Stable cycling in quasi-linkage equilibrium: fluctuating dynamics under gene conversion and selection

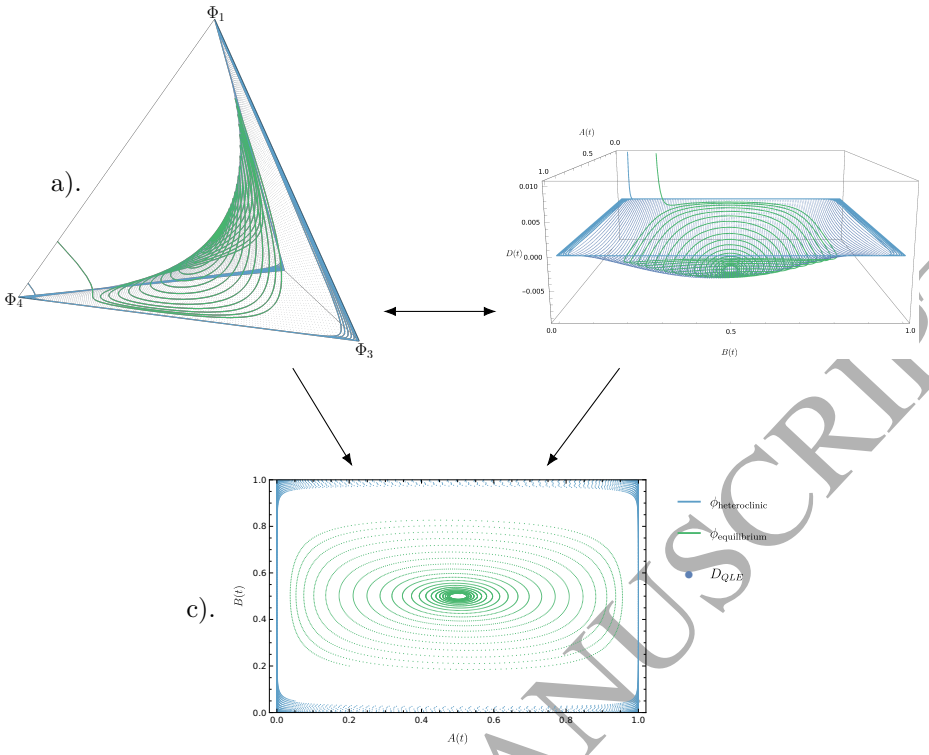
Timothy W. Russell, Matthew J. Russell, Francisco Úbeda, Vincent A.A. Jansen

PII: S0022-5193(19)30237-1
DOI: https://doi.org/10.1016/j.jtbi.2019.06.010
Reference: YJTBI 9933

To appear in: *Journal of Theoretical Biology*

Received date: 7 December 2018
Revised date: 15 April 2019
Accepted date: 12 June 2019

Please cite this article as: Timothy W. Russell, Matthew J. Russell, Francisco Úbeda, Vincent A.A. Jansen, Stable cycling in quasi-linkage equilibrium: fluctuating dynamics under gene conversion and selection, *Journal of Theoretical Biology* (2019), doi: https://doi.org/10.1016/j.jtbi.2019.06.010



Highlights

- We build a selection-recombination model conjectured to have stable cycling
- A separation of time-scales is exploited, using a biologically interpretable change of variables to achieve an explicit fast variable
- An approximate closed-form for a well-known surface, known as the quasilinkage equilibrium manifold is extracted
- The dynamics are simplified by exploiting the surface and stability of the heteroclinic cycle is deduced analytically

Stable cycling in quasi-linkage equilibrium: fluctuating dynamics under gene conversion and selection

Timothy W. Russell[a,*], Matthew J. Russell[b], Francisco Úbeda[a], Vincent A.A. Jansen[a]

[a]*School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK*
[b]*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*

Abstract

Genetic systems with multiple loci can have complex dynamics. For example, mean fitness need not always increase and stable cycling is possible. Here, we study the dynamics of a genetic system inspired by the molecular biology of recognition-dependent double strand breaks and repair as it happens in recombination hotspots. The model shows slow-fast dynamics in which the system converges to the quasi-linkage equilibrium (QLE) manifold. On this manifold, sustained cycling is possible as the dynamics approach a heteroclinic cycle, in which allele frequencies alternate between near extinction and near fixation. We find a closed-form approximation for the QLE manifold and use it to simplify the model. For the simplified model, we can analytically calculate the stability of the heteroclinic cycle. In the discrete-time model the cycle is always stable; in a continuous-time approximation, the cycle is always unstable. This demonstrates that complex dynamics are possible under quasi-linkage equilibrium.

Keywords:
Quasi-linkage equilibrium, Slow manifold, Lyapunov function, Global stability, Multiple time-scales

1. Introduction

*Corresponding author
Email address: timothy.russell.2015@rhul.ac.uk (Timothy W. Russell)

Genetic equilibrium, the idea that gene frequencies are the same from one generation to the next, was the focus of early work on population genetics. The attention shifted when it was discovered that one-locus viability models can exhibit cycling behaviour and genetic equilibrium does not have to be achieved (Kimura, 1958; Hadeler and Liberman, 1975; Asmussen and Feldman, 1977; Cressman, 1988). Further investigation showed that two-locus viability models with recombination can also exhibit cycling behaviour (Akin, 1979; Hastings, 1981; Akin, 1982, 1983, 1987).

The discrete-time selection-recombination equations (Lewontin and Kojima, 1960; Bürger, 2000) have provided a determinsitic model for changes in the genetic make up of a population. Despite the fact that these equations are often used to study the properties of stable equilibria, they are inherently nonlinear, meaning even the most simple formulations of the equations can have complex dynamics. Examples include limit cycles (Akin, 1983) and heteroclinic cycles (Haig and Grafen, 1991; Úbeda et al., 2019). Whether the cycles are maintained indefinitely or eventually die out (i.e. their stability properties) is mathematically challenging and of significant biological importance. This is the focus of the research we present here.

Many genetic processes within an interacting population of individuals can be captured by the selection-recombination equations, as they allow for arbitrary selection regimes defined by model-specific fitness matrices. Here, we investigate the stability of cycles in two-locus genetic systems characterised by a specific interaction between selection, gene conversion and crossover. This interaction corresponds to a model of the evolution of recombination hotspots (Úbeda et al., 2019). However, we re-write this model in standard selection-recombination equations form by noticing that the effect of conversion in Úbeda et al. (2019) can be split into its effect on selection (and incorporated to the selection component of the standard selection-recombination equation) and its effect on formation of double heterozygotes (and incorporated into the recombination component of the standard selection-recombination equation). Furthermore, while the model in Úbeda et al. (2019) assumes that the values taken by the selection-recombination parameters are constrained by their biological interdependence,

here we assume that the parameter values are independent and not limited by biological constraints. In doing so, we allow for multiple forms of interaction between selection, conversion and crossover, provided they produce the same equations. This formulation allow us to focus on the mathematical properties of the generalised model.

Biologically, the processes in our model are initiated by recognition between a protein formed by a modifier gene and a target locus, whereby the protein interacts with the target, initiating conversion and potentially crossover (Úbeda and Wilkins, 2011; Úbeda et al., 2019). Other than the evolution of recombination hotspots (Úbeda and Wilkins, 2011; Úbeda et al., 2019), examples of similar recognition-initiated interactions producing sustained cycling include: the evolution of homing endonucleases (Yahara et al., 2009), the evolution of meiotic drive (Haig and Grafen, 1991), the evolution of host-parasite interactions (Sasaki et al., 2002) and the evolution of altruism via tag based recognition (Jansen and Van Baalen, 2006).

If selection is weak, stable cycling cannot occur within the two-locus selection-recombination equations if the equilibria are hyperbolic (Nagylaki et al., 1999; Pontz et al., 2018). These conditions produce dynamics which converge to a stable equilibrium. Under weak selection, the argument by Nagylaki et al. (1999) uses the existence of an invariant stable manifold which attracts the dynamics. On this attracting manifold, the dynamics are gradient-like and converge to equilibrium (Pugh et al., 1977). This manifold is known in genetics as the *quasi-linkage equilibrium* (QLE) manifold (Kimura, 1965). It is the set of states defined by the property that linkage disequilibrium changes an order of magnitude slower than the allele frequencies (Kimura, 1965).

In geometrical terms, this means that the dynamics approach a manifold after a short initial time. If an approximate expression for such a manifold can be found, it can be exploited mathematically to simplify the system (Constable and McKane, 2017). This is usually done by assuming that selection in the model is weak (Barton, 1995; Nagylaki et al., 1999; Kirkpatrick et al., 2002; Lion, 2018). We identify the linkage disequilibrium as a fast variable in our model, isolate it using a coordinate transformation and find an approximation of the surface to which the dynamics converge. Here we show that the

existence of a time-scale separation between variables and hence attraction to the QLE manifold is not exclusively associated with simple dynamics which are characterised by gradient-like convergence to an interior equilibrium.

The model presented here has complex dynamics, such as bistability and a global bifurcation. We show that, in such a system, it is still possible to find an approximate yet accurate explicit expression for the QLE manifold. For analytical tractability, following standard methods in population genetics, we derive a continuous-time approximation to our discrete-time model (Nagylaki et al., 1999; Bürger, 2000; Pontz et al., 2018). We use this continuous-time approximation to find an expression for the QLE manifold. We go on to use this to constrain the dynamics analytically to this surface, reducing the dimension of the system. We are then able to calculate the stability of the now-planar heteroclinic cycle that exists in our model within certain parameter regimes. Constraining the dynamics is a powerful step as it allows for the use of the only known analytic heteroclinic stability condition in discrete-time for planar cycles (Hofbauer and Schlag, 2000). In the vicinity of this heteroclinic cycle, strong fluctuations are possible on the QLE manifold.

Finally, we numerically assess the accuracy of our approximation of the QLE manifold against both sources of error: the *quasi steady-state assumption* and the use of the continuous-time derived manifold within the discrete-time system. We find that the manifold is a good approximation for the discrete-time system for both damped oscillations towards the unique interior equilibrium and the approach towards the heteroclinic cycle.

2. The model

We investigate the dynamics of haplotype frequencies of two alleles at two interacting loci, in an infinite population, undergoing a specific selection regime (uniquely defining the fitness matrix W), recombination and random union of gametes (panmixia). Once the fitness matrix and the parameter δ are defined, the system of equations in question is fully defined (A.1). First, we describe the biological processes which justify our selection

regime, then we present the resulting fitness matrix (A.5).

Our model describes the evolution of recombination hotspots by following the dynamics between a modifier gene — producing a recombinogenic protein — and a target gene, on which the protein binds to, causing a double-strand break and initiating recombination (Úbeda et al., 2019). This model is here re-written as a system of selection-recombination equations. This system describes the following general processes: a fitness benefit derived from recognition between modifier and target (β), a fitness cost derived from gene conversion (γ) and the reshuffling of alleles in double heterozygotes caused by gene conversion and crossover (δ) (Úbeda et al., 2019). Our original formulation of the model included another parameter α, which we have normalised to one (without loss of generality) for simplicity.

The dynamics of the matching process between homozygotes and gene conversion leads to the following system of equations describing the frequency of each haplotype in the next generation

$$\begin{aligned} x_1' &= \frac{1}{\bar{w}}\Big(x_1[1+\beta x_1-\gamma x_2]-\delta D\Big),\\ x_2' &= \frac{1}{\bar{w}}\Big(x_2[1-\beta x_2+\gamma x_1]+\delta D\Big),\\ x_3' &= \frac{1}{\bar{w}}\Big(x_3[1-\beta x_3+\gamma x_4]+\delta D\Big),\\ x_4' &= \frac{1}{\bar{w}}\Big(x_4[1+\beta x_4-\gamma x_3]-\delta D\Big), \end{aligned} \tag{1}$$

where the linkage disequilibrium between alleles is

$$D = x_1x_4 - x_2x_3, \tag{2}$$

and the population mean fitness is

$$\bar{w} = x_1 + x_2 + x_3 + x_4 + \beta\left(x_1^2 - x_2^2 - x_3^2 + x_4^2\right). \tag{3}$$

Superscript primes indicate the value of the variable in the next generation. The population mean fitness, $\bar{w}$, ensures that the sum of the haplotype frequencies remains constant in time. To ensure the right hand side of the difference equations does not become negative, which would imply that the number of gametes produced is negative, we require

	A_1	A_2
B_1	x_1	x_3
B_2	x_2	x_4

Table 1: Relations between the haplotype frequencies, x_1, x_2, x_3, x_4, the alleles controlling the recombinogenic protein type, A_1, A_2, and the alleles controlling the target site sequence, B_1, B_2. The table indicates that the allele frequencies are obtained by summing over the haplotype frequencies in the corresponding row or column. Explicitly, $A_1 = x_1 + x_2$, $A_2 = x_3 + x_4$, $B_1 = x_1 + x_3$ and $B_2 = x_2 + x_4$.

that the parameters β, γ can only take values between 0 and 1. This can be justified by the fact parameters represent probabilities in the context of the selection-recombination equations. The parameter δ can only take values between 0 and $\frac{1}{2}$.

Our fitness matrix and therefore our model has similarities with that of (Karlin et al., 1970). They study symmetric viability, meaning they impose a symmetric fitness matrix. Ours is perhaps superficially similar but has a crucial difference; our matrix is not symmetric. Our matrix results in certain local symmetries within the resulting equations — symmetries which are a hallmark of heteroclinic cycles. In that sense, our model is closer to the ones of Haig and Grafen (1991) who also studied a process with a non-symmetric fitness matrix also finding a heteroclinic cycle. We choose a specific example to study for mathematical tractability and to link it to specific biological examples.

3. Analysis and Results

The model has two different qualitative behaviours: convergence to equilibrium and sustained oscillations. In both cases, the rate-of-change of D tends towards zero on a faster time scale than the rate-of-change of the allele frequencies (see Figure 1). This suggests that the system has two separate time scales and that the dynamics converge towards the QLE manifold. We will find an approximate expression for this manifold.

For brevity, we introduce $A = A_1$ and $B = B_1$ to denote the frequency of the first recombinogenic protein and its matching target allele, respectively. The frequency of the second recombinogenic protein and its target allele can then be written as $A_2 = 1 - A$

and $B_2 = 1 - B$.

subsectionChange of variables

The first step towards finding an approximation of the QLE manifold is changing coordinates so that they describe the allele frequencies and linkage disequilibrium. We achieve this by transforming variables from haplotype frequencies to allele frequencies using

$$\begin{aligned} A &= x_1 + x_2, \\ B &= x_1 + x_3, \\ D &= x_1 x_4 - x_2 x_3, \end{aligned} \tag{4}$$

where A and B take values on the interval $[0, 1]$. D represents linkage disequilibrium between alleles and takes values on $[-\frac{1}{4}, \frac{1}{4}]$. If we consider (4) to be the forward transformation, we arrive at the backward transformation

$$\begin{aligned} x_1 &= AB + D, \\ x_2 &= A(1 - B) - D, \\ x_3 &= (1 - A)B - D, \\ x_4 &= (1 - A)(1 - B) + D. \end{aligned} \tag{5}$$

Transforming using (4), the discrete-time model becomes

$$\begin{aligned} A' &= \frac{1}{\bar{w}} \beta A(1 - A)(2B - 1) + A, \\ B' &= \frac{1}{\bar{w}} \Big[(\gamma - \beta) B(2A - 1)(B - 1) + \gamma(2B - 1)D \Big] + B, \\ D' &= \frac{1}{\bar{w}^2} \Big[(A - 1)A(B - 1)B(\beta - \gamma) + \\ &\quad D\Big(\beta[2A(A - 1)(B^2 - B)(\gamma + \beta) + \\ &\quad A(A - 1)\gamma - (2A - 1)(\delta - 1)(2B - 1)] - \delta + 1 \Big) + \\ &\quad D^2 \Big(\beta(\beta + \gamma)(2A - 1)(2B - 1) + \beta(-2\delta + 3) + \gamma \Big) + \\ &\quad 2\beta D^3(\beta + \gamma) \Big]. \end{aligned} \tag{6}$$

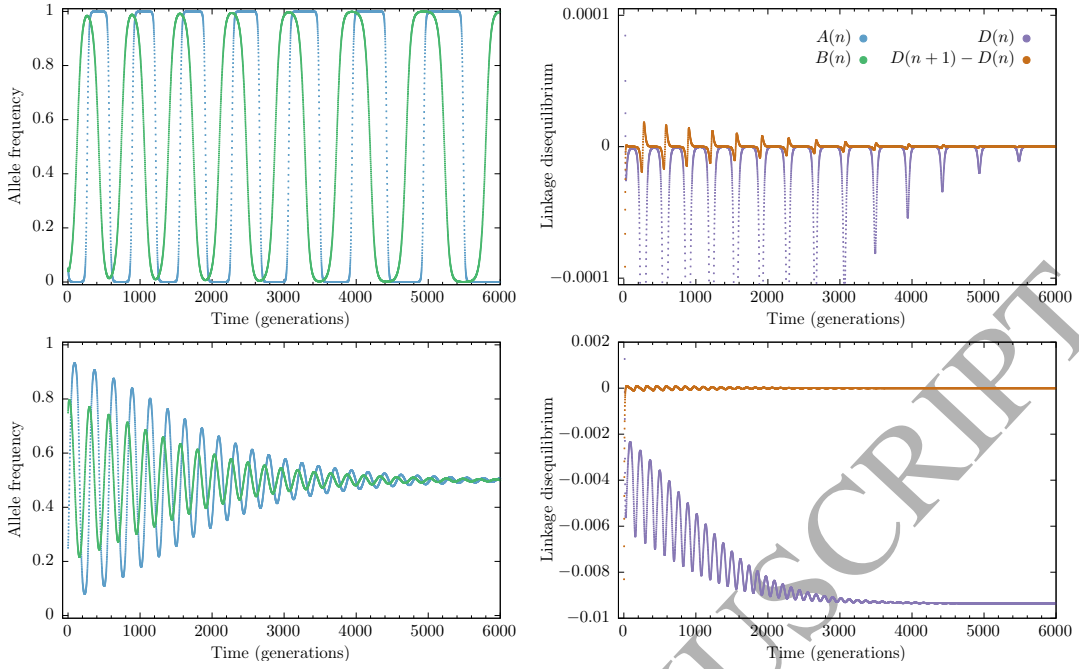


Figure 1: **Time series showing examples of the two types of behaviour of the discrete-time model** (6). The examples in the top row have initial conditions: $A(0) = 0.05$, $B(0) = 0.95$, $D(0) = 0.0005$ and those in the bottom row have initial conditions $A(0) = 0.25$, $B(0) = 0.75$, $D(0) = 0.0005$. Trajectories in both rows were solved with the same set of parameters: $\beta = 0.1$, $\gamma = 0.13$, $\delta = 0.2$. The top row shows a typical trajectory nearby the heteroclinic cycle. It also shows that after an initial period of rapid change, the linkage disequilibrium eventually changes relatively slowly (D' becomes approximately constant in time), indicating the convergence of the dynamics to QLE manifold. The bottom row shows a typical orbit exhibiting damped oscillations and convergence to the asymptotically stable interior equilibrium (9).

Additionally, $\bar{w}$ is transformed into

$$\bar{w} = 1 + \beta(2A - 1)(2B - 1) + 2\beta D. \tag{7}$$

As these coordinates include linkage disequilibrium (D) explicitly, they allow for a simple interpretation of the surface of total linkage equilibrium: the *Wright manifold*. This surface can now be written as the part of state space where $D = 0$ (Rice, 2004).

3.1. Equilibria and local stability

The system has a maximum of ten solutions when solving for potential equilibria. Five of these live within the positive state space of the model and are therefore biologically feasible. Four of the five biologically realistic equilibria are located at the four vertices of the tetrahedron that forms the 3-simplex (in haplotype coordinates). These corner equilibria, in allelic coordinates (A, B, D), are

$$\begin{aligned} \Phi_1 &= (1, 1, 0), \\ \Phi_2 &= (1, 0, 0), \\ \Phi_3 &= (0, 1, 0), \\ \Phi_4 &= (0, 0, 0). \end{aligned} \tag{8}$$

We analysed the linear stability of these equilibria in Úbeda et al. (2019) and we summarise the main results here. For our choice of parameters the equilibria Φ_2 and Φ_3 are always unstable. Moreover, if $\beta < \gamma$ these equilibria are saddles. The equilibria Φ_1 and Φ_4 are stable if $\beta > \gamma$ and are saddles, and thus unstable, if $\beta < \gamma$. Note that if A or B take values of either 0 or 1 then $D = 0$. Upon inspection of the transformed models, we find that the lines connecting the equilibria Φ_1 to Φ_2 ($A = 1, D = 0$), Φ_2 to Φ_4 ($B = 0, D = 0$), Φ_4 to Φ_3 ($A = 0, D = 0$) and Φ_3 to Φ_1 ($B = 1, D = 0$) are all invariant. When all these equilibria are saddles (i.e. when $\beta < \gamma$) a heteroclinic connection exists:

$$\cdots \to \Phi_1 \to \Phi_2 \to \Phi_4 \to \Phi_3 \to \Phi_1 \to \cdots .$$

The fifth equilibrium is positioned in the interior of the simplex. For this interior equilibrium it is easily verified that $\dot{A} = 0$ and $\dot{B} = 0$ for $A = B = \frac{1}{2}$. The interior equilibrium, in allelic coordinates, is

$$\Phi_5 = (\tfrac{1}{2}, \tfrac{1}{2}, D^*), \tag{9}$$

where D^* is the negative root of

$$(\gamma - \beta)D^{*2} - \delta D^* - \tfrac{1}{16}(\gamma - \beta) = 0, \tag{10}$$

given by

$$D^* = \frac{\delta - \sqrt{\delta^2 + \frac{1}{4}(\gamma - \beta)^2}}{2(\gamma - \beta)}. \tag{11}$$

The positive root is larger than $\frac{1}{4}$ for $\delta > 0$ and therefore the corresponding equilibrium has negative haplotype frequencies.

The multipliers of the discrete-time model (6) at the interior equilibrium Φ_5 are given by

$$\begin{aligned}\lambda_1 &= 1 + \frac{\gamma D^* + \sqrt{(\gamma D^*)^2 + \frac{1}{4}\beta(\beta-\gamma)}}{\bar{w}^*},\\ \lambda_2 &= 1 + \frac{\gamma D^* - \sqrt{(\gamma D^*)^2 + \frac{1}{4}\beta(\beta-\gamma)}}{\bar{w}^*},\\ \lambda_3 &= 1 - \frac{\delta + 2D^*(\beta-\gamma)}{\bar{w}^*},\end{aligned} \tag{12}$$

where $\bar{w}^* = 1 + 2\beta D^*$ denotes the value of $\bar{w}$ evaluated at the interior equilibrium (Úbeda et al., 2019). The eigenvalues $\hat{\lambda}_i$ of the interior equilibrium of the continuous-time approximation are given by $\hat{\lambda}_i = \lambda_i - 1$.

If $\beta > \gamma$ then $D^* > 0$ and $\bar{w}^* > 0$. Therefore, in this region of parameter space, it is relatively easy to see that the interior equilibrium is a saddle (both in the discrete and the continuous-time models). Specifically, λ_1 and λ_3 are always negative, and for $0 < \delta < \frac{1}{2}$, $\lambda_3 > -1$. λ_2 is always positive. If $\beta < \gamma$ then $D^* < 0$. Eigenvalues λ_1 and λ_2 can now form a conjugate pair of complex eigenvalues. For the equilibrium to be locally stable in the discrete-time model we require $|\lambda_{1,2}| < 1$. This leads to the conditions for local stability

$$2\gamma\bar{w}^* D^* < \frac{1}{4}\beta(\beta-\gamma). \tag{13}$$

If $\delta < \frac{1}{2}$ this condition is always fulfilled (Úbeda et al., 2019). This stability condition (13) applies only to the discrete-time model as its continuous-time approximation (15) is always locally stable (for $\beta < \gamma$).

3.2. Global stability: A Lyapunov function and heteroclinic cycle

3.2.1. A continuous-time approximate model

These results on asymptotic local stability leave the question of what the global dynamics are and, in particular, if the heteroclinic connection is an attractor, or whether orbits move away from it. While the focus of this paper is to analyse the global stability

properties of the discrete-time model (1), we introduce the following continuous-time approximation of the discrete-time model (Nagylaki et al., 1999; Bürger, 2000) to aid us in this matter significantly

$$\begin{aligned}
\dot{x}_1 &= \frac{1}{\bar{w}}\Big(x_1[1+\beta x_1-\gamma x_2]-\delta D\Big)-x_1,\\
\dot{x}_2 &= \frac{1}{\bar{w}}\Big(x_2[1-\beta x_2+\gamma x_1]+\delta D\Big)-x_2,\\
\dot{x}_3 &= \frac{1}{\bar{w}}\Big(x_3[1-\beta x_3+\gamma x_4]+\delta D\Big)-x_3,\\
\dot{x}_4 &= \frac{1}{\bar{w}}\Big(x_4[1+\beta x_4-\gamma x_3]-\delta D\Big)-x_4,
\end{aligned} \tag{14}$$

where derivatives with respect to time t are denoted by a dot above a variable. The expressions for $\bar{w}$ and D are given by (2) and (3), the same as in the discrete-time model. The continuous-time model written in the transformed variables is

$$\begin{aligned}
\dot{A} &= \frac{1}{\bar{w}}\beta A(1-A)(2B-1),\\
\dot{B} &= \frac{1}{\bar{w}}\Big[(\gamma-\beta)B(2A-1)(B-1)+\gamma(2B-1)D\Big],\\
\dot{D} &= \frac{1}{\bar{w}}\Big[(\gamma-\beta)\left[D^2-AB(1-A)(1-B)\right]-\beta D(2A-1)(2B-1)-\delta D\Big].
\end{aligned} \tag{15}$$

It is easy to show that the equilibria for the discrete-time model and its continuous-time approximation are the same (Bürger, 2000). Similarly, it is easy to show that the eigenvalues of the Jacobian at each equilibrium in the continuous-time model equal the discrete-time eigenvalues minus unity — a consequence of the fixed time-step in the discrete-time system. We use the continuous-time model in two ways: introducing a Lyapunov function for the interior equilibrium, showing it to be globally stable; using it to find an analytically tractable version of the approximate QLE manifold, as the expression is significantly simpler when derived from the continuous-time model.

3.2.2. Lyapunov function

For the continuous-time model it is relatively easy to show that the heteroclinic cycle repels orbits using a Lyapunov function. Before we show this, we first observe that for any solution of (15) as long as $D \leq 0$ at some point in time, $D \leq 0$ onwards if $\beta < \gamma$, and

with equality only if the solution lives on the heteroclinic connection. This can easily be seen by inspecting the right hand side of the differential equation describing the change in D when $\beta < \gamma$, which is negative everywhere on the Wright manifold, apart from on the heteroclinic connection, where it is zero. Therefore, if $D(t_0) < 0$, then $D(t) < 0$ for all $t > t_0$. This means that trajectories can pass through the Wright manifold where $D = 0$ in only one direction, and are then confined to the region where $D \leq 0$ once they have done so.

With this established, we now consider the function

$$V(A, B) = [A(1 - A)]^{\gamma-\beta}[B(1 - B)]^{\beta}. \tag{16}$$

This function (16) serves as a natural candidate for a Lyapunov function of system (14) as it retains invariance of the system along the boundaries (where either $A = 0$, $A = 1$, $B = 0$ or $B = 1$). Indeed, for $\beta < \gamma$ this function takes the value $V = 0$ along the heteroclinic connection, and takes positive values anywhere else in or on the simplex. The continuous-time model with D set to zero (15) is equivalent to the replicator equations for 2×2 games and our Lyapunov function (16) is equivalent to that of this system, serving as its constant of motion (Hofbauer and Sigmund, 1998).

The candidate function V is a Lyapunov function if $\beta < \gamma$ for orbits which at some point pass through the Wright manifold. To show this, we inspect its time derivative along solutions of (15):

$$\dot{V} = -\beta\gamma \frac{D\,(1 - 2B)^2}{\bar{w}\; B(1 - B)} V. \tag{17}$$

The right hand side of (17) is always less than or equal to zero if $D \leq 0$, meaning V is a Lyapunov function within this region. For orbits starting in the forward invariant part of state space where $D < 0$ the value of V will thus increase or stay constant over time. The ω-limit of these orbits must therefore be invariant sets for which either $D = 0$ or $B = \frac{1}{2}$. If $\beta < \gamma$ the only invariant part of the Wright manifold $D = 0$ is the heteroclinic connection, where $V = 0$. As the value of V cannot decrease and is positive for all points in or on the simplex that are not part of the heteroclinic connection, the heteroclinic

connection cannot be an ω-limit of these orbits, within which the only other candidates are the invariant sets contained with $B = \frac{1}{2}$, which is the interior equilibrium Φ_5. Any orbits starting within the parts of the simplex where $D < 0$ will therefore move towards the interior equilibrium.

A corollary of this observation is that arbitrarily close to the heteroclinic connection, where $D = 0$, there will be points that are within the region of the simplex where $D < 0$. The Lyapunov function (16) shows that orbits starting at these points will move away from the heteroclinic connection, towards the interior equilibrium. The heteroclinic connection is therefore not stable. The interior equilibrium clearly is stable and must be the attractor for all initial points in the interior of the simplex for which initially $D < 0$. This shows that in the continuous-time model the heteroclinic cycle is unstable. Simulations suggest that the interior equilibrium is a global attractor within the simplex.

3.2.3. Discrete-time heteroclinic cycle

The Lyapunov argument does not carry over to the discrete-time model. In the discrete-time model, does the heteroclinic connection attract or repel? We analytically investigate this using the approximate QLE manifold in section 3.5. We also numerically investigate the regions of initial condition space in which the cycle is attracting, and the results are plotted in Figure 2. In the diagram we can distinguish two regions in parameter space with qualitatively different behaviour, and the boundary between them:

1. Within the first region, $\beta < \gamma$, the interior equilibrium is stable and attracts nearby orbits. Within this region the heteroclinic connection also attracts. Between the two attractors we find the boundary of the basins of attraction. The basin boundary moves towards the heteroclinic connection for small β.

2. Within the second region $\beta > \gamma$. All trajectories converge to one of the corner equilibria, Φ_1 or Φ_4, apart from orbits starting exactly at the unstable interior equilibrium Φ_5.

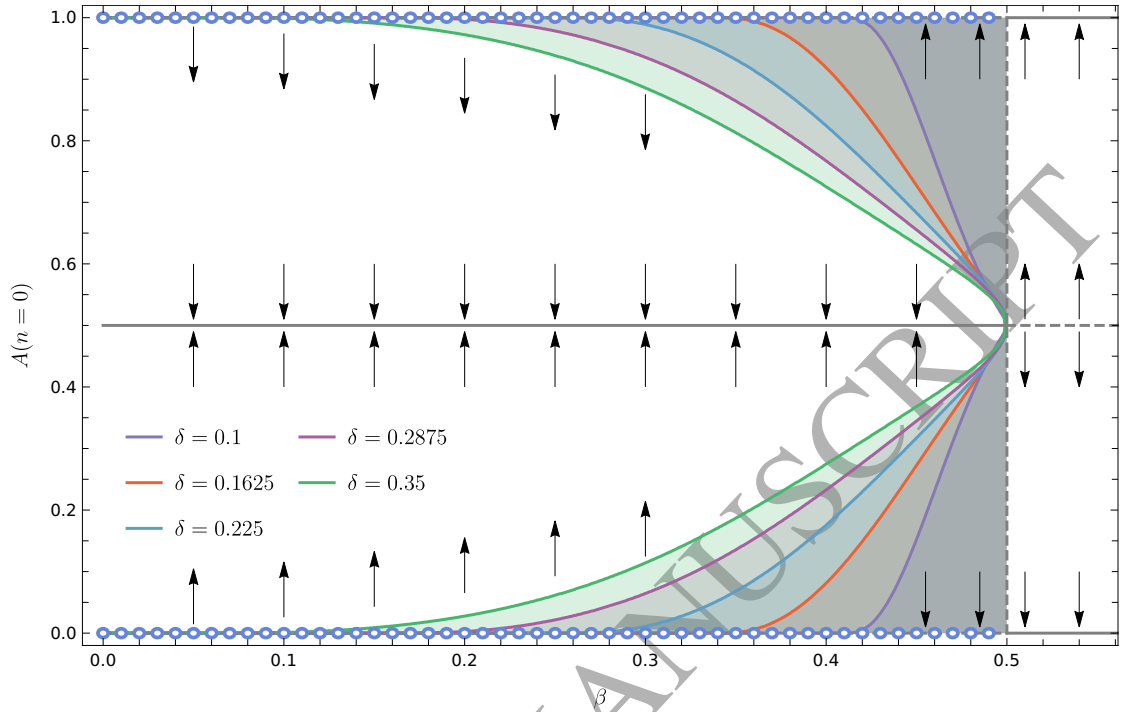


Figure 2: **The basin of attraction of the heteroclinic cycle against β for the discrete-time model**. The diagram shows the different qualitative behaviours of the model resulting from different initial conditions. The arrows point towards the different attractors. The shaded regions show the basins of attraction of heteroclinic cycle for varying values of δ (see legend). The diagram was constructed by starting orbits at different initial conditions, sampled at equally spaced intervals along the line connecting the equilibria Φ_1 and Φ_4 for which $A = B$ and $D = A(1-A)$ in allelic coordinates, or $(x_1, 0, 0, 1-x_1)$ in gametic coordinates. We determine whether a specific orbit reaches interior equilibrium or a heteroclinic cycle numerically: if an orbit reaches within $\epsilon = 10^{-12}$ distance from the equilibrium, it is assumed to be at equilibrium. The first trajectory moving along the line of initial conditions which does not tend towards equilibrium is taken to be on the basin boundary. The heteroclinic cycle exists on the left of the vertical dashed line at $\beta = \gamma = 0.5$. At this point both the interior equilibrium and heteroclinic cycle lose stability and all trajectories tend toward one of the corner equilibria, Φ_1 or Φ_4. Parameters: $\gamma = 0.5$, δ as indicated in figure. Dashed lines represent unstable equilibria, drawn lines represent stable equilibria and small blue circles represent the heteroclinic cycles.

3. Between these two regions $\beta = \gamma$, all trajectories converge to the Wright manifold. On the Wright manifold there is a line of unstable equilibria for which $B = \frac{1}{2}$, $D = 0$. Orbits starting on the Wright manifold with $B < \frac{1}{2}$ converge to the line $A = 0$, $D = 0$, and those starting with $B > \frac{1}{2}$ converge to the line $A = 1$, $D = 0$.

These results show that the heteroclinic connection in the discrete-time model can be stable. To find out how general this is we will next analytically determine the stability of the heteroclinic connection in the discrete-time model. First, we approximate the QLE manifold towards which the trajectories converge.

3.3. The QLE manifold

If $\beta = \gamma$ the interior equilibrium is degenerate: in the discrete-time model the equilibrium has two real multipliers at unity (whilst the interior equilibrium of the continuous-time model has two eigenvalues at zero). Because there are two eigenvalues at unity (zero), the equilibrium will have a two dimensional center manifold. If $\beta = \gamma$ the center manifold is the Wright manifold, the part of state space where $D = 0$, and where the gamete frequencies are in linkage equilibrium. The third eigenvalue has a modulus smaller than one (smaller than zero for the continuous-time model) and the associated stable manifold is given by the line $A = B = \frac{1}{2}$. Orbits on this stable manifold move towards the center manifold.

If $\beta < \gamma$ these two multipliers become a complex pair with real part smaller than one (or negative real part for the continuous-time model). The equilibrium within this region is hyperbolic (for all $0 < \delta < \frac{1}{2}$) for the ODE (15). The same is true for the map (6) when there is not equality in the stability condition (13). The center manifold morphs into a two dimensional invariant manifold that is different from the Wright manifold and contains the interior equilibrium (9). On this manifold, orbits cycle around the equilibrium. The invariant manifold containing the third eigenvector, the line on which $A = B = \frac{1}{2}$, remains in existence. Over this line, orbits quickly converge towards the equilibrium and as they approach the linkage disequilibrium, D changes rapidly while the allele frequencies A and B remain unchanged. Other orbits show a similar behaviour (see

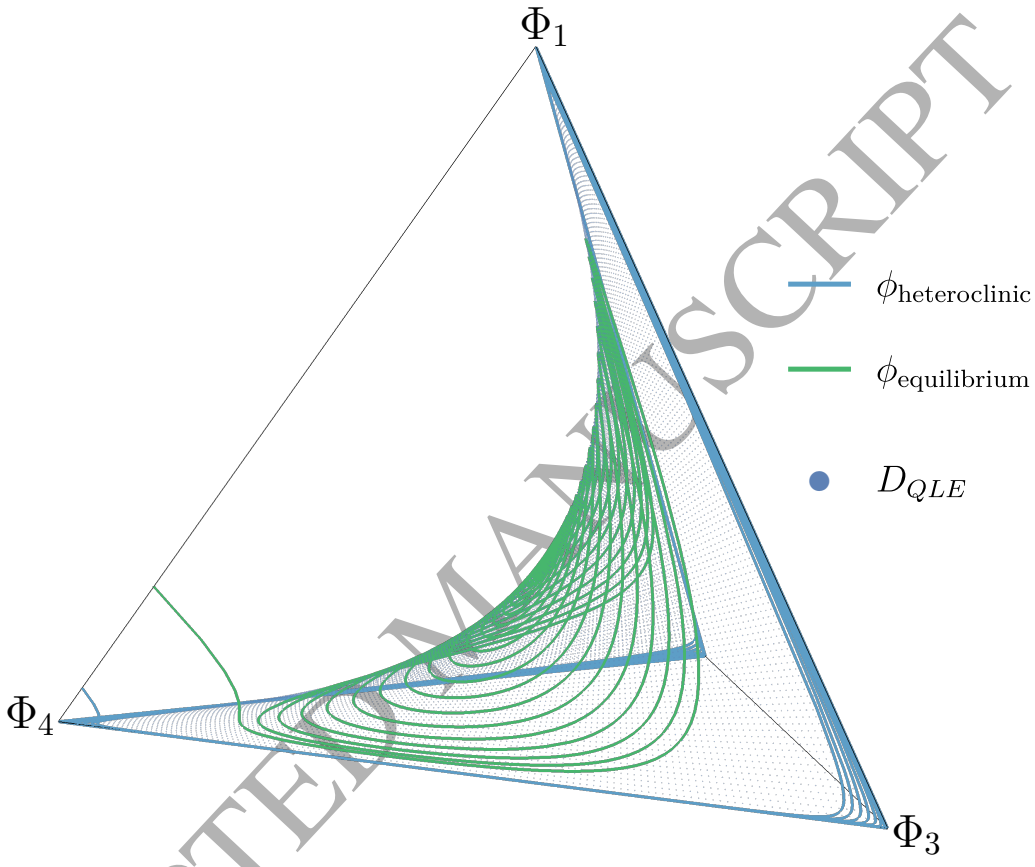


Figure 3: **The approximate quasi-linkage equilibrium manifold, and the approach to it by two typical trajectories of the discrete-time model.** Two trajectories, $\phi_{\text{heteroclinic}}$ and $\phi_{\text{equilibrium}}$, differing only in initial conditions, of the transformed discrete-time (1) system within the tetrahedron, both converging quickly to a slow manifold. Here, the small dots are points on the manifold D_{QLE}, given by (18). As can be seen, the trajectories converge quickly to this manifold. Parameters and initial conditions as in Figure 1.

Figure 3): orbits generally converge towards the two dimensional manifold. Once orbits are close to this manifold the orbits move slowly towards either the interior equilibrium or the heteroclinic cycle, depending on the initial conditions (see Figure 2).

To approximate the QLE manifold, we will use a quasi-steady state argument. Specifically, we say that the change in linkage disequilibrium $D(t)$ occurs on a much faster time scale than changes in the allele frequencies and will therefore settle on a quasi-equilibrium. This means that we can assume that the allele frequencies A and B are effectively constant, as D settles. With this assumption, we then solve the equilibrium equation for D as a function of the allele frequencies, $D_{QLE}(A, B)$. It turns out that this gives a good approximation for the QLE manifold for the discrete-time model as well as the continuous-time approximation.

Simulations suggest that the gamete frequencies are attracted towards the manifold where they are in quasi-linkage equilibrium. We approximate the QLE manifold by

$$\begin{aligned} D_{QLE}(A,B) = &\frac{\beta(2A-1)(2B-1)+\delta}{2(\gamma-\beta)} \\ &- \sqrt{\left(\frac{\beta(2A-1)(2B-1)+\delta}{2(\gamma-\beta)}\right)^2 + AB(1-A)(1-B)}. \end{aligned} \tag{18}$$

As we show in Appendix B the relevant slow time-scale is proportional to $(\gamma-\beta)^{-\frac{1}{2}}$.

3.4. Simplification by reducing to allele frequencies

Given the tendency of the haplotype frequencies to settle on the QLE, one would expect that if $\gamma > \beta$, the dynamics proceed to the QLE manifold, and that the allele frequencies then change slowly, either towards, or away from the interior equilibrium. This is indeed what happens in the vicinity of the interior equilibrium. Further away from equilibrium, and in particular in the vicinity of the heteroclinic cycle, this is not necessarily true. It is possible that the manifold $D = D_{QLE}(A, B)$ is situated outside the simplex in which all gamete frequencies are positive. If that is the case, the dynamics will be constrained by the edges of the simplex.

Inside the simplex, $D_{QLE} \leq 0$ if $\gamma > \beta$. If the manifold, D_{QLE}, cuts through the sides of the simplex, it can only be on the faces where $D \leq 0$, which is when $x_1 \leq 0$

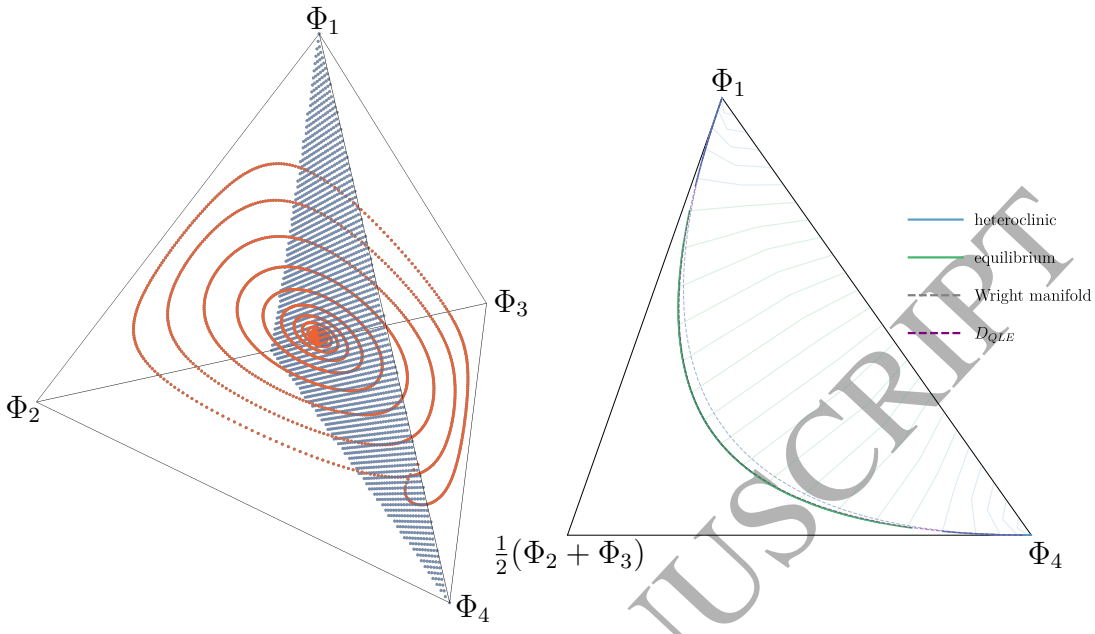


Figure 4: **The fast approach to the QLE manifold shown using a Poincar section**. The dynamics of our model has two different times scales and shows slow-fast dynamics. (a) A typical trajectory of the model (1), simulated using $\beta = 0.1, \gamma = 0.13$ and $\delta = 0.11$ and initial conditions $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.24, 0, 0, 0.76)$. To visualise the slow-fast dynamics we following the Poincaré section $x_2 = x_3$ $(=A = B)$ and record every instance where the orbit (shown in red) cuts through this section. (b) The intersection points for a orbit plotted on the Poincaré section. The points of intersection of 22 trajectories are shown. The trajectories have initial conditions equally spaced on the line connecting Φ_1 to Φ_4. The parameters used are $\beta = 0.3, \gamma = 0.35$ and $\delta = 0.2$. The figure shows the fast approach towards the slow manifold (the thin, drawn lines connect the points of intersection from the same initial condition). The slow manifold is visible as the accumulation of points forming a curve. Although the true slow manifold (blue and green filled lines) and our approximation, DQLE, (purple dashed line) are distinct from the Wright manifold (dashed grey line) apart from at the corners, where they intersect, they are very close and the purple curve is covered by the blue and green line in most of the figure. Green dots are from orbits that end up in the interior equilibrium, Φ_5, blue dots from orbits going towards the heteroclinic cycle. The gap on the slow manifold between the blue and green points contains the basin boundary. There will be an invariant closed curve located on the slow manifold in the middle of this gap.

or $x_4 \leq 0$. In terms of allele frequencies (A, B, D), that is when $D = -AB$ or when $D = -(1-A)(1-B)$. The approximate manifold to which the dynamics are drawn is thus given by $D = D_S(A, B)$, where

$$D_S(A,B) = \max\Big[D_{QLE}(A,B), -AB, -(1-A)(1-B)\Big], \tag{19}$$

and we will use this to simplify the dynamics; in particular we will use it to determine the stability of the heteroclinic cycle.

The system constrained to the attracting manifold is given by just two equations, describing the frequencies of A and B on the slow manifold,

$$\begin{aligned} A' &= \frac{1}{\bar{w}}\beta A(1-A)(2B-1) + A, \\ B' &= \frac{1}{\bar{w}}\Big[(\gamma-\beta)B(2A-1)(B-1) + \gamma(2B-1)D_S(A,B)\Big] + B, \end{aligned} \tag{20}$$

where

$$\bar{w} = \beta(2A-1)(2B-1) + 2\beta D_S(A,B) + 1. \tag{21}$$

The dimensionality is now reduced and the system is significantly simplified. We can now study and depict our model as a two dimensional system (Figure 5). The stability of the heteroclinic cycle is governed by the magnitude of the eigenvalues in the connected saddles that make up the cycle. In the planar system this is relatively simple to do.

3.5. Stability of heteroclinic cycle in the discrete-time model

To study the stability of our heteroclinic cycle, we use the condition derived in Hofbauer and Schlag (2000) which determines whether a planar discrete-time heteroclinic cycle is attracting or not. The condition involves the product of the ratio of the logarithm of the expanding (e_i) eigenvalues and the absolute value of the logarithm of the contracting eigenvalues (c_i) at the saddle equilibria (Φ_i where $i = 1, ..., 4$) the heteroclinic cycle travels between. We follow their notation and use ρ_i to denote each individual ratio and

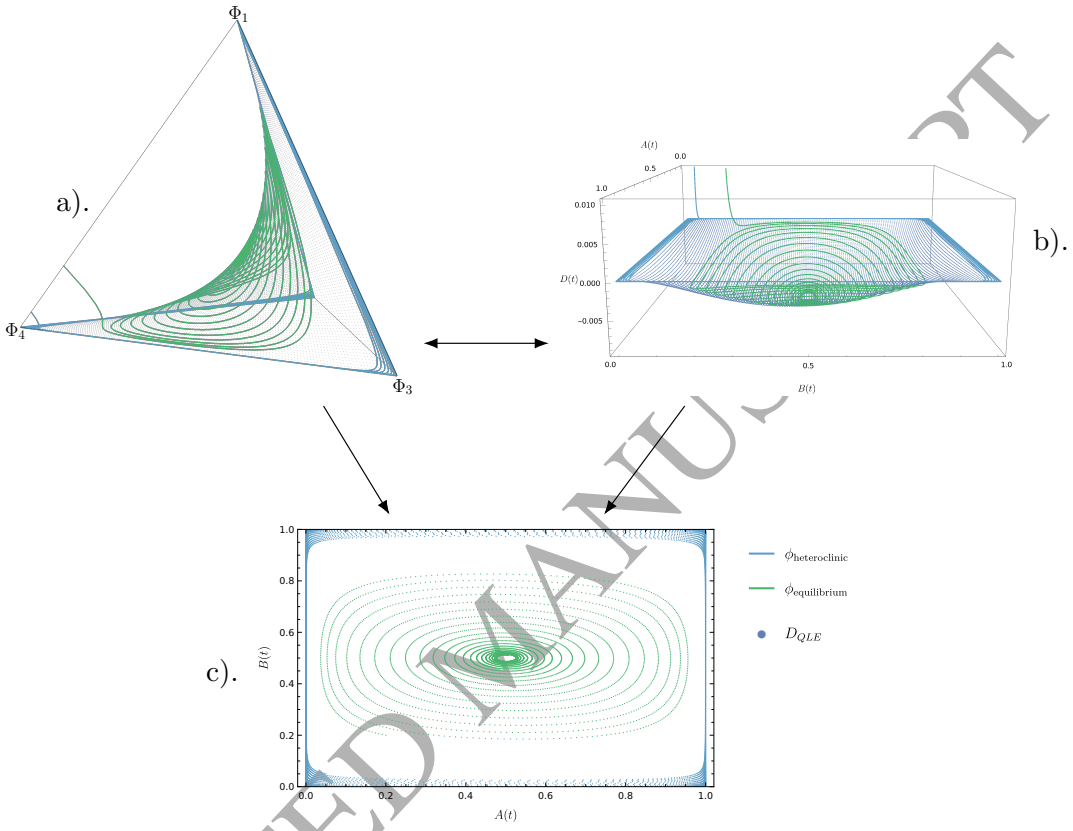


Figure 5: **The simplification of the system by using the approximate slow manifold, $\mathbf{D_{QLE}}$.** (a) The trajectories of our model represented gamete frequencies as given by eqns (1), plotted on the 3-simplex. The QLE manifold, $D = D_{QLE}$, is also plotted with a grid of equally spaced points. (b) The same trajectories and the attracting manifold plotted for the transformed model (20); in both panels (a) and (b) the fast approach to the slow manifold is visible. (c) The same trajectories but plotted on the QLE manifold. The system is reduced to a planar system in the allele coordinates. Parameters and initial conditions as in Figures 1 and 3. Panel (a) is a re-use of Figure 3.

ρ to denote the product of the ρ_i,

$$\begin{aligned} \rho &= \prod_{i=1}^{n} \rho_i, \\ \rho_i &= \frac{\log e_i}{|\log c_i|}, \qquad i = 1, ..., n. \end{aligned} \tag{22}$$

For our model, $n = 4$ and therefore $\rho = \rho_1\rho_2\rho_3\rho_4$. We are then able to state the stability condition: a planar discrete-time heteroclinic cycle is asymptotically stable if $\rho < 1$ and is unstable if $\rho > 1$ (Hofbauer and Schlag, 2000). The specific eigenvalues for the equilibria and their type are given in Table 2. Their derivation can be found in Appendix C.

Eigenvalue	$\frac{1}{1+\beta}$	$\frac{1+\gamma}{1+\beta}$	$\frac{1}{1-\beta}$	$\frac{1-\gamma}{1-\beta}$
Type	c_1, c_4	e_1, e_4	e_2, e_3	c_2, c_3
Equilibria	Φ_1 & Φ_4		Φ_2 & Φ_3	

Table 2: The eigenvalues of the saddle equilibria between which the heteroclinic cycle travels, used to determine the asymptotic stability of the heteroclinic cycle in discrete-time. Eigenvalues of type c are contracting (incoming), ones of type e are expanding (outgoing). Due to the symmetries in our system, the eigenvalues at Φ_1 and at Φ_4 are equal and the eigenvalues at Φ_2 and at Φ_3 are equal.

Calculating ρ using the eigenvalues in Table 2, we arrive at the condition for stability of the heteroclinic cycle

$$\left(\frac{\log \frac{1+\gamma}{1+\beta}}{|\log \frac{1}{1+\beta}|} \frac{\log \frac{1}{1-\beta}}{|\log \frac{1-\gamma}{1-\beta}|} \right)^2 < 1, \tag{23}$$

which, if $\beta < \gamma$, can be rewritten as

$$\frac{\log(1+\beta)}{\log(1-\beta)} < \frac{\log(1+\gamma)}{\log(1-\gamma)}. \tag{24}$$

In this form, it is readily seen that (23) is always satisfied if $\beta < \gamma$. Therefore, in our discrete-time model constrained to the QLE manifold (20), the heteroclinic cycle is always asymptotically stable if it exists.

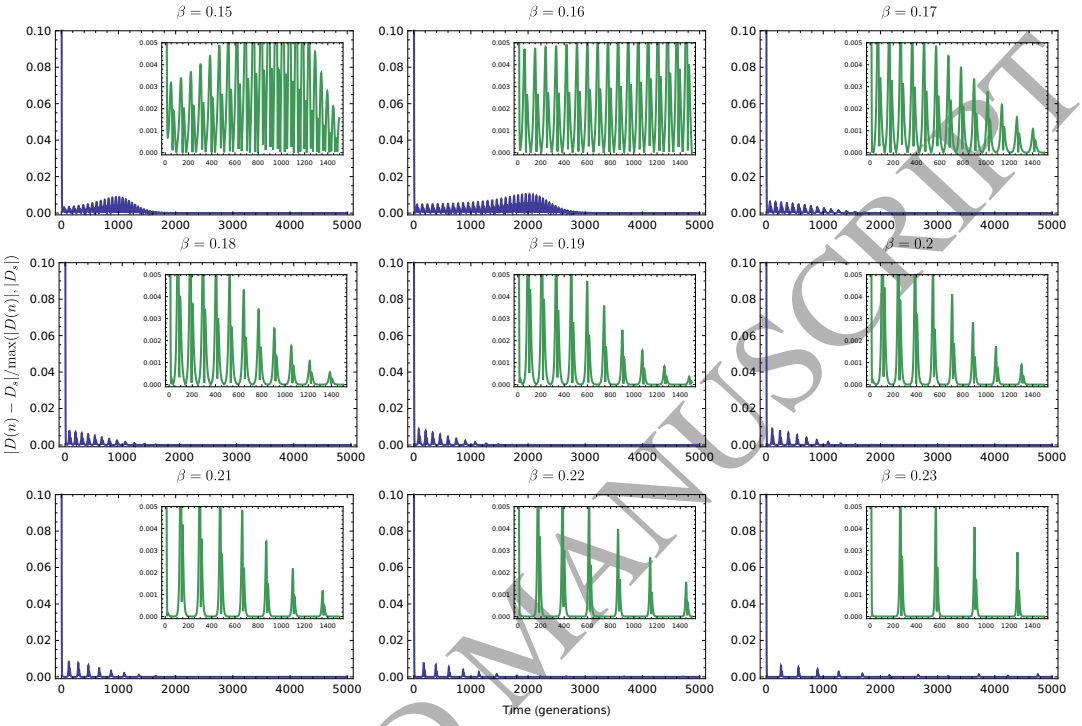


Figure 6: **Relative error of our approximate manifold D_S.** To justify the use of the manifold derived from the continuous-time system, D_S, we numerically compute the relative error between the manifold and the D component of an orbit of the discrete time system close to heteroclinic cycle. We compute both the manifold expression and the orbit at the generation times of the discrete-time model, n and plot the following error expressions $|D(n) - D_s| / \max(|D(n), |D_s|)$. Parameters were set to: $\gamma = 0.25$, $\delta = 0.3$, $A(0) = 0.9$, $B(0) = 0.9$, $D(0) = 0.05$ and the values of β are indicated in the plot titles. The The insets show the same curves but with finer grain x-axis and y-axis scales allowing the bursts to be seen in more detail. The magnitude of error is always very low.

3.6. Justifying the use of D_S derived from the continuous-time model

In Figure 6, we show the relative error between the value of $D(n)$, the linkage disequilibrium within the discrete-time model (1), and $D_S(t)$, the approximate slow manifold derived using the continuous-time approximation of the discrete-time system, finding the difference to be small. The error is computed using

$$E = \frac{|D(n) - D_S|}{\max\left(|D(n), |D_S|\right)}, \tag{25}$$

a modified form of the relative error between the approximate manifold D_S, and the D-component of a trajectory of the discrete-time system, which aims to avoid division by zero when one of the quantities is very small. The standard relative error expression could be problematic in this case, since the orbits are close to the manifold. We produce a time series of the distance between the D-component of the discrete-time orbit and the value of D_S evaluated at the values of the other variables along the orbit. This indicates that the continuous-time manifold, D_S, provides a good approximation for the discrete-time dynamics.

4. Discussion

We studied a genetic system with viability selection and gene conversion that encompass a wide range of variants where selection can be derived from different aspects of the recombinational process (Úbeda and Wilkins, 2011; Úbeda et al., 2019). We show that the selection regime associated with a fitness benefit derived from a sequence recognition (β), a fitness cost derived from a gene conversion (γ) altogether with the reshuffling of alleles in double heterozygotes induced by gene conversion and crossover (δ), can lead to stable cycling dynamics in the two-locus, two-alleles model. Our model is most similar to that of Haig and Grafen (1991), because in both models the often assumed symmetry of the fitness matrix (Karlin et al., 1970) is broken. The fluctuations that feature in the model are caused by selection for one allele burning out a target sequence followed by selection for an alternative allele that can burn out the sequence that replaced the old one. This pattern can repeat indefinitely and the resulting dynamics form a heteroclinic

cycle (Úbeda et al., 2019). To find out if sustained fluctuations are possible in either of our model variants we investigated whether the heteroclinic cycle attracts or repels (Hofbauer and Schlag, 2000).

We found that haplotype frequencies settle quickly on a state depending on the allele frequencies in the population, and the allele frequencies change on a slower time scale than the linkage disequilibrium (Kimura, 1965). After identifying the linkage disequilibrium D as a good candidate for the fast variable, we performed the nonlinear change of variables from haplotype to allele frequencies, which introduces $D(t)$ as an explicit variable. We then apply a quasi-steady state assumption to $D(t)$ and solve the resulting algebraic equation for D, which we use to reduce the dimension of our system by removing dependency on D altogether (Figure 5) (Kuehn, 2015). We find that the dynamics don't necessarily converge to a single stable interior (polymorphic) equilibrium. We thus provide a biological example of a doubly degenerate system that admits cycling.

After reducing the dimensionality, we found explicit conditions for stability of the heteroclinic cycles. Namely, the discrete-time model allows a heteroclinic cycle that is stable if $\beta < \gamma$; on the other hand, its continuous-time approximation has a heteroclinic cycle that is always unstable and the dynamics eventually settle on an equilibrium. Furthermore, we established numerically the basin of attraction for the heteroclinic cycle and studied the accuracy of the closed-form approximation D_S of the QLE manifold used to constrain the dynamics (Figure 6).

The equilibria of the discrete and continuous-time models are the same (Bürger, 2000). However, the stability of the heteroclinic cycle differs between the two models: the discrete-time model can have an attracting heteroclinic cycle and a stable equilibrium, and thus has a region of bistability in parameter space; however, its continuous-time approximation has, in the same region of parameter space, $\beta < \gamma$, a globally attracting interior equilibrium point. From a dynamical systems point of view this is not a surprise: it is well known that similar nonlinear discrete and continuous-time models can differ in various ways (May, 1976).

However, preliminary results show that if the population in the model is finite and

multinomial sampling is used to pick the individuals who mate and are replaced (Wright, 1969; Úbeda et al., 2019) — producing a stochastic and more biologically realistic version of our model — we see the gap between the discrete-time model and continuous-time approximation bridged. Indeed, similar oscillatory behaviour is now observed in both models. In fact, we see the two models behaving almost identically when the population is finite, just differing in time scale. We also observe that the deterministic slow manifold, D_{QLE}, is a good approximation for the dynamics of the stochastic model, as shown to be possible in some systems by (Constable and McKane, 2017). An in depth analysis of the stochastic model however, is beyond the scope of this paper. Further work could use D_{QLE} to simplify the dynamics of the stochastic implementation of the model. Globally attracting invariant QLE manifolds have recently been found to exist under certain parameter regimes in the continuous-time two locus-two allele selection-recombination equations by Baigent and Seymenoglu (2018).

Similar analyses using quasi-equilibria involving variables other than linkage disequilibrium have been conducted (Van Baalen and Rand, 1998; Day et al., 2011; Lion and Gandon, 2016; Lion, 2018). These models are evolutionary-ecological rather than population genetic models, and rely on the weak selection approximation, but they still observe a rapid convergence to quasi-linkage equilibrium. Our approach to studying the QLE manifold is very general, applicable to any system showing a significant separation of time-scales. Any genetic system of this sort converges to quasi-linkage equilibrium and therefore under an appropriate transformation of variables — one which isolates the fast subsystem — can be analysed in a similar fashion. Therefore, treating the QLE manifold as an slow manifold and using linkage disequilibrium as a coordinate to approximate this surface explicitly, is a powerful technique for other genetic systems and even evolutionary ecological models.

Multi-locus models can have complex dynamics (Hastings, 1981; Hofbauer and Iooss, 1984; Haig and Grafen, 1991; Úbeda et al., 2019). It appears that most analyses of multi-locus models have been carried out under weak selection assumptions, in which case the dynamics are relatively simple: stable cycling is generally not possible and the dynamics

go to an equilibrium (Nagylaki et al., 1999). The weak selection assumption allows for general analytic results (Akin, 1982; Hofbauer, 1985; Barton, 1995; Nagylaki et al., 1999; Kirkpatrick et al., 2002), often invoking the use of the QLE. Under weak selection, stable cycling and complex dynamics do not occur if the equilibria are not degenerate and therefore complex dynamics are not observed under QLE. This association of QLE with weak selection and stability might have led to the impression that complex dynamics are not compatible with quasi-linkage equilibrium (Pomiankowski and Bridle, 2004). What we have shown here is that complex dynamics are possible and, furthermore, are played out in a state of *quasi-linkage equilibrium* showing the association between QLE and convergence to equilibrium to not be true in general: it is possible to find continued fluctuations and sudden changes in the genetic make up in a population at quasi-linkage equilibrium.

References

Akin, E., 1979. The geometry of population genetics. lect. Notes Biomath 31.

Akin, E., 1982. Cycling in simple genetic systems. Journal of Mathematical Biology 13 (3), 305–324.

Akin, E., 1983. Hopf bifurcation in the two locus genetic model. Vol. 284. American Mathematical Soc.

Akin, E., 1987. Cycling in simple genetic systems: Ii. the symmetric cases. In: Dynamical Systems. Springer, pp. 139–153.

Asmussen, M. A., Feldman, M. W., 1977. Density dependent selection 1: A stable feasible equilibrium may not be attainable. Journal of Theoretical Biology 64 (4), 603–618.

Baigent, S., Seymenoglu, B., 2018. Competitive selection-recombination dynamics: A new approach to studying the quasilinkage equilibrium manifold.
URL `https://www.ucl.ac.uk/~zcahge7/files/SRpaper.pdf`

Barton, N. H., 1995. A general model for the evolution of recombination. Genetics Research 65 (2), 123–144.

Bürger, R., 2000. The mathematical theory of selection, recombination, and mutation. John Wiley & Sons.

Constable, G. W. A., McKane, A. J., 2017. Exploiting fast-variables to understand population dynamics and evolution. Journal of Statistical Physics, 1–41.

Cressman, R., 1988. Frequency-dependent viability selection (a single-locus, multiphenotype model). Journal of theoretical biology 130 (2), 147–165.

Day, T., Alizon, S., Mideo, N., 2011. Bridging scales in the evolution of infectious disease life histories: theory. Evolution: International Journal of Organic Evolution 65 (12), 3448–3461.

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

Haig, D., Grafen, A., 1991. Genetic scrambling as a defence against meiotic drive. Journal of theoretical Biology 153 (4), 531–558.

Hastings, A., 1981. Stable cycling in discrete-time genetic models. Proceedings of the National Academy of Sciences 78 (11), 7224–7225.

Hofbauer, J., 1985. The selection mutation equation. Journal of mathematical biology 23 (1), 41–53.

Hofbauer, J., Iooss, G., 1984. A Hopf bifurcation theorem for difference equations approximating a differential equation. Monatshefte für Mathematik 98 (2), 99–113.

Hofbauer, J., Schlag, K. H., 2000. Sophisticated imitation in cyclic games. Journal of Evolutionary Economics 10 (5), 523–543.

Hofbauer, J., Sigmund, K., 1998. Evolutionary games and population dynamics. Cambridge university press.

Jansen, V. A. A., Van Baalen, M., 2006. Altruism through beard chromodynamics. Nature 440 (7084), 663.

Karlin, S., Feldman, M. W., et al., 1970. Linkage and selection: two locus symmetric viability model. Theoretical population biology 1 (1), 39–71.

Kimura, M., 1958. On the change of population fitness by natural selection 2 3. Heredity 12 (2), 145.

Kimura, M., 1965. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics 52 (5), 875–890.

Kirkpatrick, M., Johnson, T., Barton, N., 2002. General models of multilocus evolution. Genetics 161 (4), 1727–1750.

Kuehn, C., 2015. Multiple time scale dynamics. Vol. 191. Springer.

Lewontin, R. C., Kojima, K.-i., 1960. The evolutionary dynamics of complex polymorphisms. Evolution 14 (4), 458–472.

Lion, S., 2018. From the Price equation to the selection gradient in class-structured populations: a quasi-equilibrium route. Journal of theoretical biology 447, 178–189.

Lion, S., Gandon, S., 2016. Spatial evolutionary epidemiology of spreading epidemics. Proc. R. Soc. B 283 (1841), 20161170.

May, R. M., 1976. Simple mathematical models with very complicated dynamics. Nature 261 (5560), 459.

Nagylaki, T., Hofbauer, J., Brunovskỳ, P., 1999. Convergence of multilocus systems under weak epistasis or weak selection. Journal of mathematical biology 38 (2), 103–133.

Pomiankowski, A., Bridle, J., 2004. Evolutionary genetics: no sex please we're at QLE (Quasi-Linkage Equilibrium).

Pontz, M., Hofbauer, J., Bürger, R., 2018. Evolutionary dynamics in the two-locus two-allele model with weak selection. Journal of mathematical biology 76 (1-2), 151–203.

Pugh, C. C., Shub, M., Hirsch, M. W., 1977. Invariant manifolds. Lecture Notes in Mathematics, Springer, New York.

Rice, S. H., 2004. Evolutionary theory: mathematical and conceptual foundations. Sinauer Associates Sunderland, MA.

Sasaki, A., Hamilton, W. D., Ubeda, F., 2002. Clone mixtures and a pacemaker: new facets of red-queen theory and ecology. Proceedings of the Royal Society of London. Series B: Biological Sciences 269 (1493), 761–772.

Ubeda, F., Haig, D., 2005. On the evolutionary stability of mendelian segregation. Genetics 170 (3), 1345–1357.

Úbeda, F., Russell, T. W., Jansen, V. A. A., 2019. PRDM9 and the evolution of recombination hotspots. Theoretical population biology 126, 19–32.

Úbeda, F., Wilkins, J. F., 2011. The Red Queen theory of recombination hotspots. Journal of evolutionary biology 24 (3), 541–553.

Van Baalen, M., Rand, D. A., 1998. The unit of selection in viscous populations and the evolution of altruism. Journal of theoretical biology 193 (4), 631–648.

Wright, S., 1969. Evolution and the genetics of populations: Vol. 2. The theory of gene frequencies.

Yahara, K., Fukuyo, M., Sasaki, A., Kobayashi, I., 2009. Evolutionary maintenance of selfish homing endonuclease genes in the absence of horizontal transfer. Proceedings of The National Academy of Sciences, pnas–0908404106.

Appendix A. Deriving the discrete-time model

Our model (Úbeda et al., 2019) can be written as a particular case of the model known as the selection-recombination equations presented in (Lewontin and Kojima, 1960; Nagylaki et al., 1999; Bürger, 2000; Ubeda and Haig, 2005) and many other papers (Nagylaki et al., 1999). In the general model, haplotype frequencies evolve according to

$$\bar{w}x_i'(n) = \sum_{j=1}^{m} w_{i,j}x_i x_j + \epsilon_i \delta \left(w_{1,4}x_1 x_4 - w_{2,3}x_2 x_3 \right), \tag{A.1}$$

where x_i denotes the frequency of haplotype i, m is the number of alleles and $n \in \mathbb{N}_+$ represents the discrete time step. The recombination terms $\delta \left(w_{1,4}x_1 x_4 - w_{2,3}x_2 x_3 \right)$ have different signs depending on the haplotype, provided by ϵ_i for haplotype i. Specifically, for a two-locus two-allele implementation of the model, e_i is defined as

$$\epsilon_i = \begin{cases} -1 & \text{for} \quad i = 1, 4 \\ 1 & \text{for} \quad i = 2, 3. \end{cases} \tag{A.2}$$

The marginal mean fitness of a haplotype whose frequency is x_i is given by

$$w_i = \sum_{j=1}^{n} w_{i,j}x_j, \tag{A.3}$$

and the mean fitness of the population is given by

$$\bar{w} = \sum_{j=1}^{n} w_j x_j. \tag{A.4}$$

Due to the normalisation of the right hand side of the governing equations of the model by the mean fitness of the population, the sum of the haplotype frequencies is always one. This means the state space of the model is the simplex of dimension $n^m - 1$, where n is the number of alleles and m is the number of loci.

Fitnesses for the two-locus two-allele version of our model are derived by computing all of the frequencies of offspring given by each possible mating combination. Due to the symmetries on the allele types determining when recombination occurs, the linkage disequilibrium D is the same for each haplotype and therefore can be taken out of the fitness matrix. This is clearly true in the more general versions of the model, meaning the linkage terms are separate in the statement of the general model equations (A.1).

After this, and other simplifications which are possible due to symmetries in the gene conversion process and the viability benefits derived from crossover, we arrive at the following fitness matrix for the two allele two loci version of the model

$$W = \begin{pmatrix} 1+\beta & 1-\gamma & 1 & 1 \\ 1+\gamma & 1-\beta & 1 & 1 \\ 1 & 1 & 1-\beta & 1+\gamma \\ 1 & 1 & 1-\gamma & 1+\beta \end{pmatrix}. \tag{A.5}$$

Applying our specific fitness matrix to the general model given gives the following system of equations

$$\begin{aligned} \bar{w}x_1(n+1) &= (1+\beta)x_1^2 + (1-\gamma)x_1x_2 + x_1x_3 + x_1x_4 - \delta D, \\ \bar{w}x_2(n+1) &= (1-\beta)x_2^2 + (1+\gamma)x_2x_1 + x_2x_3 + x_2x_4 + \delta D, \\ \bar{w}x_3(n+1) &= (1-\beta)x_3^2 + (1+\gamma)x_3x_4 + x_3x_1 + x_3x_2 + \delta D, \\ \bar{w}x_4(n+1) &= (1+\beta)x_4^2 + (1-\gamma)x_4x_3 + x_4x_1 + x_4x_2 - \delta D. \end{aligned} \tag{A.6}$$

Expanding the brackets in system (A.6) and applying the conservation law for the total population, $\sum_{i=1}^{4} x_i = 1$, we can simply the system to

$$\begin{aligned} \bar{w}x_1(n+1) &= x_1(n)[1 + \beta x_1(n) - \gamma x_2(n)] - \delta D, \\ \bar{w}x_2(n+1) &= x_2(n)[1 - \beta x_2(n) + \gamma x_1(n)] + \delta D, \\ \bar{w}x_3(n+1) &= x_3(n)[1 - \beta x_3(n) + \gamma x_4(n)] + \delta D, \\ \bar{w}x_4(n+1) &= x_4(n)[1 + \beta x_4(n) - \gamma x_3(n)] - \delta D, \end{aligned} \tag{A.7}$$

where $\bar{w}\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x})$ and $n \in \mathbb{N}_+$ and the population mean fitness is

$$\bar{w} = \sum_{i=1}^{4} f_i(\mathbf{x}) = x_1 + x_2 + x_3 + x_4 + \beta(x_1^2 + x_4^2 - x_2^2 - x_3^2). \tag{A.8}$$

Appendix B. Isolation of the multiple time-scales

The region of parameter space for which the following arguments hold is where the heteroclinic cycle exists and is attracting in the discrete-time model, i.e. $\beta < \gamma$.

Appendix B.1. Time-scale separation nearby the interior equilibrium

We find three distinct time-scales in the dynamics of the linearised system nearby the interior equilibrium. Recall that the eigenvalues of the interior equilibrium of the continuous-time model are given by

$$\begin{aligned}\lambda_1 &= \frac{\gamma D^* + \sqrt{(\gamma D^*)^2 + \frac{1}{4}\beta(\beta-\gamma)}}{\bar{w}^*},\\ \lambda_2 &= \frac{\gamma D^* - \sqrt{(\gamma D^*)^2 + \frac{1}{4}\beta(\beta-\gamma)}}{\bar{w}^*},\\ \lambda_3 &= -\frac{\delta + 2D^*(\beta-\gamma)}{\bar{w}^*},\end{aligned} \tag{B.1}$$

where $\bar{w}^* = 1 + 2\beta D^*$. If $\beta > \gamma$ then $D^* > 0$. The interior equilibrium in that case is a saddle. If $\beta < \gamma$ then $D^* < 0$. Eigenvalues λ_1 and λ_2 then are complex with negative real parts and the interior equilibrium is always locally stable.

We introduce the parameter

$$\epsilon = \sqrt{\gamma - \beta}, \tag{B.2}$$

which is small near the boundary of the region of parameter space in which we observe time-scale separation, $\beta < \gamma$. We substitute this definition into the equations and compute the eigenvalues at the interior equilibrium (9). For $0 < \epsilon \ll 1$, the eigenvalues satisfy the identities

$$\begin{aligned}\bar{w}^*\lambda_1 &= -\epsilon^2\frac{\gamma}{8\delta} + i\epsilon\frac{\sqrt{\gamma}}{2} + O(\epsilon^3),\\ \bar{w}^*\lambda_2 &= -\epsilon^2\frac{\gamma}{8\delta} - i\epsilon\frac{\sqrt{\gamma}}{2} + O(\epsilon^3),\\ \bar{w}^*\lambda_3 &= -\delta + O(\epsilon^3).\end{aligned} \tag{B.3}$$

The dynamics of the system linearised around the interior equilibrium (9) operate on three distinct time-scales: $\bar{w}\delta^{-1}$, $2\bar{w}\epsilon^{-1}\gamma^{-\frac{1}{2}}$ and $8\delta\bar{w}\epsilon^{-2}\gamma^{-1}$. If $0 < \epsilon\sqrt{\gamma} \ll 2\delta < 1$ the time scales separate as $\delta^{-1} \ll 2\epsilon^{-1}\gamma^{-\frac{1}{2}} \ll 2\delta\left(2\epsilon^{-1}\gamma^{-\frac{1}{2}}\right)^2$. The second and third time-scales are associated with the motion within the QLE manifold, while the first relates to the approach towards the QLE manifold. Under this condition, the approach is very fast compared to the dynamics on the manifold, which justifies making a quasi-steady state

assumption. This behaviour can be observed in Figure 3 where the approach to QLE is very fast with associated time-scale $\bar{w}\delta^{-1}$, and much faster than the cyclic behaviour on the manifold, which acts on time-scale $2\bar{w}\epsilon^{-1}\gamma^{-\frac{1}{2}}$, which in turn is faster than the approach to equilibrium which acts on time-scale $8\delta\bar{w}\epsilon^{-2}\gamma^{-1}$.

Note that the separation of time-scales is a direct consequence of the double degeneracy of the interior equilibrium (9). Specifically, when $\beta = \gamma$, and hence $\epsilon = 0$, two eigenvalues are zero. If the third eigenvalue is much smaller than zero, for small ϵ and continuous dependence of the eigenvalues on ϵ, the separation of time scales follows. This implies that the existence of a two-dimensional slow manifold is a generic result in the proximity of a double degeneracy and independent of the details of the model.

Appendix B.2. Time-scale separation in the full system

We introduce the new variables

$$\begin{aligned} X &= \sqrt{\gamma-\beta}\ln\left(\frac{A}{1-A}\right) + \sqrt{\beta}\ln\left(\frac{B}{1-B}\right), \\ Y &= (\gamma-\beta)\ln\big(A(1-A)\big) + \beta\ln\big(B(1-B)\big), \\ Z &= \frac{D}{\gamma-\beta}. \end{aligned} \tag{B.4}$$

If $\gamma \neq \beta$, these definitions implicitly define A and B locally as functions of X and Y and therefore the inverse transformation exists.

Rewriting the continuous-time model (15) in the new variables (B.4),

$$\begin{aligned} \frac{\mathrm{d}X}{\mathrm{d}t} &= \frac{\sqrt{\beta(\gamma-\beta)}}{\bar{w}}\left(\sqrt{\beta}(2B-1) + \sqrt{\gamma-\beta}(2A-1) + \frac{\gamma\sqrt{\gamma-\beta}(2B-1)Z}{B(1-B)}\right), \\ \frac{\mathrm{d}Y}{\mathrm{d}t} &= -\frac{\beta(\gamma-\beta)}{\bar{w}}\gamma\frac{(1-2B)^2}{B(1-B)}Z, \\ \frac{\mathrm{d}Z}{\mathrm{d}t} &= \frac{(\gamma-\beta)^{-1}}{\bar{w}}\Big[(\gamma-\beta)\Big[(\gamma-\beta)^2Z^2 - AB(1-A)(1-B)\Big] \\ &\qquad - (\gamma-\beta)Z(\beta(2A-1)(2B-1)+\delta)\Big]. \end{aligned} \tag{B.5}$$

Using (B.2), this can be written as

$$\begin{aligned}\frac{1}{\epsilon}\frac{\mathrm{d}X}{\mathrm{d}t} &= \frac{\sqrt{\beta}}{\bar{w}}\left(\sqrt{\beta}(2B-1)+\epsilon(2A-1)+\frac{\gamma\epsilon(2B-1)}{B(1-B)}Z\right),\\ \frac{1}{\epsilon^2}\frac{\mathrm{d}Y}{\mathrm{d}t} &= -\frac{\beta\gamma}{\bar{w}}\frac{(1-2B)^2}{B(1-B)}Z,\\ \frac{\mathrm{d}Z}{\mathrm{d}t} &= \frac{1}{\bar{w}}\Big[\epsilon^4 Z^2 - AB(1-A)(1-B) - Z\left(\beta(2A-1)(2B-1)+\delta\right)\Big].\end{aligned} \tag{B.6}$$

When ϵ is small, the form of (B.6) isolates three distinct time-scales. The variable Z is changing at the fastest time-scale, and for Z small the variables X and Y (and A and B) are effectively constant. If A and B are constant, the variable Z has an equilibrium at

$$\begin{aligned}Z^* =& \epsilon^{-2}\frac{\beta(2A-1)(2B-1)+\delta}{2\epsilon^2}\\ &- \epsilon^{-2}\sqrt{\left(\frac{\beta(2A-1)(2B-1)+\delta}{2\epsilon^2}\right)^2 + AB(1-A)(1-B)}.\end{aligned} \tag{B.7}$$

The linearised dynamics around Z^* are given by

$$\frac{\mathrm{d}(Z-Z^*)}{\mathrm{d}t} = -\frac{1}{\bar{w}}(Z-Z^*)\sqrt{\left(\beta(2A-1)(2B-1)+\delta\right)^2 + 4\epsilon^4 AB(1-A)(1-B)} \tag{B.8}$$

which always converges to the equilibrium $Z = Z^*$. Based on this we choose $D_{QLE} = \epsilon^2 Z^*$. If D_{QLE} is situated outside the simplex this argument is not relevant but a similar argument can be applied for attraction to the state $Z = \epsilon^{-2} D_S$.

Appendix C. Determining the eigenvalues of the corner equilibria

In the vicinity of the origin (Φ_4), we find by Taylor expanding to second order that the QLE manifold is approximately defined by $D_{QLE}(0,0) \approx -\frac{\gamma-\beta}{\beta+\delta}AB$. The attracting manifold $D = D_S(A,B)$ in the vicinity of the origin is approximately

$$D_S(A,B) \approx \begin{cases} -AB & \text{if} \quad \delta \leq \gamma - 2\beta, \\ D_{QLE}(A,B) & \text{if} \quad \delta > \gamma - 2\beta. \end{cases} \tag{C.1}$$

We then find for the eigenvalues

Eigenvalue	$1-\frac{\beta}{1+\beta}$	$1+\frac{\gamma-\beta}{1+\beta}$	$1-\frac{\beta}{1+\beta}$	$1+\frac{\gamma-\beta}{1+\beta}$
Type	c_j	e_j	c_j	e_j
Condition	$\delta \leq \gamma - 2\beta$		$\delta > \gamma - 2\beta$	

Table C.3: The eigenvalues of the equilibria Φ_1 and Φ_4. The eigenvalues do not depend on the condition.

595 Likewise, in the vicinity of the equilibrium Φ_2 and Φ_3 the QLE manifold is approxi-
596 mately

$$D_{QLE}(A,B) \approx \begin{cases} -\frac{\beta-\delta}{\gamma-\beta} + \frac{2\beta}{\gamma-\beta}A + \frac{2\beta}{\gamma-\beta}(1-B) + (\frac{\gamma-\beta}{\beta-\delta} + \frac{4\beta}{\gamma-\beta})A(B-1) & \text{if } \delta < \beta, \\ -\frac{\gamma-\beta}{\delta-\beta}A(1-B) & \text{if } \delta > \beta, \end{cases} \tag{C.2}$$

597 and hence

$$D_S(A,B) \approx \begin{cases} \max(-AB, -(1-A)(1-B)) & \text{if } \delta \leq \beta, \\ D_{QLE}(A,B) & \text{if } \delta > \beta, \end{cases} \tag{C.3}$$

We then find for the eigenvalues

Eigenvalue	$1-\frac{\gamma-\beta}{1-\beta}$	$1+\frac{\beta}{1-\beta}$	$1-\frac{\gamma-\beta}{1-\beta}$	$1+\frac{\beta}{1-\beta}$
Type	c_j	e_j	c_j	e_j
Condition	$\delta \leq \beta$		$\delta > \beta$	

Table C.4: The eigenvalues of the equilibria Φ_2 and Φ_3. The eigenvalues do not depend on the condition

598