Nonmonotone invariant manifolds in the Nagylaki-Crow model☆

Belgin Seymenoglu[a,*], Stephen Baigent[a]

[a] *Department of Mathematics, University College London, Gower Street, London WC1E 6BT*

Abstract

We use a change of dynamical variables to prove, subject to certain conditions on the parameters, that a nonmonotone invariant manifold exists and is the graph of a convex function for the Nagylaki-Crow fertility-mortality model from population genetics. Our results are obtained without the common assumption that fertilities or death rates are additive, and are not restricted to the case that the model is competitive in the new coordinates. We also provide numerical examples demonstrating that the manifold need not be the graph of a convex function, smooth, unique or globally attracting, and that the model exhibits a sequence of nonmonotone manifolds similar to those studied by Hirsch for competitive Kolmogorov systems [1].

Keywords: Invariant manifolds, Population Genetics, Nagylaki-Crow model
2010 MSC: 34C45, 37N25, 92D10

1. Introduction

In this paper we study invariant manifolds for the Nagylaki-Crow model, a continuous-time model from population genetics, sometimes referred to as the fertility-mortality model [2], because these are the two opposing forces at play for the model. By fertility, we mean the average number of offspring produced per unit time by parents with specified genotypes. Meanwhile, mortality refers to the probability of death per unit time for a given (parental) genotype.

One of the earliest attempts at considering different fertilities for mating pairs was made by Penrose in [3]. He showed that his basic discrete-time model with additive fertilities gave essentially the same results as the usual discrete selection model. For the next few decades, most investigations into differential fertility were only made for the discrete model [4].

Then in 1961 Rucknagel and Neel produced experimental evidence of fertility differences among mating pairs for the locus corresponding to sickle cell anaemia [5], a single-locus genetic disorder affecting humans. This revived interest in differential fertility models.

Over a decade later, Nagylaki and Crow provided a derivation for a continuous-time model [6, 7], now known as the Nagylaki-Crow model. However, they restricted their attention to the case of additive fertilities when analysing the model. Another special case of this model is analysed in [8, 9] with symmetric fertilities and no deaths. Hadeler and Glas showed that all orbits for the Nagylaki-Crow model with no deaths converge to some fixed point [10]. They also proposed a change of variables, which was later used in [11]

☆Supported by the EPSRC and the Department of Mathematics, UCL.
*Corresponding author
Email addresses: belgin.seymenoglu.10@ucl.ac.uk (Belgin Seymenoglu), steve.baigent@ucl.ac.uk (Stephen Baigent)

to demonstrate that the model can have periodic orbits. In the nineties, Szucs proved the existence of an invariant manifold connecting the two heterozygotic fixation states in a two allele fertility-selection model where the fertilities were additive. He also showed that with the assumption of additivity of mortalities [12] this invariant manifold coincided with the Hardy-Weinberg manifold where the genotypic frequencies are the product of allele frequencies. Szucs and Akin also showed in [13] that additivity of positive fertilities and mortalities the Hardy-Weinberg manifold is invariant, and with additional conditions on the relative sizes of fertilities and differences in mortalities implied convergence onto the Hardy-Weinberg manifold.

Our aim here is to extend this work by showing that the Nagylaki-Crow model possesses at least one nonmonotone invariant manifold *without assuming additivity of fertilities or mortalities*, which will make our result more widely applicable than Akin and Szucs. (By nonmonotone we essentially mean that that the manifold is the graph of a decreasing function; see, for example, the three solid curves in Figure 1)

Section 2 introduces the $n-$allele model, while Section 3 discusses the two-allele case of the model and shows how to rewrite the Nagylaki-Crow model as a competitive system using a change of coordinates, although later we will drop one of the two inequalities that render the model competitive, so that we obtain results for not-necessarily competitive models. It turns out that the model always has fixed points on two corners of the triangular phase plane (axial fixed points) corresponding to heterozygotic fixation, and our numerical evidence suggests the existence of at least one nonmonotone invariant manifold Σ connecting the two fixed points in the phase plot. We analyse both axial fixed points in Section 4, and investigate their relationship with the condition for the system being competitive in the new coordinates. Finally, for Section 5, we prove that a nonmonotone invariant manifold Σ does indeed exist for a certain case of the Nagylaki-Crow model, and that it is the graph of a convex function. When the fertilities and mortalities are additive, Σ coincides with the Hardy-Weinberg manifold.

2. The model

A derivation of the panmictic Nagylaki-Crow model for diploid populations can be found in [6], where the authors consider a single locus with n alleles $A_1, \dots A_n$, and the dynamical variables of their system are the frequencies P_{ij} for the ordered genotype A_iA_j.

The Nagylaki-Crow model also features the fertilities $a_{ik,lj}$ that are defined as the product of the average number of matings of an arbitrary individual per unit time and the average number of progeny per $A_iA_k \times A_lA_j$ union. With this definition in mind, it is reasonable to assume

$$a_{ik,lj} \geq 0 \qquad \forall t \quad \forall i, j, k, l,$$

since a mating pair cannot produce a negative number of offspring. Moreover, it will be assumed that they are also time-independent for all i, j, k, l.

In addition, the model contains the mortalities d_{ij}, the probability of death per unit time for genotype A_iA_j. These are also taken to be non-negative and constant for all time.

The governing equations for the genotype frequencies P_{ij}

$$\dot{P_{ij}} = \left(\sum_{kl} a_{ik,lj} P_{ik} P_{lj} - d_{ij} P_{ij}\right) - P_{ij} \sum_{uv} \left(\sum_{kl} a_{uk,lv} P_{uk} P_{lv} - d_{uv} P_{uv}\right), \tag{2.1}$$

form a system of n^2 nonlinear first order ordinary differential equations (see [6] or [7]). Note that if $P_{ij}(0) \geq 0$ then $P_{ij}(t) \geq 0$ for all $t \geq 0$. Moreover $\sum_{i,j=1}^{n} P_{ij}(t) = 1$ for all $t \geq 0$. The marginal $\sum_{j=1}^{n} P_{ij}(t) = \sum_{j=1}^{n} P_{ji}(t) = p_i(t)$ is the frequency of allele A_i at time $t \geq 0$.

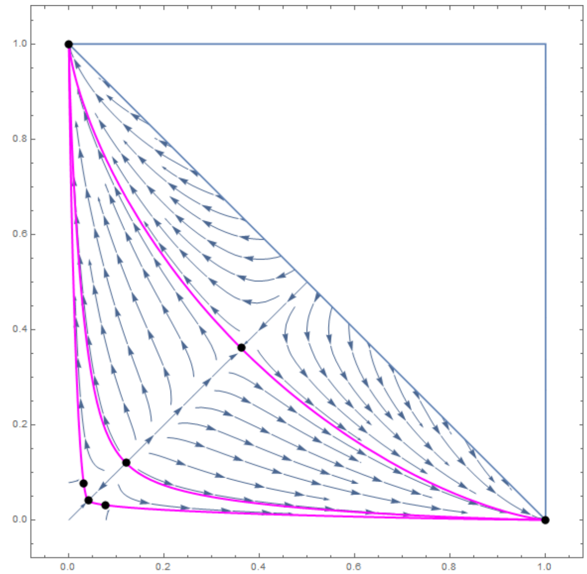


Figure 1: A case of the Nagylaki-Crow model with five interior fixed points, three of which occur on the $y = x$ with $y = 0.0402131, 0.119578$ and 0.361042 respectively. There are three different nonmonotone invariant manifolds, one of which passes through three interior fixed points.
Here, the fertilities are

$$\begin{aligned} F_{11} &= 6, & F_{33} &= 6, \\ F_{12} &= \frac{1}{2}, & F_{23} &= \frac{1}{2}, \\ F_{22} &= \frac{1}{14}, & F_{13} &= 1, \end{aligned}$$

while the mortalities are

$$D_1 = 2, \quad D_2 = 1, \quad D_3 = 2,$$

making this system competitive in (w, t) coordinates.

Figure 1 shows an example of the phase portrait for the Nagylaki-Crow model. In this example, there are five interior fixed points, which is the maximum number that the two-allele model can possess [2]. In the figure there are three different nonmonotone invariant manifolds, one of which passes through three interior fixed points.

Even for this simple model it is not possible to obtain self-contained evolution equations for the allele frequencies, which often are the variables of most interest to the geneticist. However, the presence of an attracting manifold means that differential equations can be obtained for the allele frequencies when restricted to that manifold. If an initial point is attracted to the manifold rapidly then after a short transient the equations for the allele frequencies on the manifold will be a good approximation of the true allele frequencies. The Hardy-Weinberg manifold is obtained by solving $P_{11} = x^2$, so that it is the graph of the strictly convex function $\varphi_{HW} : [0, 1] \to [0, 1]$ defined by

$$\varphi_{HW}(x) = 1 + x - 2\sqrt{x}.$$

For comparison, the Hardy-Weinberg manifold is also shown in Figure 2.

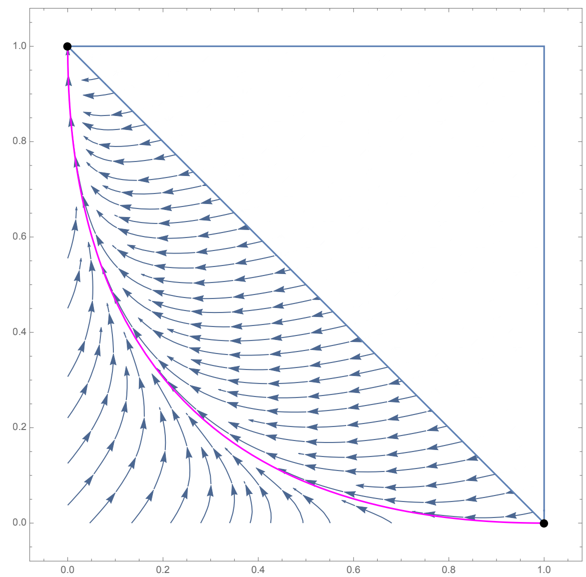


Figure 2: An example where the fertilities are additive, leading to the Hardy-Weinberg manifold φ_{HW}. Here, the fertilities are

$$F_{11} = 0.6, \qquad F_{33} = 2,$$
$$F_{12} = 0.9, \qquad F_{23} = 1.6,$$
$$F_{22} = 1.2, \qquad F_{13} = 1.3,$$

while the mortalities are

$$D_1 = 0.3, \quad D_2 = 0.25, \quad D_3 = 0.2,$$

making this system competitive in (w, t) coordinates.

The special case where the fertilities and mortalities are additive means that $a_{ik,lj} = \alpha_{ik} + \beta_{lj}$ and $d_{ij} = \mu_i + \kappa_j$, where $\alpha_{ik}, \beta_{lj}, \mu_i, \kappa_j \geq 0$. In this special case all trajectories converge to this nonmonotone invariant manifold connecting the axial fixed points [13]. Our results show that under mild conditions, when the condition of additivity of fertilities and mortalities is relaxed, there is at least one such nonmonotone manifold, and we give conditions that ensure that any such nonmonotone manifold is the graph of a convex function. A detailed analysis of the uniqueness of this manifold will be discussed elsewhere.

3. Rewriting the Nagylaki-Crow model

Even for $n = 2$ the Nagylaki-Crow is not a straightforward model to analyse and to the best of the authors' knowledge there is currently no understanding of this model for three or more alleles [2].

From now on, we will ignore the order of the genotypes, i.e. treat the genotypes A_iA_j and A_jA_i as identical.

It is assumed in [7], as is done here, that the fertilities $a_{ik,lj}$ possess the symmetries

$$a_{ij,kl} = a_{kl,ij} = a_{ji,kl}. \tag{3.1}$$

Thus the Nagylaki-Crow model has only six independent fertility parameters. Using the notation of [11] these are relabelled as follows:

$$\begin{aligned} F_{11} &= a_{11,11}, \quad & F_{33} &= a_{22,22}, \\ F_{12} &= a_{11,12}, \quad & F_{23} &= a_{12,22}, \\ F_{22} &= a_{12,12}, \quad & F_{13} &= a_{22,11}. \end{aligned} \tag{3.2}$$

Meanwhile, as $d_{12} = d_{21}$, there are only three independent mortality parameters. They are rewritten as

$$D_1 = d_{11}, \quad D_2 = d_{12}, \quad D_3 = d_{22}.$$

Moreover, since the genotypes $A_iA_j = A_jA_i$, one has

$$P_{ij} = P_{ji}.$$

With this symmetry law in mind, let

$$P_{11} = x, \quad P_{12} = P_{21} = z/2, \quad P_{22} = y. \tag{3.3}$$

Then according to [11], the Nagylaki-Crow model is reduced to

$$\dot{x} = F_{11}x^2 + F_{12}xz + \frac{1}{4}F_{22}z^2 - D_1x - x\varphi, \tag{3.4}$$

$$\dot{z} = F_{12}xz + 2F_{13}xy + F_{23}yz + \frac{1}{2}F_{22}z^2 - D_2z - z\varphi, \tag{3.5}$$

$$\dot{y} = F_{33}y^2 + F_{23}yz + \frac{1}{4}F_{22}z^2 - D_3y - y\varphi, \tag{3.6}$$

with mean fitness

$$\begin{aligned} \varphi &= \sum_{uv}\left(\sum_{kl} a_{uk,lv}P_{uk}P_{lv} - d_{uv}P_{uv}\right) \\ &= F_{11}x^2 + 2F_{12}xz + F_{22}z^2 + 2F_{23}yz + 2F_{13}xy + F_{33}y^2 \\ &\quad - D_1x - D_2z - D_3y. \end{aligned}$$

However, one also has the following condition:

$$\sum_{ij} P_{ij} = 1 \qquad \text{or} \qquad x + y + z = 1, \qquad x, y, z \geq 0 \tag{3.7}$$

which indicates that the two-allele Nagylaki-Crow model has in fact only two degrees of freedom and is fully described by just two ordinary differential equations on the phase space given by the triangle $T = \{(x, y) \in \mathbb{R}^2_+ : x + y \leq 1\}$. These ordinary differential equations can be written in the form

$$\dot{x} = f(x, y), \qquad \dot{y} = g(x, y), \tag{3.8}$$

by substituting $z = 1-x-y$ into the x and y equations. The full equations for x, y can be found in AppendixA.

As an alternative, we use the following coordinate change introduced in [10, 11]

$$w = \frac{2x}{z}, \qquad t = \frac{2y}{z}, \qquad z = 0 \quad \text{iff} \quad w, t = \infty, \tag{3.9}$$

which reduces the system to

$$\dot{w} = p(w, t), \qquad \dot{t} = q(w, t),$$

where

$$\begin{aligned} p(w, t) &= (F_{11} - F_{12} + D_2 - D_1)w^2 + (2(F_{12} + D_2 - D_1) - F_{22})w \\ &\quad + F_{22} - wt(F_{23} - D_2 + D_1 + F_{13}w), \end{aligned} \tag{3.10}$$

$$\begin{aligned} q(w, t) &= (F_{33} - F_{23} + D_2 - D_3)t^2 + (2(F_{23} + D_2 - D_3) - F_{22})t \\ &\quad + F_{22} - wt(F_{12} - D_2 + D_3 + F_{13}t). \end{aligned} \tag{3.11}$$

Here by $\dot{w}, \dot{t}$ we mean differentiation with respect to time; time will be denoted by s in the sequel.

In this set of coordinates, the phase space is the whole (non-compact) first quadrant $w \geq 0, t \geq 0$. For the boundary $w = 0$, one has $\dot{w} = F_{22} \geq 0$, which shows that $w < 0$ can never occur. Likewise, we have $\dot{t} = F_{22} \geq 0$ for the boundary at $t = 0$, therefore it is impossible to acquire $t < 0$. This indicates that the phase space $\mathbb{R}^2_+$ is forward invariant. Moreover, if $F_{22} > 0$, then $w \nrightarrow 0$ and $t \nrightarrow 0$, so the system is permanent (all interior orbits eventually enter and remain in a compact subset of the interior of $\mathbb{R}^2_+$).

The off-diagonal elements of the Jacobian for this system are

$$\begin{aligned} p_t &= (-F_{23} + D_2 - D_1)w - F_{13}w^2 \\ q_w &= (-F_{12} + D_2 - D_3)\, t - F_{13}t^2, \end{aligned}$$

which are both non-positive for all $w, t \geq 0$ if and only if

$$D_2 \leq \min(D_1 + F_{23}, D_3 + F_{12}). \tag{3.12}$$

Inequality (3.12) is a necessary and sufficient condition for equations (3.10) and (3.11) to be competitive on $\mathbb{R}^2_+ = \{(w, t) \in \mathbb{R}^2 : w \geq 0, t \geq 0\}$, i.e. $p_t, q_w \leq 0$ [11, 10]. It is known that an orbit of a planar competitive system is either unbounded or converges to a fixed point in increasing time [2, 10]. For this system, if an orbit Γ in (w, t) is unbounded, then $z \to 0$, i.e. $(x + y) \to 1$. This shows that the corresponding ω-limit set for Γ in (x, y) coordinates is a subset of the bounding line $x + y = 1$.

As shown by Hirsch [1] competitive systems often possess special codimension-1 invariant manifolds Σ that are nonmonotone, so that no two points on Σ may be ordered with respect to the partial order that defines the competitive ordering. For the standard first orthant ordering (for each $i = 1, \ldots, n$, $x \geq y$ if and only if $x_i \geq y_i$; $x > y$ if and only if $x \geq y$ and $x \neq y$, $x \gg y$ if and only if $x_i > y_i$) this means that no points $x, y \in \Sigma$ may satisfy $x < y$. In two dimensions a nonmonotone codimension-1 manifold Σ is the graph of a decreasing function. Hirsch [1] showed that a large class of competitive ordinary differential equations possess a countable sequence of nonmonotone manifolds that divide the phase space into regions, and that these manifolds are alternately repelling and attracting. Although our model is competitive, it does not satisfy the setting of Hirsch's theory since $\partial\mathbb{R}^2_+$ is not invariant. Nevertheless we find numerical evidence of similar sequences of nonmonotone manifolds. For example, Figure 1 shows three nonmonotone manifolds (solid curves) that all connect the axial fixed points. These are the only *nonmonotone* invariant manifolds in the figure. A monotone invariant manifold (where points are ordered) passes through 3 interior fixed points along the line $x = y$. Two of the nonmonotone manifolds pass through a single interior fixed point, and the third passes through 3 interior fixed points. The detailed structure of alternate repelling and attracting nonmonotone manifolds in our model will be studied elsewhere. The focus in this paper is to prove the existence of at least one nonmonotone invariant manifold and to establish convexity conditions, and it turns out that to do this it is much easier to carry out some calculations in the w, t coordinates where phase space is not compact, mapping back results to the system in (x, y) coordinates, and some in (x, y) coordinates where the phase space is compact, but there is no obvious ordering in (x, y) coordinates for which the system is monotone or competitive. Our key observation is that manifolds that are nonmonotone and graphs of convex functions in (w, t) coordinates are also nonmonotone and graphs of convex functions in (x, y) coordinates.

The first step is to prove the following lemma:

Lemma 3.1. *A strictly decreasing convex function in (w, t) coordinates is strictly decreasing and convex in (x, y) coordinates.*

Proof. Let $t = t(w)$ and $y = y(x)$ be the representations of the the two curves.

Observe that it is possible to switch back to x, y coordinates using

$$x = \frac{w}{2 + w + t}, \qquad y = \frac{t}{2 + w + t}. \tag{3.13}$$

Taking derivatives with respect to w results in

$$\begin{aligned} \frac{dx}{dw} &= \frac{1}{2 + w + t} - \frac{w(1 + \frac{dt}{dw})}{(2 + w + t)^2}, \\ \frac{dy}{dw} &= \frac{\frac{dt}{dw}}{2 + w + t} - \frac{t(1 + \frac{dt}{dw})}{(2 + w + t)^2}. \end{aligned}$$

Thus by the Chain Rule,

$$\frac{dy}{dx} = \frac{(2 + w)\frac{dt}{dw} - t}{(2 + t) - w\frac{dt}{dw}}.$$

However, recall that $w, t \geq 0$, therefore

$$\frac{dt}{dw} < 0 \quad \Rightarrow \quad \frac{dy}{dx} < 0. \tag{3.14}$$

Hence any function that is strictly decreasing in the (w, t) plane is also strictly decreasing when mapped into the triangle $T = \{(x, y) \in \mathbb{R}^2_+ : x + y \leq 1\}$.

Furthermore, by the Chain Rule,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{dw}{dx}\frac{d}{dw}\left(\frac{dy}{dx}\right) = \frac{\frac{d}{dw}\left(\frac{dy}{dx}\right)}{\frac{dx}{dw}}.$$

Hence

$$\begin{aligned}\frac{d^2y}{dx^2} &= \frac{d}{dw}\left(\frac{(2+w)\frac{dt}{dw} - t}{(2+t) - w\frac{dt}{dw}}\right)\Bigg/\left(\frac{(2+t) - w\frac{dt}{dw}}{(2+w+t)^2}\right) \\ &= \left(\frac{2(2+t+w)}{((2+t) - w\frac{dt}{dw})^2}\,\frac{d^2t}{dw^2}\right)\Bigg/\left(\frac{(2+t) - w\frac{dt}{dw}}{(2+w+t)^2}\right),\end{aligned}$$

and so

$$\frac{d^2y}{dx^2} = \frac{2(2+w+t)^3}{(2+t-w\frac{dt}{dw})^3}\,\frac{d^2t}{dw^2}. \tag{3.15}$$

Recall that the phase space is the first quadrant, so $w, t \geq 0$. Also, for a convex function in (w, t) phase plane, the second derivative of t with respect to w is non-negative. Moreover, it was assumed the function $t(w)$ is strictly decreasing and hence we deduce from Equation (3.15)

$$\frac{dt}{dw} < 0, \quad \frac{d^2t}{dw^2} \geq 0 \quad \Rightarrow \quad \frac{d^2y}{dx^2} \geq 0, \tag{3.16}$$

Thus the curve is also convex in the (x, y) plane. □

Regarding interior fixed points, there is a result which is seemingly trivial, but is worth stating.

Lemma 3.2. *An interior fixed point in the (w, t) phase plane is also an interior fixed point in (x, y) coordinates, and vice versa.*

Proof. Starting with

$$\begin{pmatrix} w \\ t \end{pmatrix} = \begin{pmatrix} \frac{2x}{1-x-y} \\ \frac{2y}{1-x-y} \end{pmatrix},$$

differentiation with respect to time yields the following:

$$\begin{pmatrix} \dot{w} \\ \dot{t} \end{pmatrix} = 2(1-x-y)^{-2}\begin{pmatrix} 1-y & x \\ y & 1-x \end{pmatrix}\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}.$$

But the matrix

$$2(1-x-y)^{-2}\begin{pmatrix} 1-y & x \\ y & 1-x \end{pmatrix}$$

has determinant $4/(1-x-y)^3$, which is non zero (as well as finite), since $x + y < 1$ at an interior fixed point. Therefore this is a nonsingular matrix, and so its homogeneous matrix equation has only the trivial solution, which in turn implies that

$$\begin{pmatrix} \dot{w} \\ \dot{t} \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = 0.$$

□

4. Steady states and stability

The two steady states $(1,0)$ and $(0,1)$, represent the cases where all members of the population are homozygotes A_1A_1 and A_2A_2 respectively. These two points are always steady states for the model regardless of the values of the parameters, and their local invariant manifolds are investigated via spectral analysis of the Jacobian

$$\mathbf{J} = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}$$

for both points.

One finds that the Jacobian corresponding to $(0,1)$ is

$$\mathbf{J^{(0,1)}} = \begin{pmatrix} D_3 - D_1 - F_{33} & 0 \\ D_1 - D_2 - 2F_{13} + F_{23} & D_3 - D_2 + F_{23} - F_{33} \end{pmatrix},$$

with eigenvalues $\lambda_1^{(0,1)} = D_3 - D_1 - F_{33}$, $\lambda_2^{(0,1)} = D_3 - D_2 + F_{23} - F_{33}$, and respective eigenvectors

$$\mathbf{v_1^{(0,1)}} = \left(-\frac{D_1 - D_2 + F_{23}}{D_1 - D_2 + F_{23} - 2F_{13}}, 1\right)^T, \quad \mathbf{v_2^{(0,1)}} = (0,1)^T, \tag{4.1}$$

which indicates that the tangent space for one of the local invariant manifolds at $(0,1)$ is always vertical. Note that for general F_{ij} and D_k the signs of the eigenvalues remain unspecified, hence it is unclear whether each local manifold is stable, unstable or centre.

There is another invariant subspace at $(0,1)$, however the positioning of the corresponding local manifold is not immediately obvious. On closer inspection, it turns out that the local manifold will lie locally in the triangular region T if and only if the gradient of its tangent space, represented by, $\mathbf{v_1^{(0,1)}}$, is no greater than (-1), i.e.

$$\begin{aligned} -\frac{D_1 - D_2 + F_{23} - 2F_{13}}{D_1 - D_2 + F_{23}} &\leq -1 && \Leftrightarrow \\ \frac{D_1 - D_2 + F_{23} - 2F_{13}}{D_1 - D_2 + F_{23}} &\geq 1 && \Leftrightarrow \\ 1 - \frac{2F_{13}}{D_1 - D_2 + F_{23}} &\geq 1 && \Leftrightarrow \\ \frac{F_{13}}{D_1 - D_2 + F_{23}} &\leq 0, \end{aligned}$$

which occurs when $F_{23} \leq (D_2 - D_1)$. Notice that this condition is satisfied if and only if $\lambda_2^{(0,1)} \leq \lambda_1^{(0,1)}$.

Meanwhile, the Jacobian at $(1,0)$ is

$$\mathbf{J^{(1,0)}} = \begin{pmatrix} D_1 - D_2 + F_{12} - F_{11} & D_3 - D_2 + F_{12} - 2F_{13} \\ 0 & D_1 - D_3 - F_{11} \end{pmatrix},$$

which has eigenvalues

$$\lambda_1^{(1,0)} = D_1 - D_3 - F_{11}, \quad \lambda_2^{(1,0)} = D_1 - D_2 + F_{12} - F_{11},$$

with corresponding eigenvectors

$$\mathbf{v_1^{(1,0)}} = \left(-\frac{D_3 - D_2 + F_{12} - 2F_{13}}{D_3 - D_2 + F_{12}}, 1\right)^T, \quad \mathbf{v_2^{(1,0)}} = (1,0)^T.$$

Again, the two eigenvalues can be generally either positive, negative or zero, hence the respective tangent spaces corresponding with the local manifolds could be stable, unstable or centre manifolds.

Hence the tangent space for one local manifold at the point is guaranteed to be horizontal at $(1, 0)$. A necessary condition for the respective tangent space of the other local manifold being inside the triangle T is that the gradient of $\mathbf{v_1^{(1,0)}}$ should be bounded by the values (-1) and 0, or equivalently,

$$\frac{F_{13}}{D_3 - D_2 + F_{12}} \leq 0.$$

This is satisfied if $F_{12} \leq (D_2 - D_3)$, which is equivalent to $\lambda_2^{(1,0)} \leq \lambda_1^{(1,0)}$.

Note that since all fertilities and death rates are taken to be real numbers, the triangular Jacobian for both steady states must always have real eigenvalues, thus these equilibria cannot have spirals or centres in their vicinity.

The system is competitive in (w, t) coordinates if and only if (3.12) holds, which is equivalent to the following inequalities combined

$$\begin{aligned} F_{23} \geq (D_2 - D_1) \quad &\Leftrightarrow \quad \lambda_2^{(0,1)} \geq \lambda_1^{(0,1)}, \\ F_{12} \geq (D_2 - D_3) \quad &\Leftrightarrow \quad \lambda_2^{(1,0)} \geq \lambda_1^{(1,0)}. \end{aligned}$$

As noted above, however, this means that the tangent spaces of the local manifolds corresponding to the eigenvector $\mathbf{v_1}$ for both $(0, 1)$ and $(1, 0)$ lie outside the (x, y) phase space. Thus competitiveness in the (w, t) phase space is equivalent to the local invariant manifolds at $(0, 1)$ being always vertical at that fixed point, and similarly, any local manifolds at $(1, 0)$ are always horizontal at that point.

All this is summarised by the following result:

Proposition 4.1. *The following are equivalent:*

1. *Both* $\lambda_2^{(0,1)} \geq \lambda_1^{(0,1)}$ *and* $\lambda_2^{(1,0)} \geq \lambda_1^{(1,0)}$ *hold.*
2. *The Nagylaki-Crow model is competitive in* (w, t) *coordinates.*
3. *The tangent spaces of the local manifolds corresponding to* $\mathbf{v_1^{(0,1)}}$ *and* $\mathbf{v_1^{(1,0)}}$ *lie outside the* (x, y) *phase space.*

5. Existence of a nonmonotone invariant manifold

All of our numerical evidence for the Nagylaki-Crow model suggests the existence of at least one nonmonotone invariant manifold Σ connecting the steady states $(0, 1)$ and $(1, 0)$ in the phase plane (see Figure 1).

The aim of this section is to prove that at least one nonmonotone invariant manifold Σ does indeed exist when

$$F_{12} > D_2 - D_3,$$

and that it is the graph of a convex function if

$$F_{11} > D_1 - D_3 > -F_{33}.$$

Here the first inequality is one half of the inequality (3.12) for competition.

We recall that the time variable is denoted by s, so as to avoid confusion with the vertical coordinate t from Section 3.

5.1. *In the original (x, y) coordinates*

In the style of [14], consider the temporal evolution of the function $\varphi : [0,1] \times \mathbb{R}_+ \to \mathbb{R}_+$ satisying $\varphi(x, 0) = \varphi_0(x)$ and

$$\varphi(1, s) = 0, \qquad \varphi(0, s) = 1 \qquad \forall s > 0. \tag{5.1}$$

We let Σ_s be the graph of $\varphi(\cdot, s)$. The boundary conditions (5.1) force the endpoints of Σ_s to remain fixed for all time. We let

$$\varphi_0(x) = (1 - x)(1 - \varepsilon x), \quad \text{where} \quad 0 < \varepsilon \ll 1. \tag{5.2}$$

Although this is not the simplest choice for the initial condition, it is nonetheless a valid choice since it satisfies the boundary condition; our reason for this condition will become apparent later. The hope is for Σ_s to converge (in the Hausdorff metric) to Σ as $s \to \infty$.

The equation governing the evolution of φ is the first order partial differential equation

$$\varphi_s + f(x, \varphi(x, s))\varphi_x = g(x, \varphi(x, s)), \tag{5.3}$$

where f and g are defined as in Equations (3.8), and $\varphi_s = \frac{\partial \varphi}{\partial s}$. Using d/ds to denote the material time derivative which follows trajectories in the phase plane, differentiating Equation (5.3) gives Lemma 2.1 from [14], which is

$$\frac{d\varphi_x}{ds} = g_x + (g_y - f_x - f_y\varphi_x)\varphi_x. \tag{5.4}$$

Note that the values of φ_x at $x = 0, 1$ do not affect the existence of Σ, since these are the endpoints of the interval which have no effect on any derivatives.

Nevertheless, it possible to investigate the right and left-sided limits of φ_x as $x \to 0, 1$ respectively, i.e. $\varphi_x(0, s)$ and $\varphi_x(1, s)$. These can be evaluated by substituting the expressions for f_x, f_y, g_x and g_y for $(x, y) = (0, 1)$ into Equation (5.4) to obtain the following ordinary differential equation for $\varphi_x(0, s)$:

$$\frac{d\varphi_x}{ds}(0, s) = (\lambda_2^{(0,1)} - \lambda_1^{(0,1)})(\varphi_x(0, s) + 1) - 2F_{13}.$$

whose initial condition is

$$\varphi_x(0, 0) = -1 - \varepsilon < -1.$$

Although this equation is separable, we will not calculate the explicit solution for $\varphi_x(0, s)$. Instead, we observe that if for some $s > 0$ we have $\varphi_x(0, s) = -1$, then

$$\frac{d\varphi_x}{ds}(0, s) = -2F_{13} \leq 0,$$

which implies that for all time $s > 0$, we have $\varphi_x(0, s) \leq -1$. A closer inspection of the explicit solution, which is outlined in AppendixB.1, reveals that in fact the solution never attains its upper bound of -1. Thus $\varphi_x(0, s) < -1$ for all $s > 0$.

Now we repeat the procedure for $(x, y) = (1, 0)$. This time we yield a differential equation for $\varphi_x(1, s)$:

$$\frac{d\varphi_x}{ds}(1, s) = -\varphi_x[(\lambda_2^{(1,0)} - \lambda_1^{(1,0)})(\varphi_x + 1) - 2F_{13}\varphi_x],$$

and its initial condition is

$$\varphi_x(1, 0) = -1 + \varepsilon \in (-1, 0).$$

Next, observe that if there exists some $s > 0$ such that $\varphi_x(0, s) = -1$, then

$$\frac{d\varphi_x}{ds}(1, s) = 2F_{13} \geq 0.$$

On the other hand, if there exists some $s > 0$ such that $\varphi_x(0, s) = 0$, then

$$\frac{d\varphi_x}{ds}(1, s) = 0,$$

and both of these findings imply that for all $s > 0$, we have $-1 \leq \varphi_x(1, s) \leq 0$. It is possible to compute the explicit solution for $\varphi_x(1, s)$ by separation of variables (see AppendixB.2), and it turns out that in fact $-1 < \varphi_x(1, s) \leq 0$ for all time.

All this indicates that the gradient of the evolving curve at the two endpoints is negative. Note, however, that there may be no lower bound for $\varphi_x(0, s)$.

Equation (5.4) can be differentiated to obtain an equivalent version of Lemma 3.1 in [14]. This governs the evolution of the convexity of φ, i.e. φ_{xx}. However, it will not be pursued in this paper, since we have found it too involved to track the sign of φ_{xx}. An approach in (x, y) coordinates does not easily lead to showing that $\phi_x < 0$, $\phi_{xx} > 0$, so we revert to (w, t) coordinates where establishing convexity is simpler via Lemma 1.

5.2. In the new (w, t) coordinates

It turns out that working with the (w, t) coordinates from Section 3 instead will simplify the calculations. Hence we will switch to the new dynamical variable.

One issue is that the phase space in (w, t) is a quadrant, which is not compact. Nonetheless, we can attempt to find a bounding curve on which the flow is directed inward and all orbits enter. We choose the initial data curve to serve as the bounding curve.

5.2.1. The set-up

For the sake of brevity, Equations (3.10) and (3.11) will be denoted as follows

$$\dot{w} = p(w, t), \qquad \dot{t} = q(w, t),$$

and, similar to the previous strategy, we seek the time evolution of a curve $\psi(w, s)$ which satisfies the first order quasilinear partial differential equation

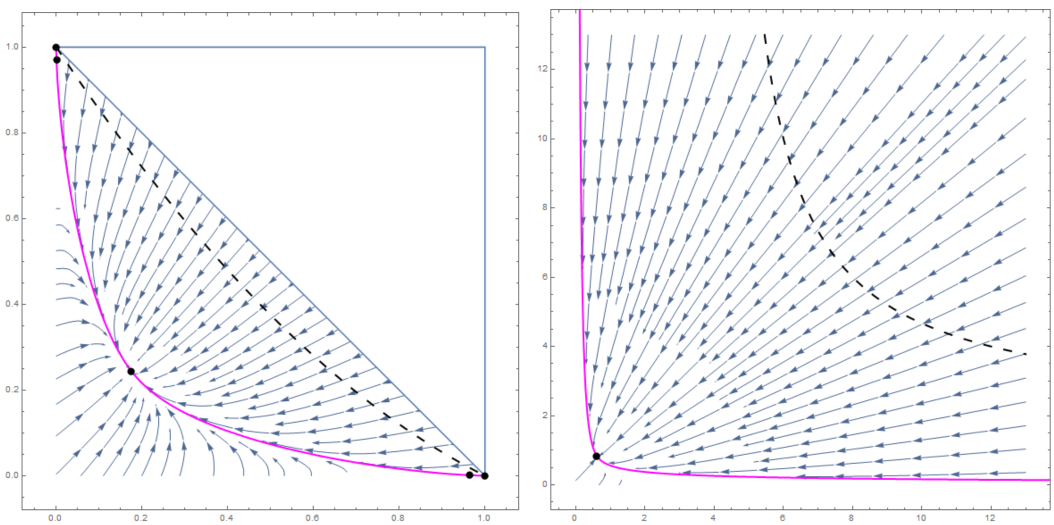
$$\psi_s + p(w, \psi(w, s))\psi_w = q(w, \psi(w, s)), \tag{5.5}$$

with the initial condition $\psi(w, 0) = \psi_0(w)$. Notice, however, that the straight line $y = 1 - x$ does not have a well-defined counterpart in (w, t) coordinates, and hence is a poor choice for an initial data curve to map onto ψ_0. Instead, we use for ψ_0 from φ_0 defined by (5.2) transformed into (w, t) coordinates (see Figure 3). With $y = \varphi_0(x)$ and $z = 1 - x - y$ in Equations (3.9) give

$$w = \frac{2}{(1 - x)\varepsilon}, \qquad t = \frac{2(1 - x\varepsilon)}{x\varepsilon}, \qquad 0 < x < 1.$$

Then by eliminating x, we obtain the hyperbola

$$t = \frac{-2(2 + w - w\varepsilon)}{2 - w\varepsilon},$$



(a) A plot of the initial data curve, the graph of φ_0, in (x, y) coordinates.

(b) A plot of the initial data curve, the graph of ψ_0, in (w, t) coordinates.

Figure 3: The initial data curve for the two coordinate systems (dashed line), which deforms in time according to the flow, converging to the invariant manifold Σ (solid line). For these plots, we take $\varepsilon = 0.5$.

which is decreasing in the first quadrant with vertical and horizontal asymptotes $w = 2/\varepsilon$ and $t = 2(1-\varepsilon)/\varepsilon$ respectively, and this is our choice for ψ_0 (see Figure 3). Since

$$(\psi_0)_w = \frac{-4}{(w\varepsilon - 2)^2} < 0 \qquad (\psi_0)_{ww} = \frac{8\varepsilon}{(w\varepsilon - 2)^3} > 0, \tag{5.6}$$

ψ_0 is both strictly decreasing and convex, as well as positive.

5.2.2. Equivalent boundary conditions

A second condition on the partial differential equation (5.5) is needed, and should be equivalent to the boundary condition (5.1). It turns out that the equivalent condition is that ψ must have a horizontal and vertical asymptote at all times, even if the positions of these asymptotes vary in time. For now, these asymptotes will be said to occur at t^* and w^* respectively.

To see this, first consider the point $(0, 1)$, i.e. $x = 0$, $y = 1$ (and $z = 0$). An attempt at substitution into Equation (3.9) indicates that t is infinite, yet w is undefined since the numerator and denominator both equal zero. The trick for determining the (w, t) counterpart of the point $(1, 0)$ is to the transformation (3.13), which is repeated here for the sake of convenience:

$$x = \frac{w}{2 + w + t}, \qquad y = \frac{t}{2 + w + t}. \tag{5.7}$$

By inspection, we see that the only means of attaining $y \to 1$ is by letting $t \to \infty$, assuming w remains finite. This directly forces $x \to 0$ anyway, implying that the point $(0, 1)$ in (x, y) coordinates corresponds to the limiting case of $t \to \infty$ in (w, t).

By a similar argument, the point $(0,1)$ turns out to represent the limiting case $w \to \infty$ in (w,t) coordinates, as long as t is kept finite. This suggests that since the graph of φ passes through $(0,1)$ and $(1,0)$ for all time, its counterpart graph of ψ should always have a horizontal and vertical asymptote respectively in $\mathbb{R}^2_+$.

Consider the one-sided derivatives for $x = 0, 1$. These will be denoted by $m_0 = \varphi_x(0,s)$ and $m_1 = \varphi_x(1,s)$ respectively.

Next, observe that the original evolving smooth curve φ can be approximated at both $x = 0$ and $x = 1$ using the following Taylor series

$$y = 1 + m_0 x + O(x^2), \tag{5.8}$$

$$y = 0 + m_1(x-1) + O((x-1)^2). \tag{5.9}$$

Then we convert to (w,t) coordinates by substituting the above two expressions for y into equations (3.13). However, as $x \to 0, 1$ respectively, the $O(x^2)$ and $O((x-1)^2)$ terms from the above Taylor series vanish. Therefore the points $(0,1)$ and $(1,0)$ correspond to the following asymptotes

$$\text{As} \quad t \to \infty, \qquad w \to w^*(s) = \frac{-2}{m_0+1}, \tag{5.10}$$

$$\text{As} \quad w \to \infty, \qquad t \to t^*(s) = \frac{-2m_1}{m_1+1}. \tag{5.11}$$

Since $m_0 < -1$ and $-1 < m_1 < 0$, we have that $w^*(s), t^*(s) > 0$. This indicates that both asymptotes will lie in the interior of our phase space, the first quadrant.

5.2.3. Investigating the gradient and convexity

Differentiating Equation (5.5) gives Lemma 2.1 from [14], which is

$$\frac{d\psi_w}{ds} = q_w + (q_t - p_w - p_t\psi_w)\psi_w. \tag{5.12}$$

Furthermore, we can repeat the procedure to obtain an equivalent version of Lemma 3.1 [14], which states that

$$\frac{d\psi_{ww}}{ds} = q_{ww} + \psi_w(2q_{wt} - p_{ww}) + \psi_w^2(q_{tt} - 2p_{wt}) - p_{tt}\psi_w^3 + \psi_{ww}(q_t - 2p_w - 3p_t\psi_w). \tag{5.13}$$

However, as the equations of motion are already simpler in (w,t), the partial derivatives of p and q are also easier to compute. In fact, we find that

$$p_{tt} = q_{ww} = 0,$$

which simplifies Equation (5.13) to

$$\frac{d\psi_{ww}}{ds} = \psi_w[(2q_{wt} - p_{ww}) + \psi_w(q_{tt} - 2p_{wt})] + \psi_{ww}(q_t - 2p_w - 3p_t\psi_w). \tag{5.14}$$

Also,

$$\begin{aligned} p_{ww} &= 2(D_2 - D_1 + F_{11} - F_{12} - F_{13}t), \\ q_{tt} &= 2(D_2 - D_3 + F_{33} - F_{23} - F_{13}w), \\ p_{wt} &= D_2 - D_1 - F_{23} - 2F_{13}w, \\ q_{wt} &= D_2 - D_3 - F_{12} - 2F_{13}t, \end{aligned}$$

and so

$$
\begin{aligned}
2q_{wt} - p_{ww} &= 2(D_1 - D_3 - F_{11} - F_{13}t), \\
q_{tt} - 2p_{wt} &= 2(D_1 - D_3 + F_{33} + F_{13}w),
\end{aligned}
$$

which are negative and positive when $F_{11} > D_1 - D_3$ and $F_{33} > D_3 - D_1$ respectively. Combining these two conditions gives the constraint

$$
F_{11} > D_1 - D_3 > -F_{33}, \tag{5.15}
$$

which is equivalent to having both $\lambda_1^{(1,0)} < 0$ and $\lambda_1^{(0,1)} < 0$. Meanwhile, note that $q_w < 0$ if and only if $D_2 > D_3 + F_{12}$ (or $\lambda_2^{(1,0)} > \lambda_1^{(1,0)}$).

Then, assuming (5.15) and $D_2 < D_3 + F_{12}$ (which is half of inequality (3.12), let $\alpha = \psi_w$ and $\beta = \psi_{ww}$ and rewrite Equations (5.12) and (5.14) as two coupled ordinary differential equations:

$$
\frac{d\alpha}{ds} = A\alpha^2 + B\alpha + C \tag{5.16}
$$

$$
\frac{d\beta}{ds} = \alpha(D + E\alpha) + \beta(F + G\alpha), \tag{5.17}
$$

where

$$
\begin{aligned}
A &= -p_t & E &= q_{tt} - 2p_{wt} > 0 \\
B &= q_t - p_w & F &= q_t - 2p_w \\
C &= q_w < 0 & G &= -3q_t \\
D &= 2q_{wt} - p_{ww} < 0 & &
\end{aligned}
$$

are all continuous (in fact, polynomial) functions of w and ψ (which replaces t). We already found that ψ_0 is strictly decreasing (and therefore nonmonotone) and convex. This will be rewritten as $\alpha_0 < 0$ and $\beta_0 > 0$, where $\alpha_0 = \alpha(w, 0)$ and $\beta_0 = \beta(w, 0)$. Now the aim is to prove

$$
\alpha < 0, \qquad \beta > 0 \qquad \forall s > 0, \tag{5.18}
$$

for all values of w for which $\psi(w, s)$ is defined. In other words, if the initial data curve is both strictly decreasing and convex in (w, t) coordinates, then it will remain that way as s increases.

Looking at the α-equation of (5.17), we would like use $C < 0$ in order to prove the following lemma (which is based on Corollary 2.2 of [14]):

Lemma 5.1. *If the smooth initial curve ψ_0 satisfies both $\psi_0 > 0$ and $(\psi_0)_w \leq 0$, then for all $s \geq 0$, $\psi(\cdot, s)$ is defined and smooth for all $w > w^*(s)$ (where $w^*(s)$ is the vertical asymptote of $\psi(\cdot, s)$ mentioned in the previous remark), with $\psi(\cdot, s) > 0$ and $\psi_w(\cdot, s) \leq 0$*

Proof. The major obstacle is that the vertical asymptote $w^*(s)$ is changing in time. To overcome this, let us rescale the asymptote by a change of variables, taking

$$
v = \frac{w}{w^*(s)}, \qquad \tau = s, \tag{5.19}
$$

then consider

$$
\tilde{\psi}(v, \tau) = \psi(w, s) = \psi(v\, w^*(\tau), \tau),
$$

whose vertical asymptote occurs at the constant $v = 1$. Now use

$$\begin{pmatrix} \psi_w \\ \psi_s \end{pmatrix} = \begin{pmatrix} v_w & \tau_w \\ v_s & \tau_s \end{pmatrix} \begin{pmatrix} \tilde{\psi}_v \\ \tilde{\psi}_\tau \end{pmatrix}.$$

By considering (5.19), the following can be deduced:

$$\begin{pmatrix} v_w & \tau_w \\ v_s & \tau_s \end{pmatrix} = \begin{pmatrix} \frac{1}{w^*(s)} & 0 \\ \frac{-ww^{*\prime}(s)}{(w^*(s))^2} & 1 \end{pmatrix},$$

hence

$$\begin{pmatrix} \psi_w \\ \psi_s \end{pmatrix} = \begin{pmatrix} \frac{1}{w^*(s)} & 0 \\ \frac{-ww^{*\prime}(s)}{(w^*(s))^2} & 1 \end{pmatrix} \begin{pmatrix} \tilde{\psi}_v \\ \tilde{\psi}_\tau \end{pmatrix}. \tag{5.20}$$

Next, using a similar argument to that used to obtain (5.12) and the Chain Rule, we have $\tilde{\psi}_v = w^*(s)\psi_w$ along with

$$\frac{d\tilde{\psi}_v}{ds} = w^* q_w + (q_t - p_w - (\ln w^*)' - \frac{p_t}{w^*}\tilde{\psi}_v)\tilde{\psi}_v. \tag{5.21}$$

But since $w^*(s) > 0$, we deduce that for $\tilde{\psi}_0 = \tilde{\psi}(\cdot, 0)$,

$$\psi_w < 0 \quad \text{for} \quad s = 0 \qquad \Rightarrow \qquad (\tilde{\psi}_0)_v < 0.$$

Similarly, we infer that $(\tilde{\psi}_0)_\tau > 0$ is also true.

Moreover, $w^* q_w < 0$, hence if $\tilde{\psi}_v = 0$ for some value of $v > 1$ and $s > 0$, then from (5.21) we have $\frac{d\tilde{\psi}_v}{ds} \leq 0$, and we deduce from that $\tilde{\psi}$ is strictly decreasing for $v > 1$.

Now we wish to prove that $\tilde{\psi}$ is defined and smooth for all $v > 1$, i.e.

$$\forall\, v \in (1, \infty) = \bigcup_{n=1}^{\infty} \left[1 + \frac{1}{n}, n\right],$$

so it suffices to prove that $\tilde{\psi}$ is defined and smooth on the interval $\left[1 + \frac{1}{n}, n\right]$ for any integer $n \geq 2$. However, each of these intervals are always closed and bounded, therefore $\tilde{\psi}$ attains a maximum and a minimum, and hence does not grow unbounded on $\left[1 + \frac{1}{n}, n\right]$ within a finite time. Thus $\tilde{\psi}$ is indeed defined for all $v > 1$. □

Invoking lemma 5.1 reveals that $\alpha < 0$ does indeed hold, since $C < 0$ (which in turn is due to $\psi = t > 0$.). Then, if we let $\theta = \alpha(D + E\alpha)$ and $\sigma = (F + G\alpha)$, the β-equation from (5.17) may be written as

$$\frac{d\beta}{ds} = \sigma\beta + \theta, \tag{5.22}$$

and $D < 0$, $E > 0$ which, combined with $\alpha < 0$, yield $\theta > 0$. But $\beta_0 > 0$ as well, hence by lemma 4.1 from [14], $\beta > 0$ whenever $s > 0$.

Thus (5.18) holds, which indicates that

$$\psi_w < 0, \qquad \psi_{ww} > 0 \qquad \forall s \geq 0,\ \forall w. \tag{5.23}$$

Now it has been established that ψ is strictly decreasing and strictly convex for at all times.

Now returning to (x, y) coordinates, and applying Lemma 3.1, we see that φ is indeed convex and decreasing in the original x, y coordinates. Moreover, φ is also obviously non-negative due to invariance of the compact phase space T. Since $\varphi(\cdot, s)$ is smooth and non-increasing for arbitrary s, by Lemma 2.7 from [14], the graph of φ is a nonmonotone Lipschitz manifold with Lipschitz constant unity. By the Arzelà-Ascoli Theorem, the space of Lipschitz functions in a compact space is itself compact. Hence a sequence of Lipschitz manifolds, such as the one constructed from φ, will always have a convergent subsequence whose limit is in turn also a nonmonotone Lipschitz manifold.

This limit indicates that at least one nonmonotone invariant manifold does indeed exist in (x, y) space. Hence we can summarise our results in the following theorem:

Theorem 5.2. *In the Nagylaki-Crow model (3.4) - (3.6) suppose that*

$$D_2 < D_3 + F_{12}. \tag{5.24}$$

Then the model has at least one nonmonotone invariant manifold that connects the fixed points corresponding to the populations consisting of all heterozygotes.

If in addition

$$F_{11} > D_1 - D_3 > -F_{33}. \tag{5.25}$$

this nonmonotone manifold is the graph of a convex function.

Observe that all of (3.12) is not needed, so existence works for not necessarily competitive models that satisfy (5.24) only. A similar result applies by interchanging w and t leading to a version of Theorem 5.2 with (5.24) replaced by $D_2 < D_1 + F_{23}$ and (5.25).

The nonmonotone invariant manifold of Theorem 5.2 is a connecting orbit that connects the two axial fixed points. When both inequalities in (3.12) hold, so that the system is competitive, and thus being planar, also monotone with the order $\geq_L$ defined by $x = (x_1, x_2) \geq_L y = (y_1, y_2)$ if and only if $x_1 \geq y_1$ and $x_2 \leq y_2$. In this case, existence of a connecting orbit (even with additional fixed points ordered by $\geq_L$) follows from [15]. Jiang [16] showed that for cooperative systems, this connecting orbit is unique if the Jacobian is irreducible at the two fixed points. Even if Jiang's result on irreducibility can be modified for planar competitive systems, in our model the Jacobian at the axial fixed points is reducible and so an alternative approach is needed to determine when the nonmonotone manifold is unique. In any case our existence result does not require a competitive model for existence, and so conditions for uniqueness of the manifold are not at all clear and will be dealt with elsewhere.

6. Discussion

We have shown the existence of a nonmonotone invariant manifold for a continuous-time differential fertility model in Population Genetics without requiring additivity of fertilities or mortalities, nor competitive dynamics.

To do this, we set up an invertible mapping between an evolving curve φ in (x, y) space and an evolving (unbounded) curve ψ in (w, t) space. Through our lemmata it then suffices to show that the graph of ψ is decreasing and convex, (as then so is the graph of φ). Then convergence is established in (x, y) space using Lipschitz and bounded sequences of the graph of φ. Thus the crucial part is setting up the map between ϕ and ψ by a change of dynamical variables.

There was no need to assume that the system was competitive in either coordinate system; nonetheless, the existence proof given in this paper for a decreasing manifold only applies when the inequality (5.24) (or its alternative in (3.12)) applies. Furthermore, when the Nagylaki-Crow model satisfies the inequalities (5.25), the invariant manifold is also convex. How far the conditions for convexity can be weakened is an open problem.

Meanwhile, it is also unknown what conditions are required for the manifold to be smooth. Even when the model is competitive or strongly competitive, results such as in [17] are not immediately applicable as $\partial\mathbb{R}^2_+$ is not invariant for our system. Since f and g are bivariate polynomials each heteroclinic connection (orbit) along Σ is an analytic invariant manifold [18], hence only the interior fixed points on Σ need to be checked for C^1-smoothness. In addition, the Stable Manifold Theorem also indicates that the stable and unstable subspaces for any saddle point in the dynamical system are both one-dimensional. Moreover Σ, which is itself one-dimensional, must be tangential to one of the aforementioned subspaces, as well as C^1-smooth, even at the saddle point. Hence all that remains is to find conditions for Σ to be also C^1-smooth at interior fixed points of the model that are not saddles.

Recall that in Figure 1 the invariant manifold Σ is not unique. In fact, numerics suggest that the model has a countable sequence of nonmonotone invariant manifolds; these are analogous to the family of manifolds described by Hirsch in Theorem 1.1 from [1]. Our model, however, is not immediately covered by Hirsch's results because the boundary of the phase space in Hirsch's system is invariant whereas in our case, the flow on the boundary points towards the interior of T. Nevertheless we believe that the difference is not problematic, as long as we have a repelling boundary for the phase space.

Acknowledgements

B. Seymenoglu is jointly funded by the EPSRC and the Department of Mathematics, UCL.

References

[1] M. W. Hirsch, Systems of differential equations which are competitive or cooperative: III. Competing species, Nonlinearity 1 (1) (1988) 51–71.

[2] J. Hofbauer, K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge University Press, Cambridge, 1998.

[3] L. S. Penrose, The meaning of "Fitness" in human populations, Journal of Mathematical Biology 31 (4) (1993) 317–349.

[4] W. F. Bodmer, Differential Fertility in Population Genetics Models, Genetics 51 (3) (1965) 411–424.

[5] D. L. Rucknagel, J. V. Neel, The hemoglobinopathies, Progress in Medical Genetics (1961) 158–260.

[6] T. Nagylaki, J. F. Crow, Continuous Selective Models, Theoretical Population Biology 5 (1974) 257–283.

[7] T. Nagylaki, Introduction to Theoretical Population Genetics, Springer-Verlag, Berlin, 1992.

[8] K. P. Hadeler, U. Liberman, Selection models with fertility differences, Journal of Mathematical Biology 2 (1) (1975) 19–32.

[9] G. J. Butler, H. I. Freedman, P. Waltman, Global dynamics of a selection model for the growth of a population with genotypic fertility differences, Journal of Mathematical Biology 14 (1) (1982) 25–35.

[10] K. P. Hadeler, D. Glas, Quasimonotone systems and convergence to equilibrium in a population genetic model, Journal of mathematical analysis and applications 95 (2) (1983) 297–303.

[11] M. Koth, F. Kemler, A one locus-two allele selection model admitting stable limit cycles, Journal of Theoretical Biology 122 (3) (1986) 263–267.

[12] J. M. Szucs, Equilibria and dynamics of selection at a diallelic autosomal locus in the Nagylaki-Crow continuous model of a monoecious population with random mating, Journal of Mathematical Biology 31 (4) (1993) 317–349.

[13] E. Akin, J. M. Szucs, Approaches to the Hardy-Weinberg manifold, Journal of Mathematical Biology 32 (7) (1994) 633–643.

[14] S. A. Baigent, Convexity-preserving flows of totally competitive planar Lotka–Volterra equations and the geometry of the carrying simplex, Proceedings of the Edinburgh Mathematical Society 55 (2012) 53–73.

[15] E. N. Dancer, P. Hess, Stability of fixed points for order-preserving discrete-time dynamical systems, Journal für die reine und angewandte Mathematik 419 (1991) 125–139.

[16] J. Jiang, On the Existence and Uniqueness of Connecting Orbits for Cooperative Systems, Acta Mathematica Sinica, New Series 8 (1986) 184–188.

[17] J. Mierczyński, Smoothness of unordered curves in two-dimensional strongly competitive systems, Applicationes Mathematicae 25 (4) (1999) 449–455.

[18] L. Perko, Differential Equations and Dynamical Systems, Third Edition, 3rd Edition, Springer-Verlag, Northern Arizona University, 2001.

AppendixA. The explicit equations for the Nagylaki-Crow model

This Appendix explicitly provides the governing equations for the three genotype frequencies, and demonstrates that only two of the three equations are required.

By substituting (3.3) into Equation (2.1), one has the following three equations of motion…

$$\begin{aligned}
\dot{x} &= \frac{1}{4}z^2 F_{22} \\
&\quad + x\left[y\left(D_3 - 2zF_{23}\right) + z\left(F_{12} + D_2\right) - y^2 F_{33} - z^2 F_{22} - D_1\right] \\
&\quad + x^2\left(-2yF_{13} - 2zF_{12} + F_{11} + D_1\right) - x^3 F_{11} \\
\dot{y} &= \frac{1}{4}z^2 F_{22} \\
&\quad + y\left[x\left(D_1 - 2zF_{12}\right) + z\left(F_{23} + D_2\right) - x^2 F_{11} - z^2 F_{22} - D_3\right] \\
&\quad + y^2\left(-2xF_{13} - 2zF_{23} + F_{33} + D_3\right) - y^3 F_{33} \\
\dot{z} &= 2xy\, F_{13} \\
&\quad + z\left[y(F_{23} + D_3 - 2xF_{13}) + x(F_{12} + D_1) - x^2 F_{11} - y^2 F_{33} - D_2\right] \\
&\quad + z^2\left(-2yF_{23} - 2xF_{12} + \frac{1}{2}F_{22} + D_2\right) - z^3 F_{22}
\end{aligned}$$

But recall that $x + y + z = 1$. Due to this, one has $\dot{x} + \dot{y} + \dot{z} = 0$, rendering the z-equation redundant. Moreover, the remaining two equations can be re-written in terms of x and y…

$$\begin{aligned}
\dot{x} &= \frac{1}{4}y^2 F_{22} - \frac{1}{2}yF_{22} + \frac{1}{4}F_{22} \\
&\quad + x\left(y\left(-F_{12} + \frac{5}{2}F_{22} - 2F_{23} - D_2 + D_3\right) + y^2(-F_{22} + 2F_{23} - F_{33})\right. \\
&\quad \left. + F_{12} - \frac{3}{2}F_{22} - D_1 + D_2\right) \\
&\quad + x^2\left(y(2F_{12} - 2F_{22} + 2F_{23} - 2F_{13}) + F_{11} - 3F_{12} + \frac{9}{4}F_{22} + D_1 - D_2\right) \\
&\quad + x^3\left(-F_{11} + 2F_{12} - F_{22}\right) \\
\dot{y} &= \frac{1}{4}x^2 F_{22} - \frac{1}{2}xF_{22} + \frac{1}{4}F_{22} \\
&\quad + y\left(x\left(-2F_{12} + \frac{5}{2}F_{22} - F_{23} + D_1 - D_2\right) + x^2(-F_{11} + 2F_{12} - F_{22})\right. \\
&\quad \left. - \frac{3}{2}F_{22} + F_{23} + D_2 - D_3\right) \\
&\quad + y^2\left(x\left(2F_{12} - 2F_{22} + 2F_{23} - 2F_{13}\right) + \frac{9}{4}F_{22} - 3F_{23} + F_{33} - D_2 + D_3\right) \\
&\quad + y^3\left(-F_{22} + 2F_{23} - F_{33}\right)
\end{aligned}$$

The attentive reader will note that a trick can be used to obtain the y-equation from the x-equation. This can be done by replacing each x with y, as well as swapping each subscript 1 with a 2 (and vice-versa).

AppendixB. Explicit solutions for the gradient at the corners

AppendixB.1. The gradient at the left corner

For this subsection, we consider the (one-sided) gradient of the evolving curve at $(0, 1)$, i.e. $\varphi_x(0, s)$. The aim is to prove that $\varphi_x(0, s) < -1$ for all $s > 0$.

Case 1: $F_{13} = \lambda_2^{(0,1)} - \lambda_1^{(0,1)} = 0$

This is a straightforward case; the solution is simply a constant:

$$\varphi_x(0, s) = -1 - \varepsilon < -1.$$

Case 2: $F_{13} = 0$**,** $\lambda_2^{(0,1)} - \lambda_1^{(0,1)} \neq 0$

The solution can be easily found using separation of variables:

$$\varphi_x(0, s) = -1 - \varepsilon \exp\left(s(\lambda_2^{(0,1)} - \lambda_1^{(0,1)})\right) < -1.$$

because ε and the exponential are both positive.

Case 3: $F_{13} > 0$

Let $\varphi_x(0, s) = -1 + \delta$, where the time-dependent function $\delta(s)$ satisfies the initial condition $\delta(0) = \varepsilon > 0$. Then the corresponding ODE for δ is

$$\frac{d\delta}{ds} = (\lambda_2^{(0,1)} - \lambda_1^{(0,1)})\delta + 2F_{13},$$

but $2F_{13} > 0$. Hence, by Lemma 4.1 from [14], $\delta(s) > 0$ for all $s > 0$. In other words, $\varphi_x(0, s) < -1$ for all $s > 0$ holds, as it does for the other two cases.

AppendixB.2. The gradient at the right corner

Now we consider the (one-sided) gradient of the evolving curve at $(1, 0)$, i.e. $\varphi_x(1, s)$. This time the intention is to show that $-1 < \varphi_x(1, s) < 0$. Assume that $\lambda_2^{(1,0)} - \lambda_1^{(1,0)} > 0$. Then we are left with only two sub-cases.

Case 1: $F_{13} = 0$

The solution can be easily found using separation of variables:

$$\varphi_x(1, s) = \frac{-1 + \varepsilon}{-1 - \varepsilon + \varepsilon \exp\left(s(\lambda_2^{(1,0)} - \lambda_1^{(1,0)})\right)},$$

but the numerator is non-zero, thus $\varphi_x(1, s) \neq 0$ for finite $s > 0$. In addition,

$$\varepsilon \exp\left(s(\lambda_2^{(1,0)} - \lambda_1^{(1,0)})\right) \neq 0 \quad \Rightarrow \quad \varphi_x(1, s) \neq -1,$$

so for all $s > 0$, $-1 < \varphi_x(1, s) < 0$.

Case 2: $F_{13} > 0$

If there exists some $s > 0$ such that $\varphi_x(0, s) = -1$, then

$$\frac{d\varphi_x}{ds}(1, s) = 2F_{13} > 0,$$

leading to a contradiction. Hence $\varphi_x(0, s) > -1$.

Let $\tilde{B} = \lambda_2^{(1,0)} - \lambda_1^{(1,0)}$. Then, using separation of variables, we obtain the explicit solution

$$\varphi_x(1, s) = \frac{\tilde{B}}{2F_{13} - \tilde{B} - e^{\tilde{B}s}\left(2F_{13} + \frac{\varepsilon B}{1-\varepsilon}\right)}, \tag{B.1}$$

which is non-zero and defined for any finite $s > 0$.

Therefore for all $s > 0$, we must have $-1 < \varphi_x(1, s) < 0$.