{b = 1/2\}$  as the (first, transversal) intersection of  $W^u(P^0)$  and  $\{b = 1/2\}$  that is at a distance of maximal  $\sigma$  away from  $P_{i-1f}$ ; in other words,  $C_{i-1f}^\sigma = W^u(P^0) \cap \{b = 1/2\} \cap \{|(b, p, w, q) - P_{i-1f}| \leq \sigma\}$ . By choosing  $\gamma(0) \in C_{i-1f}^\sigma$ , the curve  $C_{i-1f}^\sigma$  provides a parametrization of orbits  $\gamma(\xi)$  in  $W^u(P^0)$  near  $\gamma_{i-1f}(\xi)$ . In fact, the saddle structure of the fast flow around  $\mathcal{M}_\varepsilon^+$  cuts  $W^u(P^0)$ , and thus  $C_{i-1f}^\sigma$ , exactly in two along  $\gamma_{i-1f}(\xi) \subset W^s(\mathcal{M}_\varepsilon^+)$ : orbits  $\gamma(\xi) \subset W^u(P^0)$  with (by definition)  $\gamma(0) \in C_{i-1f}^{\sigma,r}$  cross through the plane  $\{p = 0\}$  near  $\mathcal{M}_\varepsilon^+$  so that their  $b$ -coordinate changes direction; the  $b$ -coordinates of  $\gamma(\xi)$ 's with  $\gamma(0)$  in  $C_{i-1f}^\sigma \setminus C_{i-1f}^{\sigma,r}$  do not change direction, these  $\gamma(\xi)$ 's pass along  $\mathcal{M}_\varepsilon^+$  without the possibility of returning to  $\mathcal{M}_\varepsilon^0$ . Thus, we may uniquely parameterize orbits  $\gamma(\xi) \subset W^u(P^0)$  that pass through  $\{p = 0\}$  near by  $\mathcal{M}_\varepsilon^+$  that are  $\sigma$ -close to  $\gamma_{i-1f}(\xi)$  by the distance  $d$  between their initial point  $\gamma(0) \in C_{i-1f}^{\sigma,r}$  and  $P_{i-1f}$  ( $\in \partial C_{i-1f}^{\sigma,r}$ ) – where we have implicitly used the fact that  $W^u(\mathcal{M}_\varepsilon^0) \supset W^u(P^0)$  is  $C^1 - \mathcal{O}(\varepsilon)$  close to its  $\varepsilon \rightarrow 0$  limit  $W^u(\mathcal{M}_0^0)$  [83, 86]. We denote these  $\gamma(\xi)$ 's by  $\gamma(\xi; d)$ .

Since the flow on  $\mathcal{M}_\varepsilon^+$  is  $\mathcal{O}(\varepsilon)$  slow, orbits  $\gamma(\xi; d)$  with  $\gamma(0) \in C_{i-1f}^{\sigma,r}$  can only follow  $\mathcal{M}_\varepsilon^+$  over an  $\mathcal{O}(1)$  distance (w.r.t.  $\varepsilon$ ) for  $d$  exponentially small; in fact it is necessary that  $d = \mathcal{O}(\exp(-\lambda_{f,+}(9/(2+9a))/\varepsilon))$ , where  $\lambda_{f,+}(w_0) = \sqrt{w_0 b_+(w_0) + 2aw_0 - 2}$ , the unstable eigenvalue of the reduced fast

flow (3.10) – with  $c = 0$  – associated to the critical point  $(b_+(w_0), 0)$ , and  $w_0 = 9/(2 + 9a)$  is the (leading order)  $w$ -coordinate of  $J_{s-d}$  (3.57). As  $d$  decreases, the ‘time’ (i.e., distance)  $\gamma(\xi; d)$  remains exponentially close to  $\mathcal{M}_\varepsilon^+$  increases monotonically (as follows from a direct perturbation analysis – for instance along the lines of [22]). Equivalently, the distance between  $J_{s-d}$  and the (projected) point on  $\mathcal{M}_\varepsilon^+$  at which  $\gamma(\xi; d)$  crosses through the  $\{p = 0\}$ -plane also increases monotonically with decreasing  $d$  – where we note that this point marks the transition between  $\gamma(\xi; d)$  approaching the ‘fast saddle’  $\mathcal{M}_\varepsilon^+$  exponentially close to  $W^s(\mathcal{M}_\varepsilon^+)$  and moving away from  $\mathcal{M}_\varepsilon^+$  exponentially close to  $W^u(\mathcal{M}_\varepsilon^+)$ . Clearly, for  $d$  ‘too large’  $\gamma(\xi; d)$  passes through  $\{p = 0\}$  before the slow flow on  $\mathcal{M}_\varepsilon^+$  – and thus  $\gamma(\xi)$  itself (since it is exponentially close to  $\mathcal{M}_\varepsilon^+$ ) – passes through  $\{q = 0\}$  – Fig. 3.11. However, by decreasing  $d$ , we can delay the passage of  $\gamma(\xi; d)$  through  $\{p = 0\}$  until after it passes through  $\{q = 0\}$ . It follows that there must be a  $d_*$  such that the associated orbit  $\gamma(\xi; d_*)$  passes through  $\{p = 0\}$  and  $\{q = 0\}$  simultaneously:  $\gamma(\xi; d_*) \cap \{p = q = 0\} \neq \emptyset$ .

The orbit  $\gamma_{s-2f}(\xi)$  coincides with  $\gamma(\xi; d_*)$ : by the reversibility symmetry (3.9) – with  $c = 0$  – it is symmetric around the ‘midpoint’ at which it passes through  $\{p = q = 0\}$ , so that indeed  $\gamma_{s-2f}(\xi) \subset W^u(P^0) \cap W^s(P^0)$ , a homoclinic 2-front orbit to  $P^0$  (asymptotically close to the skeleton  $\Gamma_{s-2f}$ ).  $\square$

The orbit  $\gamma_{s-2f}(\xi)$  is only the first of a countable family of homoclinic 2-front orbits if the slow piece of the skeleton  $\Gamma_{s-2f}$  is part of a closed orbit, i.e., if the connected part of the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$  that contains the ( $\varepsilon \rightarrow 0$  limits of the) touch down and take off points  $J_{s-d}$  and  $J_{s-0}$  is a closed orbit. In this case, the intermediate heteroclinic 1-front  $\gamma_{i-1f}(\xi) \subset W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)$  introduced in the proof of Theorem 3.11 coincides with one of the 2 heteroclinic orbits  $\gamma_{s-p}^j(\xi; \mathcal{H}_{s-d})$  between the critical point  $P^0 \in \mathcal{M}_\varepsilon^0$  and the periodic orbit  $\gamma_{p,\varepsilon}(X) \subset \mathcal{M}_\varepsilon^+$  established in Theorem 3.9. Thus, in this case  $\gamma_{i-1f}(\xi)$  passes countably many times through the plane  $\{q = 0\}$ . By steadily decreasing  $d$ , we can now determine a sequence of critical values  $d_*^i$ ,  $i = 0, 1, 2, \dots$  such that the associated orbits  $\gamma(\xi; d_*^i)$  pass through  $\{p = q = 0\}$  after  $i$  preceding passages through  $\{q = 0\}$ . Thus, the *primary* orbits  $\gamma_{s-2f}(\xi)$  of Theorem 3.11 correspond to  $\gamma(\xi; d_*^0)$  and the ‘higher order’ (stationary) homoclinic 2-front orbits  $\gamma_{s-2f}^i(\xi) \subset W^u(P^0) \cap W^s(P^0)$  coincide with the orbits  $\gamma(\xi; d_*^i)$  for  $i \geq 1$ . By the symmetry (3.9) – with  $c = 0$  – these orbits are also symmetric around the ‘midpoint’ at which they pass through  $\{p = q = 0\}$ :  $\gamma_{s-2f}^i(\xi)$  traces  $i$  full circuits over the closed orbit determined by the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$  during its passage along  $\mathcal{M}_\varepsilon^+$  – see the sketches in Fig. 3.2(c,d).

**Theorem 3.12.** *Let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f}$  such that (3.40) holds, and let  $\Gamma_{s-2f}^{\text{ext}} \subset \mathbb{R}^4$  be the extension of the singular skeleton  $\Gamma_{s-2f}$  of Theorem 3.11 that includes the entire closed orbit on  $\mathcal{M}_\varepsilon^+$  determined by the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$ . Then, for  $\varepsilon > 0$  sufficiently small, there is a countable family of symmetric 2-front*

*orbits*

$\gamma_{s-2f}^i(\xi) = (b_{s-2f}^i(\xi), p_{s-2f}^i(\xi), w_{s-2f}^i(\xi), q_{s-2f}^i(\xi)) \subset W^u(P^0) \cap W^s(P^0)$  of (3.7) with  $c = 0$  ( $i = 0, 1, 2, \dots$ ) that merges with  $\Gamma_{s-2f}^{\text{ext}}$  as  $\varepsilon \downarrow 0$  for all  $i \geq 1$  ( $\gamma_{s-2f}^0(\xi) = \gamma_{s-2f}(\xi)$  of Theorem 3.11 merges with  $\Gamma_{s-2f}$  as  $\varepsilon \downarrow 0$ ). The associated stationary patterns  $(B(x, t), W(x, t)) = (b_{s-2f}^i(x), w_{s-2f}^i(x))$  in (3.5) represent stationary localized vegetation spots embedded in bare soil with an increasing number of spatial oscillations in the vegetated area.

The construction of stationary homoclinic 2-front gap patterns – localized bare soil areas surrounded by vegetation – goes along exactly the same lines as the above construction of localized spot patterns. The main difference is that the homoclinic orbits-to-be-constructed are  $\subset W^u(P^{+,s}) \cap W^s(P^{+,s})$  so that the structure of orbits taking off and touching down now has to start out from the saddle  $P^{+,s} \in \mathcal{M}_\varepsilon^+$ . Nevertheless, the construction of the skeleton structure  $\Gamma_{g-2f}$  is completely similar to that of  $\Gamma_{s-2f}$ . Therefore, we only provide the essence of the construction of  $\Gamma_{g-2f}$ .

First, we need to assume that there is a critical point  $P^{+,s} \in \mathcal{M}_0^+$  of saddle type. The skeleton structure  $\Gamma_{g-2f}$  consists of a piece of  $W^u(P^{+,s}) \subset \mathcal{M}_0^+$  from  $P^{+,s}$  up to the ( $\varepsilon \rightarrow 0$  limit of the) take off point  $J_{g-o}^{+,0}$  from  $\mathcal{M}_0^+$ , followed by (a piece of)  $W^u(\mathcal{M}_0^+) \cap W^s(\mathcal{M}_0^0)$  (3.17) up to the ( $\varepsilon \rightarrow 0$  limit of the) touch down point  $J_{g-d}^{+,0} \in \mathcal{M}_0^0$  (that has the same  $(w, q)$ -coordinates as  $J_{g-o}^{+,0}$  in the limit  $\varepsilon \rightarrow 0$ ). Note that the take off/touch down points  $J_{g-o}^{+,0}/J_{g-d}^{+,0}$  differ essentially from their counterparts as  $J_{s-d}/J_{s-0}$  (3.57) considered so far: while  $J_{s-d}/J_{s-0}$  concerned the evolution of  $W^u(P^0) \cap W^s(\mathcal{M}_\varepsilon^+)/W^s(P^0) \cap W^u(\mathcal{M}_\varepsilon^+)$  along  $\mathcal{M}_\varepsilon^+$  in forwards/backwards ‘time’,  $J_{g-o}^{+,0}/J_{g-d}^{+,0}$  govern the orbits of  $W^u(P^{+,s}) \cap W^s(\mathcal{M}_\varepsilon^0)/W^s(P^{+,s}) \cap W^u(\mathcal{M}_\varepsilon^0)$  along  $\mathcal{M}_\varepsilon^0$ . Nevertheless, the coordinates of all take off/touch down points are at leading order determined by their  $\varepsilon \rightarrow 0$  limits (3.17) with  $w_h^\pm(0) = 9/(2 + 9a)$  (3.16). The next piece of  $\Gamma_{g-2f}$  consists of a symmetric part of a (cosh-type) orbit along  $\mathcal{M}_0^0$  of the (linear) slow reduced flow (3.20) up to the ( $\varepsilon \rightarrow 0$  limit of the) take off point  $J_{g-o}^{0,+}$ , which is again followed by a fast jump backwards along (a piece of)  $W^s(\mathcal{M}_0^+) \cap W^u(\mathcal{M}_0^0)$  to the ( $\varepsilon \rightarrow 0$  limit of the) touch down point  $J_{g-d}^{0,+} \in \mathcal{M}_0^+$ . The final piece is the symmetrical counterpart of the first piece: the flow of (3.22) along  $\mathcal{M}_0^+$  from the final touch down point back towards  $P^{+,s}$  – see Fig. 3.12a for a sketch of a projection of  $\Gamma_{s-2f}$  (without its fast jumps).

The (open) region  $\mathcal{S}_{g-2f}$  is defined by those  $(a, \Psi, \Phi, \Omega, \Theta)$ -combinations for which  $\Gamma_{g-2f}$  can be constructed. We note that there cannot be points in the intersection of  $\mathcal{S}_{s-2f}$  as defined in Theorem 3.11 and  $\mathcal{S}_{g-2f}$ : the (projections of the) take off/touch down points  $J_{s-d}/J_{s-0}$  lie on the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_{s-d}\}$  for which  $\mathcal{H}_{s-d} < \mathcal{H}_0^{+,s}$ , the value of  $\mathcal{H}_0^+(w, q)$  for  $P^{+,s}$  and its (un)stable manifolds (3.29). By construction,  $J_{g-o}^{+,0}, J_{g-d}^{+,0} \subset \{\mathcal{H}_0^+(w, q) = \mathcal{H}_0^{+,s}\}$  – compare

Fig. 3.11a to Fig. 3.12a. This also implies that  $\overline{\mathcal{S}_{s-2f}} \cap \overline{\mathcal{S}_{g-2f}} \neq \emptyset$ , in fact,  $\partial\mathcal{S}_{s-2f} \cap \partial\mathcal{S}_{g-2f} \supset \mathcal{R}_{s-1f}$  as defined in Theorem 3.8: both the homoclinic spots  $\gamma_{s-2f}(\xi)$  of Theorem 3.11 and the gaps  $\gamma_{g-2f}(\xi)$  of (upcoming) Theorem 3.13 merge with the heteroclinic cycle spanned by the standing 1-front  $\gamma_{s-1f}(\xi)$  of Theorem 3.8 and its symmetrical counterpart as  $\mathcal{S}_{s-2f} \cup \mathcal{S}_{g-2f}$  approaches  $\mathcal{R}_{s-1f}$ .

**Theorem 3.13.** *Let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{g-2f}$  and  $\Gamma_{g-2f} \subset \mathbb{R}^4$  be the singular (gap) skeleton constructed above. Then, there is for  $\varepsilon > 0$  sufficiently small a symmetric 2-front orbit*

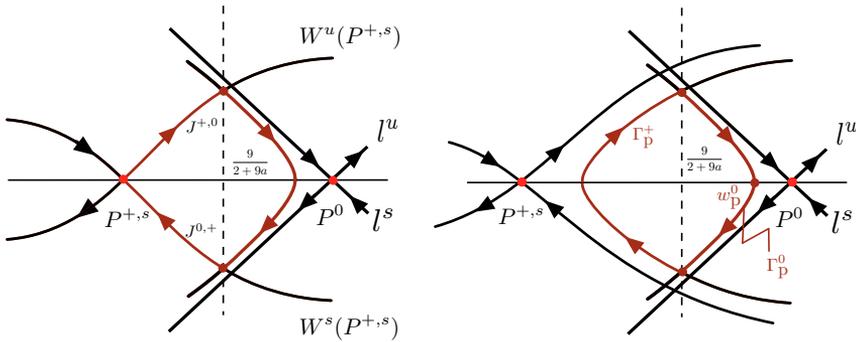
$\gamma_{g-2f}(\xi) = (b_{g-2f}(\xi), p_{g-2f}(\xi), w_{g-2f}(\xi), q_{g-2f}(\xi)) \subset W^u(P^{+,s}) \cap W^s(P^{+,s})$  of (3.7) with  $c = 0$  that merges with  $\Gamma_{g-2f}$  as  $\varepsilon \downarrow 0$ . The associated stationary pattern  $(B(x, t), W(x, t)) = (b_{g-2f}(x), w_{g-2f}(x))$  in (3.5) represents a stationary localized bare soil gap embedded in vegetation.

Of course the proof of this Theorem goes exactly along the lines of the proof of Theorem 3.11. The main difference between the cases of (stationary, symmetric, homoclinic) 2-front spots and (stationary, symmetric, homoclinic) 2-front gaps is that there cannot be periodic orbits on  $\mathcal{M}_\varepsilon^0$  – the slow reduced flow (3.20) on  $\mathcal{M}_0^0$  is linear – so that there cannot be any higher order localized gap patterns (as in Theorem 3.12 for localized spots). We refer to Figs. 3.1c and 3.15a for numerical observations of – (most likely) stable – localized 2-front spot and gap patterns in PDE (3.5).

**Remark 3.14.** *We refer to [91, 92] for studies of the process of a homoclinic 2-front orbit between a slow manifold  $\mathcal{M}_\varepsilon^1$  and a second slow manifold  $\mathcal{M}_\varepsilon^2$  detaching from  $\mathcal{M}_\varepsilon^2$  to become a slow-fast homoclinic to  $\mathcal{M}_\varepsilon^1$  that only makes 1 homoclinic excursion through the fast field (instead of 2 fast heteroclinic jumps between  $\mathcal{M}_\varepsilon^1$  and  $\mathcal{M}_\varepsilon^2$ ). The focus of [91] is on the (exchange of) stability between these 2 types of homoclinic patterns and the associated bifurcations – especially as localized stripes in 2 space dimensions. In the present work the situation is somewhat more involved than in [91, 92], since the skeleton structures as sketched in Figs. 3.11 and 3.12a must become asymptotically small in this transition – which is not necessary in the setting of [91, 92]. Notice that this implies that here the  $W$ -component of the homoclinic pulse becomes ‘small’ – a certain well-defined magnitude in  $\varepsilon$  – during this transition, but that this is not the case for the  $B$ -component since the orbit still has to make (almost) a full jump between  $\mathcal{M}_\varepsilon^0$  and  $\mathcal{M}_\varepsilon^+$ . See Figs. 3.14c, 3.15b and 3.16 in section 4.1.*

### 3.6 Spatially periodic multi-front patterns

A stationary, non-degenerate, symmetric homoclinic pulse solution of a (reversible) reaction-diffusion system (defined for  $x \in \mathbb{R}$ ) – such as the spots and gaps of Theorems 3.11 and 3.13 – must be the ‘endpoint’ (within phase space)



**Figure 3.12:** Sketches of 2 skeleton structures of multi-front patterns: (a) The stationary homoclinic 2-front gap pattern of Theorem 3.13. (b) The spatially periodic multi-front spot/gap pattern of Theorem 3.15 as (projected) closed orbit.

of a continuous family – a ‘band’ – of spatially periodic patterns (as the period/wavelength  $\rightarrow \infty$ ) – see for instance [39]. Systems (3.5) and (3.7) indeed have large families of spatially periodic solutions.

**Theorem 3.15.** *Let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f} \cup \mathcal{S}_{g-2f} \cup \mathcal{R}_{s-1f}$  (Theorems 3.11, 3.13 and 3.8) and let  $c = 0$ . Let there be a  $\rho \in \mathbb{R}$ ,  $\rho \neq 0$ , such that there is a solution of the reduced slow flow (3.20) on  $\mathcal{M}_0^0$  that connects  $(\rho, 9/(2+9a))$  to  $(-\rho, 9/(2+9a))$  over the branch  $\Gamma_\rho^0$  (by definition) and a solution of the (projected) reduced flow (3.22) on  $\mathcal{M}_0^+$  that connects  $(-\rho, 9/(2+9a))$  back to  $(\rho, 9/(2+9a))$  over  $\Gamma_\rho^+$  – see Fig. 3.12b. Then, for  $\varepsilon > 0$  sufficiently small, there is a periodic solution  $\gamma_{mf,\rho}(\xi) = (b_{mf,\rho}(\xi), \rho_{mf,\rho}(\xi), w_{mf,\rho}(\xi), q_{mf,\rho}(\xi))$  of (3.7) that merges in the limit  $\varepsilon \rightarrow 0$  with the skeleton structure spanned by  $\Gamma_\rho^0$ , the fast jump over  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  with  $w_h^+ = 9/(2+9a)$ ,  $q = -\rho$  (3.17),  $\Gamma_\rho^0$  and the fast jump back over  $W^u(\mathcal{M}_0^+) \cap W^s(\mathcal{M}_0^0)$  with  $w_h^- = 9/(2+9a)$ ,  $q = \rho$ . The associated stationary pattern  $(B(x, t), W(x, t)) = (b_{mf,\rho}(x), w_{mf,\rho}(x))$  in (3.5) represents a stationary spatially periodic multi-pulse spot/gap pattern.*

Note that if  $\gamma_{mf,\rho}(\xi)$  exists for a certain  $\rho^*$ , there clearly must be a neighborhood of  $\rho^*$  for which  $\gamma_{mf,\rho}(\xi)$  also exists: the periodic solutions of (3.5)/(3.7) indeed come in continuous families/bands [39]. Typically, there is a ‘subband’ of stable periodic patterns – see Figs. 3.1d and 3.17 for examples of numerically stable patterns  $(b_{mf,\rho}(x), w_{mf,\rho}(x))$  in (3.5).

**Proof of Theorem 3.15.** This proof can be set up very much along the lines of the proofs of similar results – the existence of spatially periodic patterns in the (generalized) Gierer-Meinhardt equation – in [44], therefore we restrict ourselves to the essential ingredients of the proof here.

### 3. The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation

98

The approach is similar to that of the proof of Theorem 3.11: we construct an orbit that intersects the plane  $\{p = q = 0\}$  and apply the reversibility symmetry (3.9) – with  $c = 0$ . However, unlike for the homoclinic orbits in Theorem 3.11, we do not ‘start out’ – as  $\xi \rightarrow -\infty$  – at the critical point  $P^0 \in \mathcal{M}_\varepsilon^0$ , but choose the initial condition of the orbit-to-be-constructed at the – exponentially short – interval  $\mathcal{I}_{\bar{b}} = \{p = q = 0, w = w_\rho^0, b = \bar{b} \in (0, \bar{b}_M)\}$ , where  $w_\rho^0$  is the midpoint of  $\Gamma_\rho^0$ , i.e.,  $(w_\rho^0, 0) = \Gamma_\rho^0 \cap \{q = 0\}$  and  $\bar{b}_M$  is exponentially small in  $\varepsilon$ , so that orbits  $\gamma_{\rho, \bar{b}}(\xi)$  with  $\gamma_{\rho, \bar{b}}(0) \in \mathcal{I}_{\bar{b}}$  remain exponentially close to  $\mathcal{M}_\varepsilon^0$  over  $\mathcal{O}(1)$  distances in  $b, p$  and  $\mathcal{O}(1/\varepsilon)$  ‘time’  $\xi$  (more precise, and as in the proof of Theorem 3.11:  $\bar{b} = \mathcal{O}(\exp(-\lambda_{f,0}(w_\rho^0)/\varepsilon)$  with  $\lambda_{f,0}(w_0) = \sqrt{1 - aw_0}$ , the unstable eigenvalue of the reduced fast flow (3.10) – with  $c = 0$  – associated to  $(0, 0)$ ). Consider the 2-dimensional ‘strip’  $\mathcal{T}_{\rho, \bar{b}}$  spanned by all  $\gamma_{\rho, \bar{b}}(\xi)$ ,  $\bar{b} \in (0, \bar{b}_M)$ : as it takes off from  $\mathcal{M}_\varepsilon^0$ , it is exponentially close to  $W^u(\mathcal{M}_\varepsilon^0)$  and thus intersects  $W^s(\mathcal{M}_\varepsilon^+)$  transversely along the orbit  $\gamma_{\rho, \bar{b}}(\xi)$  (the intersection  $\mathcal{T}_{\rho, \bar{b}} \cap W^s(\mathcal{M}_\varepsilon^+)$  is 1-dimensional and thus an orbit of (3.7)).

Clearly,  $W^s(\mathcal{M}_\varepsilon^+)$  cuts  $\mathcal{T}_{\rho, \bar{b}}$  into 2 parts – distinguished by  $\bar{b} \leq \bar{b}^*$  – due to the ‘fast saddle’ structure of  $\mathcal{M}_\varepsilon^+$  (as in the proof of Theorem 3.11). Orbits  $\gamma_{\rho, \bar{b}}(\xi) \subset \mathcal{T}_{\rho, \bar{b}}^r \subset \mathcal{T}_{\rho, \bar{b}}$  cross through  $\{p = 0\}$ , turn around (in their  $b$ -components) and return back towards  $\mathcal{M}_\varepsilon^0$ ; the  $b$ -components of orbits  $\gamma_{\rho, \bar{b}}(\xi) \subset \mathcal{T}_{\rho, \bar{b}} \setminus \mathcal{T}_{\rho, \bar{b}}^r$  – the complement of  $\mathcal{T}_{\rho, \bar{b}}^r$  – increase beyond  $\mathcal{M}_\varepsilon^+$ . It depends on the relative magnitudes of  $w_\rho^0$  and  $9/(2 + 9a)$  whether  $\mathcal{T}_{\rho, \bar{b}}^r$  is determined by  $\bar{b} \in (0, \bar{b}^*)$  or by  $\bar{b} \in (\bar{b}^*, \bar{b}_M)$ . If  $w_\rho^0 > 9/(2 + 9a)$  – as in Fig. 3.12 – orbits  $\gamma_{\rho, \bar{b}}(\xi)$  that take off ‘too soon’ – i.e., with  $\bar{b} > \bar{b}^*$  – have  $w > 9/(2 + 9a)$  at take off (Fig. 3.12). Since the unstable manifold  $W^u((0, 0))$  of the (planar) fast reduced system (3.10) with  $w_0 > 9/(2 + 9a)$  contains a closed homoclinic orbit, it follows that orbits  $\gamma_{\rho, \bar{b}}(\xi)$  with  $\bar{b} > \bar{b}^*$  follow such a homoclinic orbit through the fast field (at leading order in  $\varepsilon$ ). Hence, they pass through  $\{p = 0\}$  and turn back towards  $\mathcal{M}_\varepsilon^+$ :  $\mathcal{T}_{\rho, \bar{b}}^r$  is spanned by  $\gamma_{\rho, \bar{b}}(\xi)$  with  $\bar{b} \in (\bar{b}^*, \bar{b}_M)$ . For simplicity, we only consider this case (the arguments run along exactly the same lines in the case that  $w_\rho^0 < 9/(2 + 9a)$  and  $\mathcal{T}_{\rho, \bar{b}}^r$  is determined by  $\bar{b} \in (0, \bar{b}^*)$ ).

We can now copy the main (geometrical) argument of the proof of Theorem 3.11: if  $\gamma_{\rho, \bar{b}}(\xi) \subset \mathcal{T}_{\rho, \bar{b}}^r$  is too far removed from – but still exponentially close to –  $\gamma_{\rho, \bar{b}}(\xi) \subset W^s(\mathcal{M}_\varepsilon^+)$  – i.e., if  $\bar{b} - \bar{b}^*$  is too large – it will follow  $\mathcal{M}_\varepsilon^+$  – and thus  $\Gamma_\rho^+$  – over a relatively short distance (Fig. 3.12), take off again from  $\mathcal{M}_\varepsilon^+$  and thus cross through  $\{p = 0\}$  before reaching  $\{q = 0\}$ . By decreasing  $\bar{b} - \bar{b}^*$ , one can keep  $\gamma_{\rho, \bar{b}}(\xi)$  sufficiently long close to  $\mathcal{M}_\varepsilon^+$  that it first passes through  $\{q = 0\}$  before crossing  $\{p = 0\}$ : there is a value  $\bar{b} = b_\rho$  such that  $\gamma_{\rho, b_\rho}(L/2) \in \{p = q = 0\}$

for certain  $L > 0$  (in fact,  $L = \mathcal{O}(1/\varepsilon)$ ). It follows by (3.9) – with  $c = 0$  – that  $\gamma_{\text{mf},\rho}(\xi) = \gamma_{\rho,b_\rho}(\xi)$ , a periodic orbit with period  $L$ .  $\square$

As in the case of the homoclinic spots patterns – Theorem 3.12 – we may immediately conclude that there are countably many (families of) higher order periodic patterns if  $\Gamma_\rho^+$  (as defined in Theorem 3.15) is part of a periodic orbit on  $\mathcal{M}_0^+$  (determined by the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_\rho\}$ ) – see Fig. 3.11c. By steadily decreasing  $|\bar{b} - \bar{b}^*|$ , the orbit  $\gamma_{\rho,\bar{b}}(\xi) \subset \mathcal{T}_{\rho,\bar{b}}^+$  can be made to pass arbitrarily many times through  $\{q = 0\}$  before taking off from  $\mathcal{M}_\varepsilon^+$ .

**Corollary 3.16.** *Let  $(a, \Psi, \Phi, \Omega, \Theta) \in \mathcal{S}_{s-2f} \cup \mathcal{S}_{g-2f} \cup \mathcal{R}_{s-1f}$  such that (3.40) holds. Let  $\rho, \Gamma_\rho^0 \subset \mathcal{M}_0^0$  and  $\Gamma_\rho^+ \subset \mathcal{M}_0^+$  be as defined in Theorem 3.15, with  $\Gamma_\rho^+$  such that it is part of a closed orbit on  $\mathcal{M}_\varepsilon^+$  determined by the level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_\rho\}$ . Then, for  $\varepsilon > 0$  sufficiently small, there is a countable family of symmetric multi-front periodic orbits*

$\gamma_{\text{mf},\rho}^i(\xi) = (b_{\text{mf},\rho}^i(\xi), p_{\text{mf},\rho}^i(\xi), w_{\text{mf},\rho}^i(\xi), q_{\text{mf},\rho}^i(\xi))$  of (3.7) with  $c = 0$  ( $i = 1, 2, \dots$ ) that merges in the limit  $\varepsilon \rightarrow 0$  with the extended skeleton structure spanned by  $\Gamma_\rho^0$ , the fast jump over  $W^u(\mathcal{M}_0^0) \cap W^s(\mathcal{M}_0^+)$  with  $w_h^+ = 9/(2+9a)$ ,  $q = -\rho$  (3.17), the full closed orbit of  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}_\rho\}$  that contains  $\Gamma_\rho^+$  and the fast jump back over  $W^u(\mathcal{M}_0^+) \cap W^s(\mathcal{M}_0^0)$  with  $w_h^- = 9/(2+9a)$ ,  $q = \rho$ . The associated stationary patterns  $(B(x, t), W(x, t)) = (b_{\text{mf},\rho}^i(x), w_{s-2f}^i(x))$  in (3.5) are symmetric periodic spot/gap patterns with an increasing number of oscillations in the vegetated areas.

Finally, we note that the families of ‘higher order’ periodic patterns  $\gamma_{\text{mf},\rho}^i(\xi)$  are only the first of further – more complex – families containing periodic (and aperiodic) patterns of increasing complexity. We refer to [44] for the precise settings and proofs, here we only give a sketch of one specific example. However, this sketch provides the main ideas by which all further orbits may be constructed.

Let  $b_\rho^i \in (0, \bar{b}_M)$  be such that the  $i$ -th periodic pattern  $\gamma_{\text{mf},\rho}^i(\xi)$  of Corollary 3.16 is given by  $\gamma_{\rho,\bar{b}}(\xi)$  with  $\bar{b} = b_\rho^i$  (see the proof of Theorem 3.15). We can now choose  $\bar{b}$  so close to  $b_\rho^i$  that  $\gamma_{\rho,\bar{b}}(\xi) \subset \mathcal{T}_{\rho,\bar{b}}^+$  follows  $\gamma_{\text{mf},\rho}^i(\xi)$  along its  $i$  circuits over  $\mathcal{M}_\varepsilon^+$  – with  $i \geq 1$  – and its jump back to  $\mathcal{M}_\varepsilon^0$ . Since  $\bar{b} \neq b_\rho^i$ ,  $\gamma_{\rho,\bar{b}}(\xi)$  does not close as it passes along  $\mathcal{I}_{\bar{b}}$ , instead it keeps on following  $\gamma_{\text{mf},\rho}^i(\xi)$  as it makes its second jump towards  $\mathcal{M}_\varepsilon^+$ . By the approach of the proofs of Theorems 3.11 and 3.15, we can now tune  $\bar{b}$  so that it has its second take off from  $\mathcal{M}_\varepsilon^+$  precisely and that it passes through  $\{p = q = 0\}$  while following  $\Gamma_\rho^+$  (without making any further circuits over the periodic orbit on  $\mathcal{M}_\varepsilon^+$  that contains  $\Gamma_\rho^+$ ). It follows by the application of the reversibility symmetry (3.9) that for this value of  $\bar{b}$ ,  $\gamma_{\rho,\bar{b}}(\xi)$  is a symmetric periodic orbit that ‘starts’ at  $\Gamma_\rho^0 \subset \mathcal{M}_\varepsilon^0$ , jumps to  $\mathcal{M}_\varepsilon^+$  to make  $i$

circuits along  $\mathcal{M}_\varepsilon^+$ , jumps back again to  $\mathcal{M}_\varepsilon^0$ , follows  $\Gamma_\rho^0 \subset \mathcal{M}_\varepsilon^0$  to return again to  $\mathcal{M}_\varepsilon^+$  where it follows  $\Gamma_\rho^+$  and subsequently immediately jumps back again to  $\mathcal{M}_\varepsilon^0$  – from which it repeats the same path, etc.. Note that the associated periodic pattern in (3.5) consists of an alternating array of 2 different types of localized vegetation spots. Clearly, this procedure can be further refined to establish the existence of patterns containing arbitrary arrays of arbitrarily many different types of vegetation spots – under the conditions of Corollary 3.16.

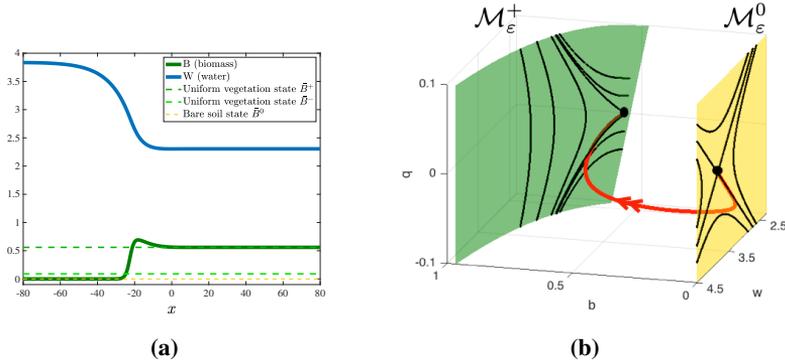
**Remark 3.17.** *We decided to focus in this chapter on stationary 2- and multi-front patterns. Of course, (3.5) also exhibits traveling multi-front patterns – see for instance Fig. 3.14d in which a vegetation spot travels towards a stationary, stable (and attracting) spot of the type established by Theorem 3.11. System (3.7) can also have homoclinic orbits to  $P^0$  for (certain specific values of)  $c \neq 0$ , i.e., vegetation spots may be traveling with constant speed (without changing shape). An approach along the lines of [45] indicates that bifurcations to traveling spots appear when the touch down manifold  $\mathcal{I}_{\text{down}}$  (Lemma 3.2) is tangent to a level set  $\{\mathcal{H}_0^+(w, q) = \mathcal{H}\}$  of the slow reduced flow on  $\mathcal{M}_0^+$  with  $c = 0$  at the (non-transversal) intersection  $\mathcal{I}_{\text{down}} \cap \{\mathcal{H}_0^+(w, q) = \mathcal{H}\}$  (recall that  $\mathcal{I}_{\text{down}}$  is parameterized by  $c$ ). A similar property holds for bifurcations of stationary spatially periodic multi-front patterns into traveling spatially periodic (wave train) patterns. A simple investigation of the relative orientations of  $\mathcal{I}_{\text{down}}$  and the various possible phase plane configurations of the slow reduced flow on  $\mathcal{M}_0^+$  shows that there indeed are parameter combinations  $(a, \Psi, \Phi, \Omega, \Theta)$  at which these bifurcations into traveling 2-/multi-fronts must occur. This bifurcation may have relevant ecological implications, nevertheless, we refrain from going into the details here (and leave this to future work) – see also section 4.2.*

## 4 Simulations and discussion

### 4.1 Simulations

The motivation for the numerical simulations presented in this section is threefold: 1) to illustrate some of the analytic results of the previous sections (without doing a systematic search for all constructed patterns) 2) to give a brief outlook beyond the worked out analysis to solution types that may be constructed by the geometric set-up developed here and, finally, 3) to give a flavor of the rich dynamics that PDE (3.1)/(3.5) exhibits. All numerical simulations have been carried using MATLAB's 'pdepe' routine. The corresponding parameter settings are specified in the captions of the figures. Almost all figures show a snapshot in time of the spatial profile of the PDE solution/pattern after it converged to a stationary or uniformly traveling solution.

The opening figure of this section, Fig. 3.13a, can be seen as a binding ele-

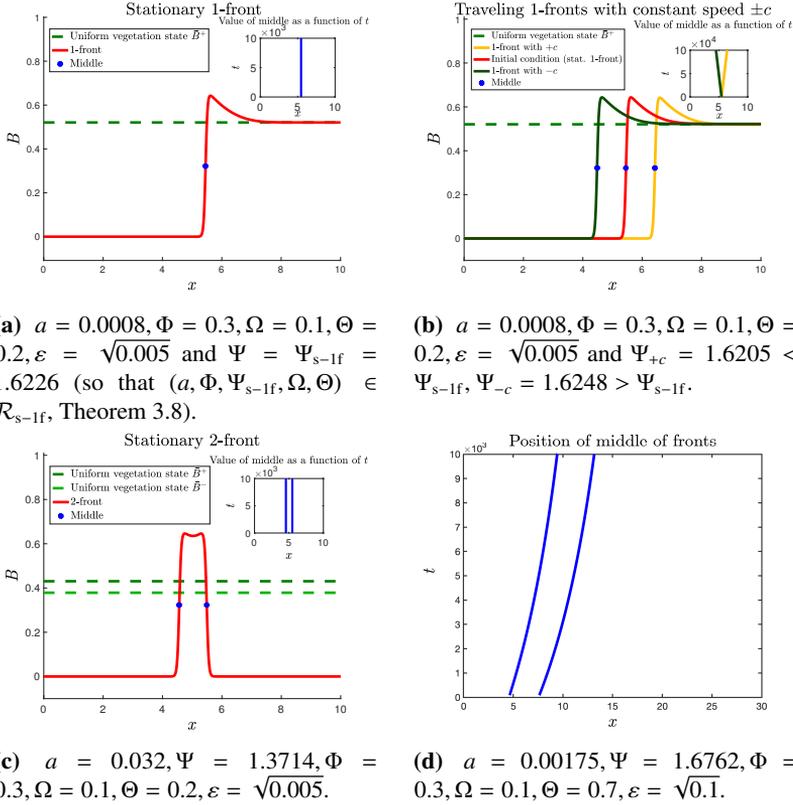


**Figure 3.13:** (a) Spatial profiles of the  $B$ - and  $W$ -components of a traveling front solution of the original, unscaled model (3.1) – corresponding to a heteroclinic orbit of (3.7) as established by Theorem 3.4 – together with the 3 background states. (b) The profile from (a) as a projection in  $(b, w, q)$ -space where  $q$  was computed by numerical differentiation. *Parameter settings* in (3.1):  $P = 180, \Lambda = 0.9, K = 0.4, E = 18, M = 15, N = 15, R = 0.7, \Gamma = 12, D_W = 150, D_B = 1.2$ , corresponding to  $\varepsilon^2 = 0.008, a \approx 0.187, \Psi \approx 3.84, \Phi = 1, \Omega \approx 0.235, \Theta \approx 1.71$  in (3.5).

ment between the chapters that have a more ecological emphasis and motivated the present work – see [174, 175] and the references therein – and the analysis here. It displays a traveling front solution as established by Theorem 3.4 for a parameter regime comparable to the one from [175] (with slight adjustment in the parameters to compensate for the choice of a 1-D model – 3.A). This profile is then shown in Fig. 3.13b as a projection in  $(b, w, q)$ -space to illustrate that it indeed starts on the slow manifold  $\mathcal{M}_\varepsilon^0$  and then jumps to the  $\mathcal{M}_\varepsilon^+$  slow manifold. As established by the analysis, the solution first follows the unstable manifold associated to the bare soil state  $(\bar{B}^0, \bar{W}^0)$  (as a solution of (3.7)), makes a fast excursion through the fast field to then touch down on  $\mathcal{M}_\varepsilon^+$  following the stable manifold associated to the uniform vegetation state  $(\bar{B}^+, \bar{W}^+)$ . Note, of course, that this figure contains two approximations: first, the manifold  $\mathcal{M}_\varepsilon^+$  is only accurate up to second order in  $\varepsilon$  – see section 2.4 – while the flows on  $\mathcal{M}_\varepsilon^0$  and  $\mathcal{M}_\varepsilon^+$  are computed numerically (using MATLAB routines).

As demonstrated in section 3, heteroclinic 1-front orbits can occur both as traveling – Theorem 3.4 – or stationary patterns – Theorem 3.8. We confirm this numerically in Figs. 3.14a and 3.14b. Note that these fronts may either represent the retreat of vegetation by the invasion of the bare soil state into the homogeneous vegetation state –  $c > 0$  – or the expansion of a (homogeneously) vegetated area into the bare soil state –  $c < 0$ . In fact, in order to find the stationary 1-front, we need to tune a single parameter –  $\Phi$  in the statement of

### 3. The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation



**Figure 3.14:** (a) A heteroclinic stationary 1-front pattern of (3.5); (b) Two traveling 1-front patterns connecting the bare soil state to a homogeneous vegetation state, one invading the bare soil ( $c < 0$ ), the other the vegetation state ( $c > 0$ ); (c) A homoclinic stationary 2-front spot pattern; (d) Evolution of the middles of the 2 interacting fronts of an evolving 2-front pattern.

Theorem 3.8 and  $\Psi$  in Fig. 3.14 – to the border point between the ranges of left-traveling and right-traveling 1-fronts (the ‘Maxwell point’ [14] described by the co-dimension 1 set  $\mathcal{R}_{s-1f}$  in Theorem 3.8).

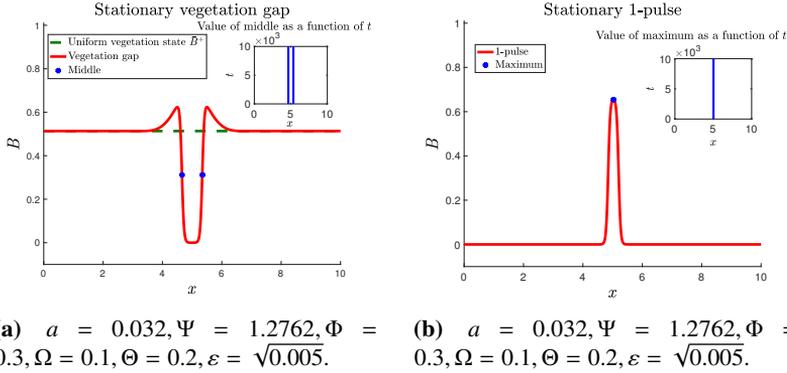
The existence of the homoclinic stationary 2-front pattern depicted in Fig. 3.14c was established in Theorem 3.11. Note that the level of vegetation on the plateau that determines the spot remains relatively far away from the value  $\bar{B}^+$  of the uniform vegetation state ( $\bar{B}^+, \bar{W}^+$ ). This is caused by the fact that the homoclinic orbit associated to the spot pattern follows an orbit on the slow manifold  $\mathcal{M}_\varepsilon^+$  of

(3.7) that does not approach the critical point associated to  $(\bar{B}^+, \bar{W}^+)$  on  $\mathcal{M}_\varepsilon^+$  – see the sketch of the skeleton structure in Fig. 3.11. From the ecological point of view, a vegetation spot benefits from soil water diffusion from the adjacent water-rich bare soil areas – see the  $W$ -profile in Fig. 3.13a – besides direct rainfall, and therefore has higher biomass density as compared with uniform vegetation. This also explains (in ecological terms) why the biomass density at the edge of a front, a spot or a gap is higher – a property also exhibited by ‘fairy circles’ [168, 175]. See also the upcoming discussion below of 2-front vegetation gaps (Fig. 3.15a). In Fig. 3.14d, we show the dynamics of the 2 interacting fronts of an evolving 2-front pattern: the distance between the fronts slowly increases while it settles into a stationary standing spot – see Remark 3.17 and the discussion in section 4.2.

One of the original motivations to analyze far-from-equilibrium patterns in the Gilad et al. model, was to gain a fundamental understanding of ‘fairy circles’ – a somewhat subtle phenomenon (for instance) observed in western Namibia [168, 175]. The homoclinic stationary 2-front gap patterns of (3.5) established by Theorem 3.13 and shown in Fig. 3.15a indeed have the strongly localized nature of observed fairy circles. Moreover, the spot/gap patterns of Theorem 3.15 represent the observed (nearly) periodic arrays of fairy circles (see Fig. 3.17a and notice that the ratio between the lengths of the vegetated state and the bare soil patches typically varies from 0 to  $\infty$  in this family (section 3.6)). As noted, fairy circle gap patterns have an excess of vegetation at the edge of the gap as distinctive feature – see for instance the images in [168, 175]. In mathematical terms, this means that the connecting fronts are non-monotonous. In the context of the present model, this non-monotonicity is caused by the orientation and curvature of the slow manifold  $\mathcal{M}_\varepsilon^+$  relative to the path traced by the gap pattern over  $\mathcal{M}_\varepsilon^+$  – see the projection in Fig. 3.16a for a representation of this ‘geometrical mechanism’ for spot patterns. We refer to [55] for a 1-component model in which the non-monotonicity of the fronts originates from nonlocal effects.

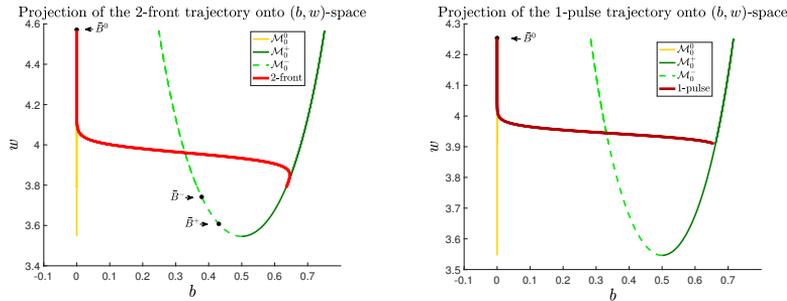
In section 3.5 – and especially in Remark 3.14 – we discussed the bifurcation of the homoclinic slow-fast-slow-fast-slow 2-front spot pattern of Theorem 3.11 into a homoclinic slow-fast-slow pulse pattern as it ‘detaches’ from  $\mathcal{M}_\varepsilon^+$ . Such a (numerically stable) ‘detached’ spot pattern of pulse type is shown Fig. 3.15b. In Fig. 3.16, the detachment process is shown by projections of both the 2-front spot pattern of Fig. 3.14c and the pulse pattern of Fig. 3.15b on the  $(w, b)$ -plane: as the parameter  $\Psi$  – which is linearly related to the rainfall  $P$  in the original model (3.6) – is decreased below a critical value  $\Psi_* = 1.2952$  the vegetated plateau disappears and the 2-front spot solution transforms into a 1-pulse solution. Note that this 1-pulse solution is of the ‘classical’ Klausmeier-Gray-Scott (and/or Gierer-Meinhardt [41]) type already discussed in the introduction of section 3.5: its existence may be established by the methods of [41, 46] and the

### 3. The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation



**Figure 3.15:** (a) A stationary homoclinic vegetation gap pattern (of fairy circle type) that is asymptotic to the stable homogeneous vegetation state  $(\bar{B}^+, \bar{W}^+)$ . (b) A stationary homoclinic spot solution of (classical) 1-pulse Gierer-Meinhardt/Gray-Scott type.

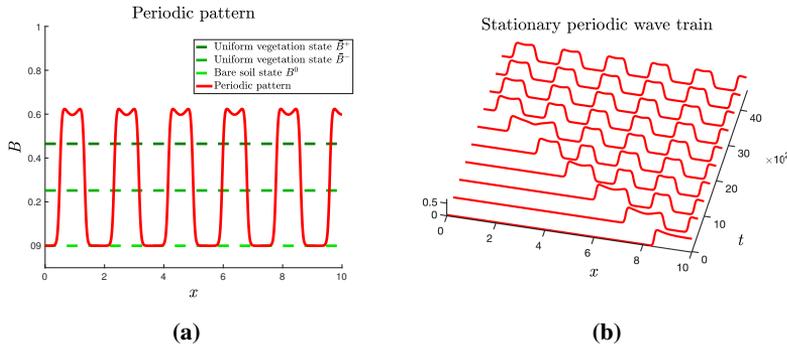
references therein. We also found that the spatially periodic spot/gap patterns



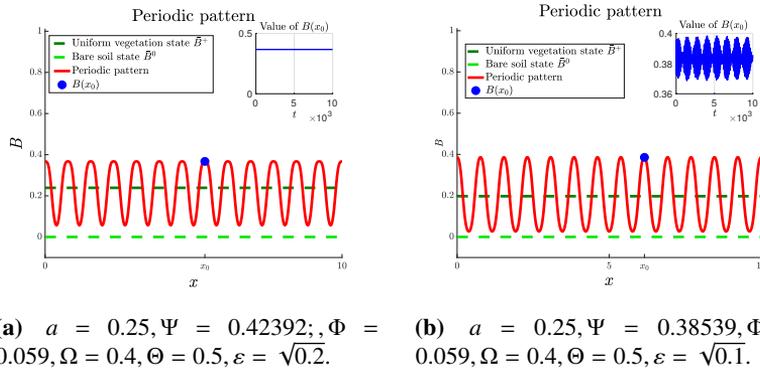
**Figure 3.16:** Projections into the  $(w, b)$ -plane of the homoclinic 2-front and 1-pulse solutions of Figs. 3.14c and 3.15b together with the slow manifolds  $\mathcal{M}_\varepsilon^0$  and  $\mathcal{M}_\varepsilon^\pm$ . Note that the trajectories are symmetric around the middle of the 2-front/1-pulse – due to the reversibility symmetry (3.9) of (3.5)/(3.7) – which results in the red branches shown.

of Theorem 3.15 may have quite a large domain of attraction: Fig. 3.17b shows the evolution of a traveling vegetation front into the bare soil state that leaves behind a spatially periodic spot/gap pattern – Fig. 3.17a. This behavior may possibly be related to the existence of a Turing bifurcation – see Remark 2.2 – of the uniform vegetation state and calls for further studies. Finally, we show in Fig. 3.18a such a numerically obtained, almost sinusoidal, small amplitude Turing pattern that bifurcated from a (destabilized) uniform vegetation state and note that there are paths through parameter space on which this Turing pattern evolves into a (periodic) multi-pulse pattern – built from homoclinic 1-pulses of Gierer-Meinhardt/Gray-Scott type as in Fig. 3.15b and typically observed in

Klausmeier-type models [133, 146, 155] – that subsequently touches down on  $\mathcal{M}_\varepsilon^+$  like the solitary pulses of Figs. 3.15b and 3.14c, to indeed take the shape of the periodic fairy circle-type spot/gap pattern of Theorem 3.15 and Fig. 3.17a. By further tuning parameters it may also happen that the stationary, spatially periodic, Turing pattern undergoes a Hopf bifurcation (in time), resulting in an oscillating pattern that is periodic both in space and in time – see Fig. 3.18b. (Note, however, that it is not clear whether this may occur for ecologically feasible parameters – see [151].)



**Figure 3.17:** (a) The standing asymptotic (for  $t \rightarrow \infty$ ) spatially periodic spot/gap pattern generated by the invasion dynamics of Fig. 3.17b. (b) A time/space plot of a vegetation front that invades the bare soil state and leaves the spatially periodic pattern of Fig. 3.17a behind. In both plots:  $a = 0.032, \Psi = 1.619, \Phi = 0.3, \Omega = 0.1, \Theta = 0.5, \varepsilon = \sqrt{0.01}$ .



(a)  $a = 0.25, \Psi = 0.42392; \Phi = 0.059, \Omega = 0.4, \Theta = 0.5, \varepsilon = \sqrt{0.2}$ . (b)  $a = 0.25, \Psi = 0.38539, \Phi = 0.059, \Omega = 0.4, \Theta = 0.5, \varepsilon = \sqrt{0.1}$ .

**Figure 3.18:** (a) A small amplitude stationary spatially periodic solution generated by a Turing bifurcation (see Remark 2.2. (b) A pattern that is periodic in space and time that appeared by decreasing  $\Psi$  from the Turing pattern of Fig. 3.18a through a Hopf bifurcation (in time).

**Remark 4.1.** *The basic configurations shown in Fig. 3.1 have been obtained by the same procedures as used in this section, with the following parameter settings. The traveling 1-front:  $a = 0.0008, \Psi = 1.6248, \Phi = 0.3, \Omega = 0.1, \Theta = 0.2, \varepsilon = \sqrt{0.005}$ ; the stationary homoclinic 2-front vegetation spot:  $a = 0.032, \Psi = 1.3714, \Phi = 0.3, \Omega = 0.1, \Theta = 0.2, \varepsilon = \sqrt{0.005}$ ; the stationary homoclinic 2-front vegetation gap:  $a = 0.032, \Psi = 1.2762, \Phi = 0.3, \Omega = 0.1, \Theta = 0.2, \varepsilon = \sqrt{0.005}$ ; the stationary spatially periodic multi-front:  $a = 0.032, \Psi = 1.619, \Phi = 0.3, \Omega = 0.1, \Theta = 0.5, \varepsilon = \sqrt{0.01}$ .*

## 4.2 Discussion

Of course, the potential relevance of the various singular slow/fast patterns constructed in this chapter is ultimately determined by their stability as solutions of PDE (3.1)/(3.5). In general, this is a seriously hard problem to study analytically. However, the singularly perturbed nature of the patterns considered here enables us to explicitly and rigorously analyze the (spectral) stability of the constructed (multi-)front patterns. In fact, the explicit ‘control’ we established on the slow-fast structure of the (multi-)fronts provides the perfect (and necessary) starting point for a spectral stability analysis along the lines of (for instance) [20, 43] and [31] (for the spatially periodic patterns). This is especially the case for the basic front/spot/gap/periodic patterns of Theorems 3.4, 3.8, 3.11, 3.13, 3.15 shown in Fig. 3.1.

However, the question whether the non-basic, ‘higher order’ patterns (for instance) sketched in Fig. 3.2 can be stable also requires novel mathematical insights and methods. All constructed higher order patterns involve the existence of persisting periodic solutions on the slow manifold  $\mathcal{M}_\varepsilon^+$  – see Theorems 3.5, 3.6, 3.9, 3.12 and Corollary 3.16. Therefore, the structure of the spectrum associated to the stability of the higher order patterns essentially depends on the preliminary question about the spectrum and stability of the persistent periodic solutions on the slow manifold of Theorem 2.4 – and especially of their homoclinic (or heteroclinic) limits also considered in Theorem 2.4. In fact, this issue is not (at all) specific for the explicit model here. We claim that higher order patterns of the type sketched in Fig. 3.2 will generically appear as potentially stable solutions in a fully general class of singularly perturbed reaction-diffusion models that includes (3.5),

$$\begin{cases} U_t &= U_{xx} + F(U, V), \\ V_t &= \frac{1}{\varepsilon^2} V_{xx} + G(U, V). \end{cases} \quad (3.60)$$

By going into a traveling framework – and thus introducing  $\xi = x - ct, U(x, t) = u(\xi), V(x, t) = v(\xi), p(\xi) = u_\xi(\xi), q = v_\xi(\xi)/\varepsilon$  as in section 1 – (3.60) is reduced into the 4-dimensional form of (3.7). By taking the  $\varepsilon \rightarrow 0$  limit, we find that the 2-dimensional (reduced) slow manifolds are determined by  $F(v_0, u) = 0$  (and

$p = 0, (v_0, q_0) \in \mathbb{R}^2$  – see section 2 – which generically determines  $J \geq 1$  branches, locally given by graphs,

$$\mathcal{M}_0^j = \{(u, p, v, q) \in \mathbb{R}^4 : u = f_j(v), p = 0\}, \quad j = 1, 2, \dots, J,$$

with  $f_j(v)$  such that  $F(f_j(v), v) \equiv 0$  (cf. (3.12) and note that  $J = 3$  for (3.5)). For those (parts of)  $\mathcal{M}_0^j$  that are normally hyperbolic,  $\mathcal{M}_0^j$  persists as  $\mathcal{M}_\varepsilon^j$ , that is approximately given by,

$$\mathcal{M}_\varepsilon^j = \{(u, p, v, q) \in \mathbb{R}^4 : u = f_j(v) + \varepsilon c q u_1^j(v) + \mathcal{O}(\varepsilon^2), p = \varepsilon q p_1^j(v) + \mathcal{O}(\varepsilon^2)\}$$

with

$$u_1^j(v) = -f_j'(v) \frac{\partial F}{\partial u}(f_j(v), v), \quad p_1^j(v) = f_j'(v).$$

(cf. (3.36), (3.37)). Thus, completely analogous to the analysis in section 2.4, we find that the slow flow on a persisting, normally hyperbolic 2-dimensional slow manifold  $\mathcal{M}_\varepsilon^j$  is given by a planar Hamiltonian system perturbed by a nonlinear friction term,

$$v_{XX} + G(f_j(v), v) + \varepsilon c q \left[ 1 - \frac{\partial G}{\partial u}(f_j(v), v) u_1^j(v) \right] = \mathcal{O}(\varepsilon^2),$$

with  $X = \varepsilon \xi$  (cf. (3.38)). Typically, the unperturbed  $\varepsilon \rightarrow 0$  limit  $v_{XX} + G(f_j(v), v)$  – i.e., the reduced slow flow on  $\mathcal{M}_0^j$  – is nonlinear and has families of periodic solutions and homoclinic or heteroclinic orbits to critical points on  $\mathcal{M}_\varepsilon^j$  that correspond to (potentially stable [39]) homogeneous background states  $(U(x, t), V(x, t)) \equiv (\bar{U}, \bar{V})$  of PDE (3.60) – as is the case for (3.22) on  $\mathcal{M}_0^+$ . Thus, indeed, the situation is completely similar to that of section 2.4: using Melnikov-type arguments persistence results equivalent to Theorem 2.4 may be deduced, also in the present general setting. The geometric framework of orbits ‘jumping up and down’ between two (normally hyperbolic) slow manifolds  $\mathcal{M}_\varepsilon^j$  and  $\mathcal{M}_\varepsilon^k$  presented in section 3.1 is based on the persistence of both the stable and unstable manifolds  $W^{s,u}(\mathcal{M}_\varepsilon^{j,k})$  of  $\mathcal{M}_\varepsilon^{j,k}$  and thus of the intersections  $W^u(\mathcal{M}_\varepsilon^j) \cap W^s(\mathcal{M}_\varepsilon^k)$  and  $W^s(\mathcal{M}_\varepsilon^j) \cap W^u(\mathcal{M}_\varepsilon^k)$ . Therefore, we may use the arguments, methods and insights of section 3 to deduce the equivalents of the ‘higher order’ existence Theorems 3.5, 3.6, 3.9, 3.12 and Corollary 3.16 in the setting of general system (3.60). Moreover, this also implies that bifurcation scenarios as sketched in Fig. 3.9 appear generically (where we notice that the sketch in Fig. 3.9 was just a first example – many other scenarios may occur). In fact, the geometrical setting allows us to (for instance) explicitly establish the existence of heteroclinic networks of orbits jumping between various slow manifolds  $\mathcal{M}_\varepsilon^j$  and (slowly) following periodic orbits on  $\mathcal{M}_\varepsilon^j$  in between its fast jumps – like the networks considered in [120, 121] and the references therein. Thus, the above noted preliminary (and essential) issue of the spectrum associated to the stability of the persisting periodic and homoclinic solutions on  $\mathcal{M}_\varepsilon^+$  of Theorem

2.4 also has a fully general – and thus fundamental – counterpart, with a similar relevance for the higher order patterns (almost) heteroclinic to these orbits. In other words, insight in the spectrum associated to the stability of the orbits on  $\mathcal{M}_\varepsilon^j$  established by a generalization of Theorem 2.4 for (3.60) is expected to yield explicit insight in the stability and bifurcations of the higher order patterns in PDE (3.60) established by the generalizations of Theorems 3.5, 3.6, 3.9, 3.12, etc. and the subsequent more complex ‘networks’.

This will be the subject of future work, both in the setting of explicit system (3.5) – which will also include a systematic numerical search for the higher order patterns sketched in Fig. 3.2 – and in the general setting of (3.60).

To optimally embed the present analysis in the ecological context, we first need to obtain insight in the ranges of the (scaled) parameters  $(a, \Psi, \Phi, \Omega, \Theta)$  of (3.5) that may correspond to ecologically realistic settings of the unscaled model. Finding such parameter ranges is possible (more than in the Klausmeier model) since (3.1) is directly linked to the more elaborate 3-component model of Gilad et al. [62] – [175], 3.A – and thus to concrete underlying ecological mechanisms [63, 106, 134]. A crucial question for the potential ecological relevance of the above discussed higher order patterns is whether there are realistic values of  $(\Lambda, K, E, M, P, N, R, \Gamma)$  for which there are 2 critical points on  $\mathcal{M}_\varepsilon^+$ , i.e., for which  $\mathcal{C}^2 - 4\mathcal{A}\mathcal{D} > 0, \mathcal{C} < 0, \mathcal{D} > 0$  (section 2.3) – where  $(\mathcal{A}, \mathcal{C}, \mathcal{D})$  is related to  $(\Lambda, K, E, M, P, N, R, \Gamma)$  in a rather nonlinear fashion by (3.3), (3.5), (3.6) (3.23), (3.24). Naturally, this will be part of our upcoming work on (3.5).

Each of the higher order invasion fronts established by Theorem 3.6 and sketched in Figs. 3.2a and 3.2b travels with a different speed – in fact, the (discrete) family may even limit on a stationary front pattern (Remark 3.10). Thus, when stable, these invasion fronts may introduce the possibility of slowing down gradual desertification. Moreover, stationary multi-front patterns may bifurcate into traveling patterns with the same structure – see Remark 3.17 for a brief sketch of the underlying geometrical mechanism. When stable, the appearance of such traveling multi-front patterns – either localized spots or spatially periodic wave trains – may have a similar ecological interpretation and relevance: localized vegetated states may even reverse desertification by invading bare soil – see [172, 174, 176]. Together, the various traveling 1-front patterns and traveling multi-spots form an interacting group of invasion patterns within the transition zone between the bare soil state and a homogeneous vegetation state; in principle all entities in this group travel with different speed. Understanding pattern formation in this zone – and especially also understanding the translation and/or expansion of this zone in terms of the parameters in the model – may have a direct ecological significance. In mathematical terms, such a study may also be performed by a front interaction analysis along the lines of [25, 26, 157] –

although the dynamics generated by (3.5) in this ecological transition zone is expected to be richer than that of the generalized FitzHugh-Nagumo model considered there.

To truly obtain ecological relevance, we must consider the model in 2 space dimensions. Clearly, the extension to more than 1 space dimension does pose fundamental challenges, moreover 2-dimensional systems show much richer dynamical behaviors associated with propagating fronts – see for instance [73, 74]. However, the results obtained here form a foundation upon which aspects of the step from 1 to 2 space dimensions can be taken – see for instance [143, 159, 162, 165] and the references therein. By extending the patterns constructed here trivially in the second spatial direction, the above mentioned stability analysis can be directly extended to include the stability (and bifurcations) of planar (multi-)fronts/interfaces (where we for simplicity neglect the (technical) fact that (3.5) takes a somewhat different form in  $\mathbb{R}^2$ , [106, 175], 3.A). Unlike in the extended-Klausmeier model [133, 143], localized stripes are of 2-front type and may thus be expected to possibly be stable – see [9] for a rigorous treatment in a generalized Klausmeier type model (posed on a sloped terrain without a diffusion term for the water component – like the original Klausmeier model [90]). Naturally, the interfaces will evolve and their curvature driven dynamics may be studied analytically along the lines of [111]. Especially in the above discussed multi-front transition region between bare soil and homogeneous vegetation, the ecosystem dynamics generated by the model may be very rich and complex – see for instance [73–75].

As a final direction of possible future research, we note that our results may be used to establish the existence – and later stability – of localized patterns in the original 3-component model of Gilad et al. [62]. Since (3.1) and thus (3.5) – is obtained from the nonlocal, 3-component model of Gilad et al. (see (3.A.61)) in a systematic way – i.e., by taking several limits ([106, 175], 3.A) – it may be expected that it is possible to establish the persistence of patterns constructed here into the nonlocal, 3-component setting, especially since these patterns are constructed geometrically through transversal intersections of invariant manifolds. Once again, this is interesting and relevant both from mathematical and ecological point of view: (asymptotically) small nonlocal and topographical terms may have a significant effects, even on the most simple – ‘basic’ – (vegetation) patterns exhibited by a model [10, 49].



# Appendices

---

## 3.A Derivation of the model equations in one spatial dimension

We follow [175] to briefly show how model (3.1) is derived from the original model introduced in [63] and given by

$$\begin{cases} \partial_T B = G_B B(1 - B/K) - MB + D_B \nabla^2 B, \\ \partial_T W = IH - N(1 - RB/K)W - G_W W + D_W \nabla^2 W, \\ \partial_T H = P - IH + D_H \nabla^2 (H^2) + 2D_H \nabla H \cdot \nabla Z + 2D_H H \nabla^2 Z, \end{cases} \quad (3.A.61)$$

where

$$G_B(\mathbf{X}, T) = \Lambda \int_{\Omega} G(\mathbf{X}, \mathbf{X}', T) W(\mathbf{X}', T) d\mathbf{X}' \quad (3.A.62a)$$

$$G_W(\mathbf{X}, T) = \Gamma \int_{\Omega} G(\mathbf{X}', \mathbf{X}, T) B(\mathbf{X}', T) d\mathbf{X}' \quad (3.A.62b)$$

$$G(\mathbf{X}, \mathbf{X}', T) = \left( \frac{1}{\sqrt{2\pi S_0^2}} \right)^2 \exp \left[ -\frac{|\mathbf{X} - \mathbf{X}'|^2}{2S_0^2(1 + EB(\mathbf{X}, T))^2} \right] \quad (3.A.62c)$$

$$I = A \frac{B(\mathbf{X}, T) + Qf}{B(\mathbf{X}, T) + Q} \quad (3.A.62d)$$

with  $\mathbf{X} = (X, Y)$  the spatial coordinates of the 2-dimensional system. The last equation in (3.A.61) describes overland water flow with  $H$  being the height of a thin layer of surface water above ground level given by the topography function  $Z$ . We consider the case of a flat terrain, for which  $Z = \text{constant}$ , and of high infiltration rates  $I$ , both in bare soil and vegetated areas (no infiltration contrast), for which  $I$  can be assumed to be independent of  $B$ . Both conditions are met in the Namibian fairy-circles ecosystems that consist of sandy soil. Since  $H$  varies on time scales much shorter than those of  $W$  and  $B$ , these conditions imply fast equilibration of surface water at  $H = P/I$ . Insertion of this equilibrium value in the equation for  $W$  results in the elimination of the surface water equation.

A further simplification we make is related to the nonlocal forms of the biomass growth rate,  $G_B$ , and the water uptake rate,  $G_W$ , in (3.A.61). We assume, consistently with the plant species in the Namibian fairy-circles ecosystems, that the roots, described by the root kernel  $G(\mathbf{X}, \mathbf{X}', T)$ , are laterally confined. We employ this assumption by taking the lateral root extension of a seedling,  $S_0$ , to be very small. Using the limit  $S_0 \rightarrow 0$  in the integrals in (3.A.62c) we obtain, for a 1-dimensional system, the simpler algebraic expressions

$$\begin{aligned} G_B(X, T) &= \Lambda \int_{\Omega} \lim_{S_0 \rightarrow 0} \frac{1}{\sqrt{2\pi S_0^2}} \exp\left[-\frac{|X - X'|^2}{2S_0^2(1 + EB(X, T))^2}\right] B(X', T) dX' \\ &= \Lambda (1 + EB(X, T)) B(X, T). \end{aligned} \quad (3.A.63)$$

Similarly,

$$G_W(X, T) = \Gamma (1 + EB(X, T)) W(X, T). \quad (3.A.64)$$

Inserting these expressions in (3.A.61) we obtain the 2-component model (3.1). Finally, we note that in [175] this reduction was performed in 2 space dimensions and that the general  $n$ -dimensional situation is considered in [106].

### 3.B The derivation of the scaled model

Introducing the scalings (3.2) into (3.1) yields,

$$\begin{cases} \alpha \delta B_t = \alpha \beta \Lambda W B (1 - \alpha B/K) (1 + \alpha EB) - \alpha M B + \alpha \gamma^2 D_B B_{xx}, \\ \beta \delta W_t = P - N \beta W (1 - \alpha R B/K) - \alpha \beta \Gamma W B (1 + \alpha EB) + \beta \gamma^2 D_W W_{xx}, \end{cases} \quad (3.B.65)$$

which can be brought into the form,

$$\begin{cases} \frac{\delta K}{\alpha^2 \beta \Lambda E} B_t = \left( \frac{K}{\alpha^2 E} W - \frac{MK}{\alpha^2 \beta \Lambda E} \right) B + \frac{K}{\alpha E} \left( E - \frac{1}{K} \right) W B^2 - W B^3 + \frac{\gamma^2 K}{\alpha^2 \beta \Lambda E} D_B B_{xx} \\ \frac{\delta K}{\alpha^2 \beta \Lambda E} W_t = \frac{PK}{\alpha^2 \beta^2 \Lambda E} - \frac{K}{\alpha^2 \beta \Lambda E} \left[ N \left( 1 - \frac{\alpha R}{K} B \right) + \alpha \Gamma B (1 + \alpha EB) \right] W + \frac{\gamma^2 K}{\alpha^2 \beta \Lambda E} D_W W_{xx}. \end{cases} \quad (3.B.66)$$

By choosing  $\delta$  and  $\gamma$  as in (3.3), we arrive at,

$$\begin{cases} B_t = \left( \frac{K}{\alpha^2 E} W - \frac{MK}{\alpha^2 \beta \Lambda E} \right) B + \frac{1}{\alpha} \left( K - \frac{1}{E} \right) W B^2 - W B^3 + B_{xx}, \\ W_t = \frac{PK}{\alpha^2 \beta^2 \Lambda E} - \frac{K}{\alpha^2 \beta \Lambda E} \left[ N \left( 1 - \frac{\alpha R}{K} B \right) + \alpha \Gamma B (1 + \alpha EB) \right] W + \frac{D_W}{D_B} W_{xx}. \end{cases} \quad (3.B.67)$$

Next, we use our freedom in  $\alpha$  and  $\beta$  to simplify the  $B$ -equation and scale the factors of the  $B$ - and  $W B^2$ -terms (to  $-1$  and to  $+1$  respectively) – which is

achieved by the choices in (3.3),

$$\begin{cases} B_t = \left( \frac{KE}{(KE-1)^2} W - 1 \right) B + WB^2 - WB^3 + B_{xx}, \\ W_t = \frac{\alpha^2 P \Lambda E}{M^2 K} - \left[ N \left( 1 - \frac{\alpha R}{K} B \right) + \alpha \Gamma B (1 + \alpha EB) \right] \frac{W}{M} + \frac{D_W}{D_B} W_{xx}. \end{cases} \quad (3.B.68)$$

This is equivalent to (3.5) by definitions (3.4) and (3.6).

Note that our choice to scale the factor of the term  $WB^2$  in the  $B$ -equation to +1 implies – together with the (implicit, natural) assumption that  $\tilde{B}$  and  $B$  have the same signs (3.2) – that we have chosen to consider  $EK > 1$ . Of course, it may happen that  $0 < EK \leq 1$ . In these cases, either the term  $WB^2$  disappears from the equation – in the ‘non-generic’ case  $EK = 1$  – or its pre-factor can be scaled to  $-1$ . All of the analysis in this work can also be performed for  $EK \leq 1$ , without any conceptual differences. However, we chose to focus of  $EK > 1$  – and thus on a  $+WB^2$  term in (3.5) – to not further complicate the necessary ‘algebra’.

### 3.C Lemma 2.6 and the Bogdanov-Takens bifurcation scenario

A planar ODE of the form

$$\begin{cases} y_{\tilde{x}} = z, \\ z_{\tilde{x}} = \beta_1 + \beta_2 y + y^2 + syz + \mathcal{G}(y, z), \end{cases}$$

for  $\beta_1, \beta_2 \in \mathbb{R}$ ,  $s = +1$  is known to possess two fixed points - a saddle and a focus - with an unstable periodic orbit (that emerged from the focus in a Hopf bifurcation in the open parameter region)

$$S_{BT} = \left\{ (\beta_1, \beta_2) \mid \beta_2 < 0, \beta_1 < -\frac{6}{25}\beta_2^2 + o(\beta_2^2) \right\},$$

The right border  $\{\beta_2 < 0, \beta_1 = 0\}$  marks the Hopf bifurcation, while the left border  $\{\beta_2 < 0, \beta_1 = -\frac{6}{25}\beta_2^2 + o(\beta_2^2)\}$  describes the region where a homoclinic orbit emerged from the periodic orbit (whose period tends to infinity towards that border).

Here, we denote the slow system (3.38) by

$$\begin{cases} w_X = q, \\ q_X = F(w) + \varepsilon c q \rho_1(w) + G(w, q) \end{cases}$$

where  $F(w) = -\mathcal{A} + (\mathcal{B} + a\Theta)w + \mathcal{C}w\sqrt{a + \frac{1}{4} - \frac{1}{w}}$  and  $G$  accounts for the higher order term, and assume that the parameters are chosen such that both fixed points are on  $\mathcal{M}_\varepsilon^+$ ,  $\mathcal{D}$  is close (but beyond) the saddle-node bifurcation point, that is,  $\mathcal{D} = \frac{\mathcal{C}^2}{4\mathcal{A}} - \sigma^2\mathcal{A}$ ,  $0 < \sigma \ll 1$  and the  $w$ -coordinate of both emerging fixed points is well within the strip  $(4/(1+4a), 1/a)$ . For  $\sigma$  sufficiently small, there is a neighborhood of  $w_0^{SN}$  such that the slow ODE has the form

$$\begin{cases} w_X = q, \\ q_X = \mu_1 + \mu_2 w + \mu_3 w^2 + \delta q + \mu_4 w q + \tilde{G}(w, q) \end{cases} \quad (3.C.69)$$

where

$$\mu_1 = F(w_0^{SN}), \quad \mu_2 = F'(w_0^{SN}), \quad \mu_3 = \frac{1}{2}F''(w_0^{SN}), \quad \delta = \varepsilon c \rho_1(w_0^{SN}), \quad \mu_4 = \varepsilon c \rho_1'(w_0^{SN}),$$

and  $\mu_j = \mu_j(\sigma^2)$ ,  $\delta = \delta(\sigma^2)$ . By assuming the non-degeneracy condition  $\mu_3(\sigma^2)\mu_4(\sigma^2) \neq 0$ , performing a shift and scaling  $q = \tilde{q} - \frac{\delta(\sigma)}{\beta_4(0)}$  (assuming that  $\delta(0) \neq 0$  for simplicity),  $w = \frac{\mu_3(\sigma^2)}{\mu_4(\sigma^2)^2}y$ ,  $\tilde{X} = \left| \frac{\mu_3(\sigma^2)}{\mu_4(\sigma^2)} \right| X$ , we bring (3.C.69) into the form

$$\begin{cases} y_{\tilde{X}} = z, \\ z_{\tilde{X}} = \beta_1 + \beta_2 y + y^2 + s y z + \mathcal{G}(y, z), \end{cases}$$

where

$$\beta_1(\sigma^2) = \frac{\mu_4^4(\sigma^2)}{\mu_3^3(\sigma^2)}\mu_1(\sigma^2), \quad \beta_2(\sigma^2) = \left( \frac{\mu_4(\sigma^2)}{\mu_3(\sigma^2)} \right)^2 \mu_2(\sigma^2),$$

and  $s = \text{sign}(\beta_3(\sigma^2)\beta_4(\sigma^2))$ . Hence, in order to conclude the corresponding scenario as described for  $\mathcal{S}_{BT}$  for our original system, it remains to analyze the mapping  $\sigma^2 \mapsto (\beta_1(\sigma^2), \beta_2(\sigma^2))$ , which we refrain from doing here.

# Outlook

---

At first glance, Chapter 2 and Chapter 3 do not appear connected in a cohesive way, but together they certainly endeavor to lay the foundations of understanding the mathematical mechanisms driving the formation and nonlinear evolution of ecosystems and the observations thereof. While achieving this, they consolidate the universality of distinct features that bridge the gap between mathematics and ecology.

This thesis constitutes a constant interplay and wavering between two parallel, yet complementary worlds: a fundamental and a realistic one. Both Chapter 2 and Chapter 3 have shown the importance of ecological background in understanding the nature of patterns exhibited by reaction-diffusion models. Additionally, the scope of patterns we were initially interested in was broadened. By posing the relevant ecological questions, the underlying mathematical mechanisms unrolled themselves into a realm of expected and unexpected patterns and geometrical constructions. Rigorous analysis of reaction-diffusion models in return demarcates the possible outcomes and sceneries an ecosystem can have as well as its spatio-temporal evolution. A real ecosystem is typically not close to onset, i.e., features parameters  $\mu$  that are not close to the critical parameter value  $\mu_c$ , thereby yielding patterns exhibited by (1.20–1.21–2.1–3.1) that are ‘far-from-equilibrium’. This allows for these patterns to be studied by the methods of geometric singular perturbation theory (as has been done in this thesis).

This is a mere recapitulation of the longer call we have stated in the introduction, that, trying to understand natural complex phenomena such as climate change and soil degradation, requires the fusing of these two fields of research. It has become clear throughout this thesis that in the study of ecosystem dynamics the knowledge of both ecology and mathematics needs to be combined in order to optimize the posed research questions and the models that are used in order to describe the processes at play. As data acquisition and sampling methods develop and expand by the minute, the areas of ecosystem modeling and model

assessment are just set out. Along this development, a multitude of follow-up research questions arise from the studies done in this thesis, from the establishment of the Busse balloon in Chapter 2 to the geometric singular perturbation theory tools of Chapter 3.

As outlined in Chapter 1.2, the models studied differ in their approach by their different ecological settings, one being more conceptual and the other trying to include more realistic parameters and environmental constraints. Studying the most conceptual model, the extended-Klausmeier model, and its Busse balloon in Chapter 2, and laying it next to real state-of-the-art acquired dryland vegetation data, confirmed that ecological systems are more resilient on a fundamental level than previously thought and that various patterns can co-exist simultaneously for the same parameter values. In an analogous way, in Chapter 3, we have found the Turing bifurcation in model (3.1) and we have constructed by means of geometric singular perturbation theory several classes of spatially periodic far-from equilibrium patterns as seen in Figures (3.17a,3.18a) of Chapter 3.6. These represent sections of the Busse balloon of the model of Chapter 3, and form a basis to build upon in completing the Busse balloon in parameter-space and thus for the existence of the Busse balloon and the multistability of periodic patterns for this model.

Along the lines of parameter-space and parameters, the more realistic model of Chapter 3, may also also have been used to analyze the data of Chapter 2 from a quantitative perspective. With the abundance of ecological parameters within this model, this would have gone beyond qualitatively establishing the fundamental phenomenon of increased resilience through multistability. Thus, this may be a relevant extension of the study of these sites, by providing a more qualitative interpretation of the latter. For this, a meticulous study of the ecological literature is required to acquire realistic values of the parameters present. In model (3.1), the Busse balloon of the model of Chapter 3 can be determined explicitly. The model of Chapter 3 can also easily incorporate an advection term to account for the topological variations (which was initially done in the 3D Gilad et al. version [61]) that were present on site. This would be an interesting next step: to put both models side by side in order to analyze and compare their predictions. Actual model comparison, to my knowledge, has not been done systematically as the parameters' choices of each model are often hard to retrace, therefore making one-on-one correlation a challenge. Nevertheless, a qualitative comparison of both models and their predictions with respect to actual observations would be a natural follow-up, in order to also fine-tune the ecological parameters that have been worked with for years within Klausmeier-type models.

A fundamental challenge that we did not touch upon is, of course, the direct relevance of the 1D model analyses of Chapters 2 and 3 in the realistic setting

2-dimensional (spatial) domains. Naturally, extrapolating the patterns found in the one dimensional case to the 2-dimensional case using the symmetry in the  $y$ -direction results in trivial straight stripes. Moreover, perfectly circular symmetric patterns can also be studied by the 1D-methods of this thesis – see for instance [158] and the references therein. However, these idealized two dimensional shapes are, for example, very different from the realistic ecological case, where we have a broad range of patterns that are far from being perfectly symmetric: fairy circles, labyrinths, gaps, curved stripes as can be seen in Figure 4.1. Thus, a mere extension of the 1D analysis of our models to a two dimensional space is too simplistic (in fact, it should even be noted that the precise nature of the interaction terms in the model may be different in 2-space dimensions – see Chapter 3, and [106]). However, the patterns exhibited in Figure 4.1 also suggest that the analysis of realistic patterns may be developed by a perturbative approach that starts out from perfectly symmetrical intrinsically 1D patterns. Moreover, even if one would restrict oneself to perhaps the most simple start-up problem, the evolution of curved interfaces, a problem that has been extensively studied in the mathematical literature (see for instance [23] and the references therein), the mathematical challenge should not be underestimated: unlike the evolving interfaces generally considered in the literature, ecologically relevant stripes and interfaces may be destabilized by their evolution or through their interactions with other stripes/interfaces – see for instance [11] for examples in one space dimension. Nevertheless, we may conclude that the 1D analysis applied and further developed in this thesis does provide a first stepping stone towards understanding the evolution and dynamics of ecologically relevant patterns in 2D.



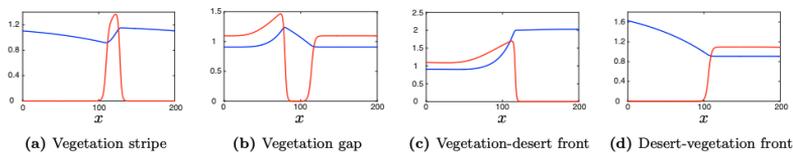
(a) Fairy circles in the Namibian desert [1].

(b) A site displaying curved banded vegetation patterns in the Western Creek Basin southwest of Newman, Australia ( $-23.5^{\circ}N, 119.5^{\circ}E$ ). Picture taken from [60].

**Figure 4.1:** Two dimensional vegetation patterns in nature.

The dimensionality of an ecosystem model is not the only place where space and spatial coordinates play a key role in the type of patterns exhibited and their behavior. A research question that has not yet been thoroughly explored is the

role of topography and its effect, even on a small scale, on the family of patterns that the models can exhibit and that can be observed. Model predictions in Chapter 2 regarding the role of slope with respect to migration speed was not corroborated by the empirical data analyzed, suggesting that the inclusion of small-scale topographical heterogeneities is a promising avenue for future model analysis, something that was further researched in [10]. Including slope into model of Chapter 3 would be a next step into incorporating the topographical effect into the analysis of ecosystem dynamics. The curvature of the slow manifolds has played a crucial role in the asymmetrical type of connections that we encountered, starting with the 'basic' connections between bare soil and vegetated state. Including advection into the model, that is, breaking of the reversibility symmetry  $(u(x, t), v(x, t)) \rightarrow (u(-x, -t), v(-x, -t))$  of the system will at the very least affect the profile of the connections established, if not completely reshape the family of periodic patterns found on the slow manifolds, as these are symmetric stationary solutions and breaking of the reversibility symmetry would disrupt these constructions. Such modifications have, for example, been researched in [9] for a 2-component extended Klausmeier-type model, the basic connections between bare soil and vegetated state will be altered from being completely symmetric in the case of a pulse and a gap to the new patterns that can be seen in Figure 4.2.



**Figure 4.2:** Shown are the different patterned solutions of an extended Klausmeier-type model. Presented figures show 1D cross-sections  $u(x)$  (blue) and  $v(x)$  (red) of direct numerical simulations. Figures from [9] – where water diffusion has been omitted for simplicity.

---

Aside from the possible model extensions and further ecological applications, a natural next step building on the existence proofs of Chapter 3 is to study the stability of the new found orbits and connections in both one and two dimensions. This is, besides its mathematical importance, of high relevance to the ecological community as the only observable patterns in nature are those that are mathematically stable. As the model of Chapter 3 encapsulates a ‘new’ variety of slow localized patterns, that are completely embedded in a slow manifold of the singularly perturbed spatial dynamical system, further stability analysis is required, by performing the corresponding spectral stability analysis. In [40], a very first step in this direction has been taken.

One overall clear takeaway from this thesis is that the practical applicability of geometrical singular perturbation theory transcends specific model formulation. Given a reaction-diffusion system that incorporates spatial scale separation, a successful geometric singular perturbation analysis will make clear which patterns exist, what they look like, what their period is (if present), and how these properties depend on the model parameters [41, 44]. Given that patterns in real ecosystems are singularly perturbed by definition [113], as biomass and water diffuse with very different speeds, geometric singular perturbation theory is an excellent tool to go with in order to investigate the different vegetation patterns exhibited. The existence and properties of orbits are directly related to the shape and transversal intersection of the geometrical objects (stable and unstable manifolds) introduced in Chapter 1 and Chapter 3 in phase space. This has, from a model perspective, two consequences. First, recent insight shows that the existence and properties of several important types of special orbits can be established [31, 40, 46] using only general properties of the reaction terms  $f(u) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  in (1.8) – that is, patterns such as pulses or periodic orbits, and their properties, can be found for general classes of reaction-diffusion systems. For a specific reaction-diffusion system, one only needs to check whether its reaction terms obey certain (mild) conditions; if so, the pattern properties are explicitly given in terms of integrals involving the reaction terms, and certain solutions to lower-dimensional differential equations. A second, and related, consequence of the geometric approach is that because the specific functional form of the reaction terms  $f(u)$  is not important, the reaction terms can be directly defined for instance through an (experimentally obtained) response curve. Only the geometric shape of this curve determines the existence and properties of pattern solutions, not the specific algebraic implementation of this shape. Therefore, patterns obtained by a geometric singular perturbation construction are structurally stable. Thus, geometric singular perturbation theory is an extremely suitable ‘tool’ by which we can understand ecosystem models – even the more realistic and thus complex ones. However, the present state-of-the-art theory is still insufficiently developed (especially concerning the stability of patterns): ecology will keep on driving the development of the theory for quite a number

of years.

# Bibliography

---

- [1] Credits to Sjors van der Stelt.
- [2] Credits to Stephan Gretzin - <https://www.namibia-eco-tours.com/fairy-circles/>.
- [3] United Nations : The Climate Crisis – A Race We Can Win - 2020 - <https://www.un.org/en/un75/climate-crisis-race-we-can-win/>.
- [4] <https://wallpapersafari.com/w/nYZjKS>.
- [5] Z. Adeel et al. (2005), *Ecosystems and human well-being : desertification synthesis : a report of the Millennium Ecosystem Assessment*. World Resources Institute.
- [6] N. Barbier, P. Couteron and V. Deblauwe (2014), *Case Study of Self-Organized Vegetation Patterning in Dryland Regions of Central Africa*, pp. 347–356. Dordrecht: Springer Netherlands.
- [7] N. Barbier, P. Couteron, J. Lejoly, V. Deblauwe and O. Lejeune (2006), Self-organized vegetation patterning as a fingerprint of climate and human impact on semi-arid ecosystems. *Journal of Ecology* **94**(3), 537–547.
- [8] N. Barbier, P. Couteron, O. Planchon and A. Diouf (2010), Multiscale comparison of spatial patterns using two-dimensional cross-spectral analysis: application to a semi-arid (gapped) landscape. *Landscape ecology* **25**(6), 889–902.
- [9] R. Bastiaansen, P. Carter and A. Doelman (2019), Stable planar vegetation stripe patterns on sloped terrain in dryland ecosystems. *Nonlinearity* **32**(8), 2759–2814.
- [10] R. Bastiaansen, M. Chirilus-Bruckner and A. Doelman (2020), Pulse solutions for an extended Klausmeier model with spatially varying coefficients. *SIAM Journal on Applied Dynamical Systems* **19**, 1–57.

- [11] R. Bastiaansen and A. Doelman (2019), The dynamics of disappearing pulses in a singularly perturbed reaction–diffusion system with parameters that vary in time and space. *Physica D* **388**, 45–72.
- [12] R. Bastiaansen, A. Doelman, M. B. Eppinga and M. Rietkerk (2020), The effect of climate change on the resilience of ecosystems with adaptive spatial pattern formation. *Ecology Letters* **23**(3), 414–429.
- [13] R. Bastiaansen, O. Jaïbi, V. Deblauwe, M. B. Eppinga, K. Siteur, E. Siero, S. Mermoz, A. Bouvet, A. Doelman, and M. Rietkerk (2018), Multistability of model and real dryland ecosystems through spatial self-organization. *Proceedings of the National Academy of Sciences* **115**(44), 11256–11261.
- [14] G. Bel, A. Hagberg and E. Meron (2012), Gradual regime shifts in spatially extended ecosystems. *Theoretical Ecology* **5**(4), 591–604.
- [15] A. Bouvet, S. Mermoz, T. L. Toan, L. Villard, R. Mathieu, L. Naidoo, and G. P. Asner (2018), An above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR. *Remote Sensing of Environment* **206**, 156 – 173.
- [16] J. Bruniquel and A. Lopes (1997), Multi-variate optimal speckle reduction in SAR imagery. *International journal of remote sensing* **18**(3), 603–627.
- [17] S. L. Brusatte, R. J. Butler, P. M. Barrett, M. T. Carrano, D. C. Evans, G. T. Lloyd, P. D. Mannion, M. A. Norell, D. J. Peppe, P. Upchurch, and T. E. Williamson (2015), The extinction of the dinosaurs. *Biological Reviews* **90**(2), 628–642.
- [18] F. Busse (1978), Non-linear properties of thermal convection. *Reports on Progress in Physics* **41**(12), 1929.
- [19] R. Canfield (1957), Reproduction and Life Span of Some Perennial Grasses of Southern Arizona. *Rangeland Ecology & Management/Journal of Range Management Archives* **10**(5), 199–203.
- [20] P. Carter, B. de Rijk and B. Sandstede (2016), Stability of traveling pulses with oscillatory tails in the FitzHugh-Nagumo system. *Journal of Nonlinear Science* **26**, 1369–1444.
- [21] P. Carter and A. Doelman (2018), Traveling stripes in the Klausmeier model of vegetation pattern formation. *SIAM Journal on Applied Mathematics* **78**(6), 3213–3237.
- [22] P. Carter and B. Sandstede (2015), Fast Pulses with Oscillatory Tails in the FitzHugh–Nagumo System. *SIAM Journal on Mathematical Analysis* **47**(5), 3393–3441.

- [23] X. Chen (1992), Generation and Propagation of Interfaces in Reaction-Diffusion Systems. *Transactions of the American Mathematical Society* **334**(2), 877–913.
- [24] Y. Chen, T. Kolokolnikov, J. Tzou and C. Gai (2015), Patterned vegetation, tipping points, and the rate of climate change. *European Journal of Applied Mathematics* **26**(6), 945–958.
- [25] M. Chirilus-Bruckner, A. Doelman, P. van Heijster and J. D. M. Rademacher (2015), Butterfly Catastrophe for Fronts in a Three-Component Reaction–Diffusion System. *Journal of Nonlinear Science* **25**(1), 87–129.
- [26] M. Chirilus-Bruckner, P. van Heijster, H. Ikeda and J. D. M. Rademacher (2019), Unfolding Symmetric Bogdanov–Takens Bifurcations for Front Dynamics in a Reaction–Diffusion System. *Journal of Nonlinear Science* **29**(6), 2911–2953.
- [27] P. Couteron (2001), Using spectral analysis to confront distributions of individual species with an overall periodic pattern in semi-arid vegetation. *Plant Ecology* **156**(2), 229–243.
- [28] P. Couteron and O. Lejeune (2001), Periodic spotted patterns in semi-arid vegetation explained by a propagation-inhibition model. *Journal of Ecology* **89**(4), 616–628.
- [29] M. C. Cross and P. C. Hohenberg (1993), Pattern formation outside of equilibrium. *Rev. Mod. Phys.* **65**, 851–1112.
- [30] J. H. P. Dawes and J. L. M. Williams (2016), Localised pattern formation in a model for dryland vegetation. *Journal of Mathematical Biology* **73**(1), 63–90.
- [31] B. de Rijk, A. Doelman and J. Rademacher (2016), Spectra and stability of spatially periodic pulse patterns: Evans function factorization via Riccati transformation. *SIAM Journal of Mathematical Analysis* **48**(1), 61–121.
- [32] V. Deblauwe, N. Barbier, P. Couteron, O. Lejeune and J. Bogaert (2008), The global biogeography of semi-arid periodic vegetation patterns. *Global Ecology and Biogeography* **17**(6), 715–723.
- [33] V. Deblauwe, P. Couteron, J. Bogaert and N. Barbier (2012), Determinants and dynamics of banded vegetation pattern migration in arid climates. *Ecological Monographs* **82**(1), 3–21.
- [34] V. Deblauwe, P. Couteron, O. Lejeune, J. Bogaert and N. Barbier (2011), Environmental modulation of self-organized periodic vegetation patterns in Sudan. *Ecography* **34**(6), 990–1001.

- [35] V. Deblauwe, V. Droissart, R. Bose, B. Sonké, A. Blach-Overgaard, J.-C. Svenning, J. Wieringa, B. Ramesh, T. Stévar, and T. Couvreur (2016), Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography* **25**(4), 443–454.
- [36] V. Deblauwe, P. Kennel and P. Couteron (2012), Testing pairwise association between spatially autocorrelated variables: A new approach using surrogate lattice data. *PloS one* **7**(11), e48766.
- [37] U. Deichmann and L. Eklundh (1991), Global Digital Datasets for Land Degradation Studies: A GIS Approach. Nairobi, Kenya: United Nations Environment Programme. *Global Resource Information Database case study series no. 4*.
- [38] H. A. Dijkstra (2011), Vegetation pattern formation in a semi-arid climate. *International Journal of Bifurcation and Chaos* **21**(12), 3497–3509.
- [39] A. Doelman (2019), Pattern formation in reaction-diffusion systems—an explicit approach. In: *Complexity Science*. World Scientific, pp. 129–182.
- [40] A. Doelman (2021), Slow localized patterns in singularly perturbed 2-component reaction-diffusion equations.
- [41] A. Doelman, R. A. Gardner and T. J. Kaper (2001), Large stable pulse solutions in reaction-diffusion equations. *Indiana Univ. Math. J* **50**(1), 443–508.
- [42] A. Doelman and P. Holmes (1996), Homoclinic explosions and implosions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **354**(1709), 845–893.
- [43] A. Doelman, D. Iron and Y. Nishiura (2004), Destabilization of fronts in a class of bi-stable systems. *SIAM Journal of Mathematical Analysis* **35**, 1420–1450.
- [44] A. Doelman, T. J. Kaper and H. van der Ploeg (2001), Spatially periodic and aperiodic multi-pulse patterns in the one-dimensional Gierer-Meinhardt equation. *Methods and Applications of Analysis* **8**(3), 387–414.
- [45] A. Doelman, P. van Heijster and T. Kaper (2009), Pulse Dynamics in a three-Component System: existence Analysis. *Journal of Dynamics and Differential Equations* **21**(1), 73–115.
- [46] A. Doelman and F. Veerman (2015), An Explicit Theory for Pulses in Two Component, Singularly Perturbed, Reaction–Diffusion Equations. *Journal of Dynamics and Differential Equations* **27**(3), 555–595.

- [47] A. K. Duraiappah et al. (2005), *Ecosystems and human well-being: biodiversity synthesis; a report of the Millennium Ecosystem Assessment*. World Resources Institute.
- [48] M. B. Eppinga, P. C. de Ruiter, M. J. Wassen and M. Rietkerk (2009), Nutrients and Hydrology Indicate the Driving Mechanisms of Peatland Surface Patterning. *The American Naturalist* **173**(6), 803–818.
- [49] D. Escaff, C. Fernandez-Oto, M. G. Clerc and M. Tlidi (2015), Localized vegetation patterns, fairy circles, and localized patches in arid landscapes. *Physical Review E* **91**, 022924.
- [50] S. Faurby, D. Silvestro, L. Werdelin and A. Antonelli (2020), Brain expansion in early hominins predicts carnivore extinctions in East Africa. *Ecology Letters* **23**(3), 537–544.
- [51] N. Fenichel (1974), Asymptotic Stability with Rate Conditions. *Indiana University Mathematics Journal* **23**(12), 1109–1137.
- [52] N. Fenichel (1977), Asymptotic Stability with Rate Conditions, II. *Indiana University Mathematics Journal* **26**(1), 81–93.
- [53] N. Fenichel (1979), Geometric singular perturbation theory for ordinary differential equations. *Journal of differential equations* **31**(1), 53–98.
- [54] N. Fenichel and J. K. Moser (1971), Persistence and Smoothness of Invariant Manifolds for Flows. *Indiana University Mathematics Journal* **21**(3), 193–226.
- [55] C. Fernandez-Oto, M. Tlidi, D. Escaff and M. G. Clerc (2014), Strong interaction between plants induces circular barren patches: fairy circles. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **372**(2027), 20140009.
- [56] C. Fernandez-Oto, O. Tzuk and E. Meron (2019), Front Instabilities Can Reverse Desertification. *Physical Review Letters* **122**, 048101.
- [57] G. Feulner (2017), Global Challenges: Climate Change. *Global Challenges* **1**(1), 5–6.
- [58] J. Fuentes, D. Varga and J. Pintó (2018), The Use of High-Resolution Historical Images to Analyse the Leopard Pattern in the Arid Area of La Alta Guajira, Colombia. *Geosciences* **8**(10), 366.
- [59] C. Funk, P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, et al. (2015), The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data* **2**, 150066.

- [60] P. Gandhi, L. Werner, S. Iams, K. Gowda and M. Silber (2018), A topographic mechanism for arcing of dryland vegetation bands. *Journal of The Royal Society Interface* **15**(147), 20180508.
- [61] S. Getzin, H. Yizhaq, B. Bell, T. E. Erickson, A. C. Postle, I. Katra, O. Tzuk, Y. R. Zelnik, K. Wiegand, T. Wiegand, and E. Meron (2016), Discovery of fairy circles in Australia supports self-organization theory. *Proceedings of the National Academy of Sciences* **113**(13), 3551–3556.
- [62] E. Gilad, J. von Hardenberg, A. Provenzale, M. Shachak and E. Meron (2004), Ecosystem Engineers: From Pattern Formation to Habitat Creation. *Physical Review Letters* **93**, 098105.
- [63] E. Gilad, J. von Hardenberg, A. Provenzale, M. Shachak and E. Meron (2007), A mathematical model of plants as ecosystem engineers. *Journal of Theoretical Biology* **244**(4), 680–691.
- [64] Y. Goto, D. Hilhorst, E. Meron and R. Temam (2011), Existence theorem for a model of dryland vegetation. *Discrete and Continuous Dynamical Systems. Series B* **16**(1), 197–224.
- [65] K. Gowda, Y. Chen, S. Iams and M. Silber (2016), Assessing the robustness of spatial pattern sequences in a dryland vegetation model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **472**(2187), 20150893.
- [66] K. Gowda, S. Iams and M. Silber (2018), Signatures of human impact on self-organized vegetation in the Horn of Africa. *Scientific Reports* **8**(1), 3622.
- [67] K. Gowda, H. Riecke and M. Silber (2014), Transitions between patterned states in vegetation models for semiarid ecosystems. *Phys. Rev. E* **89**, 022701.
- [68] D. Green, S. Sadedin and T. Leishman (2008), Self-Organization. In: *Encyclopedia of Ecology*. Oxford: Academic Press, pp. 3195 – 3203.
- [69] V. Grimm and S. F. Railsback (2012), Pattern-oriented modelling: a ‘multi-scope’ for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**(1586), 298–310.
- [70] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis (2005), Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* **310**(5750), 987–991.

- [71] J. Grodecki and G. Dial (2003), Block adjustment of high-resolution satellite images described by rational polynomials. *Photogrammetric Engineering & Remote Sensing* **69**(1), 59–68.
- [72] J. Guckenheimer and P. Holmes (1983), *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Applied Mathematical Sciences. Springer, New York.
- [73] A. Hagberg and E. Meron (1994), Complex patterns in reaction-diffusion systems: A tale of two front instabilities. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **4**, 477–484.
- [74] A. Hagberg and E. Meron (1997), The dynamics of curved fronts: beyond geometry. *Physical Review Letters* **78**, 1166–1169.
- [75] A. Hagberg and E. Meron (1998), Order parameter equations for front transitions: nonuniformly curved fronts. *Physica D* **123**, 460–473.
- [76] M. Haragus and G. Iooss (2011), *Local Bifurcations, Center Manifolds, and Normal Forms in Infinite-Dimensional Dynamical Systems*. Springer.
- [77] G. Hek (2010), Geometric singular perturbation theory in biological practice. *Journal of Mathematical Biology* **60**(3), 347–386.
- [78] C. F. Hemming (1965), Vegetation Arcs in Somaliland. *Journal of Ecology* **53**(1), 57–67.
- [79] R. HilleRisLambers, M. Rietkerk, F. van den Bosch, H. Prins and H. de Kroon (2001), Vegetation pattern formation in semi-arid grazing systems. *Ecology* **82**, 50–61.
- [80] C. S. Holling (1973), Resilience and stability of ecological systems. *Annual review of ecology and systematics* **4**(1), 1–23.
- [81] T. P. Hughes, C. Linares, V. Dakos, I. A. van de Leemput and E. H. van Nes (2013), Living dangerously on borrowed time during slow, unrecognized regime shifts. *Trends in Ecology & Evolution* **28**(3), 149–155.
- [82] O. Jaïbi, A. Doelman, M. Chirilus-Bruckner and E. Meron (2020), The existence of localized vegetation patterns in a systematically reduced model for dryland vegetation. *Physica D: Nonlinear Phenomena* **412**, 132637.
- [83] C. K. R. T. Jones (1995), Geometric singular perturbation theory. In: *Dynamical systems (Montecatini Terme, 1994)*, Vol. 1609 of *Lecture Notes in Math*. Springer, Berlin, pp. 44–118.
- [84] C. K. R. T. Jones, T. J. Kaper and N. Kopell (1996), Tracking Invariant Manifolds up to Exponentially Small Errors. *SIAM Journal on Mathematical Analysis* **27**(2), 558–577.

- [85] A. Kamb, M. Weir, B. Rudy, H. Varmus and C. Kenyon (1989), Identification of genes from pattern formation, tyrosine kinase, and potassium channel families by DNA amplification. *Proceedings of the National Academy of Sciences* **86**(12), 4372–4376.
- [86] T. J. Kaper (1999), An introduction to geometric methods and dynamical systems theory for singular perturbation problems. In: *Analyzing multiscale phenomena using singular perturbation methods* (Baltimore, MD, 1998), Vol. 56 of *Proceedings of Symposia in Applied Mathematics*. American Mathematical Society, pp. 85–131.
- [87] S. Kéfi, V. Dakos, M. Scheffer, E. H. Van Nes and M. Rietkerk (2013), Early warning signals also precede non-catastrophic transitions. *Oikos* **122**(5), 641–648.
- [88] S. Kéfi, M. Rietkerk, C. L. Alados, Y. Pueyo, V. P. Papanastasis, A. ElAich, and P. C. de Ruiter (2007), Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* **449**(7159), 213–217.
- [89] T. H. Keitt (2000), Spectral representation of neutral landscapes. *Landscape Ecology* **15**(5), 479–494.
- [90] C. A. Klausmeier (1999), Regular and Irregular Patterns in Semiarid Vegetation. *Science* **284**(5421), 1826–1828.
- [91] D. Kok, R. Bastiaansen and A. Doelman (2021), Existence, stability and bifurcations of stripe patterns in a generalized Gierer-Meinhardt model. *To appear*.
- [92] T. Kolokolnikov, M. J. Ward, W. Sun and J. Wei (2006), The stability of a stripe for the Gierer-Meinhardt model and the effect of saturation. *SIAM Journal on Applied Dynamical Systems* **5**, 313–363.
- [93] L. Larsen, C. Thomas, M. Eppinga and T. Coulthard (2014), Exploratory modeling: Extracting causality from complexity. *Eos, Transactions American Geophysical Union* **95**(32), 285–286.
- [94] L. G. Larsen, M. B. Eppinga, P. Passalacqua, W. M. Getz, K. A. Rose, and M. Liang (2016), Appropriate complexity landscape modeling. *Earth-Science Reviews* **160**(Supplement C), 111 – 130.
- [95] W. K. Lauenroth and P. B. Adler (2008), Demography of perennial grassland plants: survival, life expectancy and life span. *Journal of Ecology* **96**(5), 1023–1032.

- [96] K. J. Lee, W. D. McCormick, Q. Ouyang and H. L. Swinney (1993), Pattern Formation by Interacting Chemical Fronts. *Science* **261**(5118), 192–194.
- [97] R. Lefever and O. Lejeune (1997), On the origin of tiger bush. *Bull. Math. Biol.* **59**, 263–294.
- [98] O. Lejeune, T. M. and R. Lefever (2004), Vegetation spots and stripes: dissipative structures in arid landscapes. *Int. J. Quantum Chem.* **98**, 261–271.
- [99] S. A. Levin (1992), The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology* **73**(6), 1943–1967.
- [100] Y. Lin, G. Han, M. Zhao and S. X. Chang (2010), Spatial vegetation patterns as early signs of desertification: a case study of a desert steppe in Inner Mongolia, China. *Landscape Ecology* **25**(10), 1519–1527.
- [101] J. A. Ludwig, B. P. Wilcox, D. D. Breshears, D. J. Tongway and A. C. Imeson (2005), VEGETATION PATCHES AND RUNOFF-EROSION AS INTERACTING ECOHYDROLOGICAL PROCESSES IN SEMI-ARID LANDSCAPES. *Ecology* **86**(2), 288–297.
- [102] W. A. Macfadyen (1950), Soil and Vegetation in British Somaliland. *Nature* **165**, 121.
- [103] R. M. May (1977), Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* **269**(5628), 471–477.
- [104] J. D. Meiss (2007), *Differential Dynamical Systems*. Society for Industrial and Applied Mathematics.
- [105] S. Mermoz, T. Le Toan, L. Villard, M. Réjou-Méchain and J. Seifert-Granzin (2014), Biomass assessment in the Cameroon savanna using ALOS PALSAR data. *Remote sensing of environment* **155**, 109–119.
- [106] E. Meron (2015), *Nonlinear Physics of Ecosystems*. CRC Press.
- [107] E. Meron (2018), From Patterns to Function in Living Systems: Dryland Ecosystems as a Case Study. *Annual Review of Condensed Matter Physics* **9**, 79–103.
- [108] E. Meron (2019), Vegetation pattern formation: The mechanisms behind the forms. *Physics Today* **72**(11), 30–36.
- [109] E. Meron, H. Yizhaq and E. Gilad (2007), Localized structures in dryland vegetation: Forms and functions. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **17**(3), 037109.

- [110] C. A. Mgbemene, C. C. Nnaji and C. Nwozor (2016), Industrialization and its backlash: focus on climate change and its consequences. *Journal of Environmental Science and Technology* **9**(4), 301.
- [111] I. R. Moyles and M. J. Ward (2017), Existence, Stability, and Dynamics of Ring and Near-Ring Solutions to the Saturated Gierer–Meinhardt Model in the Semistrong Regime. *SIAM Journal on Applied Dynamical Systems* **16**(1), 597–639.
- [112] E. N. Mueller, J. Wainwright, A. J. Parsons and L. Turnbull (eds.) (2014), *Patterns of Land Degradation in Drylands : Understanding Self-Organised Ecogeomorphic Systems*. Springer, Netherlands.
- [113] J. Murray (2013), *Mathematical Biology II: Spatial Models and Biomedical Applications*, Interdisciplinary Applied Mathematics. Springer New York.
- [114] G. Nicolis and I. Prigogine (1977), *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*. Wiley.
- [115] I. Noy-Meir (1975), Stability of Grazing Systems: An Application of Predator-Prey Graphs. *Journal of Ecology* **63**(2), 459–481.
- [116] N. Otsu (1979), A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* **9**(1), 62–66.
- [117] J. E. Pearson (1993), Complex Patterns in a Simple System. *Science* **261**(5118), 189–192.
- [118] G. G. Penny, K. E. Daniels and S. E. Thompson (2013), Local properties of patterned vegetation: quantifying endogenous and exogenous effects. *Phil. Trans. R. Soc. A* **371**(2004), 20120359.
- [119] B. Purinton and B. Bookhagen (2017), Validation of digital elevation models (DEMs) and comparison of geomorphic metrics on the southern Central Andean Plateau. *Earth Surface Dynamics* **5**(2), 211.
- [120] J. Rademacher (2005), Homoclinic orbits near heteroclinic cycles with one equilibrium and one periodic orbit. *Journal of Differential Equations* **218**, 390–443.
- [121] J. Rademacher (2010), Lyapunov-Schmidt reduction for unfolding heteroclinic networks of equilibria and periodic orbits with tangencies. *Journal of Differential Equations* **249**, 305–348.
- [122] J. D. Rademacher, B. Sandstede and A. Scheel (2007), Computing absolute and essential spectra using continuation. *Physica D: Nonlinear Phenomena* **229**(2), 166–183.

- [123] E. Renshaw and E. Ford (1984), The description of spatial pattern using two-dimensional spectral analysis. *Vegetatio* **56**(2), 75–85.
- [124] M. Rietkerk, M. C. Boerlijst, F. van Langevelde, R. HilleRisLambers, J. de Koppel, L. Kumar, H. H. T. Prins, and A. de Roos (2002), Self-organization of vegetation in arid ecosystems. *The American Naturalist* **160**(4), 524–530.
- [125] M. Rietkerk, S. C. Dekker, P. C. de Ruiter and J. van de Koppel (2004), Self-Organized Patchiness and Catastrophic Shifts in Ecosystems. *Science* **305**(5692), 1926–1929.
- [126] M. Rietkerk and J. van de Koppel (2008), Regular pattern formation in real ecosystems. *Trends in Ecology & Evolution* **23**(3), 169–175.
- [127] M. Rietkerk, F. van den Bosch and J. van de Koppel (1997), Site-Specific Properties and Irreversible Vegetation Changes in Semi-Arid Grazing Systems. *Oikos* **80**(2), 241–252.
- [128] D. Ruiz-Reynés, D. Gomila, T. Tomàs Sintès, E. Hernández-García, N. Marbà, and C. M. Duarte (2017), Fairy circle landscapes under the sea. *Science Advances* **3**, e1603262.
- [129] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara (2009), Early-warning signals for critical transitions. *Nature* **461**(7260), 53.
- [130] M. Scheffer, S. Carpenter, J. A. Foley, C. Folke and W. B. (2001), Catastrophic shifts in ecosystems. *Nature* **413**, 53–59.
- [131] M. Scheffer and S. R. Carpenter (2003), Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in ecology & evolution* **18**(12), 648–656.
- [132] J. Schmidt, I. S. Evans and J. Brinkmann (2003), Comparison of polynomial models for land surface curvature calculation. *International Journal of Geographical Information Science* **17**(8), 797–814.
- [133] L. Sewalt and A. Doelman (2017), Spatially Periodic Multipulse Patterns in a Generalized Klausmeier-Gray-Scott Model. *SIAM Journal on Applied Dynamical Systems* **16**(2), 1113–1163.
- [134] E. Sheffer, J. von Hardenberg, H. Yizhaq, M. Shachak and E. Meron (2013), Emerged or imposed: a theory on the role of physical templates and self-organisation for vegetation patchiness. *Ecology Letters* **16**(2), 127–139.

- [135] J. A. Sherratt (2010), Pattern solutions of the Klausmeier Model for banded vegetation in semi-arid environments I. *Nonlinearity* **23**(10), 2657–2675.
- [136] J. A. Sherratt (2011), Pattern solutions of the Klausmeier model for banded vegetation in semi-arid environments II: patterns with the largest possible propagation speeds. *Proc. R. Soc. A* **467**, 3272–3294.
- [137] J. A. Sherratt (2012), Turing Patterns in Deserts. In: S. B. Cooper, A. Dawar, and B. Löwe (eds.): *How the World Computes*. Springer Berlin Heidelberg, pp. 667–674.
- [138] J. A. Sherratt (2013), History-dependent patterns of whole ecosystems. *Ecological Complexity* **14**, 8–20.
- [139] J. A. Sherratt (2013), Pattern Solutions of the Klausmeier Model for Banded Vegetation in Semiarid Environments IV: Slowly Moving Patterns and Their Stability. *SIAM Journal on Applied Mathematics* **73**(1), 330–350.
- [140] J. A. Sherratt (2015), Using wavelength and slope to infer the historical origin of semiarid vegetation bands. *Proceedings of the National Academy of Sciences* **112**(14), 4202–4207.
- [141] J. A. Sherratt and G. J. Lord (2007), Nonlinear dynamics and pattern bifurcations in a model for vegetation stripes in semi-arid environments. *Theoretical population biology* **71**(1), 1–11.
- [142] E. Siero (2020), Resolving soil and surface water flux as drivers of pattern formation in Turing models of dryland vegetation: A unified approach. *Physica D: Nonlinear Phenomena* **414**, 132695.
- [143] E. Siero, A. Doelman, M. B. Eppinga, J. D. M. Rademacher, M. Rietkerk, and K. Siteur (2015), Striped pattern selection by advective reaction-diffusion systems: Resilience of banded vegetation on slopes. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**(3), 036411.
- [144] K. Siteur (2016), Off the beaten track. How ecosystems fail to respond to environmental change. *PhD dissertation*.
- [145] K. Siteur, M. B. Eppinga, A. Doelman, E. Siero and M. Rietkerk (2016), Ecosystems off track: rate-induced critical transitions in ecological models. *Oikos* **125**(12), 1689–1699.
- [146] K. Siteur, E. Siero, M. B. Eppinga, J. D. Rademacher, A. Doelman, and M. Rietkerk (2014), Beyond Turing: The response of patterned ecosystems to environmental change. *Ecological Complexity* **20**, 81–96.

- [147] T. Tadono, H. Nagai, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto (2016), GENERATION OF THE 30 m-MESH GLOBAL DIGITAL SURFACE MODEL BY ALOS PRISM. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* **41**.
- [148] J. Takaku, T. Tadono, K. Tsutsui and M. Ichikawa (2016), Validation of 'AW3D' global DSM generated from ALOS prism. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* **3**(4).
- [149] C. Toulmin and K. Brock (2016), *Desertification in the Sahel: Local Practice Meets Global Narrative*, pp. 37–63. Springer Berlin Heidelberg.
- [150] A. M. Turing (1952), The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **237**(641), 37–72.
- [151] O. Tzuk, S. Ujjwal, C. Fernandez-Oto, M. Seifan and E. Meron (2019), Interplay between exogenous and endogenous factors in seasonal vegetation oscillations. *Scientific Reports* **9**, 354.
- [152] C. Valentin, J. d'Herbès and J. Poesen (1999), Soil and water components of banded vegetation patterns. *CATENA* **37**(1), 1–24.
- [153] J. van de Koppel and C. M. Crain (2006), Scale Dependent Inhibition Drives Regular Tussock Spacing in a Freshwater Marsh. *The American Naturalist* **168**(5), E136–E147.
- [154] E. van der Maarel and J. Franklin (2012), *Vegetation Ecology, 2nd Ed.* Wiley-Blackwell, UK.
- [155] S. Van Der Stelt, A. Doelman, G. Hek and J. D. M. Rademacher (2013), Rise and fall of periodic patterns for a generalized Klausmeier-Gray-Scott model. *Journal of Nonlinear Science* **23**(1), 39–95.
- [156] K. J. van Groenigen, X. Qi, C. W. Osenberg, Y. Luo and B. A. Hungate (2014), Faster Decomposition Under Increased Atmospheric CO<sub>2</sub> Limits Soil Carbon Storage. *Science* **344**(6183), 508–509.
- [157] P. van Heijster, A. Doelman, T. Kaper and K. Promislow (2010), Front Interactions in a Three-Component System. *SIAM Journal of Applied Dynamical Systems* **9**, 292–332.
- [158] P. van Heijster and B. Sandstede (2011), Planar Radial Spots in a Three-Component FitzHugh–Nagumo System. *Journal of Nonlinear Science* **21**(5), 705–745.

- [159] P. van Heijster and B. Sandstede (2014), Bifurcations to travelling planar spots in a three-component FitzHugh-Nagumo system. *Physica D* **275**, 19–34.
- [160] P. Villa Martín, J. A. Bonachela, S. A. Levin and M. A. Muñoz (2015), Eluding catastrophic shifts. *Proceedings of the National Academy of Sciences* **112**(15), E1828–E1836.
- [161] J. von Hardenberg, E. Meron, M. Shachak and Y. Zarmi (2001), Diversity of vegetation patterns and desertification. *Physical Review Letters* **87**(19), 198101.
- [162] M. J. Ward (2018), Spots, traps, and patches: asymptotic analysis of localized solutions to some linear and nonlinear diffusive systems. *Nonlinearity* **31**(8), R189–R239.
- [163] A. S. Watt (1947), Pattern and Process in the Plant Community. *Journal of Ecology* **35**(1/2), 1–22.
- [164] D. Webb (1954), Is the classification of plant communities either possible or desirable. *Botanisk Tidsskrift* **51**, 362–370.
- [165] J. Wei and M. Winter (2014), *Mathematical Aspects of Pattern Formation in Biological Systems*, Vol. 189 of *Applied Mathematical Sciences*. Springer-Verlag, London.
- [166] R. H. Whittaker (1967), GRADIENT ANALYSIS OF VEGETATION\*. *Biological Reviews* **42**(2), 207–264.
- [167] G. E. Wickens (1998), *Arid and Semi-arid Environments of the World*, pp. 5–15. Springer Berlin Heidelberg.
- [168] T. Wiegand, J. von Hardenberg, K. Wiegand, E. Meron, S. Getzin, and H. Yizhaq (2014), Adopting a spatially explicit perspective to study the mysterious fairy circles of Namibia. *Ecography* **38**(1), 1–11.
- [169] S. Wiggins (2003), *Introduction to applied nonlinear dynamical systems and chaos*. Springer-Verlag New York.
- [170] M. Woolfson (2000), The origin and evolution of the solar system. *Astronomy & Geophysics* **41**(1), 1.12–1.19.
- [171] R. G. Wright and G. M. Van Dyne (1976), Environmental factors influencing semidesert grassland perennial grass demography. *The Southwestern Naturalist* pp. 259–273.
- [172] Y. R. Zelnik, P. Gandhi, E. Knobloch and E. Meron (2018), Implications of tristability in pattern-forming ecosystems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28**(3), 033609.

- 
- [173] Y. R. Zelnik, S. Kinast, H. Yizhaq, G. Bel and E. Meron (2013), Regime shifts in models of dryland vegetation. *Philosophical Transactions R. Soc. A* **371**, 20120358.
- [174] Y. R. Zelnik and E. Meron (2018), Regime shifts by front dynamics. *Ecological Indicators* **94**, 544–552.
- [175] Y. R. Zelnik, E. Meron and G. Bel (2015), Gradual regime shifts in fairy circles. *Proceedings of the National Academy of Sciences* **112**(40), 12327–12331.
- [176] Y. R. Zelnik, H. Uecker, U. Feudel and E. Meron (2017), Desertification by front propagation? *Journal of Theoretical Biology* **418**, 27–35.



# Samenvatting

---

Klimaatverandering heeft flinke gevolgen voor zowel natuur als mens. Een van de verschijnselen hiervan is het proces van verwoestijning, dat in vele, grote gebieden met weinig neerslag plaatsvindt. Deze gebieden bedekken nu al meer dan 40% van de landoppervlakte op de wereld en huisvesten meer dan 2 miljard mensen. Verwoestijning is een vorm van landdegradatie waar onder meer de kwaliteit van de bodem sterk afneemt evenals de hoeveelheid aanwezige vegetatie. Deze afname van de vegetatie hoeft niet op een egale manier te gebeuren, waardoor er bijvoorbeeld gaten of strepen zonder vegetatie kunnen ontstaan. Dit leidt tot patronen in het landschap.

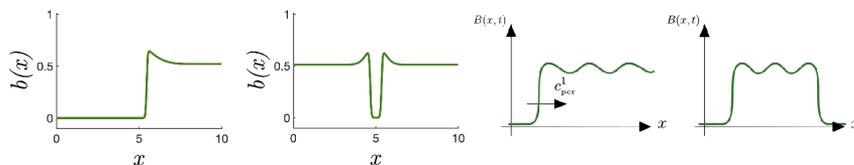
Het bestuderen van deze vegetatiepatronen op een wiskundige manier kan helpen bij het begrijpen van het verwoestijningsproces. Hiervoor wordt het proces vertaald naar een model, in meerdere of mindere mate van realisme en complexiteit. Het proces kan hiermee in grote lijnen beschreven worden, op een conceptuele manier, of op een realistischere manier, door meer factoren te beschouwen in het model. Door het proces op een conceptuele manier te benaderen kunnen de voornaamste gedragingen van het proces of van een ecologisch systeem in kaart worden gebracht en daarmee al enige voorspellingen worden gedaan over fundamentele aspecten van het gedrag in de werkelijkheid. De ecologische achtergrond van zowel een conceptueel als een realistisch model worden geïntroduceerd in hoofdstuk 1, evenals de wiskundige concepten en technieken die gebruikt worden bij het bestuderen hiervan.

Hoofdstuk 2 buigt zich over een conceptueel model waarbij er gekeken wordt naar het intrinsieke gedrag van een ecosysteem. Dit gedrag is vergeleken met satellietmetingen aan vegetatiepatronen, in het bijzonder strepen, in Somalië. Zo werd er gekeken naar de theoretische voorspellingen van het model betreffende de variatie van afstand tussen de vegetatiestrepen (periode), de relatieve aanwezige hoeveelheid vegetatie (biomassa) en verandering van patronen door verplaatsing of verdwijning (migratiesnelheid) en daarna vergeleken met de data. Een nieuw inzicht van deze studie is het concept van 'multistabiliteit': de bestudeerde gebieden worden niet gekenmerkt door een specifiek patroon met een unieke periode maar er bestaat een breed spectrum aan verschillende patronen die tegelijkertijd bestaan. Dit betekent dat een verandering van het

klimaat of de omgeving van een ecosysteem minder drastische gevolgen heeft dan tot nu toe werd aangenomen en dat het landschap minder snel verandert in een kale woestijn. Dit komt doordat het ecosysteem zich makkelijker kan aanpassen door, bij een verstoring, over te gaan van het ene patroon in het andere, omdat er meer dan één specifiek patroon mogelijk is. De multistabiliteit van de bestudeerde gebieden betekent dat deze veel veerkrachtiger zijn dan eerder gedacht.

In hoofdstuk 3 wordt er gekeken naar een realistischer model, waarbij de manier waarop het ecosysteem gemodelleerd is complexer en gedetailleerder is. Hierin lag de focus bij het wiskundig bewijzen van de existentie van verschillende vegetatiepatronen die in de natuur geobserveerd zijn en ook in numerieke simulaties van het model terug te vinden zijn. Stap voor stap wordt er in hoofdstuk 3 ‘gebouwd’ aan verschillende mogelijke structuren en patronen. Door middel van geometrische technieken wordt er gekeken naar de kenmerken van het verwoestijningsproces en de verschillende vegetatiepatronen die het model vertoont. Zo is er begonnen met een simpele overgang van gebied met vegetatie (begroeid) naar een kaal gebied (onbegroeid), naar een gat in de vegetatie, enzovoort (zie linkerhelft bijgevoegd figuur 4.3). Deze patronen worden daarna gebruikt als bouwstenen om de existentie van nieuwe vegetatiepatronen die ook in algemene dynamische systemen voorkomen te bewijzen (zie rechterhelft figuur 4.3). Hiermee wordt een groot spectrum aan bestaande en nieuwe vegetatiepatronen bestudeerd en wordt de basis voor verder onderzoek hiernaar gelegd.

Tot slot worden in hoofdstuk 4 de overeenkomsten tussen en verdere uitbreidingen en onderzoeksvragen omtrent beide voorgaande hoofdstukken besproken.



**Figure 4.3:** *Linkerhelft* : Twee basis vegetatiepatronen van het bestudeerde model van hoofdstuk 3: een overgang van begroeid naar kaal gebied en een gat in een begroeid gebied. *Rechterhelft*: Twee schetsen van ‘nieuwe’ gebouwde gelocaliseerde vegetatiepatronen: een enkele overgang tussen een kaal gebied en een begroeid gebied en een lapje begroeiing omringd door kale grond.

# Curriculum vitae

---

Olfa Jaïbi was born on the 11<sup>th</sup> of August 1989. From 2001 to 2008 she attended Lycée Carthage Byrsa in Carthage, Tunisia, where she graduated from high school *summa cum laude*. In the fall of 2008, she emigrated to the Netherlands to enroll at Leiden University and follow both a bachelor program in Mathematics and in Physics. She completed both bachelor degrees in 2013. As a middle ground, she pursued her studies in Theoretical Physics. She worked on her master thesis entitled ‘Topology and geometry in a quantum condensed matter system: Weyl semimetals.’ under the supervision of Dr. Robert-Jan Slager and Prof. Jan Zaanen. Olfa then graduated *cum laude* in the beginning of 2016. Ten days after her master degree graduation, she started with a PhD on the other side of the road (still at Leiden University) under the supervision of Prof. Arjen Doelman and Dr. Martina Chirilus Brucker on a, back-then, ill-defined project, revolving around ‘fairy circles’ and pattern formation in ecosystems. The PhD project turned out to be diverse, merging the different fields of mathematics and ecology. During her PhD, she had the opportunity to present her work at different major SIAM conferences like ‘Snowbird’ and ‘EquaDiff’.

During her studies and PhD, Olfa was also active within the Physics and Mathematical Institute. She worked as a "mentor" for first-year students and promoted physics in Leiden, being the front figure of the ‘voorlichtingscommissie’ as a student ambassador. She also played an active role within both the studies’ union de Leidsche Flesch as well as the student’s union Catena. She was a board member of the Leiden University Pride organization in 2018-2019 and has been a board member of the European Women in Mathematics-The Netherlands branch since 2017 until now. She has also organized the NDNS+ PhD Days, the PhD colloquia at the Mathematical Institute and the group meetings of the Analysis Group. Besides her student life, she has led an active role standing for her personal beliefs, doing lots of activism within and alongside animal and human rights organizations as well as within the vegan community. She is also a board member of the plant shelter of Leiden.

---

---