Strategy and sub-strategy dynamics in an evolutionary game

Hinke M. Osinga

Department of Mathematics, The University of Auckland
Private Bag 92019, Auckland 1142, New Zealand (H.M.Osinga@auckland.ac.nz)

James A. R. Marshall*

Department of Computer Science, University of Sheffield
Sheffield S1 4DP, United Kingdom (James.Marshall@sheffield.ac.uk)

Abstract

It has long been known in the field of population genetics that the equilibria attained by selection on a trait are not independent of the genetic bases of that trait. Whether one chooses to model selection acting on a single locus or multiple loci does matter. In evolutionary game theory, analysis of a simple and general game involving distinct roles for the two players has shown how correlated asymmetries can arise, in which one action is favoured in one role, but a different action in the other. The results of analysis of this game at the level of the entire strategy, or at the level of independent sub-strategies for the different roles, are in agreement, however. This is curious given the aforementioned population genetical results on the importance of the genetic bases of traits. Here we present a novel dynamical systems analysis of our game with roles, and show that while the stable equilibria in this game are unchanged according to whether one models selection on entire strategies, or independent selection for 'genes' for different components of the strategies, equilibrium selection, however, may differ under the two modelling approaches. Our results are an evolutionary game theory counterpart to existing results from the population genetics literature.

Keywords: adaptive topographies, phenotypic gambit, evolutionary game theory, evolutionary stability, correlated asymmetries, linkage disequilibrium, multilocus population genetics, dynamical systems theory

*Corresponding author, tel.: +44 114 222 1832, fax.: +44 114 222 1810

1 Introduction

Fitness often depends on complicated phenotypes, which are presumed to depend on multiple genetic loci. This raises an interesting modelling dilemma. At one extreme, one may model selection acting on phenotypes as if they were under simple genetic control at a single haploid locus; this is the 'phenotypic gambit' (Grafen 1984) widely used in evolutionary modelling, and referred to as evolutionary game theory when applied to model social interactions (Maynard Smith 1982). If multiple loci do underly a phenotype then such models should capture inter-locus fitness interactions, yet they can be of much greater complexity, having to account for a number of phenotypes that may be exponential in the number of loci involved. At the other extreme, a very simple model may be formulated that considers selection acting independently on frequencies of different alleles at different loci. Such a model would be more tractable, but neglects important quantities such as linkage disequilibrium between loci. Hence, it may give incorrect predictions. An intermediate solution is also possible, through the adoption of multilocus population genetics (Kimura 1965, Kirkpatrick., Johnson and Barton 2002, Gardner, West and Barton 2007).

In this paper, we examine the consequences of the two extreme approaches to modelling a simple, general and classical problem; interactions in a social game where the players are assigned distinct roles (Maynard Smith and Parker 1976). Such interactions occur in many contexts, such as interactions where one individual possesses a territory and the other does not (Maynard Smith and Parker 1976), or between adult reproductives and helpers (Queller 1996), or between parents and offspring (Trivers 1974). Even where payoffs are the same from both individuals' perspectives, 'uncorrelated asymmetries' can lead to different behaviours being stable in the distinct roles, and these have previously been analysed in terms of evolutionary stable strategies at the genotype level (Maynard Smith and Parker 1976, Maynard Smith 1982). Recently, a new analysis of a social game with roles played between relatives has taken the independent gene-level view, and has shown that this gives the same attracting equilibria as the genotype-level view (Marshall 2009). This is intriguing on several fronts. First, modelling selection at the genotype-level is akin to modelling selection acting on a larger number of genes competing for a single locus. Yet, it is known that 'adaptive topographies' that take no account of the underlying genetic-basis of fitness do not exist; changing the number of genes involved in representing a behaviour can lead to different evolutionary outcomes (Moran 1964). Second, the dimensions of the phase spaces of the two dynamical systems describing these different modelling levels are different, which means that one should not expect their behaviour to be the same. We show in this paper that a projection of the higher-dimensional system onto the phase space of the other still does not lead to a topologically equivalent system. In particular, we show that they do not have equal numbers of equilibria, but for both models there are always at most two stable coexisting equilibria, and the same stable equilibria exist in both models. Yet, despite this, seemingly equivalent initial conditions in the two systems can lead to the selection of different equilibria, demonstrating the fallacy

Table 1: Payoffs for the non-additive donation game

	C	**D**
C	$b-c+d$	$-c$
D	b	0

of the 'adaptive topography' concept for this particular system. Thus, while assumptions can be made that simplify the modelling while apparently leaving the equilibrium behaviour of the system unaffected, the equilibrium selection behaviour itself can be different. Our approach thus complements that of Moran (1964) who showed how, for sexual models, population equilibria themselves can be changed according to the genetic details underlying fitness. Here, we show how for the asexual models of evolutionary game theory, different strategy representations leave equilibria in a simple social interaction unchanged, but do change the equilibrium selection process.

2 Donation games with roles played between relatives

We consider the donation game with potentially non-additive payoffs as presented in Table 1. Interactions are structured such that there is an 'uncorrelated asymmetry' (Maynard Smith 1982); that is, players occupy distinct behavioural roles, and have different strategies according to the role they occupy. Interactions are further structured such that they occur between genetic relatives (Marshall 2009). There are two different ways of modelling such a game as a dynamical system. On the one hand, the dynamical system can describe the evolution of the frequencies of all possible genotypes. The set of all possible genotypes for the donation game, denoted $\mathcal{G}$, contains four elements, namely, $\mathcal{G} = \{\mathrm{CC},\ \mathrm{CD},\ \mathrm{DC},\ \mathrm{DD}\}$. The genotype dynamics is modeled by the rates of change of the frequencies $f_\bullet$ of these four genotypes, with $\bullet \in \mathcal{G}$. The equations are of the form

$$\dot{f}_\bullet = f_\bullet\,(w_\bullet - \overline{w})\,, \quad \bullet \in \mathcal{G}. \tag{1}$$

Here, $w_\bullet$ is the inclusive fitness of a given genotype and $\overline{w}$ is the mean fitness defined as

$$\overline{w} = f_{\mathrm{CC}}\,w_{\mathrm{CC}} + f_{\mathrm{CD}}\,w_{\mathrm{CD}} + f_{\mathrm{DC}}\,w_{\mathrm{DC}} + f_{\mathrm{DD}}\,w_{\mathrm{DD}}.$$

The inclusive fitnesses of the different genotypes are given in the equivalent neighbour-modulated fitness form (Taylor and Frank 1996) by

$$\begin{aligned}
w_{\mathrm{CC}} &= r\,(b+d) - c + (1-r)\,[f_{\mathrm{CC}} + \tfrac{1}{2}\,f_{\mathrm{CD}} + \tfrac{1}{2}\,f_{\mathrm{DC}}]\,(b+d), && (2)\\
w_{\mathrm{CD}} &= \tfrac{1}{2}\,(r\,b - c + (1-r)\,[f_{\mathrm{CC}}\,(2b+d) + f_{\mathrm{CD}}\,b + f_{\mathrm{DC}}\,(b+d)])\,, && (3)\\
w_{\mathrm{DC}} &= \tfrac{1}{2}\,(r\,b - c + (1-r)\,[f_{\mathrm{CC}}\,(2b+d) + f_{\mathrm{CD}}\,(b+d) + f_{\mathrm{DC}}\,b])\,, && (4)\\
w_{\mathrm{DD}} &= (1-r)\,[f_{\mathrm{CC}} + \tfrac{1}{2}\,f_{\mathrm{CD}} + \tfrac{1}{2}\,f_{\mathrm{DC}}]\,b. && (5)
\end{aligned}$$

Here, $0 \leq r \leq 1$ is the degree of relatedness within the population, giving the probability that interacting individuals have identical genotypes over-and-above that given by the population frequencies of these genotypes (Grafen 1985). In the formulation above, we used the fact that the sum of the frequencies is 1, that is, $f_{\mathrm{DD}} = 1 - (f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}})$.

As discussed by Marshall (2009), an alternative model is based on gene dynamics. Here, one describes the rate of change in the frequency of each allele for each role. There are two roles (players) in the donation game and two different alleles, namely, cooperative C and defective D. Since the frequencies in both roles again add up to 1, we only consider the frequencies ϕ_{C1} and ϕ_{C2} of cooperative alleles occuring in each of the two roles; the frequencies of defective alleles, and corresponding equations, follow from the equalities $\phi_{\mathrm{D1}} = 1 - \phi_{\mathrm{C1}}$ and $\phi_{\mathrm{D2}} = 1 - \phi_{\mathrm{C2}}$. The replicator dynamics is then defined as

$$\dot{\phi}_{\mathrm{Ci}} = \phi_{\mathrm{Ci}} \, (1 - \phi_{\mathrm{Ci}}) \, (\omega_{\mathrm{Ci}} - \omega_{\mathrm{Di}}) \,, \tag{6}$$

where $i, j \in \{1, 2\}$ and $i \neq j$. The inclusive fitnesses of cooperative and defective alleles for each case are now given by

$$\begin{aligned}
\omega_{\mathrm{C1}} &= -c + (b + d)\, \phi_{\mathrm{C2}} + r[b + (-c + d)\, \phi_{\mathrm{C2}}], && (7)\\
\omega_{\mathrm{D1}} &= b\, \phi_{\mathrm{C2}} - r\, c\, \phi_{\mathrm{C2}}, && (8)\\
\omega_{\mathrm{C2}} &= -c + (b + d)\, \phi_{\mathrm{C1}} + r[b + (-c + d)\, \phi_{\mathrm{C1}}], && (9)\\
\omega_{\mathrm{D2}} &= b\, \phi_{\mathrm{C1}} - r\, c\, \phi_{\mathrm{C1}}. && (10)
\end{aligned}$$

Both formulations describe the evolution of four different frequencies, but the dynamical systems are not the same. In particular, note that the state of system (1) is determined by three of the four frequencies, since $f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}} + f_{\mathrm{DD}} = 1$; this means that the phase space of system (1) has dimension three. The state of system (6), however, is already determined by two of the four frequencies, since $\phi_{\mathrm{C1}} + \phi_{\mathrm{D1}} = 1$ and $\phi_{\mathrm{C2}} + \phi_{\mathrm{D2}} = 1$, and the dimension of its phase space is only two. Due to this difference in dimensions, the two systems cannot be topologically equivalent (Guckenheimer and Holmes 1983) and one shoud expect that the behaviour of the two systems is not the same. One may be tempted to believe that the higher-dimensional system (1) implies the behaviour of system (6), because ϕ_{C1} and ϕ_{C2} should evolve in the same way as $f_{\mathrm{CC}} + f_{\mathrm{CD}}$ and $f_{\mathrm{CC}} + f_{\mathrm{DC}}$, respectively. However, it is not hard to show that also in this sense the two systems are not topologically equivalent. While the proof is straightforward, it is rather tedious and not very insightful. Therefore, we provide this proof in A.

Despite this lack of topological equivalence between systems (1) and (6), it might be assumed that the two systems have the same number of stable equilibria and predictions of the long-term behaviour made using either model give the same results. In this paper, we explain in detail that systems (1) and (6) can, in fact, give conflicting predictions for the long-term behaviour. We discuss the equilibria and stability properties for system (1) and compare predictions from system (1) with predictions from system (6) by defining $\phi_{\mathrm{C1}} = f_{\mathrm{CC}} + f_{\mathrm{CD}}$ and $\phi_{\mathrm{C2}} = f_{\mathrm{CC}} + f_{\mathrm{DC}}$. In the following section, we first consider system (6).

3 Analysis of equilibrium states for the gene dynamics model

A detailed analysis of the equilibria for system (6) in their most general form has already been provided by Marshall (2009). We present here the analysis as is standard in dynamical systems theory (Guckenheimer and Holmes 1983, Kuznetsov 1995) and determine stability properties using the Jacobian matrix; this same approach will also be used in Sec. 4 for the analysis of system (1). The two-dimensional system (6) can be rewritten explicitly in terms of the two variables $\phi_{\rm C1}$ and $\phi_{\rm C2}$ and the parameters b, c, d and r as

$$\left\{ \begin{array}{rcl} \dot{\phi}_{\rm C1} & = & \phi_{\rm C1}\,(1-\phi_{\rm C1})\,(r\,b-c+[1+r]\,d\,\phi_{\rm C2})\,, \\ \dot{\phi}_{\rm C2} & = & \phi_{\rm C2}\,(1-\phi_{\rm C2})\,(r\,b-c+[1+r]\,d\,\phi_{\rm C1})\,. \end{array} \right.$$

Recall that the dynamics for $\phi_{\rm D1}$ and $\phi_{\rm D2}$ can readily be deduced from the relationships $\phi_{\rm D1}=1-\phi_{\rm C1}$ and $\phi_{\rm D2}=1-\phi_{\rm C2}$. We focus here on the cases $d>0$ and $d<0$, where we assume $b,c>0$ and $0<r<1$.

Equilibria are found as solutions that simultaneously satisfy $\dot{\phi}_{\rm C1}=0$ and $\dot{\phi}_{\rm C2}=0$. The equality $\dot{\phi}_{\rm C1}=0$ holds if

$$\phi_{\rm C1}=0, \quad \phi_{\rm C1}=1 \quad \text{or} \quad \phi_{\rm C2}=\frac{c-r\,b}{(1+r)\,d}:=e^*.$$

Similarly, the equation $\dot{\phi}_{\rm C2}=0$ is satisfied if $\phi_{\rm C2}=0$, $\phi_{\rm C2}=1$ or $\phi_{\rm C1}=e^*$. Hence, system (6) always has the four equilibria $(\phi_{\rm C1},\phi_{\rm C2})=(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$, and there exists a fifth equilibrium

$$(\phi_{\rm C1},\phi_{\rm C2})=(e^*,e^*),$$

provided $0<e^*<1$, that is,

$$0<\frac{c-r\,b}{(1+r)\,d}<1 \Leftrightarrow \left\{ \begin{array}{lll} d>0 & \text{and} & \dfrac{c-d}{b+d} \;<\; r \;<\; \dfrac{c}{b}, \\ d<0 & \text{and} & \dfrac{c}{b} \;<\; r \;<\; \dfrac{c-d}{b+d}, \end{array} \right.$$

where we assume $b>c$ and $b+d>0$. Note that these assumptions are satisfied for the standard altruistic donation game, where $b>c>0$, regardless of the value of d.

The stability of these equilibria is determined by the eigenvalues of the Jacobian matrix, obtained from linearizing system (6) about the respective equilibria. Let us define

$$e(\phi):=(r\,b-c+[1+r]\,d\,\phi)\,.$$

The Jacobian matrix at an equilibrium $(\phi_{\rm C1},\phi_{\rm C2})$ is then defined as

$$\mathrm{Jac}(\phi_{\rm C1},\phi_{\rm C2})=\left(\begin{array}{cc} (1-2\,\phi_{\rm C1})\,e(\phi_{\rm C2}) & \phi_{\rm C1}\,(1-\phi_{\rm C1})\,e'(\phi_{\rm C2}) \\ \phi_{\rm C2}\,(1-\phi_{\rm C2})\,e'(\phi_{\rm C1}) & (1-2\,\phi_{\rm C2})\,e(\phi_{\rm C1}) \end{array} \right).$$

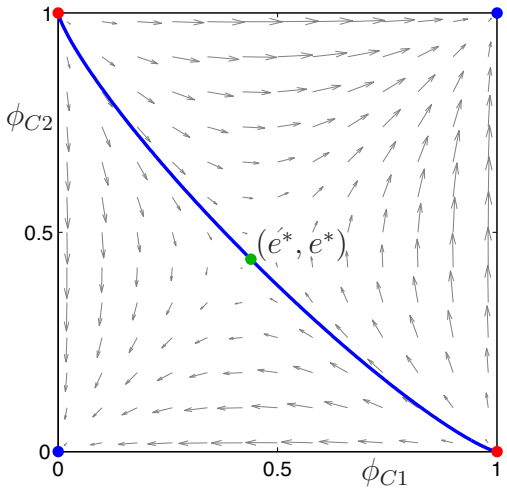

Figure 1: *Phase portrait illustrating bistability of the equilibria* $(0,0)$ *and* $(1,1)$ *for the gene system* (6) *with parameters* $b = 2$, $c = 0.5$, $d = 0.25 > 0$, *and* $r = 0.185$. *The (solid blue) curve through the saddle equilibrium* $(e^*, e^*) \approx (0.4388, 0.4388)$ *is the stable manifold of* (e^*, e^*) *that separates the two basins of attraction for* $(0,0)$ *and* $(1,1)$.

Hence, the Jacobian matrices for $(\phi_{\rm C1}, \phi_{\rm C2}) = (0,0)$ and $(\phi_{\rm C1}, \phi_{\rm C2}) = (1,1)$ are diagonal matrices with double eigenvalues $e(0) = r\,b - c$ and $-e(1) = c - d - r\,(b+d)$, respectively. Therefore, the origin is stable if and only if $e(0) < 0 \Leftrightarrow r < c/b$. If we again assume $b + d > 0$, then $(1,1)$ is stable if and only if $-e(1) < 0 \Leftrightarrow r > (c-d)/(b+d)$. We conclude that $(0,0)$ and $(1,1)$ are both stable precisely when (e^*, e^*) exists. This equilibrium $(\phi_{\rm C1}, \phi_{\rm C2}) = (e^*, e^*)$, has the anti-diagonal Jacobian matrix

$$\mathrm{Jac}(1,0) = \begin{pmatrix} 0 & e^*\,[1-e^*]\,(1+r)\,d \\ e^*\,[1-e^*]\,(1+r)\,d & 0 \end{pmatrix},$$

with eigenvalues $\pm e^*\,[1-e^*]\,(1+r)\,d$. Hence, (e^*, e^*) is always a saddle equilibrium. Finally, the Jacobian matrices for $(\phi_{\rm C1}, \phi_{\rm C2}) = (1,0)$ and $(\phi_{\rm C1}, \phi_{\rm C2}) = (0,1)$ are diagonal matrices with both the same eigenvalues, namely, $-e(0) = c - r\,b$ and $e(1) = r\,(b+d) - (c-d)$. Therefore, $(1,0)$ and $(0,1)$ are sources in the parameter regime where $(0,0)$ and $(1,1)$ are both stable. Otherwise, they are typically saddles, because stability of $(1,0)$ and $(0,1)$ requires $(c-d)/(b+d) > c/b$ and this can only hold if $d < 0$; see also (Marshall 2009).

To illustrate the behaviour of the gene dynamics model (1), let us consider an example of a situation where the equilibrium (e^*, e^*) exists; as parameters, we choose $b = 2$, $c = 0.5$, $d = 0.25 > 0$, and $r = 0.185$. Figure 1 shows the phase portrait for this case in the (ϕ_{C1}, ϕ_{C2})-plane. The (gray) arrows indicate the direction of the flow and clearly show a situation of bistability, with both $(0,0)$ and $(1,1)$ (blue dots) attracting nearby

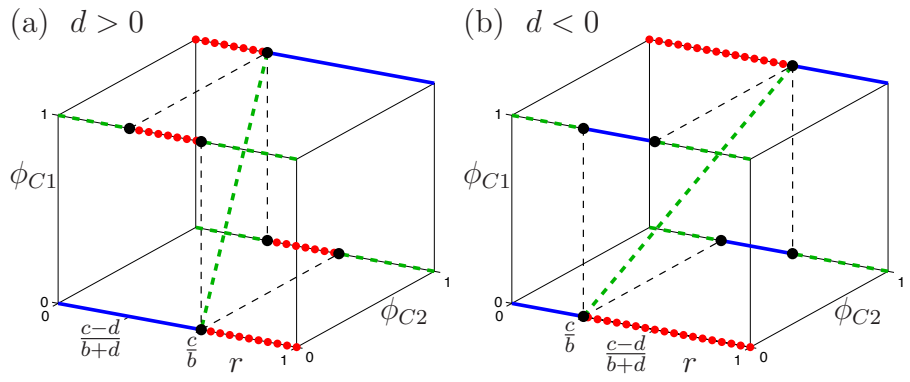

Figure 2: *Bifurcation diagrams with* $0 < r < 1$ *of the gene model with* $d > 0$ *(a) and* $d < 0$ *(b); here, we asssume* $c > b > 0$ *are such that* $0 < (c-d)/(b+d) < 1$ *(which is automatically satisfied if* $d > 0$*, but not if* $d < 0$*). The stability of the equilibria is indicated by solid (blue), dashed (green) and dotted (red) lines for attractors, saddles and repellors, respectively.*

points. Note that $(1,0)$ and $(0,1)$ (red dots) are both sources, because nearby points all move away from these two equilibria. The basins of the two attracting equilibria are separated by two trajectories of points that flow from the respective two sources to the saddle equilibrium $(e^*, e^*) \approx (0.4388, 0.4388)$; all other points near (e^*, e^*) flow away from (e^*, e^*). These two special trajectories form the stable manifold of (e^*, e^*) that acts as a separatrix for the two attractors at $(0,0)$ and $(1,1)$.

Behaviour of the gene dynamics model as r varies from 0 to 1

We are primarily interested in how the stability of the equilibrium states change as the degree r of relatedness varies between 0 and 1. We again refer to the results in (Marshall 2009) for comparison. The two cases $d > 0$ and $d < 0$ are different and we first consider the case $d > 0$. If $d > 0$ then $(c-d)/(b+d) < c/b$. The existence and stability properties of the equilibria are illustrated in Fig. 2(a). For $r < (c-d)/(b+d)$, only the origin is attracting (solid blue line), $(1,1)$ is a repellor (dotted red line), $(1,0)$ and $(0,1)$ are saddles (dashed green lines), and (e^*, e^*) does not exist. When $r = (c-d)/(b+d)$ a bifurcation occurs and the equilibrium (e^*, e^*) emerges from the $(1,1)$-branch; this bifurcation is a transcritical bifurcation, but it is degenerate due to the symmetries of the model and not only the stability of $(1,1)$, but also of $(1,0)$ and $(0,1)$ changes. For $(c-d)/(b+d) < r < c/b$, both the origin and $(1,1)$ are attractors (solid blue lines), $(1,0)$ and $(0,1)$ are repellors (dotted red lines), and (e^*, e^*) is a saddle (dashed green line). At $r = c/b$, another transcritical bifurcation occurs, which is similarly degenerate, where (e^*, e^*) merges with the origin and again $(1,0)$ and $(0,1)$ change stability as well. For $r > c/b$, the origin is a repellor (dotted red line), $(1,1)$ is an attractor (solid blue line), $(1,0)$ and $(0,1)$ are again saddles (dashed green lines), and (e^*, e^*) no longer exists.

The situation for $d < 0$ is quite different for the parameter regime where (e^*, e^*) exists, because it gives rise to bistability of the off-diagonal equilibria $(1, 0)$ and $(0, 1)$. The corresponding bifurcation diagram in shown in Fig. 2(b). The equilibrium (e^*, e^*) can only exist if b and c are such that $0 < c/b < (c-d)/(b+d) < 1$. As before, the origin is the only attracting equilibrium as long as $r < c/b$; the equilibrium $(1, 1)$ is a repellor, $(1, 0)$ and $(0, 1)$ are saddles, and (e^*, e^*) does not exist. There are again two (degenerate) transcritical bifurcations, one at $r = c/b$ and one at $r = (c-d)/(b+d)$, where the equilibrium (e^*, e^*) merges in opposite order with $(0, 0)$ and $(1, 1)$, respectively. This means that both the origin and $(1, 1)$ are repellors when $c/b < r < (c - d)/(b + d)$, and the equilibria $(1, 0)$ and $(0, 1)$ are the attractors. In this regime, (e^*, e^*) is again a saddle. As for the case $d > 0$, if $r > (c - d)/(b + d)$ then the origin is a repellor, $(1, 1)$ is an attractor and $(1, 0)$ and $(0, 1)$ are saddles; the equilibrium (e^*, e^*) no longer exists.

Let us mention here that the equilibrium (e^*, e^*) only occurs if $d \neq 0$. If $d = 0$ then $(1, 0)$ and $(0, 1)$ are always saddles, and the origin is an attractor for $r < c/b$, with $(1, 1)$ a repellor, while it is a repellor for $r > c/b$, when $(1, 1)$ is an attractor. The bifurcation at $r = c/b$ is highly degenerate in this case.

4 Analysis of equilibrium states for the genotype model

As we already mentioned in Sec. 2, it might be assumed that stable equilibria of the gene dynamics model (6) should correspond to stable equilibria of the genotype model (1) and, more importantly, in the case of bistability, both systems should have the same predictions for corresponding initial conditions. Therefore, we now compare our findings in Sec. 3 for the gene dynamics model with a similar equilibrium analysis for the genotype model (1). Recall that the genotype dynamics is modeled as

$$\dot{f}_\bullet = f_\bullet \, (w_\bullet - \overline{w}) \, ,$$

with $\bullet \in \mathcal{G} = \{\mathrm{CC}, \mathrm{CD}, \mathrm{DC}, \mathrm{DD}\}$. In its most general form, this system is fully determined by the dynamics of the frequencies f_{CC}, f_{CD} and f_{DC}, with f_{DD} given by the relationship $f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}} + f_{\mathrm{DD}} = 1$. Its equilibria satisfy

$$f_\bullet = 0 \quad \text{or} \quad w_\bullet = \overline{w} \tag{11}$$

for all combinations $\bullet \in \mathcal{G}$. Note that all $f_\bullet = 0$, that is, the origin $(0, 0, 0, 0)$ is not a solution, because we require $f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}} + f_{\mathrm{DD}} = 1$. We can show that there also exist no equilibria with all $f_\bullet \neq 0$ and we have the following Lemma.

Lemma 1 *If* $(f_{\mathrm{CC}}, f_{\mathrm{CD}}, f_{\mathrm{DC}}, f_{\mathrm{DD}})$ *is an equilibrium of* (1) *with* $0 < r < 1$ *and* $d \neq 0$ *then*

$$f_{\mathrm{CC}} \cdot f_{\mathrm{CD}} \cdot f_{\mathrm{DC}} \cdot f_{\mathrm{DD}} = 0,$$

that is, at least one of its coordinates is zero.

The proof of Lemma 1 is given in B.1.

Equation (11) provides a systematic way to derive all possible equilibria of (1). Furthermore, we can use the number of nonzero coordinates as a guide to ensure we find all of them. This leads to the following complete classification of equilibria of (1).

Theorem 2 *Consider the genotype dynamics modeled as system* (1) *with* $d \neq 0$ *and* $0 < r < 1$*. There are at most eight equilibria, which can all coexist for a small range of* r *depending on the choice of the parameters* $b, c > 0$*. Based on their numbers of nonzero coordinates, we distinguish three classes:*

(i) There are four equilibria with a single nonzero coordinate. These are $(1,0,0,0)$*,* $(0,1,0,0)$*,* $(0,0,1,0)$*, and* $(0,0,0,1)$*, which exist for all* $0 < r < 1$*.*

(ii) There are two equilibria with two nonzero coordinates, namely, $\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$*, which exists for all* $0 < r < 1$*, and*

$$E_{23} := \left(\frac{c - r\,(b+d)}{(1-r)\,d},\, 0,\, 0,\, \frac{d - c + r\,b}{(1-r)\,d} \right). \tag{12}$$

The equilibrium E_{23} *only exists if* $c - b < d$ *(for either* $d > 0$ *or* $d < 0$*), and its bounds of existence are defined by*

$$r_{23}^b := \max\left(0,\, \frac{c-d}{b}\right) < 1,$$

and

$$r_{23}^e := \frac{c}{b+d} \in (0,1).$$

If $d > 0$ *then* E_{23} *exists for* $r_{23}^b < r < r_{23}^e$*; this range becomes* $r_{23}^e < r < r_{23}^b$ *if* $d < 0$*.*

(iii) There are also two equilibria with three nonzero coordinates, but these must have $f_{\mathrm{CD}} = f_{\mathrm{DC}}$*; here, either* $f_{\mathrm{CC}} = 0$ *or* $f_{\mathrm{DD}} = 0$*. The first equilibrium is*

$$E_1 := \left(0,\, \frac{c - r\,b}{(1-r)\,d},\, \frac{c - r\,b}{(1-r)\,d},\, \frac{d - 2c + r\,(2b-d)}{(1-r)\,d}\right). \tag{13}$$

If we assume $2b - d > 0$*, then* E_1 *can only exist if* $b > c$ *and its bounds of existence become*

$$r_1^b := \max\left(0,\, \frac{2c-d}{2b-d}\right) < 1,$$

and

$$r_1^e := \frac{c}{b} < 1.$$

As before, E_1 *exists only for* $r_1^b < r < r_1^e$ *if* $d > 0$ *and only for* $r_1^e < r < r_1^b$ *if* $d < 0$*. For the case* $2b - d < 0$*, we must have* $d > 0$ *and* E_1 *can exist only if*

$c < \frac{1}{2}\, d$, with bounds $0 < r < r_1^e < 1$ if $b > c$ and $0 < r < r_1^b < 1$ if $b < c$. The only other possible equilibrium is

$$E_4 := \left(\frac{2c - d - r\,(2b + 3d)}{(1-r)\,d},\ \frac{d - c + r\,(b+d)}{(1-r)\,d},\ \frac{d - c + r\,(b+d)}{(1-r)\,d},\ 0 \right). \tag{14}$$

Existence of this equilibrium requires $\frac{1}{2}\,(c-b) < d$ and the bounds on r become

$$r_4^b := \max\left(0,\ \frac{c-d}{b+d}\right) < 1,$$

and

$$r_4^e := \max\left(0,\ \frac{2c-d}{2b+3d}\right) < 1.$$

Again, if $d > 0$ then E_4 exists only for $r_4^b < r < r_4^e$ and the range becomes $r_4^e < r < r_4^b$ if $d < 0$.

The proof of Theorem 2 is given in B.2.

Contradicting predictions from the gene dynamics and genotype models

As an illustration of the behaviour of the genotype model (1), let us consider the same parameter values used for the gene dynamics model (6) in Fig. 1, namely, $b = 2$, $c = 0.5$, $d = 0.25 > 0$, and $r = 0.185$. The gene dynamics model (6) has five equilibria for this choice of parameters, which is the largest possible number of equilibria for this two-dimensional model. For the genotype model (1) we find six co-existing equilibria. This model is three dimensional and the $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$-coordinates of these equilibria are $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, along with $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ and $E_{23} \approx (0.4110, 0, 0)$. A phase portrait is shown in Fig. 3. Note that the phase space is confined to the tetrahedron bounded by the three coordinate planes $f_{\rm CC} = 0$, $f_{\rm CD} = 0$, $f_{\rm DC} = 0$, and the plane $f_{\rm CC} + f_{\rm CD} + f_{\rm DC} = 1$. We find that two of the equilibria are stable, namely, $(0,0,0)$ and $(1,0,0)$. The equilibria $(0,1,0)$ and $(0,0,1)$ are sources, and $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ and E_{23} are saddles. The equilibrium E_{23} is the only equlibrium with a two-dimensional stable manifold and it is this surface that separates the basins of the two attracting equilibria in phase space. We computed the stable manifold of E_{23} using the two-point boundary value solver AUTO (Doedel 1981, Doedel and Oldeman 2007); the formulation of this computational method is described in (Krauskopf, Osinga, Doedel, Henderson, Guckenheimer, Vladimirsky, Dellnitz and Junge 2005, Krauskopf and Osinga 2007). Our computation hows that the stable manifold of E_{23} is an almost planar surface. It intersects the tetrahedron that defines the phase space of the gene dynamics model (6) in two curves along the sides $f_{CD} = 0$ and $f_{DC} = 0$, and the closure of this two-dimensional stable manifold includes the straight line $f_{CD} + f_{DC} = 1$ on the side $f_{CC} = 0$.

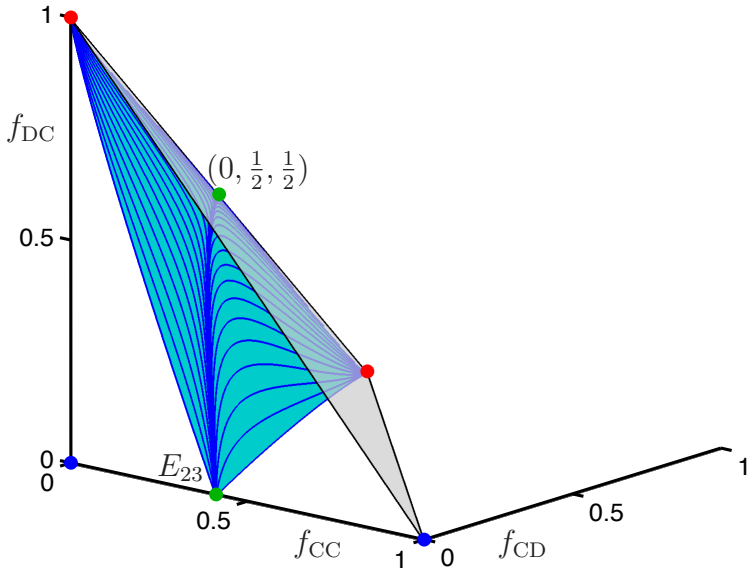

Figure 3: *Phase portrait in* $(f_{\rm CC}, f_{\rm CD}), f_{\rm DC})$*-space illustrating bistability of the equilibria* $(0,0,0)$ *and* $(1,0,0)$ *for the genotype model* (1) *with parameters* $b = 2$*,* $c = 0.5$*,* $d = 0.25 > 0$*, and* $r = 0.185$*. The (blue) surface emanating from the saddle equilibrium* $E_{23} \approx (0.4110, 0, 0)$ *is the stable manifold of* E_{23} *that separates the two basins of attraction for* $(0,0,0)$ *and* $(1,0,0)$*; compare also Fig. 1.*

Despite the fact that there are more equilibria than for the gene dynamics model (6), the phase portrait of the genotype model (1) in Fig. 3 seems rather similar: comparing Fig. 1, there are two attracting equilibria separated by the stable manifold of a saddle equilibrium; since the equilibrium $(0, \frac{1}{2}, \frac{1}{2})$ of the gene dynamics model (6) is contained in the closure of the separatrix, its role in the dynamics is determined by the stable manifold of E_{23}. Furthermore, the equilibria (e^*, e^*) and E_{23} are also roughly in the middle between the two attracting equilibria. In order to compare the dynamics of these two systems (6) and (1) more precisely, we define the variables $f_{\rm C1} := f_{\rm CC} + f_{\rm CD}$ and $f_{\rm C2} := f_{\rm CC} + f_{\rm DC}$ as given by system (1). The two systems could be considered equivalent if any trajectory for system (1) would give rise to a projection onto $(f_{\rm C1}, f_{\rm C2})$-coordinates that maps one-to-one to a trajectory for system (6). Note that there is a one-to-one correspondence between the equilibria $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ of (1) in class (i) and the equilibria $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$ of (6). However, none of the equilibria of (1) map to the equilibrium (e^*, e^*) of (6). This can have dramatic consequences for the behaviour of the two systems. In particular, it should be possible to choose an initial condition in $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$-space that lies in the basin of $(1,0,0)$, that is, to the right of $W^s(E_{23})$, such that its projection onto the

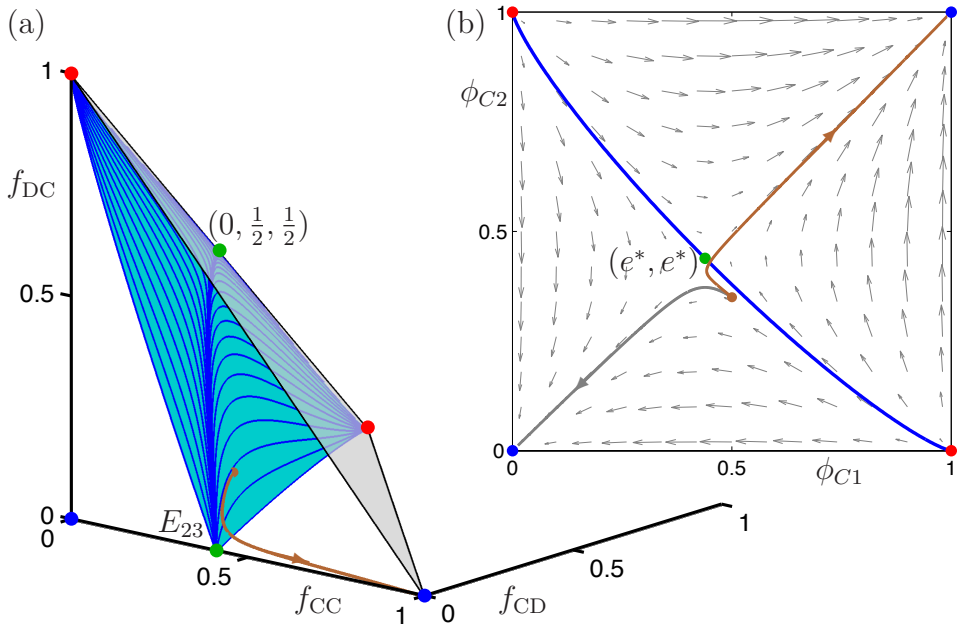

Figure 4: *The initial conditions for a trajectory of the genotype model* (1) *that converges to the attracting equilibrium* $(1,0,0)$ *in* $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$*-space can project onto the two-dimensional phase plane* $f_{\rm C1} = f_{\rm CC} + f_{\rm CD}$ *and* $f_{\rm C2} = f_{\rm CC} + f_{\rm DC}$ *such that the corresponding trajectory for the gene dynamics model* (6) *converges to the equilibrium* $(0,0)$*. Panel (a) shows the trajectory for* (1) *in* $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$*-space (brown curve) and panel (b) shows the corresponding projection overlayed on the phase portrait for the gene dynamics model* (6)*; the expected trajectory as defined by* (6) *is shown in grey.*

$(f_{\rm C1}, f_{\rm C2})$-plane lies in the basin of $(0,0)$. An example to this effect is given in Fig. 4. Here, we consider again the parameters $b = 2$, $c = 0.5$, $d = 0.25 > 0$, and $r = 0.185$, and choose the initial condition $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC}) = (0.25, 0.25, 0.1)$. Under the flow of (1), this point converges to $(1,0,0)$, as indicated by the (brown) curve in Fig. 4(a). However, the projection onto the $(f_{\rm C1}, f_{\rm C2})$-plane of this trajectory starts from the initial condition $(0.5, 0.35)$, which lies in the basin of $(0,0)$ with respect to the flow of (6); the projected trajectory (brown curve) and the trajectory as dictated by (6) (grey curve) are shown in Fig. 4(b).

The example discussed above is a numerical illustration of the following important conjecture.

Proposition 3 *Suppose the parameters* $b, c > 0$*,* $d \neq 0$ *and* $0 < r < 1$ *are chosen such that the gene dynamics model* (6) *exhibits bistability between attracting equilibria* A_1 *and* A_2*. Consider the basin of attraction of* A_1*, denoted* $\mathcal{B}(A_1)$*, and an initial condition* $(\phi_{\rm C1}, \phi_{\rm C2}) \in \mathcal{B}(A_1)$*, that is, the trajectory through* $(\phi_{\rm C1}, \phi_{\rm C2})$ *converges to* A_1*. Let* $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$ *be an initial condition of the genotype system* (1) *with* $f_{\rm CC} + f_{\rm CD} = \phi_{\rm C1}$

and $f_{\rm CC} + f_{\rm DC} = \phi_{\rm C2}$*; here, we use the same values for* b*,* c*,* d *and* r *as for* (6)*. It is possible to choose* $(\phi_{\rm C1}, \phi_{\rm C2}) \in \mathcal{B}(A_1)$ *such that the trajectory associated with the flow of* (1) *does not converge to an attractor that corresponds to* A_1*.*

Proposition 3 is motivated by the fact that there is no candidate equilibrium of (1) that corresponds to the equilibrium (e^*, e^*) of (6). This equilibrium is important, because its stable manifold $W^s(e^*, e^*)$ acts as a separatrix between the basins of A_1 and A_2. If we assume that the genotype system (1) also exhibits bistability for the chosen parameter values, then there must exist an equilibrium of (1) and corresponding stable manifold that acts as a separatrix in a similar way. The projection of this stable manifold onto the $(\phi_{\rm C1}, \phi_{\rm C2})$-plane will not be the same as $W^s(e^*, e^*)$ and the mismatch causes the possible differences in dynamics of the two systems.

Stability properties of genotype equilibria in class (i) as r varies from 0 to 1

The example illustrated in Fig. 4 does not constitute a proof of Proposition 3, but clearly indicates its validity for a particular choice of parameters. Here, we give a detailed analysis for a class of parameters, where we only consider the case $d > 0$ and $b > c$; the case $d < 0$ can be obtained in a similar fashion.

For $d > 0$ and $b > c$, system (6) has coexisting equilibria $(\phi_{\rm C1}, \phi_{\rm C2}) = (0, 0)$ and $(1, 1)$ that are both stable in the regime

$$r_4^b = \frac{c-d}{b+d} < r < \frac{c}{b} = r_1^e,$$

where we used the notation r_4^b and r_1^e from Theorem 2. For $r < r_4^b$ the equilibrium $(\phi_{\rm C1}, \phi_{\rm C2}) = (1, 1)$ is a source instead of a sink, and for $r > r_1^e$ the equilibrium $(\phi_{\rm C1}, \phi_{\rm C2}) = (0, 0)$ is a source instead of a sink. Let us now investigate the stability properties of the corresponding equilibria $(0, 0, 0)$ and $(1, 0, 0)$ of (1).

The stability of equilibria of (1) is determined by the eigenvalues of the Jacobian matrix

$$\mathrm{Jac} = \begin{pmatrix} \dfrac{\partial f_{\rm CC}\,(w_{\rm CC} - \overline{w})}{\partial f_{\rm CC}} & \dfrac{\partial f_{\rm CC}\,(w_{\rm CC} - \overline{w})}{\partial f_{\rm CD}} & \dfrac{\partial f_{\rm CC}\,(w_{\rm CC} - \overline{w})}{\partial f_{\rm DC}} \\ \dfrac{\partial f_{\rm CD}\,(w_{\rm CD} - \overline{w})}{\partial f_{\rm CC}} & \dfrac{\partial f_{\rm CD}\,(w_{\rm CD} - \overline{w})}{\partial f_{\rm CD}} & \dfrac{\partial f_{\rm CD}\,(w_{\rm CD} - \overline{w})}{\partial f_{\rm DC}} \\ \dfrac{\partial f_{\rm DC}\,(w_{\rm DC} - \overline{w})}{\partial f_{\rm CC}} & \dfrac{\partial f_{\rm DC}\,(w_{\rm DC} - \overline{w})}{\partial f_{\rm CD}} & \dfrac{\partial f_{\rm DC}\,(w_{\rm DC} - \overline{w})}{\partial f_{\rm DC}} \end{pmatrix}. \qquad (15)$$

Note that $f_{\rm CC} + f_{\rm CD} + f_{\rm DC} + f_{\rm DD} = 1$ induces a dependency of $f_{\rm DD}$ on all three coordinates. In particular, this means that the partial derivatives must be calculated using the formulation for $\overline{w} = f_{\rm CC}\, w_{\rm CC} + f_{\rm CD}\, w_{\rm CD} + f_{\rm DC}\, w_{\rm DC} + f_{\rm DD}\, w_{\rm DD}$ in terms of $f_{\rm CC}$, $f_{\rm CD}$ and $f_{\rm DC}$ only. The evaluation at an equilibrium simplifies a lot due to the fact that $w_\bullet - \overline{w} = 0$ for any $f_\bullet \neq 0$.

For the equilibrium $(0,0,0)$ almost all terms drop out and we get

$$\begin{aligned}\mathrm{Jac}(0,0,0) &= \left.\begin{pmatrix} w_{\mathrm{CC}}-\overline{w} & 0 & 0 \\ 0 & w_{\mathrm{CD}}-\overline{w} & 0 \\ 0 & 0 & w_{\mathrm{DC}}-\overline{w} \end{pmatrix}\right|_{(0,0,0,1)} \\ &= \left.\begin{pmatrix} w_{\mathrm{CC}}-w_{\mathrm{DD}} & 0 & 0 \\ 0 & w_{\mathrm{CD}}-w_{\mathrm{DD}} & 0 \\ 0 & 0 & w_{\mathrm{DC}}-w_{\mathrm{DD}} \end{pmatrix}\right|_{(0,0,0)}.\end{aligned}$$

Hence, the eigenvalues of $\mathrm{Jac}(0,0,0)$ are on the diagonal and using (2)–(5) with $(f_{\mathrm{CC}}, f_{\mathrm{CD}}, f_{\mathrm{DC}}) = (0,0,0)$, we find

$$\begin{aligned} w_{\mathrm{CC}}-w_{\mathrm{DD}} &= r\,(b+d)-c, \\ w_{\mathrm{CD}}-w_{\mathrm{DD}} = w_{\mathrm{DC}}-w_{\mathrm{DD}} &= \tfrac{1}{2}\,(r\,b-c)\,. \end{aligned}$$

An equilibrium is stable if and only if all its eigenvalues have negative real part. Since we assume $d>0$ and $b>c$, we find that the stability interval for $(0,0,0)$ is

$$0 < r < \frac{c}{b+d} < \frac{c}{b} \Leftrightarrow 0 < r < r_{23}^e < r_1^e.$$

Note that E_{23} merges with $(0,0,0)$ when $r = r_{23}^e$; this is a transcritical bifurcation that renders two of the three eigenvalues of $(0,0,0)$ unstable. The saddle $(0,0,0)$ becomes a source at a second transcritical bifurcation when E_1 merges with it at $r = r_1^e$. We conclude that $(0,0,0)$ of (1) destabilises at an r-value below the r-value at which $(0,0)$ of (6) destabilises.

Let us now consider the equilibrium $(1,0,0)$. The Jacobian matrix becomes

$$\mathrm{Jac}(1,0,0) = \left.\begin{pmatrix} \dfrac{\partial(w_{\mathrm{CC}}-\overline{w})}{\partial f_{\mathrm{CC}}} & \dfrac{\partial(w_{\mathrm{CC}}-\overline{w})}{\partial f_{\mathrm{CD}}} & \dfrac{\partial(w_{\mathrm{CC}}-\overline{w})}{\partial f_{\mathrm{DC}}} \\ 0 & w_{\mathrm{CD}}-\overline{w} & 0 \\ 0 & 0 & w_{\mathrm{DC}}-\overline{w} \end{pmatrix}\right|_{(1,0,0)}. \tag{16}$$

Analogous to the case for $(0,0,0)$, we have $\overline{w} = w_{\mathrm{CC}}$. Due to the upper triangular structure, the eigenvalues of $\mathrm{Jac}(1,0,0)$ are also on the diagonal, with the first one determined by

$$\begin{aligned} &\frac{\partial(w_{\mathrm{CC}}-\overline{w})}{\partial f_{\mathrm{CC}}} \\ =\ &\frac{\partial\left[(1-f_{\mathrm{CC}})\,(w_{\mathrm{CC}}-w_{\mathrm{DD}}) - f_{\mathrm{CD}}\,(w_{\mathrm{CD}}-w_{\mathrm{DD}}) - f_{\mathrm{DC}}\,(w_{\mathrm{DC}}-w_{\mathrm{DD}})\right]}{\partial f_{\mathrm{CC}}} \\ =\ &-(w_{\mathrm{CC}}-w_{\mathrm{DD}}). \end{aligned}$$

Here, we used the fact that $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC}) = (1, 0, 0)$. Hence, using (2)–(5), we find that the eigenvalues of $(1, 0, 0)$ are given by

$$\begin{aligned} w_{\rm DD} - w_{\rm CC} &= (1-r)\,b - [r(b+d) - c + (1-r)(b+d)] \\ &= c - d - r\,b \\ w_{\rm CD} - w_{\rm CC} = w_{\rm DC} - w_{\rm CC} &= \tfrac{1}{2}\left[c - r\,(b+2d) - (1-r)\,d\right] \\ &= \tfrac{1}{2}\left[c - d - r\,(b+d)\right]. \end{aligned}$$

Therefore, $(1, 0, 0)$ is stable if and only if

$$r > \frac{c-d}{b} > \frac{c-d}{b+d} \Leftrightarrow r > r_{23}^b > r_4^b,$$

provided $c - d > 0$. For $r > 0$ small enough, $(1, 0, 0)$ is a source; it becomes a saddle with one stable eigenvalue when r increases past $r = r_4^b$, which causes the emergence of E_4 in a transcritical bifurcation. As r increases further, $(1, 0, 0)$ becomes stable in a second transcritical bifurcation; this time, two eigenvalues change sign simultaneously (due to the symmetry $f_{\rm CD} = f_{\rm DC}$ and the bifurcation gives rise to the equilibrium E_{23}. We conclude that $(1, 0, 0)$ of (1) stabilises at $r = r_{23}^b = (c - d)/b$, which lies above the r-value $r = r_4^b = (c - d)/(b + d)$ at which $(1, 1)$ of (6) stabilises.

We can utilise this mismatch in stability intervals to illustrate Proposition 3 for a range of r-values with $d > 0$ and $b > c$. Consider $r_{23}^e < r < r_1^e$ and let $(0, 0)$ of (6) the attractor A_1 of Proposition 3. Then almost any initial condition $(\phi_{\rm C1}, \phi_{\rm C2}) \in \mathcal{B}(A_1)$ of (6) satisfies the conditions of Proposition 3: almost all initial conditions $(f_{\rm CC}, f_{\rm CD}, f_{\rm DC})$ of (1) with $f_{\rm CC} + f_{\rm CD} = \phi_{\rm C1}$ and $f_{\rm CC} + f_{\rm DC} = \phi_{\rm C2}$ will not converge to $(0, 0, 0)$, because $(0, 0, 0)$ is not stable. (The only exceptions are initial conditions that lie on the one-dimensional stable manifold of the saddle $(0, 0, 0)$.) Similarly, for $r_4^b < r < r_{23}^b$, the equilibrium $(1, 1)$ of (6) is stable, but $(1, 0, 0)$ of (1) is not and Proposition 3 applies.

5 Discussion

Early results in population genetics demonstrated that 'adaptive topographies', independent of the genetic bases underlying phenotypes, cannot exist; by changing the genetic representation of traits, the stable equilibria under the action of natural selection change (Moran 1964). Two extreme approaches exist, one in which all the frequencies of all gene combinations under selection are accounted for, and one in which selection at the different genetic loci is treated as completely independent; in between these two extremes, multilocus population genetics approaches such as quasi-linkage equilibrium manage to simplify analyses without completely ignoring interactions between loci (Kimura 1965, Kirkpatrick. et al. 2002, Gardner et al. 2007).

Interestingly, recent analysis has shown that, for a simple case of social evolution where phenotypes are controlled by two genetic loci, the stable equilibria of natural selection are the same regardless of whether one considers selection acting on the entire

genotype (Maynard Smith and Parker 1976, Maynard Smith 1982), or acting on independent 'genes' (Marshall 2009). Thus, it may be tempting to assume that, for certain kinds of sufficiently simple model, multilocus and independent-locus approaches yield equivalent answers. Here we have presented an analysis using the tools of dynamical systems theory, for the simple case of asymmetric non-additive donation games played between relatives. This analysis reveals the following main points: first, the gene dynamics and the genotype dynamics cannot be made topologically equivalent in a dynamical systems sense, since the dimensions of the respective phase spaces are different. It is also not possible to 'slave' the dynamics of the higher-dimensional gene dynamics model to the genotype model, because the two models differ in their number of equilibria and in the locations of some of these equilibria. Second, we find additional equilibria for the genotype model to those previously found using techniques from evolutionary game theory, since we find unstable equilibria as well as the previously-discovered stable equilibria. Third, by observing that the unstable equilibria under the gene and the genotype dynamics are different, we show that although the stable equilibria are the same in the two systems, initial conditions always exist in which the population equilibria that result under natural selection in each system are different. That is, for the same starting population the two different model analyses predict different evolutionary outcomes.

Our results are the evolutionary game theory counterpart of earlier results from the population genetics literature, that the genetic bases of traits under selection affect population equilibria. These population genetics approaches, briefly reviewed in (Feldman 2009), rest on analysis of sexual models. In particular, analysis of population genetics models shows that the concept of an 'adaptive landscape' independent of genetic details is incorrect (Moran 1964). Our analysis is an evolutionary game theory one, which is inherently asexual; strategies, or strategy components ('genes') reproduce directly. Our analysis of the particular social game presented here also demonstrates a different effect, since here the stable equilibria are the same, but the selected equilibria can differ. The fact that it is only equilibrium selection, rather than the stable equilibria themselves, that is affected by using the analytically simpler model of this game may be of interest.

Acknowledgments

J.A.R.M. thanks J.M. McNamara for drawing Moran's paper on 'adaptive topographies' to his attention. H.M.O. was supported by an Advanced Research Fellowship of the Engineering and Physical Sciences Research Council (EPSRC), UK.

A Topological non-equivalence of the two models

Note that the relationships $\phi_{\rm C1} + \phi_{\rm D1} = 1$ and $\phi_{\rm C2} + \phi_{\rm D2} = 1$ for system (6) mean that we only need to consider the cases $\dot{\phi}_{\rm Ci}$ with $i \in \{1, 2\}$. We show here that $\dot{\phi}_{\rm C1} \neq \dot{f}_{\rm CC} + \dot{f}_{\rm CD}$, where $\dot{f}_{\rm CC}$ and $\dot{f}_{\rm CD}$ are given by equation (1); the case for $\dot{\phi}_{\rm C2}$ is similar.

Let us consider the equation for $\dot{\phi}_{\rm C1}$. We are given

$$\dot{\phi}_{\rm C1} = \phi_{\rm C1}\ (1 - \phi_{\rm C1})\ (\omega_{\rm C1} - \omega_{\rm D1})\,,$$

where

$$\omega_{\rm C1} - \omega_{\rm D1} = r\,b - c + (1 + r)\,d\,\phi_{\rm C2}.$$

We have a similar equation for $\dot{\phi}_{\rm C2}$, but let us assume that $\phi_{\rm C2} = \phi_{\rm C1}$, that is, we consider the diagonal dynamics only, as given by the single equation

$$\dot{\phi} = \phi\ (1 - \phi)\ [r\,b - c + (1 + r)\,d\,\phi] := \phi\ (1 - \phi)\ G(\phi). \tag{17}$$

Since $\phi_{\rm C1} = f_{\rm CC} + f_{\rm CD}$ and $\phi_{\rm C2} = f_{\rm CC} + f_{\rm DC}$, the equality $\phi_{\rm C2} = \phi_{\rm C1}$ implies $f_{\rm DC} = f_{\rm CD}$. Under this assumption, the mean fitness becomes

$$\overline{\omega} = (r\,b - c + [1 - r]\,b + [1 - r]\,d\,[f_{\rm CC} + f_{\rm CD}])\ [f_{\rm CC} + f_{\rm CD}] + r\,d\,f_{\rm CC}.$$

Hence, we have

$$\begin{aligned} \dot{f}_{\rm CC} &= f_{\rm CC}\ (w_{\rm CC} - \overline{\omega}) \\ &= f_{\rm CC}\ \{(1 - [f_{\rm CC} + f_{\rm CD}])\,(r\,b - c + [1 - r]\,d\,[f_{\rm CC} + f_{\rm CD}]) \\ &\qquad + r\,d\,(1 - f_{\rm CC})\} \end{aligned} \tag{18}$$

and

$$\begin{aligned} \dot{f}_{\rm CD} &= f_{\rm CD}\ (w_{\rm CD} - \overline{\omega}) \\ &= \tfrac{1}{2}\,f_{\rm CD}\ \{(1 - 2\,[f_{\rm CC} + f_{\rm CD}])\,(r\,b - c + [1 - r]\,d\,[f_{\rm CC} + f_{\rm CD}]) \\ &\qquad - 2\,r\,d\,f_{\rm CC}\} \end{aligned} \tag{19}$$

If we write $g = f_{\rm CC} + f_{\rm CD}$, then (18) and (19) give

$$\begin{aligned} \dot{g} &= \dot{f}_{\rm CC} + \dot{f}_{\rm CD} \\ &= g\,(1 - g)\,(r\,b - c + [1 - r]\,d\,g) \\ &\qquad + r\,d\,(1 - f_{\rm CC})\,g - \tfrac{1}{2}\,f_{\rm CD}\,(r\,b - c + [1 - r]\,d\,g) - r\,d\,f_{\rm CD} \\ &= g\,(1 - g)\,G(g) - \tfrac{1}{2}\,G(g)\,f_{\rm CD} - r\,d\,[g\,(1 - g) + 2\,g\,f_{\rm CC} + f_{\rm CD}]. \end{aligned}$$

Comparing with (17), we observe that already the diagonal dynamics does not satisfy $\dot{\phi}_{\rm C1} = \dot{f}_{\rm CC} + \dot{f}_{\rm CD}$.

□

B Analysis of equilibrium states for the genotype model (1)

We provide here a detailed analysis of the equilibria for the genotype model (1) in their most general form. Let us begin with the proof of lemma 1.

B.1 Proof of lemma 1

Since all $f_\bullet \neq 0$, we must have $w_\bullet = \overline{w}$ for all $\bullet \in \{\mathrm{CC}, \mathrm{CD}, \mathrm{DC}, \mathrm{DD}\}$. This means that

$$\overline{w} = w_{\mathrm{CC}} = w_{\mathrm{CD}} = w_{\mathrm{DC}} = w_{\mathrm{DD}},$$

so we must have equality of all inclusive fitnesses. Equations (2) and (3) give

$$\begin{aligned} w_{\mathrm{CC}} = w_{\mathrm{CD}} &\Leftrightarrow 2\,(w_{\mathrm{CC}} - w_{\mathrm{CD}}) = 0 \\ &\Leftrightarrow r\,(b+2d) - c + (1-r)\,[f_{\mathrm{CC}} + f_{\mathrm{CD}}]\,d = 0 \\ &\Leftrightarrow f_{\mathrm{CC}} + f_{\mathrm{CD}} = \frac{c - r\,(b+2d)}{(1-r)\,d}. \end{aligned} \tag{20}$$

Here, we used the assumption $d \neq 0$. Due to symmetry, even without requiring $f_{\mathrm{CD}} = f_{\mathrm{DC}}$, we also have

$$w_{\mathrm{CC}} = w_{\mathrm{DC}} \Leftrightarrow f_{\mathrm{CC}} + f_{\mathrm{DC}} = \frac{c - r\,(b+2d)}{(1-r)\,d}. \tag{21}$$

Similarly, (2) and (5) give

$$\begin{aligned} w_{\mathrm{CC}} = w_{\mathrm{DD}} &\Leftrightarrow w_{\mathrm{CC}} - w_{\mathrm{DD}} = 0 \\ &\Leftrightarrow r\,(b+d) - c + (1-r)\,[f_{\mathrm{CC}} + \tfrac{1}{2}\,f_{\mathrm{CD}} + \tfrac{1}{2}\,f_{\mathrm{DC}}]\,d = 0 \\ &\Leftrightarrow [f_{\mathrm{CC}} + \tfrac{1}{2}\,f_{\mathrm{CD}} + \tfrac{1}{2}\,f_{\mathrm{DC}}] = \frac{c - r\,(b+d)}{(1-r)\,d}. \end{aligned} \tag{22}$$

Using (3) and (5) leads to

$$\begin{aligned} w_{\mathrm{DD}} = w_{\mathrm{CD}} &\Leftrightarrow 2\,(w_{\mathrm{CD}} - w_{\mathrm{DD}}) = 0 \\ &\Leftrightarrow r\,b - c + (1-r)\,[f_{\mathrm{CC}} + f_{\mathrm{DC}}]\,d = 0 \\ &\Leftrightarrow f_{\mathrm{CC}} + f_{\mathrm{DC}} = \frac{c - r\,b}{(1-r)\,d}. \end{aligned} \tag{23}$$

Similarly, (4) and (5) give

$$w_{\mathrm{DD}} = w_{\mathrm{DC}} \quad \Leftrightarrow \quad f_{\mathrm{CC}} + f_{\mathrm{CD}} = \frac{c - r\,b}{(1-r)\,d}. \tag{24}$$

Finally, (3) and (4) give

$$w_{\mathrm{CD}} = w_{\mathrm{DC}} \Leftrightarrow f_{\mathrm{CD}}\,d = f_{\mathrm{DC}}\,d \Leftrightarrow d = 0 \text{ or } f_{\mathrm{CD}} = f_{\mathrm{DC}}. \tag{25}$$

It is clear that (20)–(25) can be satisfied simultaneously only if $r = 0$; for example, $w_{\mathrm{CC}} = w_{\mathrm{CD}}$ and $w_{\mathrm{DD}} = w_{\mathrm{DC}}$ require (20) and (24), that is

$$f_{\mathrm{CC}} + f_{\mathrm{CD}} = \frac{c - r\,(b + 2d)}{(1-r)\,d} = \frac{c - r\,b}{(1-r)\,d}$$
$$\Leftrightarrow \quad \frac{2r\,d}{(1-r)\,d} = \frac{2r}{(1-r)} = 0$$

Since $0 < r < 1$, this proves the Lemma.

□

B.2 Proof of theorem 2

Lemma 1 implies that any equilibrium of (1) must have at least one of its coordinates equal to zero. Furthermore, $f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}} + f_{\mathrm{DD}} = 1$, so the equilibria of (1) can indeed all be classified by the classes listed in Theorem 2. Let us begin with class (i).

Class (i):
The equality $f_{\mathrm{CC}} + f_{\mathrm{CD}} + f_{\mathrm{DC}} + f_{\mathrm{DD}} = 1$ implies that only $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$, and $(0,0,0,1)$ are possible candidates for this class. These four points are equilibria of (1) if the equilibrium condition (11) is satisfied for each of their coordinates. Clearly, we only need to check (11) for the single nonzero coordinate $f_\bullet = 1$, for which we require $w_\bullet = \overline{w}$. However, the mean fitness,

$$\overline{w} = f_{\mathrm{CC}}\,w_{\mathrm{CC}} + f_{\mathrm{CD}}\,w_{\mathrm{CD}} + f_{\mathrm{DC}}\,w_{\mathrm{DC}} + f_{\mathrm{DD}}\,w_{\mathrm{DD}},$$

simply reduces to $w_\bullet$ if three of the four frequencies are zero. Hence, $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$, and $(0,0,0,1)$ are all equilibria and there are no restrictions on r for their existence.

Class (ii):
This class contains all equilibria with two coordinates equal to zero. Suppose $f_{\mathrm{CC}} = 0$ and $f_{\mathrm{CD}} = 0$, while $f_{\mathrm{DC}}, f_{\mathrm{DD}} \neq 0$. Then (24) must hold in order to satisfy (11), but $f_{\mathrm{CC}} + f_{\mathrm{CD}} = 0$, so there is no (generic) solution. Similarly, if we assume $f_{\mathrm{CD}} \neq 0$ and $f_{\mathrm{DC}} = 0$, then (23) implies

$$\frac{c - r\,b}{(1-r)\,d} = 0 \Leftrightarrow r = \frac{c}{b},$$

which is not generic. At the special value $r = \frac{c}{b}$ a two-dimensional continuum of equilibria $(0, 0, f_{\mathrm{DC}}, f_{\mathrm{DD}})$ and another two-dimensional continuum of equilibria $(0, f_{\mathrm{CD}}, 0, f_{\mathrm{DD}})$ exist that are both not persistent under variations in r. Hence, a generic equilibrium from class (ii) with $f_{\mathrm{CC}} = 0$ must have $f_{\mathrm{DD}} = 0$. Then (25) holds, which gives the candidate $\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$. Since $w_{\mathrm{DC}} = w_{\mathrm{CD}}$, the mean fitness becomes

$$\overline{w} = \frac{1}{2}\,w_{\mathrm{CD}} + \frac{1}{2}\,w_{\mathrm{DC}} = w_{\mathrm{CD}} = w_{\mathrm{DC}},$$

so $\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$ is indeed an equilibrium. Note that this equilibrium exists without restrictions on r.

The only other option for equilibria in this class are equilibria with two zero coordinates and $f_{\rm CC} \neq 0$. If we assume that the other nonzero coordinate is $f_{\rm CD} \neq 0$, then (20) implies

$$\frac{c - r\,(b + 2d)}{(1 - r)\,d} = 1 \Leftrightarrow r = \frac{c - d}{b + d},$$

because $f_{\rm DC} = f_{\rm DD} = 0$, so that $f_{\rm CC} + f_{\rm CD} = 1$. This is again not generic. The same applies to the case $f_{\rm CD} = f_{\rm DD} = 0$, using (21), and the only remaining candidate is an equilibrium with $f_{\rm CD} = f_{\rm DC} = 0$. For this case (22) applies and we find

$$f_{\rm CC} = \frac{c - r\,(b + d)}{(1 - r)\,d}.$$

The value for $f_{\rm DD}$ follows from the remainder $f_{\rm DD} = 1 - f_{\rm CC}$. The equality of the inclusive fitnesses for all nonzero frequencies again implies $\overline{w} = w_{\rm CC} = w_{\rm DD}$. Hence, the candidate E_{23} as given in (12) is indeed an equilibrium.

The existence interval of E_{23} is determined by the fact that all coordinates of E_{23} must lie in the interval $[0, 1]$; it suffices to check this for the $f_{\rm CC}$-coordinate of E_{23}, since $f_{\rm CC} + f_{\rm DD} = 1$ then implies $0 \leq f_{\rm DD} \leq 1$ as well. Let us first consider the case with $d > 0$; we have:

$$0 \leq \frac{c - r\,(b + d)}{(1 - r)\,d} \leq 1 \quad \Leftrightarrow \quad \begin{cases} c - r\,(b + d) & \geq & 0 \\ c - r\,(b + d) & \leq & (1 - r)\,d, \end{cases} \quad \text{and}$$

$$\Leftrightarrow \quad \begin{cases} r & \leq & \dfrac{c}{b + d} \\ r & \geq & \dfrac{c - d}{b}. \end{cases} \quad \text{and}$$

Hence, the existence interval is $\frac{c-d}{b} \leq r \leq \frac{c}{b+d}$, which only makes sense if

$$\frac{c - d}{b} < \frac{c}{b + d} \Leftrightarrow c - b < d. \tag{26}$$

The bounds r_{23}^b and r_{23}^e defined in Theorem 2 take into account that one could have $c - d < 0$, in which case $0 < r < \frac{c}{b+d}$.

The case $d < 0$ is analogous, with '$\leq$' replaced by '$\geq$' and vice versa as soon as the inequality is multiplied by $(1 - r)\,d$. Note that we must consider the possibility $b+d < 0$, but this leads to $r < 0$, which is not acceptable. Hence, we have $b+d > 0$ and the bounds r_{23}^b and r_{23}^e simply swap places. We now need $r_{23}^e < r_{23}^b$, which leads to the same condition $c - b < d$ as derived in (26) for $d > 0$; note that $c - b < d \Rightarrow b + d > 0$.

Class (iii):

The final class consists of equilibria with three nonzero coordinates. We obtain E_1

given by (13) if we assume $f_{\rm CC} = 0$. Indeed, for this case, (23), (24) and (25) must hold, which requires

$$f_{\rm CD} = f_{\rm DC} = \frac{c - r\,b}{(1-r)\,d},$$

and the value for $f_{\rm DD}$ follows from the fact that all frequencies sum up to one. As before, the equality $w_{\rm CD} = w_{\rm DC} = w_{\rm DD}$ implies that $\overline{w}$ is equal to each of these inclusive fitnesses and E_1 is, indeed, an equilibrium.

The existence interval of E_1 is then determined by the values of r for which $f_{\rm CD} = f_{\rm DC} \in [0, \frac{1}{2}]$; this automatically implies $f_{\rm DD} \in [0, 1]$. Let us first consider the case $d > 0$. If we assume $2b - d > 0$ then we have:

$$0 \le \frac{c - r\,b}{(1-r)\,d} \le \frac{1}{2} \quad \Leftrightarrow \quad \begin{cases} c - r\,b & \ge & 0 \\ 2\,(c - r\,b) & \le & (1-r)\,d, \end{cases} \text{ and}$$

$$\Leftrightarrow \quad \begin{cases} r & \le & \dfrac{c}{b} \\ r & \ge & \dfrac{2c-d}{2b-d}. \end{cases} \text{ and}$$

These bounds lead to an r-interval if $\frac{2c-d}{2b-d} < \frac{c}{b}$, which holds if $b > c$; note that the additional condition $0 < r < 1$ defines the bounds r_1^b and r_1^e given in Theorem 2. If d is large and $2b - d < 0$ then E_1 exists for $0 < r < \frac{c}{b}$, if $b > c$, and for $0 < r < \frac{2c-d}{2b-d}$, if $b < c$.

The case $d < 0$ is again analogous, and we get $r_1^e < r < r_1^b$. The condition $r_1^e < r_1^b$ leads to the requirement $b > c$, which automatically ensures that this r-interval is contained in $[0, 1]$.

Let us now consider the possible existence of an equilibrium with $f_{\rm CD} = 0$ and all other coordinates nonzero. This means that (21), (22) and (24) must hold. Since $f_{\rm CD} = 0$, equation (24) defines $f_{\rm CC}$, and combined with (21), this gives

$$\begin{aligned} f_{\rm DC} = \frac{c - r\,(b+2d)}{(1-r)\,d} - f_{\rm CC} &= \frac{c - r\,(b+2d)}{(1-r)\,d} - \frac{c - r\,b}{(1-r)\,d} \\ &= \frac{-2d\,r}{(1-r)\,d} = \frac{-2r}{(1-r)} < 0; \end{aligned}$$

here, we used the fact that $0 < r < 1$. Hence, there is no admissible equilibrium in class (iii) that satisfies $f_{\rm CD} = 0$. A similar argument holds for the case with $f_{\rm DC} = 0$.

The only other possibility is an equilibrium with all nonzero coordinates except for $f_{\rm DD} = 0$. We must satisfy (20), (21) and (25), which implies

$$f_{\rm CC} + f_{\rm CD} = f_{\rm CC} + f_{\rm DC} = \frac{c - r\,(b+2d)}{(1-r)\,d}.$$

Furthermore, $f_{\rm CC} + f_{\rm CD} + f_{\rm DC} = 1$, so

$$f_{\rm CD} = f_{\rm DC} = 1 - \frac{c - r\,(b+2d)}{(1-r)\,d} = \frac{d - c + r\,(b+d)}{(1-r)\,d},$$

which fixes $f_{\rm CC}$ as well. Hence, E_4 as defined in (14) is an equilibrium of system (1).

As before, we find the existence interval of the equilibrium E_4 using the condition $f_{\rm CD} = f_{\rm DC} \in [0, \frac{1}{2}]$. Let us first consider the case $d > 0$, which leads to:

$$\begin{aligned} & 0 \leq \frac{d - c + r\,(b+d)}{(1-r)\,d} \leq \frac{1}{2} \\ \Leftrightarrow & \begin{cases} \quad d - c + r\,(b+d) & \geq & 0 \\ 2\,(d - c + r\,(b+d)) & \leq & (1-r)\,d, \end{cases} \quad \text{and} \\ \Leftrightarrow & \begin{cases} r & \geq & \dfrac{c-d}{b+d} \\ r & \leq & \dfrac{2c-d}{2b+3d}. \end{cases} \quad \text{and} \end{aligned}$$

As for the other equilibria, we must show that these bounds lead to a nontrivial r-interval. We have

$$\frac{c-d}{b+d} < \frac{2c-d}{2b+3d} \Leftrightarrow \frac{1}{2}(c-b) < d, \tag{27}$$

so E_4 can only exist for $d > 0$ if $\frac{1}{2}(c-b) < d$; the bounds r_4^b and r_4^e defined in Theorem 2 take into account that $0 < r < 1$ as well.

For the case $d < 0$ we have $(1-r)\,d < 0$ and we find the existence interval $r_4^e < r < r_4^b$, provided the same bound $\frac{1}{2}(c-b) < d$ from (27) is satisfied; here we assume $b + d > 0$ and $2b + 3d > 0$. The case $b + d < 0$ leads to an interval with $r < 0$, which is not admissible; the case $b + d > 0$, but $2b + 3d < 0$ also requires $r < 0$. Note that the condition $\frac{1}{2}(c-b) < d$ implies

$$b + d > b + \frac{1}{2}(c-b) = \frac{1}{2}(c+b) > 0,$$

and

$$2b + 3d > 2b + \frac{3}{2}(c-b) = \frac{1}{2}(3c+b) > 0.$$

This concludes the investigation of all possible equilibria for system (1). In total, we found the eight equilibria listed in Theorem 2 and there are no other equilibria.

□

References

Doedel, E. J.: 1981, AUTO: A program for the automatic bifurcation analysis of autonomous systems, *Congressus Numerantium* **30**, 265–284.

Doedel, E. J. and Oldeman, B. E.: 2007, *AUTO-07P: Continuation and bifurcation software for ordinary differential equations*, Concordia University, Montréal. with

major contributions from Champneys, A. R., Fairgrieve, T. F., Kuznetsov, Yu. A, Paffenroth, R. C., Sandstede, B., Wang, X. J., and Zhang, C.; available via `http://cmvl.cs.concordia.ca/auto/`.

Feldman, M. W.: 2009, Sam karlin and multi-locus population genetics, *Theoretical Population Biology* **75**(4), 233–235.

Gardner, A., West, S. A. and Barton, N. H.: 2007, The relation between multilocus population genetics and social evolution theory, *The American Naturalist* **169**, 207–226.

Grafen, A.: 1984, Natural selection, kin selection and group selection, *in* J. R. Krebs and N. B. Davies (eds), *Behavioural Ecology: an Evolutionary Approach*, Sinauer Associates Inc, Sunderland, pp. 62–84.

Grafen, A.: 1985, A geometric view of relatedness, *Oxford Surveys in Evolutionary Biology* **2**(237), 28–90.

Guckenheimer, J. and Holmes, P.: 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Applied Mathematical Sciences **42**, Springer-Verlag, New York.

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

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

Krauskopf, B. and Osinga, H. M.: 2007, Computing invariant manifolds via the continuation of orbit segments, *in* B. Krauskopf, H. M. Osinga and J. Galán-Vioque (eds), *Numerical Continuation Methods for Dynamical Systems: Path Following and Boundary Value Problems*, Springer-Verlag, Berlin, pp. 117–154.

Krauskopf, B., Osinga, H. M., Doedel, E. J., Henderson, M. E., Guckenheimer, J., Vladimirsky, A., Dellnitz, M. and Junge, O.: 2005, A survey of methods for computing (un)stable manifolds of vector fields, *International Journal of Bifurcation and Chaos* **15**, 763–791.

Kuznetsov, Y. A.: 1995, *Elements of Applied Bifurcation Theory*, Applied Mathematical Sciences **112**, second edn, Springer-Verlag, New York.

Marshall, J. A. R.: 2009, The donation game with roles played between relatives, *Journal of Theoretical Biology* **260**, 386–391.

Maynard Smith, J.: 1982, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.

Maynard Smith, J. and Parker, G. A.: 1976, The logic of asymmetric contests, *Animal Behaviour* **24**(1), 159–175.

Moran, P. A.: 1964, On the non-existence of adaptive topographies, *Annals of Human Genetics* **27**, 383–393.

Queller, D. C.: 1996, The measurement and meaning of inclusive fitness, *Animal Behaviour* **51**(1), 229–232.

Taylor, P. D. and Frank, S. A.: 1996, How to make a kin selection model, *Journal of Theoretical Biology* **180**(1), 27–37.

Trivers, R. L.: 1974, Parent-offspring conflict, *Integrative and Comparative Biology* **14**(1), 249–264.