Pattern solutions of the Klausmeier Model for banded vegetation in semi-arid environments I

Jonathan A Sherratt

Department of Mathematics and Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK

E-mail: jas@ma.hw.ac.uk

Abstract

In many semi-arid environments, vegetation cover is sparse, and is self-organized into large-scale spatial patterns. In particular, banded vegetation is typical on hillsides. Mathematical modelling is widely used to study these banded patterns, and many models are effectively extensions of a coupled reaction–diffusion–advection system proposed by Klausmeier (1999 *Science* **284** 1826–8). However, there is currently very little mathematical theory on pattern solutions of these equations. This paper is the first in a series whose aim is a comprehensive understanding of these solutions, which can act as a springboard both for future simulation-based studies of the Klausmeier model, and for analysis of model extensions. The author focusses on a particular part of parameter space, and derives expressions for the boundaries of the parameter region in which patterns occur. The calculations are valid to leading order at large values of the 'slope parameter', which reflects a comparison of the rate of water flow downhill with the rate of vegetation dispersal. The form of the corresponding patterns is also studied, and the author shows that the leading order equations change close to one boundary of the parameter region in which there are patterns, leading to a homoclinic solution. Conclusions are drawn on the way in which changes in mean annual rainfall affect pattern properties, including overall biomass productivity.

1. Introduction

In many semi-arid environments, vegetation cover is sparse, and is self-organized into spatial patterns. For trees and shrubs, pattern wavelengths of about 1 km are typical, with bands of

vegetation separated by gaps of bare ground; shorter wavelengths are observed for grasses (see Valentin *et al* (1999) and Rietkerk *et al* (2004) for review). The patterns are most striking on gentle slopes, where they typically consist of stripes running along the contours. Such banded vegetation is hard to detect on the ground, and it was first observed in aerial photographs of sub-Saharan Africa (MacFadyen 1950, Hemming 1965, Wickens and Collier 1971). Banded vegetation patterns are now known to also be widespread in Australia (Mabbutt and Fanning 1987, Dunkerley and Brown 2002) and Mexico (Montaña *et al* 1990, Montaña 1992), with more limited occurrence in South America, Asia and the Middle East (see table 1 and figure 3 of Valentin *et al* (1999)).

The geographical remoteness of most instances of banded vegetation, coupled with their physical harshness, make field studies difficult and expensive. Moreover, there are no laboratory replicates of the phenomenon. Therefore, theoretical models are an important tool for studying these patterns, and a range of models have been developed (see Borgogno *et al* (2009) for a detailed review). Early models were of cellular automaton type (Mauchamp *et al* 1994, Thiéry *et al* 1995, Dunkerley 1997), but partial differential equations are now the standard formalism. These can be broadly divided into models that assume plant–plant interactions to be based on their structural properties, such as root and crown geometry (e.g. Lefever and Lejeune 1997, Couteron and Lejeune 2001, Barbier *et al* 2008, Lefever *et al* 2009), and those centred around water redistribution. This latter class is currently in the most widespread use, and all of the models are effectively extensions of a model proposed by Klausmeier (1999), whose dimensionless form is

$$\partial u / \partial t = \overbrace{wu^2}^{\substack{\text{plant} \\ \text{growth}}} - \overbrace{Bu}^{\substack{\text{plant} \\ \text{loss}}} + \overbrace{\partial^2 u / \partial x^2}^{\substack{\text{plant} \\ \text{dispersal}}}, \tag{1a}$$

$$\partial w / \partial t = \underbrace{A}_{\substack{\text{rain-} \\ \text{fall}}} - \underbrace{w}_{\substack{\text{evap-} \\ \text{oration}}} - \underbrace{wu^2}_{\substack{\text{uptake} \\ \text{by plants}}} + \underbrace{\nu \partial w / \partial x}_{\substack{\text{flow} \\ \text{downhill}}}. \tag{1b}$$

Here $u(x, t)$ is plant density, $w(x, t)$ is water density, t is time and x is space in a one-dimensional domain of constant slope, with the positive direction being uphill. The (dimensionless) parameters A, B and ν reflect rainfall, plant loss and slope gradient, respectively. For full details of the dimensional model and nondimensionalization, see Klausmeier (1999), Sherratt (2005) or Sherratt and Lord (2007).

Most extensions of the Klausmeier model involve separate variables for soil and surface water (e.g. HilleRisLambers *et al* 2001, von Hardenberg *et al* 2001, Rietkerk *et al* 2002, Gilad *et al* 2007, Ursino 2007, 2009). Some authors have also incorporated rainfall variability (Ursino and Contarini 2006, Guttal and Jayaparakash 2007, Kletter *et al* 2009), a herbivore population (van de Koppel *et al* 2002), more realistic representations of plant dispersal (Pueyo *et al* 2008), and explicit representation of the dependence of model parameters on temperature and atmospheric carbon dioxide (Kefi *et al* 2008). These various studies have made important contributions to the ongoing ecological debates on the causes of vegetation patterns, and the most effective management strategies. However, they are almost exclusively simulation-based, with the mathematical theory available to guide the simulations being extremely limited. Indeed, even for the original model of Klausmeier (1999), there has been little mathematical work. This paper is the first in a series whose objective is a detailed understanding of pattern solutions of the Klausmeier model (1a) and (1b).

Patterned solutions of (1a) and (1b) move in the positive x direction (uphill) at a constant speed. Ecologically this migration arises because of higher moisture levels on the uphill edge of the bands compared with their downhill edge, which is reflected in significant differences in
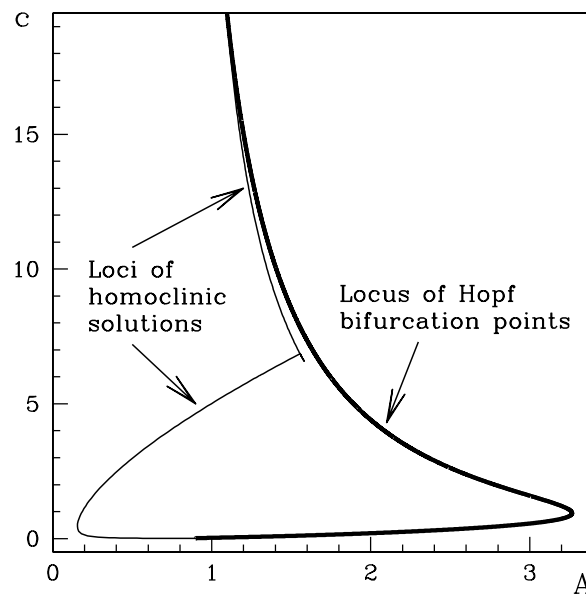
Figure 1. A typical example of the part of the A–c parameter plane in which there are patterned solutions of (1a) and (1b), which corresponds to limit cycles in (2a) and (2b). I plot the loci of Hopf bifurcation points (——) and homoclinic solutions (——) in (2a) and (2b), which bound the pattern region. The other parameters are $B = 0.45$ and $\nu = 182.5$. The plot is truncated at $c \approx 20$: patterns actually exist for values of c up to about 50. The numerical solutions were performed using AUTO (Doedel 1981, Doedel *et al* 1991, 2006). The loci of homoclinic orbits are approximations; they are in fact the loci of solutions of a fixed but very long wavelength (3000). Further details of the numerical continuation approach are given in Sherratt and Lord (2007).

plant death and seedling density (Montaña *et al* 2001, Tongway and Ludwig 2001). Estimates of between 0.2 and 1 m year^{-1} are typical for the migration rate, although accurate data spanning an appropriately long time period are limited (see table 5 of Valentin *et al* 1999). Mathematically, migration is a consequence of the advection term in (1b), and it makes the solution ansatz $u(x, t) = U(z)$, $w(x, t) = W(z)$ appropriate, where $z = x - ct$ with c being the migration speed. Substituting these solution forms into (1a) and (1b) gives the travelling wave equations

$$\mathrm{d}^2 U/\mathrm{d}z^2 + c\,\mathrm{d}U/\mathrm{d}z + WU^2 - BU = 0, \tag{2a}$$

$$(\nu + c)\,\mathrm{d}W/\mathrm{d}z + A - W - WU^2 = 0. \tag{2b}$$

Patterned solutions correspond to limit cycles in this system of ODEs.

In Klausmeier's original paper (Klausmeier 1999), mathematical analysis was limited to the linear stability of spatially homogeneous solutions of (1a) and (1b). The only studies to go beyond this investigate (2a) and (2b) using numerical bifurcation analysis (Sherratt and Lord 2007; Wang *et al* 2009). A typical result is illustrated in figure 1, which shows the region of the A–c parameter plane in which patterns occur, for fixed values of B and ν. For a given value of the migration speed c, patterns occur for a range of rainfall parameter values A, bounded above by a Hopf bifurcation point and below by a homoclinic solution; details of the relevant homogeneous equilibria are given in section 2. Intuitively, sufficiently high rainfall levels give rise to only homogeneous vegetation, while very low rainfall levels imply a complete desert. Intermediate rainfall levels are too low to maintain homogeneous plant cover, but are compatible with vegetation bands.

The dimensionless parameter ν depends on the ratio of the advection rate of water and the (square root of the) plant diffusion coefficient (Klausmeier 1999, Sherratt 2005); as a result it is very large, with a value of 200 being typical. In contrast, estimates of the plant loss parameter B are between about 0.05 and 2 (Klausmeier 1999, Rietkerk *et al* 2002), with the rainfall parameter A varying between about 0.1 and 3 depending on geographical region and plant type (Klausmeier 1999). Therefore a natural mathematical approach is to investigate the leading order asymptotic form of the patterns for large ν. When doing this, I follow the approach of Sherratt and Lord (2007) and regard A and c as the primary parameters: in applications, one is mainly concerned with the implications of significant changes in rainfall level, set against the background of a relatively constant rate of plant loss. Note, however, that small changes in B due to altered grazing and cropping may be important as a tool for ecosystem management (Freudenberger and Hiernaux 2001, Noble *et al* 2001).

A major complication for this asymptotic approach is that the region in the A–c parameter plane in which patterns occur (see figure 1) spans a range of order of magnitude dependences of both A and c on ν. For example, the maximum rainfall level A at which patterns occur is $O_s(\nu)$, with the corresponding speed c being $O_s(1)$ (Sherratt 2005), while the homoclinic and Hopf bifurcation loci intersect at $A = O_s(1)$ and $c = o(1)$ (Sherratt and Lord 2007; this point is visible near the $c = 0$ axis in figure 1). Note that the notation $f = O_s(g)$ denotes $f = O(g)$ and $f \neq o(g)$. As a result, a number of different cases must be considered, for different regions of the parameter plane. This paper is the first in a series in which I will consider these cases, documenting the boundaries of parameter space in which patterns occur, and the form of these patterns. My overall objective in this programme of work is a comprehensive understanding of patterned solutions of the Klausmeier model, which can act as a springboard both for simulation-based studies of that model and for analysis of model extensions. This requires essentially separate mathematical studies for different parts of the parameter region giving patterns. In this paper I focus on the 'tusk-shaped' region that occurs for c greater than about 6.5 in figure 1; I will show that this occurs when $\nu^{1/2} \ll c \ll \nu$ and $A^2 c = O_s(\nu)$.

In section 2 I discuss spatially homogeneous steady states and their stability. In section 3 I study the equations that govern pattern solutions to leading order as $\nu \to \infty$, stating a theorem on the existence of a homoclinic solution that I prove in section 4. In section 5 I show that the equations studied in sections 3 and 4 cease to be the leading order equations for patterns in a thin strip in the parameter plane, close to one boundary and in section 6 I study the patterns that occur for parameters in this strip. In section 7 I summarize the main conclusions and discuss their ecological implications.

2. Spatially homogeneous steady states

For all parameter values, (1a) and (1b) has a stable trivial steady state $u = 0$, $w = A$, corresponding to bare ground, without vegetation. When $A \geqslant 2B$ there are also two other homogeneous steady states:

$$u = u_u \equiv \frac{A - \sqrt{A^2 - 4B^2}}{2B} \qquad w = w_u \equiv \frac{2B^2}{A - \sqrt{A^2 - 4B^2}}, \qquad (3)$$

$$u = u_s \equiv \frac{A + \sqrt{A^2 - 4B^2}}{2B} \qquad w = w_s \equiv \frac{2B^2}{A + \sqrt{A^2 - 4B^2}}. \qquad (4)$$

The first of these is always unstable to homogeneous perturbations, while the second is stable to homogeneous perturbations provided B is not too large (Klausmeier 1999, Sherratt 2005); $B < 2$ is a sufficient condition for this, and holds for any realistic parameter values for semi-arid

environments (Klausmeier 1999). For larger values of B, more complicated local dynamics occur when A is small, but these are not relevant in applications, and henceforth I assume that $B < 2$.

The region of the A–c parameter plane in which patterns occur is bounded on one side by the locus of Hopf bifurcation points of the steady state (u_s, w_s) in the travelling wave equations (2a) and (2b) (illustrated in figure 1). For $v \gg 1$, this locus can be calculated explicitly. Linearizing (2a) and (2b) about (u_s, v_s) gives the eigenvalue equation

$$(v + c)\lambda^3 + (cv + c^2 - 1 - u_s^2)\lambda^2 + (Bv + Bc - c - cu_s^2)\lambda + B(u_s^2 - 1) = 0.$$

Comparing this with the generic form of a cubic eigenvalue equation at Hopf bifurcation

$$0 = (\lambda^2 + C_1^2)(C_2\lambda + C_3) = C_2\lambda^3 + C_3\lambda^2 + C_1^2 C_2\lambda + C_1^2 C_3$$

gives the conditions for Hopf bifurcation as

$$(v + c)(Bu_s^2 - B) - (cv + c^2 - 1 - u_s^2)(Bv + Bc - c - cu_s^2) = 0, \qquad (5a)$$

and

$$(u_s^2 - 1) \cdot (cv + c^2 - 1 - u_s^2) > 0. \qquad (5b)$$

Condition (5a) and (5b) is a cubic equation for c, and the solution relevant to the 'tusk-shaped' part of the pattern region is the largest of three real roots. Expanding c as a power series in v shows that this largest root is $c = B^3 v/A^2 + o(v)$, provided that $1 \ll c \ll v$ (i.e. $1 \ll A \ll v^{1/2}$). Moreover, $u_s^2 = A^2/B^2 + o(1)$ as $A \to \infty$, so that $cv + c^2 - 1 - u_s^2 \sim cv - Bv/c \sim cv$ as $v \to \infty$, and thus (5b) is satisfied.

3. Leading order pattern equations

In this paper I restrict attention to the 'tusk-shaped' part of the pattern region in the A–c parameter plane, which occurs for c greater than about 6.5 in figure 1. The discussion in section 2 shows that this is bounded on one side by the Hopf bifurcation locus $cA^2 = B^3 v$, to leading order as $v \to \infty$. To investigate further the patterns that occur within the region, it is convenient to rescale the equations. The appropriate rescalings are indicated via the Hopf bifurcation locus; straightforward calculation shows that along this locus $u_s \sim A/B$ and $w_s \sim B^2/A$ as $v \to \infty$, with the neutrally stable oscillations having period $2\pi c/B$. Therefore, I set

$$\widetilde{U} = (B/A)U \qquad \widetilde{W} = (A/B^2)W \qquad \tilde{z} = (B/c)z. \qquad (6)$$

Substituting these into (2a) and (2b) gives

$$(B/c^2)\, d^2\widetilde{U}/d\tilde{z}^2 + d\widetilde{U}/d\tilde{z} + \widetilde{U}^2\widetilde{W} - \widetilde{U} = 0, \qquad (7a)$$

$$(v + c) \cdot (B^3/cA)\, d\widetilde{W}/d\tilde{z} + A - (B^2/A)\widetilde{W} - A\,\widetilde{U}^2\widetilde{W} = 0. \qquad (7b)$$

Since $A \gg 1$ and $1 \ll c \ll v$, the leading order form of these equations as $v \to \infty$ is

$$d\widetilde{U}/d\tilde{z} = \widetilde{U} - \widetilde{U}^2\widetilde{W}, \qquad (8a)$$

$$d\widetilde{W}/d\tilde{z} = \sigma^2(\widetilde{U}^2\widetilde{W} - 1), \qquad (8b)$$

where $\sigma = Ac^{1/2}/B^{3/2}v^{1/2}$, which is $O_s(1)$ as $v \to \infty$ near the Hopf bifurcation curve. The rest of this section and section 4 are concerned with limit cycle solutions of (8a) and (8b).

Equations (8a) and (8b) have a unique steady state: $\widetilde{U} = \widetilde{W} = 1$. Explicit calculation of eigenvalues shows that this steady state is stable if and only if $\sigma < 1$, with a Hopf bifurcation at $\sigma = 1$; as expected, this condition is the same as the leading order condition for Hopf bifurcation in the full travelling wave equations (2a) and (2b). Standard calculation
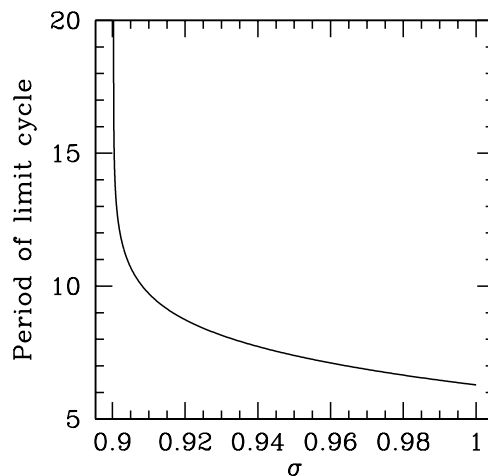
Figure 2. Numerical continuation of the branch of limit cycles emanating from the Hopf bifurcation in (8a) and (8b) at $\sigma = 1$. I plot the period of the limit cycle against the parameter σ. The limit cycle branch proceeds monotonically with σ, terminating at a homoclinic orbit at $\sigma \approx 0.9003$. The computation was performed using the numerical continuation software package AUTO (Doedel 1981, Doedel *et al* 1991, 2006).

(e.g. section 3.4 of Guckenheimer and Holmes 1983) shows that the Hopf bifurcation is subcritical, so that there are limit cycles (patterns) for σ just below 1. I have been unable to determine analytically the behaviour of the limit cycle branch away from the neighbourhood of the Hopf bifurcation point, but an important indicator of the way in which the branch terminates is given by the following result.

Theorem. *For equations (8a) and (8b) there is exactly one value of $\sigma = \sigma^* \in \left(\sqrt{16/27}, 1\right)$ for which there is a homoclinic solution.*

The proof of this theorem is presented in section 4, and is constructive in the sense that it gives a detailed picture of the form of the limit cycles for σ just above σ^*; the form is unusual and will be important in determining the behaviour of limit cycles of the full equations (2a) and (2b) as $\sigma \to \sigma^{*+}$.

Numerical continuation of the limit cycle branch of (8a) and (8b) is straightforward; I used the software package AUTO (Doedel 1981, Doedel *et al* 1991, 2006). This shows that the solution branch proceeds monotonically in σ from the Hopf bifurcation point at $\sigma = 1$, terminating at the homoclinic orbit identified in the theorem (illustrated in figure 2). Numerical computations show that $\sigma^* \approx 0.9003$.

At first glance, it appears that the investigation of the 'tusk-shaped' part of the pattern region in the A–c plane is complete, since (8a) and (8b) have limit cycle solutions for parameter values between the loci of a Hopf bifurcation point and a homoclinic orbit, exactly as illustrated for the full equations (2a) and (2b) in figure 1. However, it transpires that the homoclinic orbit in (2a) and (2b) is quite different from that in (8a) and (8b). This is because when σ is sufficiently close to σ^*, (8a) and (8b) cease to be the leading order form of (7a) and (7b) for large ν. Explanation of this, and determination of the actual form of the homoclinic solution of (2a) and (2b), requires a fuller understanding of the homoclinic solution of (8a) and (8b), which is provided by the proof of the theorem.
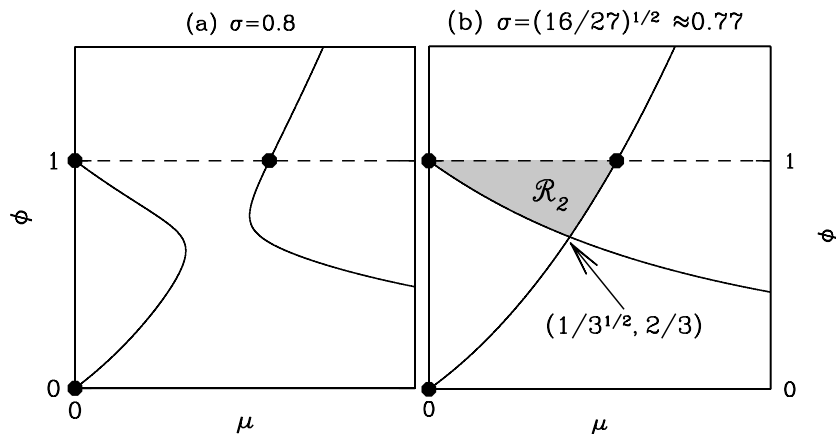
Figure 3. An illustration of the form of the ϕ nullcline for equations (9a) and (9b). (a) A typical example for $\sigma > \sqrt{16/27}$: there are two nullcline branches. (b) The case of $\sigma = \sqrt{16/27}$, when the two branches intersect. The three large dots in each picture indicate the three steady states $(0, 0)$, $(\sigma, 1)$ and $(0, 1)$. In (b), the shaded region is $\mathcal{R}_2$, defined in section 4.4. The nullcline is given explicitly by $\mu = [\sigma \pm (\sigma^2 - 4\phi^2(1-\phi))^{1/2}]/(2\phi)$.

4. Proof of the theorem

4.1. Alternative formulation of the equations

Direct investigation of the homoclinic solution of (8a) and (8b) is difficult because, as I will show, the solution is actually homoclinic to a point at infinity. Therefore, it is convenient to rewrite the equations, replacing $\widetilde{W}$ by $\phi = \widetilde{U}\widetilde{W}$. For algebraic simplicity I also replace $\widetilde{U}$ by $\mu = \sigma\widetilde{U}$, giving

$$d\mu/d\tilde{z} = \mu - \mu\phi, \tag{9a}$$

$$d\phi/d\tilde{z} = \mu^2\phi - \sigma\mu + \phi - \phi^2. \tag{9b}$$

Equations (9a) and (9b) have three steady states: $(\mu, \phi) = (0, 0)$, $(\sigma, 1)$ and $(0, 1)$. The first of these plays no role in the proof, and explicit calculation of eigenvalues shows that the second is a stable focus for all $\sigma \in (\sqrt{16/27}, 1)$. The steady state $(0, 1)$ is key to the proof, since I will show that it is this steady state to which the homoclinic solution connects; note that this steady state corresponds to a point at infinity in the original $\widetilde{U}$–$\widetilde{W}$ formulation. $(0, 1)$ is non-hyperbolic. It has a stable eigenvector $(0, 1)$, and an unstable centre manifold $\phi = 1 - \sigma\mu + \mu^2 + O(\mu^3)$ for $\mu > 0$, along which $\mu = -1/(\sigma\tilde{z}) + O(1/\tilde{z}^2)$ and $\phi = 1 + 1/\tilde{z} + O(1/\tilde{z}^2)$ as $\tilde{z} \to -\infty$.

The nullclines play an important role later in the proof, and are easily calculated. The μ nullclines are $\mu = 0$ and $\phi = 1$, independent of σ. However, the ϕ nullcline depends on σ, and for $\sigma \in (\sqrt{16/27}, 1)$, there are two nullcline branches, whose form is illustrated in figure 3(a); at $\sigma = \sqrt{16/27}$ these two branches meet (figure 3(b)).

4.2. Classification of the phase plane into three cases

There is exactly one trajectory, $\mathcal{T}$ say, leaving the steady state $(0, 1)$, and it does so along the centre manifold $\phi = 1 - \sigma\mu + O(\mu^2)$, entering the strip $0 < \phi < 1$. Since $d\mu/d\tilde{z} > 0$

throughout this strip, the phase plane can be classified into one of three cases.

Case 1. $\mathcal{T}$ leaves $0 < \phi < 1$ through the line $\phi = 0$.

Case 2. $\mathcal{T}$ remains within $0 < \phi < 1$. This requires $d\phi/d\tilde{z} \to 0$ as $\tilde{z} \to \infty$, so that $\mathcal{T}$ must approach the right-hand branch of the ϕ nullcline as $\tilde{z} \to \infty$, with $\mu \to \infty$.

Case 3. $\mathcal{T}$ leaves $0 < \phi < 1$ through the line $\phi = 1$.

I will show that cases 1 and 3 apply when $\sigma = 1$ and $\sigma = \sqrt{16/27}$, respectively, and I will then show that this implies that case 2 holds for exactly one intermediate value of σ, and that this corresponds to a homoclinic solution.

An important general consideration underlying the argument in various places is the location of $\mathcal{T}$ relative to the ϕ nullcline. Explicit calculation shows that near the steady state $(0, 1)$, the ϕ nullcline is $\phi = 1 - \sigma \mu + (1 - \sigma^2)\mu^2 + O(\mu^3)$; recall that $\mathcal{T}$ leaves $(0, 1)$ along the centre manifold $\phi = 1 - \sigma \mu + \mu^2 + O(\mu^3)$. Therefore for all $\sigma > 0$, $\mathcal{T}$ lies between the ϕ nullcline and the line $\phi = 1$, sufficiently close to $(0, 1)$. Moreover, $d\mu/d\tilde{z} > 0$ when $0 < \phi < 1$, implying that all trajectories crossing the ϕ nullcline with $0 < \phi < 1$ do so in the direction of increasing μ. Thus $\mathcal{T}$ must remain to the right of the left-hand branch of the ϕ nullcline, when $\sigma \in \left(\sqrt{16/27}, 1\right)$.

4.3. Behaviour when $\sigma = 1$

Fixing $\sigma = 1$, I define $\mathcal{L}_1$ to be the line segment $\phi = 1 - \mu/2$, $0 < \mu < 2$. I further define $\mathcal{R}_1$ to be the triangular region bounded by $\mathcal{L}_1$ and the line segments $\phi = 0$, $0 < \mu < 2$ and $\mu = 0$, $0 < \phi < 1$ (illustrated in figure 4(a)). Since the slope of $\mathcal{T}$ at $(0, 1)$ is -1, $\mathcal{T}$ enters $\mathcal{R}_1$. Now along $\mathcal{L}_1$,

$$d\phi/d\mu = \frac{\phi}{\mu} + \frac{\mu\phi - 1}{1 - \phi} = -\frac{1}{2} - \frac{(\mu - 1)^2}{\mu} \leqslant -\frac{1}{2},$$

implying that all trajectories crossing $\mathcal{L}_1$ do so into $\mathcal{R}_1$. Therefore $\mathcal{T}$ cannot exit $\mathcal{R}_1$ through $\mathcal{L}_1$, and neither can it leave through $\mu = 0$, $0 < \phi < 1$, which is a solution trajectory. Moreover, there are no steady states in the interior of $\mathcal{R}_1$, and $\mathcal{T}$ cannot terminate at the vertex $(0, 0)$ (which is a steady state) since $d\mu/d\tilde{z} > 0$ throughout the interior of $\mathcal{R}_1$. Therefore $\mathcal{T}$ must exit $\mathcal{R}_1$ through $\phi = 0$, i.e. case 1 applies.

4.4. Behaviour when $\sigma = \sqrt{16/27}$

When $\sigma = \sqrt{16/27}$, the two branches of the ϕ nullcline meet (at $(\mu, \phi) = (1/\sqrt{3}, 2/3)$), and I define $\mathcal{R}_2$ to be the region bounded by the line segment $\phi = 1$, $0 < \mu < \sqrt{16/27}$ and the part of the ϕ nullcline with $1 > \phi > 2/3$ (illustrated in figure 3(b)). In section 4.2 I showed that $\mathcal{T}$ enters $\mathcal{R}_2$. Since $d\phi/d\tilde{z} < 0$ on $\phi = 1$, $0 < \mu < \sqrt{16/27}$, and $d\mu/d\tilde{z} > 0$ in the interior of $\mathcal{R}_2$, $\mathcal{T}$ must exit $\mathcal{R}_2$ through the part of the ϕ nullcline with positive slope, i.e. at a value of $\mu \in \left(1/\sqrt{3}, \sqrt{16/27}\right)$. Subsequently, μ and ϕ both increase with $\tilde{z}$ along $\mathcal{T}$, until the trajectory crosses the line $\phi = 1$, i.e. case 3 applies.

4.5. Case 3 for $\sigma = \sigma_0$ implies case 3 for $\sigma < \sigma_0$

I now show that if case 3 applies for some $\sigma = \sigma_0 \in \left(\sqrt{16/27}, 1\right)$, then it also applies for all $\sigma \in \left(\sqrt{16/27}, \sigma_0\right)$. I denote by $\mathcal{T}_0$ the curve in the μ–ϕ plane traced out by the trajectory $\mathcal{T}$ when $\sigma = \sigma_0$, and I define $\mathcal{R}_3$ to be the region enclosed by the segment of $\mathcal{T}_0$ between $(0, 1)$
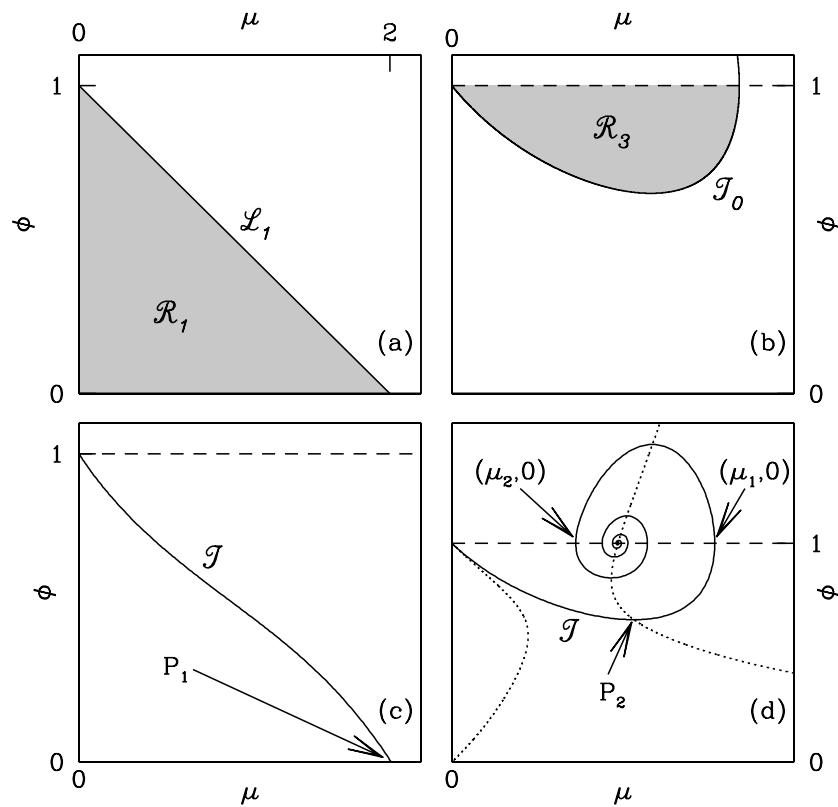
Figure 4. An illustration of various lines, regions and points in the μ–ϕ plane that are referred to in the proof of the theorem. (*a*) The line $\mathcal{L}_1$ and region $\mathcal{R}_1$, defined in section 4.3. (*b*) The region $\mathcal{R}_3$, defined in section 4.5. (*c*) The point P_1, defined in section 4.6. (*b*) The point P_2, defined in section 4.6, and the points $(\mu_1, 0)$ and $(\mu_2, 0)$, defined in section 4.7; the dotted lines are the ϕ nullclines.

and its first crossing of $\phi = 1$, and the segment of the line $\phi = 1$ between $(0, 1)$ and this crossing point (illustrated in figure 4(*b*)). The form of the centre manifold at $(0, 1)$ implies that the slope of $\mathcal{T}$ at $(0, 1)$ is $-\sigma$, and thus for all $\sigma < \sigma_0$, $\mathcal{T}$ enters $\mathcal{R}_3$. Now $\mathrm{d}\phi/\mathrm{d}\mu$ is a decreasing function of σ in the strip $0 < \phi < 1$, with $\mathrm{d}\mu/\mathrm{d}\tilde{z} > 0$; thus all trajectories crossing $\mathcal{T}_0$ do so into $\mathcal{R}_3$. Moreover, there are no steady states in the interior of $\mathcal{R}_3$. Therefore for any $\sigma < \sigma_0$, $\mathcal{T}$ must exit $\mathcal{R}_3$ through $\phi = 1$, i.e. case 3 applies.

4.6. Application of the continuity of $\mathcal{T}$

The results in sections 4.3–4.5 together imply the existence of $\sigma^* \in \left[\sqrt{16/27}, 1 \right]$ such that case 3 applies for $\sigma < \sigma^*$, while case 1 or case 2 apply for $\sigma > \sigma^*$. I now define two special points in the μ–ϕ plane. When case 1 applies, I denote by P_1 the point at which $\mathcal{T}$ first crosses $\phi = 0$ (illustrated in figure 4(*c*)), and when case 3 applies, I denote by P_2 the point at which $\mathcal{T}$ first crosses the ϕ nullcline (illustrated in figure 4(*d*)). Note that I have shown in section 4.1 that $\mathcal{T}$ lies to the right of the left-hand branch of the ϕ nullcline, implying that P_2 lies on the right-hand branch.

Now $\mathcal{T}$ varies continuously with σ. Therefore, if case 1 were to apply when $\sigma = \sigma^*$, the locus of P_2 would have to terminate (as $\sigma \to \sigma^{*-}$) at $P_1|_{\sigma=\sigma^*}$, which would be a finite point;

this is impossible. Similarly case 3 cannot apply when $\sigma = \sigma^*$, leaving only the possibility that case 2 applies. Continuity is then possible, with $\lim_{\sigma \to \sigma^{*+}} P_1 = \lim_{\sigma \to \sigma^{*-}} P_2 = (+\infty, 0)$. Note that I have already shown that $\sigma^* \in \left[\sqrt{16/27}, 1 \right]$, and since cases 1 and 3 apply at $\sigma = 1$ and $\sigma = \sqrt{16/27}$, respectively, I can now conclude further that $\sigma^* \in \left(\sqrt{16/27}, 1 \right)$.

As an aside, I mention that numerical calculations suggest that P_1 moves monotonically along the line $\phi = 0$ as σ increases above σ^*. However, I have not proved this, and thus I cannot exclude the possibility that case 2 applies for values of $\sigma > \sigma^*$. This is not an issue for the proof of the theorem, though, which does not require uniqueness of case 2.

4.7. The form of $\mathcal{T}$ as $\sigma \to \sigma^{*-}$

For $\sigma < \sigma^*$, I define μ_1 and μ_2 to be the values of μ at which $\mathcal{T}$ crosses the line $\phi = 1$ for the first and second times (illustrated in figure 4(d)). Now $\phi = 1$ is a μ nullcline, on which $d\phi/d\tilde{z}$ has the same sign as $\mu - \sigma$. Therefore $0 < \mu_2 < \sigma$, so that $\lim_{\sigma \to \sigma^{*-}} \mu_2$ exists and lies in $[0, \sigma^*]$. I will show that the limiting value is zero, by contradiction. Suppose that $\lim_{\sigma \to \sigma^{*-}} \mu_2 \in (0, \sigma^*]$. Then fixing $\sigma = \sigma^*$ and integrating (9a) and (9b) backwards in $\tilde{z}$ shows that μ_1 is finite, since $d\mu/d\tilde{z}$ and $d\phi/d\tilde{z}$ remain bounded throughout the integration. But $d\mu/d\tilde{z} > 0$ in $0 < \phi < 1$, and thus $\mu|_{P_2} < \mu_1$ when $\sigma = \sigma^*$. This contradicts the result in section 4.6 that $P_2 \to (+\infty, 0)$ as $\sigma \to \sigma^{*-}$.

Therefore $\mu_2 \to 0$ as $\sigma \to \sigma^{*-}$, implying that the solution trajectory $\mathcal{T}$ approaches an orbit that is homoclinic to $(0, 1)$ as $\sigma \to \sigma^{*-}$. This completes the proof of the theorem.

5. Breakdown of the leading order equations (8a) and (8b)

Numerical continuation (e.g. figure 2) indicates that the branch of limit cycles of (9a) and (9b) emanating from the Hopf bifurcation ($\sigma = 1$) proceeds monotonically in σ, terminating at the homoclinic solution ($\sigma = \sigma^* \approx 0.9003$). I have been unable to prove this, but in the rest of the paper I assume it to be true, on the basis of strong numerical evidence. The theorem then gives a clear picture of the form of limit cycles when σ is close to (and above) σ^*. In particular, the maxima of μ and ϕ on the limit cycle both $\to \infty$ as $\sigma \to \sigma^{*+}$. For $\sigma \in (\sigma^*, 1)$, I define $\mu_{\max}(\sigma)$ to be the maximum value of μ on the limit cycle of (9a) and (9b); recall that $\mu = \sigma \widetilde{U}$. Then scaling arguments indicate that the maximum value of ϕ on the limit cycle of (9a) and (9b) is also $O_s(\mu_{\max})$ as $\sigma \to \sigma^{*+}$, with a scaling of $\tilde{z}$ needed in order to balance the terms in the equations. Specifically, substituting

$$\hat{\mu} = \mu/\mu_{\max} \qquad \hat{\phi} = \phi/\mu_{\max}^2 \qquad \hat{z} = \mu_{\max}^2 \tilde{z} \qquad (10)$$

into (9a) and (9b) gives

$$d\hat{\mu}/d\hat{z} = -\hat{\mu}\hat{\phi} + O(\mu_{\max}^{-2}),$$
$$d\hat{\phi}/d\hat{z} = \hat{\mu}^2\hat{\phi} - \hat{\phi}^2 + O(\mu_{\max}^{-2})$$

as $\sigma \to \sigma^{*+}$. These imply that $\hat{\phi} = \hat{\mu}(\hat{k} - \hat{\mu})$ to leading order, for some constant of integration $\hat{k}$, and this parabolic shape is evident in numerical solutions of (9a) and (9b) for σ just above σ^*.

The scalings (10) have two key implications for the limit cycle solution of (8a) and (8b)/(9a) and (9b). Firstly, as $\sigma \to \sigma^{*+}$, part of the solution changes over a $\tilde{z}$ length scale that is $O(\mu_{\max}^{-2})$, and secondly the maximum value of $\widetilde{W} \to \infty$ in proportion to $\mu_{\max}$ (recall that $\widetilde{W} = \sigma\phi/\mu$). Both of these invalidate the derivation of (8a) and (8b) as the leading order form of (7a) and (7b) for large ν. A sufficiently rapid (in $\tilde{z}$) change in the solution means that the term $(B/c^2)d^2\widetilde{U}/d\tilde{z}^2$ in (7a) and (7b) cannot be neglected, while sufficiently large $\widetilde{W}$ causes the term $(B^2/A)\widetilde{W}$ to be comparable in size to A, and hence non-negligible.
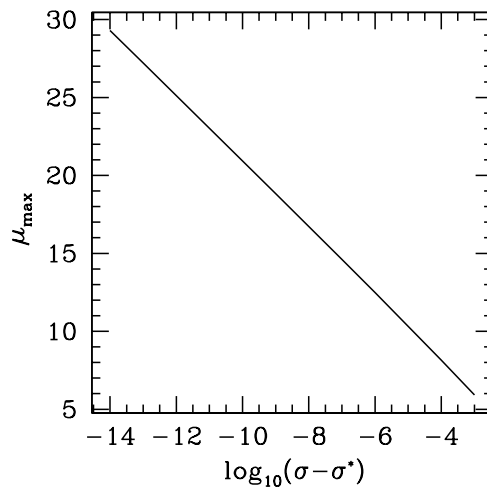
Figure 5. The increase in $\mu_{\max}$, the maximum value of μ in the limit cycle solution of (9a) and (9b), as σ approaches σ^* (from above). Computations are done with 16-byte precision. These results suggest that $\mu_{\max} = O_s(\log(\sigma - \sigma^*))$ as $\sigma \to \sigma^{*+}$. The numerical procedure involves two distinct stages. The first stage is to determine accurately the value of σ^*. I did this by numerical integration of (9a) and (9b) starting on the centre manifold: σ is below/above σ^* if the solution crosses $\mu = 1/\mu = 0$. The second stage of the calculation is to integrate (9a) and (9b) for a sequence of values of σ just above σ^*. In each case I use initial conditions close to the equilibrium $(\sigma, 1)$, which lies inside the limit cycle, integrating backwards in $\tilde{z}$ for a long interval, and then recording the next two points at which the right-hand side of (9a) is zero: the larger of the two μ values provides the estimate of $\mu_{\max}$. The equations were solved numerically using the routine DLSODAR (Hindmarsh 1983, Petzold 1983), which is part of the ODEPACK collection, and is freely available at www.netlib.org. This routine incorporates root-finding capability, enabling easy calculation of the zeros of the right-hand side of (9a).

Therefore, the 'tusk-shaped' part of the region of c–A parameter space giving patterns cannot be studied as a single unit. In the bulk of the region, as $\nu \to \infty$, patterns approach the limit cycle solutions of (9a) and (9b), for which sections 3 and 4 provide a detailed account. However, there is a thin strip in parameter space, close to the curve $c = \sigma^{*2}B^3\nu/A^2$ (i.e. $\sigma = \sigma^*$), in which different scalings apply. The width of this strip depends on the rate at which $\mu_{\max} \to \infty$ as $\sigma \to \sigma^{*+}$. I have been unable to determine this rate analytically, but a numerical study is relatively straightforward (figure 5). Details of the numerical method are given in the legend to figure 5; note that the required level of accuracy necessitates 16-byte numerical precision. The results in figure 5 suggest that $\mu_{\max} = O_s(\log(\sigma - \sigma^*))$ as $\sigma \to \sigma^{*+}$. Note, however, that this is not assumed in the subsequent analysis.

The term $(B/c^2)\mathrm{d}^2\widetilde{U}/\mathrm{d}\tilde{z}^2$ in (7a) can be neglected to leading order as $\nu \to \infty$ provided that $\mu_{\max} \ll c$, while the condition for (8b) to be the leading order form of (7b) is $\mu_{\max} \ll A^2$. Throughout the region of parameter space that I am studying, $\sigma = O_s(1)$ as $\nu \to \infty$, $\Rightarrow A^2c = O_s(\nu)$. Therefore if $\nu^{1/2} \ll c \ll \nu$, the initial breakdown of (8a) and (8b) occurs because the term $(B/A)\widetilde{W}$ in (8b) must be retained to leading order. I will show (section 6) that the resulting equations contain a homoclinic orbit which is reached with the assumption $\mu_{\max} \ll c$ remaining valid. The scaling range $\nu^{1/2} \ll c \ll \nu$ defines the bulk of the 'tusk-shaped' parameter region that I am considering. When $c = O_s(\nu^{1/2})$ and $A = O_s(\nu^{1/4})$, (8a) and (8b) can become invalid simultaneously as the leading order forms of (7a) and (7b). This determines the location of the base of the 'tusk'. Its top occurs at a finite value of c that

is $O_s(\nu)$. Investigation of both of these cases are major studies in themselves, which I will publish separately.

I conclude this section with a final point of clarification. I have shown (section 4) that equations (8a) and (8b) have a homoclinic solution, at $\sigma = \sigma^*$. Nothing in the preceding discussion negates this. Rather, in this section I have additionally shown that this homoclinic solution is not the limit for large ν of any solution of (7a) and (7b). I will show (section 6) that for $\nu^{1/2} \ll c \ll \nu$, the homoclinic solution of (7a) and (7b) has a quite different form. The significance of the homoclinic solution of (8a) and (8b) is that an understanding of its form is an essential precursor to studying the homoclinic solution of (7a) and (7b).

6. Limit cycle solutions near $\sigma = \sigma^*$

I have shown that equations (8a) and (8b) cease to be the leading order form of (7a) and (7b) as $\nu \to \infty$ in a thin strip in the c–A parameter plane, adjacent to the curve $\sigma = \sigma^*$. To study behaviour in this strip, it is convenient to rewrite (7a) and (7b) as

$$(B/c^2)\,\mathrm{d}^2\widetilde{U}/\mathrm{d}\tilde{z}^2 + \mathrm{d}\widetilde{U}/\mathrm{d}\tilde{z} + \widetilde{U}^2\widetilde{W} - \widetilde{U} = 0, \tag{11a}$$

$$(1 + c/\nu)\,\mathrm{d}\widetilde{W}/\mathrm{d}\tilde{z} + \sigma^2(1 - \widetilde{U}^2\widetilde{W}) - \epsilon\widetilde{W} = 0, \tag{11b}$$

where $\epsilon = B^2\sigma^2/A^2 = c/(B\nu)$. (Recall that $\sigma^2 = A^2c/(B^3\nu)$). Note that as $\nu \to \infty$, $A \gg 1$; since σ is bounded, this implies that $\epsilon \ll 1$. My objective is to investigate limit cycle solutions of (11a) and (11b), for large ν. I will derive such solutions in the form of four matched layers, each of which has different scalings of the variables with respect to ν. I will show that my solutions are valid provided that $\nu^{1/2} \ll c \ll \nu$; since $A^2c = O_s(\nu)$ in the parameter region that I am studying in this paper, this is equivalent to $1 \ll A \ll \nu^{1/4}$, i.e. $\nu^{-1/2} \ll \epsilon \ll 1$. I present the layers in a clockwise order going round the limit cycle in the $\widetilde{U}$–$\widetilde{W}$ plane, which is in the direction of decreasing $\tilde{z}$; all rescalings can be given just in terms of the parameter ϵ.

Layer 1. Here there is no rescaling. Therefore provided that $\nu \gg c \gg 1$, the leading order equations are simply (8a) and (8b). For notational consistency with what follows, I write $\widetilde{U} = \widetilde{U}_1$, $\widetilde{W} = \widetilde{W}_1$, $\tilde{z} = \tilde{z}_1$ in this layer.

Layer 2. This is a 'thick' layer, corresponding to slow changes with respect to $\tilde{z}$. The appropriate rescalings are

$$\widetilde{U}_2 = \widetilde{U}/\epsilon \qquad \widetilde{W}_2 = \epsilon\widetilde{W} \qquad \tilde{z}_2 = \epsilon\tilde{z}, \tag{12}$$

$$\Rightarrow (B\epsilon^2/c^2)\mathrm{d}^2\widetilde{U}_2/\mathrm{d}\tilde{z}_2^2 + \epsilon\,\mathrm{d}\widetilde{U}_2/\mathrm{d}\tilde{z}_2 + \widetilde{U}_2^2\widetilde{W}_2 - \widetilde{U}_2 = 0,$$

$$(1 + c/\nu)\mathrm{d}\widetilde{W}_2/\mathrm{d}\tilde{z}_2 + \sigma^2\left(1 - \epsilon\widetilde{U}_2^2\widetilde{W}_2\right) - \widetilde{W}_2 = 0.$$

Therefore provided that $\nu \gg c \gg \epsilon$, the leading order equations are

$$\widetilde{U}_2\left(\widetilde{U}_2\widetilde{W}_2 - 1\right) = 0 \qquad \mathrm{d}\widetilde{W}_2/\mathrm{d}\tilde{z}_2 + \sigma^2 - \widetilde{W}_2 = 0. \tag{13}$$

Equations (13) imply that

$$\widetilde{U}_2 = (\sigma^2 - k_2\mathrm{e}^{\tilde{z}_2})^{-1} \qquad \widetilde{W}_2 = \sigma^2 - k_2\mathrm{e}^{\tilde{z}_2} \tag{14}$$

to leading order, where k_2 is a constant of integration. (The alternative possibility of $\widetilde{U}_2 \equiv 0$ is not relevant.) Because of the different scalings for $\tilde{z}$ in this layer and layer 1, I require that the behaviour of $(\widetilde{U}_1, \widetilde{W}_1)$ as $\tilde{z}_1 \to -\infty$ matches that of $(\widetilde{U}_2, \widetilde{W}_2)$ at some finite value of $\tilde{z}_2$, which I arbitrarily take to be zero. Expanding (14) in power series about $\tilde{z}_2 = 0$ and then undoing the scalings (12) shows that the behaviour of $\widetilde{W}_1$ as $\tilde{z}_1 \to -\infty$ must match $(\sigma^2k_2)/\epsilon - k_2\tilde{z}_1 + \cdots$, which requires $k_2 = \sigma^2$. Then matching of $\widetilde{U}_1$ and $\widetilde{U}_2$ requires $\widetilde{U}_1 \sim 1/(-\sigma^2\tilde{z}_1)$ as $\tilde{z}_1 \to -\infty$.

The substitutions $\mu = \sigma \widetilde{U}_1$ and $\phi = \widetilde{U}_1 \widetilde{W}_1$ convert the leading order equations in layer 1 to (9a) and (9b), and the matching conditions become $\mu \sim -1/(\sigma \tilde{z}_1)$ and $\phi \to 1$ as $\tilde{z}_1 \to -\infty$. Therefore, the required solution in layer 1 corresponds to the unique trajectory in (9a) and (9b) leaving the steady state $(0, 1)$ along the centre manifold (see section 4; the required trajectory is $\mathcal{T}$).

Layer 3: Intuitively, layer 2 reflects the approach of the limit cycle to its minimum value of $\widetilde{U}$, and it is a 'thick' layer, corresponding to the slow (algebraic) approach to $(\mu, \phi) = (0, 1)$ of the trajectory of (9a) and (9b) relevant to layer 1. Layer 3 corresponds to the increase in $\widetilde{U}$ away from its minimum value, and no rescaling of $\tilde{z}$ is required: this corresponds intuitively to (9a) and (9b) having an unstable eigenvector at $(\mu, \phi) = (0, 1)$. However, $\widetilde{U}$ and $\widetilde{W}$ must be rescaled to reflect the proximity to $(\mu, \phi) = (0, 1)$. Formally I set

$$\widetilde{U}_3 = \widetilde{U}/\epsilon \qquad \widetilde{W}_3 = \epsilon \widetilde{W} \qquad \tilde{z}_3 = \tilde{z},$$

$$\Rightarrow (B/c^2)\mathrm{d}^2\widetilde{U}_3/\mathrm{d}\tilde{z}_3^2 + \mathrm{d}\widetilde{U}_3/\mathrm{d}\tilde{z}_3 + \widetilde{U}_3^2\widetilde{W}_3 - \widetilde{U}_3 = 0,$$

$$(1 + c/\nu) \cdot (1/\epsilon)\mathrm{d}\widetilde{W}_3/\mathrm{d}\tilde{z}_3 + \sigma^2\left(1 - \epsilon\widetilde{U}_3^2\widetilde{W}_3\right) - \widetilde{W}_3 = 0.$$

Therefore provided that $\nu \gg c \gg 1$, the leading order solution is

$$\widetilde{U}_3 = \left[k_3 - h_3\exp(-\tilde{z}_3)\right]^{-1} \qquad \widetilde{W}_3 = k_3,$$

where k_3 and h_3 are constants of integration. Matching between this layer and layer 2 requires that $\lim_{\tilde{z}_2 \to -\zeta}(\widetilde{U}_2, \widetilde{W}_2) = \lim_{\tilde{z}_3 \to +\infty}(\widetilde{U}_3, \widetilde{W}_3)$ for some finite $\zeta > 0$, i.e. $k_3 = \sigma^2(1 - \mathrm{e}^{-\zeta})$.

Layer 4: Layers 1 and 3 are linked by a transition layer in which variation with respect to $\tilde{z}$ is rapid. The appropriate scalings are

$$\widetilde{U}_4 = \epsilon\widetilde{U} \qquad \widetilde{W}_4 = \epsilon\widetilde{W} \qquad \tilde{z}_4 = \tilde{z}/\epsilon^2,$$

$$\Rightarrow \left(B/(c^2\epsilon^2)\right)\mathrm{d}^2\widetilde{U}_4/\mathrm{d}\tilde{z}_4^2 + \mathrm{d}\widetilde{U}_4/\mathrm{d}\tilde{z}_4 + \widetilde{U}_4^2\widetilde{W}_4 - \epsilon^2\widetilde{U}_4 = 0,$$

$$(1 + c/\nu)\mathrm{d}\widetilde{W}_4/\mathrm{d}\tilde{z}_4 + \sigma^2\left(\epsilon^3 - \widetilde{U}_4^2\widetilde{W}_4\right) - \epsilon^2\widetilde{W}_4 = 0.$$

Therefore provided that $\nu \gg c \gg 1/\epsilon$, the leading order solution is

$$\sigma^2\widetilde{U}_4 + \widetilde{W}_4 = k_4 \qquad \frac{k_4^2\tilde{z}_4}{\sigma^2} + h_4 = \frac{k_4}{\sigma^2\widetilde{U}_4} + \log\left(\frac{k_4 - \sigma^2\widetilde{U}_4}{\widetilde{U}_4}\right), \qquad (15)$$

where k_4 and h_4 are constants of integration. I require that the limiting form of this solution as $\tilde{z}_4 \to +\infty$ matches the behaviour of the layer 3 solution as $\tilde{z}_3$ approaches an arbitrary finite matching point, which I take to be zero. This holds provided that $k_4 = k_3 = h_3$. The constant h_4 corresponds to an expected arbitrary $o(1)$ translation in $\tilde{z}$, and can be taken as zero without loss of generality.

Matching layers 1 and 4. There remains one outstanding constant of integration, for which ζ is the most convenient choice. The final step in the construction of the limit cycle solution is to match the behaviour of the layer 4 solution as $\tilde{z}_4 \to -\infty$ with that of the layer 1 solution as $\tilde{z}_1$ approaches an (arbitrary) finite matching value; this matching will determine the relationship between ζ and σ. The implicit solution (15) implies that $(\widetilde{U}_4, \widetilde{W}_4) \to (k_4/\sigma^2, 0)$ as $\tilde{z}_4 \to -\infty$. From the difference in the scalings of $\widetilde{U}$ in layers 1 and 4, it follows that $\widetilde{U}_1$ must contribute a term $(k_4/\sigma^2) \cdot (1/\epsilon)$ at the matching point. Moreover, a detailed investigation (omitted for brevity) shows that the first term in the asymptotic expansion of $\widetilde{W}_4$ with a non-zero limit as $\tilde{z}_4 \to -\infty$ is $O_s(\epsilon^3)$, implying that $\widetilde{W}_1 = o(1)$ is required at the matching point (as $\nu \to \infty$).
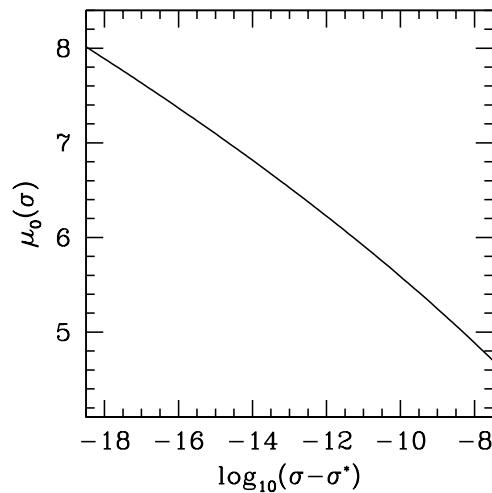
Figure 6. Numerically calculated values of $\mu_0(\sigma)$, the point at which the trajectory $\mathcal{T}$ of equations (9a) and (9b) (defined in section 4) crosses the $\phi = 0$ axis. Note that I plot $\log_{10}(\sigma - \sigma^*)$ on the horizontal axis; μ_0 is defined only for $\sigma \in (\sigma^*, 1)$, and the results in section 4 imply that $\mu_0 \to \infty$ as $\sigma \to \sigma^{*+}$. Computations were done with 16-byte precision. I calculated the value of σ^* to high accuracy as described in the legend to figure 5, and then integrated (9a) and (9b) for a sequence of values of σ just above σ^*. In each case I used initial conditions close to the equilibrium $(0, 1)$ and lying on the centre manifold at that equilibrium. I used the ODE solver DLSODAR (Hindmarsh 1983, Petzold 1983), which is part of the ODEPACK collection, and is freely available at www.netlib.org. This routine incorporates root-finding capability, enabling easy calculation of the point at which ϕ becomes zero.

These conditions on $\widetilde{U}_1$ and $\widetilde{W}_1$ determine the relationship between ζ and σ. In section 4, I defined $\mathcal{T}$ to be the unique trajectory of (9a) and (9b) leaving $(\mu, \phi) = (0, 1)$, and I showed that for $\sigma \in (\sigma^*, 1)$, ϕ decreases along $\mathcal{T}$ until the trajectory crosses the $\phi = 0$ axis. I now define $\mu_0(\sigma)$ to be the value of μ at this crossing; leading order matching of $\widetilde{W}_1$ and $\widetilde{W}_4$ determines this as the matching point. Note that μ_0 is defined only for $\sigma \in (\sigma^*, 1)$, and that $\mu_0 \to \infty$ as $\sigma \to \sigma^{*+}$. Since $\widetilde{U}_1 = \mu/\sigma$, matching between layers 1 and 4 requires that $\mu_0(\sigma)/\sigma = (k_4/\sigma^2) \cdot (1/\epsilon)$. But $k_4 = \sigma^2(1 - e^{-\zeta})$, and thus the required relationship between ζ and σ is

$$\mu_0(\sigma)/\sigma = \left(1 - e^{-\zeta}\right) B\nu/c. \tag{16}$$

Note in particular that since $c \ll \nu$ as $\nu \to \infty$, (16) implies that for any fixed value of ζ, $\mu_0(\sigma) \to \infty$ as $\nu \to \infty$, $\Rightarrow \sigma \to \sigma^{*+}$.

The solution that I have derived is valid in a thin strip in the A–c parameter plane, adjacent to the curve $\sigma = \sigma^*$. The width of this strip depends on the rate at which $\mu_0(\sigma) \to \infty$ as $\sigma \to \sigma^{*+}$, which I have been unable to determine analytically. However, numerical calculations (illustrated in figure 6) suggest that $\mu_0(\sigma) = O_s(\log(\sigma - \sigma^*))$ as $\sigma \to \sigma^{*+}$, which would imply that the approach of σ to σ^* is exponential in ν for any fixed ζ. It would then follow that the width is exponentially small in ν as $\nu \to \infty$, which is consistent with the numerical calculations in section 5.

Away from $\sigma = \sigma^*$, it is the limit cycle solution of (8a) and (8b) that is relevant for patterns. In the thin strip, the period of the limit cycles is dominated by layer 2 ('thick'), and is ζ/ϵ to leading order. Therefore, as expected, the layer structure breaks down as $\zeta \to 0^+$,

corresponding to the solution approaching a limit cycle solution of (8a) and (8b). As $\zeta \to +\infty$, the period of the limit cycle $\to \infty$, so that the limit cycle approaches a homoclinic solution. This corresponds to the homoclinic solution in the full model (2a) and (2b), and it has quite a different form from the homoclinic solution of (8a) and (8b). Since $\tilde{z}_2 = -\zeta$ is the matching point between layers 2 and 3, the solution is homoclinic to the limiting location of this point as $\zeta \to \infty$, i.e.

$$(\widetilde{U}, \widetilde{W}) = \lim_{\zeta \to \infty} \left(\frac{\epsilon}{\sigma^2(1 - e^{-\zeta})}, \frac{\sigma^2(1 - e^{-\zeta})}{\epsilon} \right) = (\epsilon/\sigma^2, \sigma^2/\epsilon) = (B^2/A^2, A^2/B^2).$$

Using (6), this corresponds to $(U, W) = (B/A, A)$, which is exactly the leading order form (as $\nu \to \infty$) of (u_u, w_u), defined in (3). This shows that on the part of the boundary of the pattern region that I am considering, the pattern solution is homoclinic to (u_u, w_u).

In equations (8a) and (8b), the limit cycles become unbounded in both $\widetilde{U}$ and $\widetilde{W}$ as they approach the homoclinic limit. However, this is not the case for the limit cycle solutions I have constructed in this section. On these limit cycles, the maximum/minimum of $\widetilde{U}$ is given by the matching points between layers 1 and 4/layers 2 and 3, respectively. Therefore, on these limit cycles

$$\epsilon / \left[\sigma^2 (1 - e^{-\zeta}) \right] \leqslant \widetilde{U} \leqslant (1 - e^{-\zeta})/\epsilon,$$

$$\text{i.e.} \quad \frac{B^2/A^2}{1 - e^{-\zeta}} \leqslant \widetilde{U} \leqslant (1 - e^{-\zeta}) B\nu/c,$$

$$\text{i.e.} \quad \frac{B/A}{1 - e^{-\zeta}} \leqslant U \leqslant (1 - e^{-\zeta}) A\nu/c$$

to leading order as $\nu \to \infty$, using (6). Hence the limit cycle solution for U remains both bounded and bounded away from zero, up to and including the homoclinic limit ($\zeta = \infty$), for which U varies between B/A and $A\nu/c$. The variation in W can be found in the same way, although the minimum value depends on calculation of the $O_s(\epsilon^3)$ term in the asymptotic expansion of $\widetilde{W}_4$, which I have omitted for brevity. This shows that $c^2/(A\nu^2) \leqslant W \leqslant A$.

To test these results numerically, I tracked the branch of limit cycle solutions of (2a) and (2b) by numerical continuation. I used the software package AUTO (Doedel 1981, Doedel *et al* 1991, 2006), with A as the principle continuation parameter. I decreased A from a large initial value, first detecting the Hopf bifurcation point of (u_s, w_s) and then following the limit cycle branch until numerical continuation fails, close to the homoclinic limit. Figure 7 shows that $\max_z\{U(z)\}c/(A\nu)$ increases with the period up to a maximum value, and then remains approximately constant as the period increases further. (Note that the period increases with A). In terms of my notation above, the transition to a constant value of $\max_z\{U(z)\}c/(A\nu)$ would be expected to occur at a value of ζ that is at least approximately independent of ν. Now the period of the limit cycle is $\zeta/\epsilon = \zeta B\nu/c$, to leading order as $\nu \to \infty$. Moreover $\nu \gg c$, and thus the value of the period at which the transition occurs should increase with ν, as seen in figure 7. Moreover, as the parameter ν is increased, figure 7 shows that the limiting maximum value of $\max_z\{U(z)\}c/(A\nu)$ approaches 1, exactly as predicted by my calculations.

I conclude this section with a brief comment on its validity. The solution structure that I have derived in this section is valid only if c satisfies $\nu \gg c \gg 1/\epsilon$. Since $1/\epsilon = B\nu/c$, the condition for validity is $\nu \gg c \gg \nu^{1/2}$. Qualitatively different behaviours occur when $c = O_s(\nu)$ and $c = O_s(\nu^{1/2})$, and as mentioned previously, I will publish details of these cases separately.
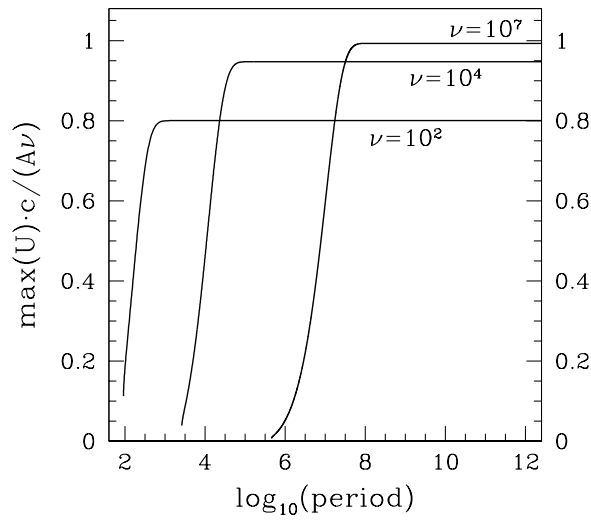
Figure 7. Numerical verification of the predictions of the matched asymptotic expansion constructed in section 6 for the limit cycle solutions of (2a) and (2b) when σ is close to σ^*. Using the software package AUTO (Doedel 1981, Doedel *et al* 1991, 2006), I performed a numerical continuation of limit cycle solutions of (2a) and (2b). I used A as the principle continuation parameter, and decreased A from a large initial value, determining the Hopf bifurcation point of (u_s, w_s), and then following the limit cycle branch emanating from this. In the figure, I plot $\max_z\{U(z)\}c/(A\nu)$ against the period along this branch, for three different values of ν. The key features of these plots are that $\max_z\{U(z)\}c/(A\nu)$ approaches a limiting value as the period $\to \infty$, that the period at which this limiting value is attained increases with ν, and that this limiting value approaches 1 as $\nu \to \infty$, as predicted by the analysis in section 6. The other parameter values are $B = 0.45$ and $c = 2B^3\nu^{3/4}$, with ν as indicated on the curves.

7. Discussion

The Klausmeier model (1a) and (1b) is one of the most generic mathematical models for vegetation patterning in semi-arid environments. Nevertheless, there has been very little previous analysis of the model equations. My objective in this paper has been a detailed investigation of pattern solutions, in one particular region of parameter space. My specific results, which are valid to leading order for large values of the slope parameter ν, are as follows.

(1) For any migration speed c in the range $\nu^{1/2} \ll c \ll \nu$, there are patterns for values of the rainfall parameter A satisfying

$$A \cdot \left(\frac{c}{B^3\nu}\right)^{1/2} \in \left(\sigma^*, 1\right),$$

where $\sigma^* \approx 0.9003$ is defined in section 3.

(2) The upper limit on A in #1 corresponds to a Hopf bifurcation in equations (2a) and (2b), so that the pattern amplitude approaches zero at this boundary, while the limiting spatial period is $2\pi c/B$.

(3) The lower limit on A in #1 corresponds to a homoclinic solution in equations (2a) and (2b), that is homoclinic to the steady state (u_u, w_u), defined in (3). In this solution, the

plant density u varies between B/A and $A\nu/c$, while the water density w varies between $c^2/(A\nu^2)$ and A.

(4) Denote by L the spatial period of a pattern solution. Then patterns for which L increases with ν such that either $L = O_s(\nu/c)$ or $L \gg \nu/c$, with $\nu^{1/2} \ll c \ll \nu$, occur at a value of the rainfall parameter A satisfying

$$\mu_0(Ac^{1/2}B^{-3/2}\nu^{-1/2}) = (A\nu^{1/2}B^{-3/2}c^{-1/2}) \cdot (1 - e^{-cL/(B\nu)}),$$

where $\mu_0(.)$ is defined in section 6. This follows from equation (16), and the condition $L = O_s(\nu/c)$ is equivalent to $\zeta = O_s(1)$. If $L \ll \nu/c$, then for sufficiently large ν the solution lies outside the thin strip in which the calculations in section 6 apply, and the leading order relationship between A and L is simply that between the period of the limit cycle solution of (9a) and (9b) and the parameter $\sigma = Ac^{1/2}B^{-3/2}\nu^{-1/2}$.

(5) The equations governing the leading order form of the patterns as $\nu \to \infty$ have one form (see sections 3 and 4) for values of A away from the lower limit in #1, and another form (see section 6) close to this lower limit. Numerical calculations suggest that the latter case applies only when A is exponentially close (in ν) to the lower limit.

The ecological literature contains a large amount of data on banded vegetation patterns. For example, the comprehensive review of Valentin *et al* (1999) lists 41 'main references' to field studies (see their table 1 and see section 1 of Borgogno *et al* (2009) for more recent references). However, few if any of these empirical data sets are complete, and an important role for theoretical modelling in this area is to provide estimates for missing measurements. In particular, there are only a handful of studies containing estimates of uphill migration speed (see table 5 of Valentin *et al* 1999). My results show clearly that migration speed can be estimated using the model (1a) and (1b), on the basis of data on mean annual rainfall, the slope of the study site, and the level of herbivory, together with one observational statistic on pattern type, such as wavelength. In older studies, wavelength is usually available from aerial photographs, and in more recent cases its estimation is straightforward, via online repositories of satellite images.

Another ecologically important quantity is the net biomass productivity, which corresponds to the mean value of the plant density $u(x, t)$. My analysis shows that in the parameter region I am considering, productivity varies between u_s (defined in (4)) at the upper limit on A for patterns, and u_u (defined in (3)) at the lower limit. The latter follows from the exponential approaches in $\tilde{z}$ of the homoclinic solution constructed in section 6 to the equilibrium (u_u, w_u),[1] which imply that the 'pulse' makes no contribution to the mean. For a given value of the migration speed c, the transition between these two limiting values can be calculated numerically via continuation of the limit cycle solutions of (2a) and (2b), and a typical result is illustrated in figure 8. As one would expect intuitively, the productivity decreases with the rainfall parameter A. Through most of the pattern region, the productivity remains relatively close to the plant density at the coexistence equilibrium, u_s, dropping sharply to u_u near the homoclinic limit. This reflects the rapid change in pattern form close to $\sigma = \sigma^*$. An important implication of my analysis is that away from this boundary in parameter space, and to leading order for large ν, the ratio of productivity in banded vegetation to the corresponding uniform vegetation density (u_s) is a function of the single parameter grouping $A^2c/(B^3\nu)$.

---

[1] The homoclinic limit for the solution in section 6 is $\zeta \to \infty$. The 'pulse' of the solution has layer 2 at one side and layer 3 at the other. In both of these layers, in the limit of $\zeta \to \infty$, the leading order solution decays exponentially to its asymptotic steady state.
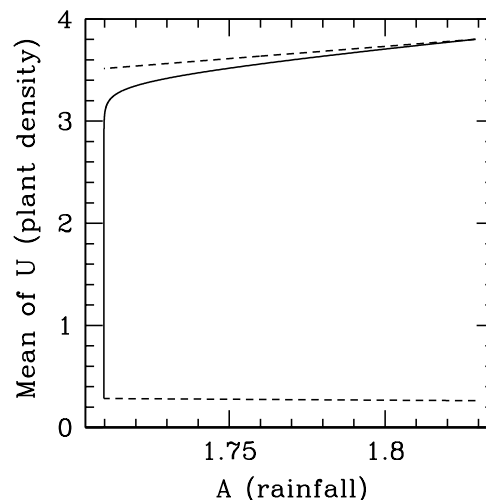
Figure 8. An illustration of the mean value of plant density U (the biomass productivity) as a function of mean annual rainfall A (solid curve), for a fixed value of migration speed c. This was calculated by numerical continuation of limit cycle solutions of (2a) and (2b), as described in the legend to figure 7. The upper and lower dashed lines are u_s and u_u, respectively. The parameter values are $B = 0.45$, $\nu = 10^6$, $c = 3 \times 10^4$. The horizontal axis shows the full range of A for which patterns occur, for these values of the other parameters.

Acknowledgments

References

Barbier N, Couteron P, Lefever R, Deblauwe V and Lejeune O 2008 Spatial decoupling of facilitation and competition at the origin of gapped vegetation patterns *Ecology* **89** 1521–31
Borgogno F, D'Odorico P, Laio F and Ridolfi L 2009 Mathematical models of vegetation pattern formation in ecohydrology *Rev. Geophys.* **47** RG1005
Couteron P and Lejeune O 2001 Periodic spotted patterns in semi-arid vegetation explained by a propagation-inhibition model *J. Ecol.* **89** 616–28
Doedel E J, Govaerts W, Kuznetsov Y A and Dhooge A 2006 Numerical continuation of branch points of equilibria and periodic orbits *Modelling and Computations in Dynamical Systems* ed E J Doedel *et al* (Singapore: World Scientific) pp 145–64
Doedel E J, Keller H B and Kernévez J P 1991 Numerical analysis and control of bifurcation problems: I. Bifurcation in finite dimensions *Int. J. Bifurcation Chaos* **1** 493–520
Doedel E J 1981 AUTO, a program for the automatic bifurcation analysis of autonomous systems *Congr. Numer.* **30** 265–384
Dunkerley D L 1997 Banded vegetation: development under uniform rainfall from a simple cellular automaton model *Plant Ecol.* **129** 103–11
Dunkerley D L and Brown K J 2002 Oblique vegetation banding in the Australian arid zone: implications for theories of pattern evolution and maintenance *J. Arid Environ.* **52** 163–81
Freudenberger D O and Hiernaux P 2001 Productivity of patterned vegetation *Banded Vegetation Patterning in Arid and Semi-Arid Environments* ed D J Tongway *et al* (New York: Springer) pp 198–209
Gilad E, von Hardenberg J, Provenzale A, Shachak M and Meron E 2007 A mathematical model of plants as ecosystem engineers *J. Theor. Biol.* **244** 680–91
Guckenheimer J and Holmes P 1983 *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields* (Berlin: Springer-Verlag)

Guttal V and Jayaprakash C 2007 Self-organisation and productivity in semi-arid ecosystems: implications of seasonality in rainfall *J. Theor. Biol.* **248** 490–500

Hemming C F 1965 Vegetation arcs in Somaliland *J. Ecol.* **53** 57–67

HilleRisLambers R, Rietkerk M, van de Bosch F, Prins H H T and de Kroon H 2001 Vegetation pattern formation in semi-arid grazing systems *Ecology* **82** 50–61

Hindmarsh A C 1983 ODEPACK a systematized collection of ODE solvers *Scientific Computing* ed R S Stepleman *et al* (North-Holland: Amsterdam) pp 55–64

Kefi S, Rietkerk M and Katul G G 2008 Vegetation pattern shift as a result of rising atmospheric CO_2 in arid ecosystems *Theor. Pop. Biol.* **74** 332–44

Klausmeier C A 1999 Regular and irregular patterns in semiarid vegetation *Science* **284** 1826–8

Kletter A Y, von Hardenberg J, Meron E and Provenzale A 2009 Patterned vegetation and rainfall intermittency *J. Theor. Biol.* **256** 574–83

Lefever R, Barbier N, Couteron P and Lejeune O 2009 Deeply gapped vegetation patterns: on crown/root allometry, criticality and desertification *J. Theor. Biol.* **261** 194–209

Lefever R and Lejeune O 1997 On the origin of tiger bush *Bull. Math. Biol.* **59** 263–94

Mabbutt J A and Fanning P C 1987 Vegetation banding in arid Western Australia *J. Arid Environ.* **12** 41–59

MacFadyen W 1950 Vegetation patterns in the semi-desert plains of British Somaliland *Geograph. J.* **115** 199–211

Mauchamp A, Rambal S and Lepart J 1994 Simulating the dynamics of a vegetation mosaic: a spatialized functional model *Ecolo. Modelling* **71** 107–30

Montaña C, Seghieri J and Cornet A 2001 Vegetation dynamics: recruitment and regeneration in two-phase mosaics *Banded Vegetation Patterning in Arid and Semi-Arid Environments* ed D J Tongway *et al* (New York: Springer) pp 132–45

Montaña C, Lopez-Portillo J and Mauchamp A 1990 The response of two woody species to the conditions created by a shifting ecotone in an arid ecosystem *J. Ecol.* **78** 789–98

Montaña C 1992 The colonization of bare areas in two-phase mosaics of an arid ecosystem *J. Ecol.* **80** 315–27

Noble J C, Peltier R, Montagne P and Mahamane El E H 2001 Toward improved management of arid and semiarid banded landscapes *Banded Vegetation Patterning in Arid and Semi-Arid Environments* ed D J Tongway *et al* (New York: Springer) pp 210–27

Petzold L R 1983 Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations *SIAM J. Sci. Stat. Comput.* **4** 136–48

Pueyo Y, Kefi S, Alados C L and Rietkerk M 2008 Dispersal strategies and spatial organization of vegetation in arid ecosystems *Oikos* **117** 1522–32

Rietkerk M, Dekker S C, de Ruiter P C and van de Koppel J 2004 Self-organized patchiness and catastrophic shifts in ecosystems *Science* **305** 1926–9

Rietkerk M, Boerlijst M C, van Langevelde F, HilleRisLambers R, van de Koppel J, Prins H H T and de Roos A 2002 Self-organisation of vegetation in arid ecosystems *Am. Nat.* **160** 524–30

Sherratt J A 2005 An analysis of vegetation stripe formation in semi-arid landscapes *J. Math. Biol.* **51** 183–97

Sherratt J A and Lord G J 2007 Nonlinear dynamics and pattern bifurcations in a model for vegetation stripes in semi-arid environments *Theor. Pop. Biol.* **71** 1–11

Thiéry J M, D'Herbès J-M and Valentin C 1995 A model simulating the genesis of banded vegetation patterns in Niger *J. Ecol.* **83** 497–507

Tongway D J and Ludwig J A 2001 Theories on the origins, maintenance, dynamics, and functioning of banded landscapes *Banded Vegetation Patterning in Arid and Semi-Arid Environments* ed D J Tongway *et al* (New York: Springer) pp 20–31

Ursino N 2007 Modeling banded vegetation patterns in semiarid regions: inter-dependence between biomass growth rate and relevant hydrological processes *Water Resources Res.* **43** W04412

Ursino N 2009 Above and below ground biomass patterns in arid lands *Ecol. Modelling* **220** 1411–18

Ursino N and Contarini S 2006 Stability of banded vegetation patterns under seasonal rainfall and limited soil moisture storage capacity *Adv. Water Resources* **29** 1556–64

van de Koppel J *et al* 2002 Spatial heterogeneity and irreversible vegetation change in semiarid grazing systems *Am. Nat.* **159** 209–18

von Hardenberg J, Meron E, Shachak M and Zarmi Y 2001 Diversity of vegetation patterns and desertification *Phys. Rev. Lett.* **87** 198101

Valentin C, d'Herbès J M and Poesen J 1999 Soil and water components of banded vegetation patterns *Catena* **37** 1–24

Wang R H, Liu Q X, Sun G Q, Jin Z and van de Koppel J 2009 Nonlinear dynamic and pattern bifurcations in a model for spatial patterns in young mussel beds *J. R. Soc. Interface* **6** 705–18

Wickens G E and Collier F W 1971 Some vegetation patterns in the Republic of Sudan *Geoderma* **6** 43–59