

Lorens A. Imhof · Martin A. Nowak

Evolutionary game dynamics in a Wright-Fisher process

Received: 5 May 2004 / Revised version: 2 December 2005 /
Published online: 7 February 2006 – © Springer-Verlag 2006

Abstract. Evolutionary game dynamics in finite populations can be described by a frequency dependent, stochastic Wright-Fisher process. We consider a symmetric game between two strategies, A and B . There are discrete generations. In each generation, individuals produce offspring proportional to their payoff. The next generation is sampled randomly from this pool of offspring. The total population size is constant. The resulting Markov process has two absorbing states corresponding to homogeneous populations of all A or all B . We quantify frequency dependent selection by comparing the absorption probabilities to the corresponding probabilities under random drift. We derive conditions for selection to favor one strategy or the other by using the concept of total positivity. In the limit of weak selection, we obtain the 1/3 law: if A and B are strict Nash equilibria then selection favors replacement of B by A , if the unstable equilibrium occurs at a frequency of A which is less than 1/3.

1. Introduction

Evolutionary game theory is the study of frequency dependent selection. The relative fitness of two (or more) phenotypes (strategies) is not constant, but depends on the composition of the population. Traditionally, evolutionary game dynamics are studied by deterministic differential equations describing infinitely large populations. A wide-spread system is the celebrated replicator dynamic which was introduced by Taylor & Jonker (1978) and Hofbauer, Schuster & Sigmund (1979). Other game dynamics include imitation dynamics, best-response dynamics, monotone selection dynamics and adjustment dynamics, see e.g. Hofbauer & Sigmund (1998). All of these are deterministic descriptions applying to infinitely large populations. Papers that deal with stochastic modifications include those of Foster & Young (1990); Fudenberg & Harris (1992); Corradi & Sarin (2000); Dostálková & Kindlmann (2004) and Imhof (2005a). For recent reviews and comprehensive treatments of evolutionary game dynamics see Fudenberg & Tirole (1991); Binmore (1993, 1998); Weibull (1995); Samuelson (1997); Fudenberg & Levine (1998);

L.A. Imhof: Statistische Abteilung, Universität Bonn, Adenauerallee 24-42, 53113 Bonn, Germany. e-mail: limhof@uni-bonn.de

M.A. Nowak: Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA

Key words or phrases: Frequency dependent process – Stochastic game dynamics – Finite populations

Hofbauer & Sigmund (1998, 2003); Gintis (2000); Nowak & Sigmund (2004) and Imhof (2005b).

Imagine two strategies A and B engaged in a symmetric game with payoff matrix

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array} \quad (1)$$

Consider an infinite population and denote by x the frequency of A and by $1 - x$ the frequency of B . The payoff of A is $f_A = ax + b(1 - x)$. The payoff of B is $f_B = cx + d(1 - x)$. The average payoff of the population is $\bar{f} = xf_A + (1 - x)f_B$. The replicator dynamics assume that strategies reproduce proportional to their payoff. We have

$$\frac{dx}{dt} = x(f_A - \bar{f}).$$

This can be written as

$$\frac{dx}{dt} = x(1 - x)(f_A - f_B).$$

If $a > c$ and $b > d$ then A dominates B ; the only stable equilibrium is $x = 1$. If $a < c$ and $b < d$ then B dominates A ; the only stable equilibrium is $x = 0$. If $a > c$ and $b < d$ then A and B are bi-stable; both $x = 0$ and $x = 1$ are stable equilibria; there is an unstable equilibrium at $x^* = (d - b)/(a - b - c + d)$. If $a < c$ and $b > d$ then A and B co-exist; both $x = 0$ and $x = 1$ are unstable equilibria; the only stable equilibrium is given by $x^* = (d - b)/(a - b - c + d)$.

If $a > c$ then A is a strict Nash equilibrium. The equilibrium $x = 1$ is stable; an infinitesimally small amount of B cannot invade. A closely related concept is evolutionary stability (Maynard Smith 1982). A is an evolutionarily stable strategy (ESS) if either $a > c$ or both $a = c$ and $b > d$. Again, if A is an ESS, then strategy B cannot invade. Note that strict Nash implies ESS, but the converse does not hold. In both cases, however, a homogeneous population of A is protected by natural selection against invasion by B .

The replicator dynamics and the uninvadability of a strict Nash equilibrium (and an ESS) hold in the limit of infinitely large population size. It is natural to study evolutionary game dynamics in finite populations. There are various approaches to studying game dynamics in finite populations (Schaffer 1988; Kandori et al. 1993; Fogel et al. 1998; Ficici & Pollack 2000; Schreiber 2001). In one such approach, a frequency dependent Moran process was investigated (Nowak et al. 2004; Taylor et al. 2004; Fudenberg et al. 2004). Again, let us consider the interaction between two strategies A and B as given by payoff matrix (1). In each step of the stochastic process, one individual is chosen for reproduction with probability proportional to fitness. The offspring of this individual will replace a randomly chosen individual. The total population size is constant and given by N . In the absence of mutation, there are two absorbing states, corresponding to all A or all B . Denote by ρ_{AB} the

probability that a single A player in a population of $N - 1$ B players will generate a lineage that will take over the whole population. We say that selection favors A replacing B if $\rho_{AB} > 1/N$, because for neutral drift the corresponding fixation probability would be $1/N$.

If A and B are strict Nash equilibria, then there is an unstable equilibrium of the replicator dynamics at a frequency of A given by $x^* = (d - b)/(a - b - c + d)$. For the frequency dependent Moran process, it can be shown that selection favors A replacing B if $x^* < 1/3$. This surprisingly simple ‘1/3 law’ holds in the limit of weak selection and sufficiently large population size N . In particular, it turns out that the conditions of a strict Nash equilibrium or an evolutionarily stable strategy (ESS) do not imply protection by natural selection against invasion and replacement in finite populations. Recent papers that deal with the frequency dependent Moran process include Wild & Taylor (2004), Antal & Scheuring (2005) and Imhof et al. (2005).

The Moran process is well known in population genetics (Moran 1962; Bürger 2000; Ewens 2004) where it is normally used to study the dynamics of constant selection in finite populations. The Moran process describes a biological population with asynchronous reproduction. At any one time step a single individual is chosen for reproduction. There are overlapping generations. In contrast, the Wright-Fisher process describes a biological population with discrete generations. All individuals reproduce at the same time. They generate a pool of offspring from which the next generation is chosen (Ewens 2004). Both synchronous and asynchronous reproduction occur in biological populations, but in population genetics the Wright-Fisher process is more widespread than the Moran process (Ewens 2004). Therefore, we would like to extend our analysis of evolutionary game dynamics in finite populations also to the Wright-Fisher process.

This is a difficult endeavour. The Moran process is a birth-death process and consequently allows simple explicit solutions for the absorption probabilities. In contrast, the Wright-Fisher process that we study here is a Markov process nearly all of whose one-step transition probabilities are strictly positive, and the process does not allow such explicit solutions. Nevertheless we can derive analytic results and outline similarities and differences between the two processes. In particular, it turns out that the same ‘1/3 law’ holds for the Wright-Fisher process.

A common approach to analysing Wright-Fisher-type processes for large finite populations is to derive mathematically more tractable diffusion approximations. In the present paper we do not rely on such an approximation and study the Markov chain directly. In so doing we take advantage of the fact that the transition matrix is totally positive, which gives useful information on the fixation probabilities. Most of our results apply to every finite population size. Starting from these general results, we also derive two assertions for sufficiently large finite populations (Theorems 4 and 5). It is, however, straightforward to check whether a given population size is indeed large enough in order that these assertions hold. This would not be the case had we used the diffusion approximation.

2. A Frequency Dependent Wright-Fisher Model

Consider a finite population consisting of N individuals, each playing either A or B . If i players use strategy A , every A -player faces $N - i$ opponents using strategy B and $i - 1$ opponents using A . Under random mating, the expected payoff to A -players is therefore given by

$$\frac{a(i - 1) + b(N - i)}{N - 1}.$$

Similarly, the expected payoff to B -players is

$$\frac{ci + d(N - i - 1)}{N - 1}.$$

We assume that the degree to which the payoffs contribute to fitness is specified by a parameter $w \in [0, 1]$. Thus the fitness of A - and B -players is

$$\begin{aligned} f_i &= 1 - w + w \frac{a(i-1)+b(N-i)}{N-1}, \\ g_i &= 1 - w + w \frac{ci+d(N-i-1)}{N-1}, \end{aligned}$$

respectively. Both in the deterministic replicator model for infinite populations and in the stochastic replicator model of Fudenberg & Harris (1992), the evolution of the population is determined by the difference between the fitness of each strategy and the average fitness. Consequently, for these models the parameter w would affect only the speed of evolution, but would not influence the long-run behavior as long as $w > 0$. For finite population models, however, the parameter w does have an impact on the long-run behavior. Note that the case $w = 0$ corresponds to neutral selection.

We now define a Wright-Fisher process with frequency dependent fitness to describe the evolution of the finite population. We assume throughout that all payoffs are positive. Suppose that in the current generation i individuals use strategy A . Then the composition of the next generation is determined through N independent binomial trials, where in each trial, the probability of producing an A -player is given by $if_i / (if_i + (N - i)g_i)$. Let $X(n)$ denote the number of A -players in the n th generation. Then $\{X(n)\}$ is a discrete-time Markov chain with state space $\{0, \dots, N\}$ and transition probabilities

$$p_{ij} = \binom{N}{j} \left(\frac{if_i}{if_i + (N - i)g_i} \right)^j \left(\frac{(N - i)g_i}{if_i + (N - i)g_i} \right)^{N-j}. \quad (2)$$

The states $1, \dots, N - 1$ are transient, and the states 0 and N are absorbing. For every initial configuration of the population, the process $\{X(n)\}$ will reach one of the absorbing states in finite time and will then stay there forever. That is, within finite time, the whole population will use the same strategy.

Let x_i denote the probability that the process, starting from $X(0) = i$, ends up in state N . Of particular interest is $\rho_{AB} = x_1$, the probability that a single invading

A -player can take over a population of B -players. If x_1 is larger than the corresponding fixation probability for a neutral mutant, we say that selection favors A replacing B .

To emphasize the dependence on w , we write $x_i = x_i(w)$. If $w = 0$, the process $\{X(n)\}$ is a martingale, and the absorption probabilities are given by

$$x_i(0) = \frac{i}{N}. \quad (3)$$

Thus selection favors A replacing B if $\rho_{AB} = x_1(w) > 1/N$, and selection favors B replacing A if $\rho_{BA} = 1 - x_{N-1}(w) > 1/N$. For every $w \in [0, 1]$, the probabilities $x_i(w)$ are the unique solution of the linear equations

$$x_i(w) - \sum_{j=1}^{N-1} p_{ij}x_j(w) = p_{iN}, \quad i = 1, \dots, N-1. \quad (4)$$

We now describe the relation between the fixation probabilities ρ_{AB} , ρ_{BA} and the fitness differences $f_i - g_i$. While the fixation probabilities involve global aspects of the dynamics and are difficult to evaluate, the fitness differences reflect only local properties and are simple functions of the underlying game matrix and the population size. Note that

$$f_1 - g_1 = w \left(b - d + \frac{d - c}{N - 1} \right), \quad f_{N-1} - g_{N-1} = w \left(a - c + \frac{b - a}{N - 1} \right).$$

We say that selection favors A invading B if $f_1 > g_1$, and say that selection favors B invading A if $f_{N-1} < g_{N-1}$. The next theorem holds for any fixed $w \in (0, 1]$ and every population size $N \geq 2$.

Theorem 1. *a) If $f_1 \geq g_1$ and $f_{N-1} \geq g_{N-1}$ with at least one inequality being strict, then $x_k > k/N$ for all $k = 1, \dots, N-1$, in particular, $\rho_{AB} > 1/N$ and $\rho_{BA} < 1/N$.*

b) If $f_1 \leq g_1$ and $f_{N-1} \leq g_{N-1}$ with at least one inequality being strict, then $x_k < k/N$ for all $k = 1, \dots, N-1$, in particular, $\rho_{AB} < 1/N$ and $\rho_{BA} > 1/N$.

Thus if selection favors A invading B and opposes B invading A , then for every non-degenerate initial composition of the population, the probability of fixation at all A is larger than the benchmark obtained by neutral selection. We then say that selection favors A . In particular, in this case, A replacing B is favored whereas B replacing A is opposed by selection.

Proof of Theorem 1. We are interested in the signs of the differences

$$\delta_i = x_i(w) - x_i(0) = x_i(w) - \frac{i}{N}, \quad i = 1, \dots, N-1.$$

From (4),

$$\begin{aligned} \delta_i - \sum_{j=1}^{N-1} p_{ij}\delta_j &= -\frac{i}{N} + \sum_{j=0}^N p_{ij} \frac{j}{N} = -\frac{i}{N} + \frac{if_i}{if_i + (N-i)g_i} \\ &= \frac{i(N-i)(f_i - g_i)}{N(if_i + (N-i)g_i)}. \end{aligned} \quad (5)$$

Every entry of the matrix $\tilde{P} = (p_{ij})_{i,j=1}^{N-1}$ is positive and since every row sum of \tilde{P} is strictly less than 1, so is the spectral radius of \tilde{P} . If $f_1 \geq g_1$ and $f_{N-1} \geq g_{N-1}$ with at least one inequality being strict, then the expression on the right-hand side of (5) is non-negative and does not vanish identically. It now follows by Theorem 2.1 in Seneta (1981) that the solution satisfies $\delta_i > 0$ for all i , proving a). Part b) follows by symmetry. \square

Corollary 1. *Suppose $c < b$. Then there exists a number $N_0 \geq 2$, such that if the population size satisfies $N \leq N_0$, selection favors A in the sense that $x_i(w) > i/N$, $i = 1, \dots, N-1$.*

This is obvious from Theorem 1 and the fact that if $N = 2$, then $f_1 - g_1 = f_{N-1} - g_{N-1} = w(b - c)$. Since Theorem 1 gives only a sufficient condition for selection to favor A, the maximal N_0 for which that condition is satisfied will in general be smaller than the maximal population size for which selection favors A. If selection is weak, the precise threshold can be obtained from Theorem 3 a) in the next section.

We say that selection favors change if it favors both A replacing B and B replacing A. Selection opposes change if neither A replacing B nor B replacing A is favored. The next theorem gives necessary conditions for selection to favor or oppose change.

Theorem 2. *Let $N \geq 3$.*

- a) *If $\rho_{AB} < 1/N$ and $\rho_{BA} < 1/N$, then $f_1 < g_1$ and $f_{N-1} > g_{N-1}$.*
- b) *If $\rho_{AB} > 1/N$ and $\rho_{BA} > 1/N$, then $f_1 > g_1$ and $f_{N-1} < g_{N-1}$.*

In words, if selection opposes change, then selection must already oppose A and B invading each other. Selection can favor change only if mutual invasion is favored. Note that if $N = 2$, then $\rho_{AB} = 1 - \rho_{BA}$, so that neither case a) nor case b) of Theorem 2 can occur. This explains why only populations of size $N \geq 3$ are considered here.

The proof of Theorem 2 uses the concept of total positivity, see Karlin (1968). A matrix is called totally positive if all its minors of every order are non-negative.

Lemma 1. *The transition matrix $(p_{ij})_{i,j=0}^N$ with entries p_{ij} given by (2) is totally positive.*

The proof of the lemma is in the appendix.

Proof of Theorem 2. We prove only a); the proof of b) is similar. Thus let

$$\rho_{AB} < \frac{1}{N}, \quad \rho_{BA} < \frac{1}{N}. \quad (6)$$

The fitness difference $f_i - g_i$ is an affine function of i , that is, $f_i - g_i$ is of the form $\alpha + \beta i$. Therefore, if $f_1 = g_1$ and $f_{N-1} = g_{N-1}$, then $f_i = g_i$ for every i . This would imply $\rho_{AB} = \rho_{BA} = 1/N$, contradicting (6). Hence $f_1 \neq g_1$ or $f_{N-1} \neq g_{N-1}$. It now follows from (6) and Theorem 1 that

$$(f_1 < g_1 \text{ and } f_{N-1} > g_{N-1}) \quad \text{or} \quad (f_1 > g_1 \text{ and } f_{N-1} < g_{N-1}). \quad (7)$$

For an indirect argument assume that $f_1 > g_1$ and $f_{N-1} < g_{N-1}$. Let $\delta_i = x_i(w) - i/N$. It was shown in the proof of Theorem 1 that $\delta = (\delta_1, \dots, \delta_{N-1})^T$ is the solution of $[I - \tilde{P}]\delta = h$, where $\tilde{P} = (p_{ij})_{i,j=1}^{N-1}$, $h = (h_1, \dots, h_{N-1})^T$ and $h_i = i(N-i)(f_i - g_i)/(N(if_i + (N-i)g_i))$, $i = 1, \dots, N-1$. As the spectral radius of \tilde{P} is strictly less than 1, $[I - \tilde{P}]^{-1} = \sum_{k=0}^{\infty} \tilde{P}^k$. Hence

$$\delta = \sum_{k=0}^{\infty} q^{(k)}, \quad (8)$$

where $q^{(k)} = (q_1^{(k)}, \dots, q_{N-1}^{(k)})^T := \tilde{P}^k h$.

It follows from Lemma 1 that \tilde{P} , a submatrix of the transition matrix, is totally positive. Thus, in view of the Binet-Cauchy formula, every power \tilde{P}^k is totally positive as well. The assumption that $f_1 > g_1$ and $f_{N-1} < g_{N-1}$ implies that the sequence h_1, \dots, h_{N-1} has exactly one sign change from + to -. It follows from the theorem on variation-diminishing properties of totally positive matrices (Karlin 1968, page 233) that for every k , the sequence $q_1^{(k)}, \dots, q_{N-1}^{(k)}$ has either no sign change or exactly one, which must be from non-negative to non-positive. Specifically, for every k , $q_i^{(k)} \geq 0$ for all i , or $q_i^{(k)} \leq 0$ for all i , or else there exists some $i_0 = i_0(k)$ such that $q_i^{(k)} \geq 0$ for $i \leq i_0$ and $q_i^{(k)} \leq 0$ for $i > i_0$. Let

$$k_0 = \inf \left\{ k \in \mathbb{N}_0 : \begin{aligned} & q_i^{(k)} \geq 0 \text{ for all } i = 1, \dots, N-1 \text{ or} \\ & (q_i^{(k)} \leq 0 \text{ for all } i = 1, \dots, N-1) \end{aligned} \right\},$$

where $\inf \emptyset = \infty$. Then, for every $k < k_0$, $q_1^{(k)} \geq 0$ and $q_{N-1}^{(k)} \leq 0$. Thus if $k_0 = \infty$, then, by (8), $\delta_1 \geq 0$, that is, $\rho_{AB} \geq 1/N$, contradicting (6). Therefore, $k_0 < \infty$. If $q_i^{(k_0)} \geq 0$ for all i , then for every $k \geq k_0$, $q^{(k)} = \tilde{P}^{k-k_0} q^{(k_0)}$ is a non-negative vector. In particular, $q_1^{(k)} \geq 0$ for all $k \geq 0$, so that again $\rho_{AB} \geq 1/N$, contradicting (6). A similar argument shows that if $q_i^{(k_0)} \leq 0$ for all i , then $\rho_{BA} \geq 1/N$, contradicting (6). Thus the assumption $f_1 > g_1$ and $f_{N-1} < g_{N-1}$ must have been false, and it follows from (7) that $f_1 < g_1$ and $f_{N-1} > g_{N-1}$. \square

3. Weak Selection

Let us study the fixation probabilities when selection is weak, that is, when w is close to zero. Observe that if the derivative $x'_i(0)$ is positive, then selection favors A replacing B for $w > 0$ sufficiently small. Set $y_i = x'_i(0)$ and let f'_i and g'_i denote the derivatives of f_i and g_i with respect to w . Then y_1, \dots, y_{N-1} are the unique solution of the linear system

$$y_i - \sum_{j=1}^{N-1} p_{ij}(0)y_j = (f'_i - g'_i) \frac{i}{N} \left(1 - \frac{i}{N} \right), \quad i = 1, \dots, N-1. \quad (9)$$

Indeed, from (4) and (3),

$$y_i - \sum_{j=1}^{N-1} p_{ij}(0)y_j = \sum_{j=1}^{N-1} p'_{ij}(0)x_j(0) + p'_{iN}(0) = \frac{1}{N} \sum_{j=0}^N j p'_{ij}(0).$$

But $\sum_{j=0}^N j p_{ij}(w)$ is just the expectation of a binomial random variable with parameters N and $i f_i / (i f_i + (N - i) g_i)$, so that

$$\frac{1}{N} \sum_{j=0}^N j p'_{ij}(0) = \frac{\partial}{\partial w} \frac{i f_i}{i f_i + (N - i) g_i} \Big|_{w=0} = (f'_i - g'_i) \frac{i}{N} \left(1 - \frac{i}{N}\right).$$

This shows that y_1, \dots, y_{N-1} are a solution of (9). The solution is unique since the spectral radius of the matrix $(p_{ij}(0))_{i,j=1}^{N-1}$ is strictly less than 1. It is shown in the appendix that the solution is given by

$$x'_i(0) = \frac{i(N-i)}{N-1} \left\{ \frac{a-b-c+d}{3N-2} (N+i-1) - \left(d-b + \frac{a-d}{N} \right) \right\}. \quad (10)$$

The next theorem gives, for weak selection, a complete classification of the fixation probabilities ρ_{AB} and ρ_{BA} in terms of a, b, c, d and N .

Theorem 3. *Let*

$$\begin{aligned} \alpha &= 4(b-c) + (a+6b-4c-3d)(N-2) + (a+2b-c-2d)(N-2)^2, \\ \beta &= 4(b-c) + (3a+4b-6c-d)(N-2) + (2a+b-2c-d)(N-2)^2. \end{aligned}$$

The following assertions hold for $w > 0$ sufficiently small.

- a) *If $\alpha > 0$ and $\beta > 0$, then selection favors A: $x_i(w) > i/N$ for all $i = 1, \dots, N-1$. If $\alpha < 0$ and $\beta < 0$, then selection favors B: $x_i(w) < i/N$ for all $i = 1, \dots, N-1$.*
- b) *If $\alpha > 0$ and $\beta < 0$, then selection favors change: $\rho_{AB} > 1/N$ and $\rho_{BA} > 1/N$.*
- c) *If $\alpha < 0$ and $\beta > 0$, then selection opposes change: $\rho_{AB} < 1/N$ and $\rho_{BA} < 1/N$.*

Proof. We have $\alpha = (3N-2)N x'_1(0)$ and $\beta = (3N-2)N x'_{N-1}(0)$. Thus if $\alpha > 0$ and $\beta > 0$, then $x'_1(0) > 0$ and $x'_{N-1}(0) > 0$. It is obvious from (10) that this implies that $x'_i(0) > 0$ for all $i = 1, \dots, N-1$. Therefore, for $w > 0$ sufficiently small, $x_i(w) > x_i(0) = i/N$ for all i . The other cases are similar. \square

We now turn to games that have two strict Nash equilibria and apply Theorem 3 to analyse the Wright-Fisher process for large but finite populations. The result is surprising and surprisingly simple: Let $x^* = (d-b)/(a-b-c+d)$, which is the invasion barrier in the infinite population case. If $x^* < 1/3$, selection favors A; if $x^* > 2/3$, selection favors B; otherwise selection opposes change. The last statement remains true even if not just one mutant but a whole group of mutants tries to invade, provided the population is large enough in comparison to the group. Moreover, whether selection opposes change from A to B more strongly than change from B to A depends simply on whether $x^* < 1/2$.

Write $\rho_{AB}(i) = x_i$ and $\rho_{BA}(i) = 1 - x_{N-i}$.

Theorem 4. *Suppose $a > c$ and $d > b$. Let i_0 be a fixed group size. Then there exists $N_0 \geq 2$ such that for every population size $N \geq N_0$ the following holds, provided selection is sufficiently weak, that is, $0 < w \leq w_0(N)$.*

a) If $x^* < 1/3$, then

$$\rho_{BA}(i) < \frac{i}{N} < \rho_{AB}(i), \quad i = 1, \dots, N-1.$$

b) If $1/3 < x^* < 1/2$, then

$$\rho_{BA}(i) < \rho_{AB}(i) < \frac{i}{N}, \quad i = 1, \dots, i_0.$$

c) If $1/2 < x^* < 2/3$, then

$$\rho_{AB}(i) < \rho_{BA}(i) < \frac{i}{N}, \quad i = 1, \dots, i_0.$$

d) If $2/3 < x^*$, then

$$\rho_{AB}(i) < \frac{i}{N} < \rho_{BA}(i), \quad i = 1, \dots, N-1.$$

Proof. Note first that $x^* < 1/3$ if and only if the leading coefficient $a + 2b - c - 2d$ of the polynomial defining α in Theorem 3 is positive; and $x^* < 2/3$ if and only if the leading coefficient of β is positive. Thus if $x^* < 1/3$, then both $\alpha > 0$ and $\beta > 0$ for N sufficiently large. Part a) is therefore a consequence of Theorem 3 a). Part d) follows along similar lines.

To prove b) assume that $1/3 < x^* < 1/2$. This is equivalent to $a + 2b - c - 2d < 0 < a + b - c - d$. Using the explicit expression (10) for $x'_i(0)$ we obtain that for every fixed i ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \left\{ \frac{d}{dw} \rho_{AB}(i) \Big|_{w=0} \right\} &= \lim_{N \rightarrow \infty} x'_i(0) = \frac{i}{3}(a + 2b - c - 2d) < 0, \\ \lim_{N \rightarrow \infty} \left\{ \frac{d}{dw} [\rho_{AB}(i) - \rho_{BA}(i)] \Big|_{w=0} \right\} &= \lim_{N \rightarrow \infty} x'_i(0) + x'_{N-i}(0) \\ &= i(a + b - c - d) > 0. \end{aligned}$$

Consequently, there exists a number N_0 such that $(d/dw)\rho_{AB}(i)|_{w=0} < 0$ and $(d/dw)[\rho_{AB}(i) - \rho_{BA}(i)]|_{w=0} > 0$ for every $N \geq N_0$ and every $i = 1, \dots, i_0$. For these N and i , $\rho_{AB}(i) < 1/N$ and $\rho_{AB}(i) - \rho_{BA}(i) > 0$, provided $0 < w \leq w_0(N)$. This proves b). The proof of c) is analogous. \square

Suppose that in the situation of Theorem 4, $x^* < 1/3$. By solving two quadratic equations one may explicitly determine a number N_0 such that for every $N \geq N_0$ both α and β are positive. The assertion under part a) of Theorem 4 holds as soon as $N \geq N_0$. In the remaining cases it is likewise straightforward to compute explicitly a value N_0 such that the inequalities for the fixation probabilities hold whenever $N \geq N_0$.

Example. Consider a game with payoff matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 22 & 2 \\ 1 & 12 \end{bmatrix}.$$

Then $b > c$ and $a + 2b > c + 2d$, so that $x^* < 1/3$. Under weak selection, the frequency dependent Moran process introduced in Nowak et al. (2004) will favor A replacing B for every population size N . In the Wright-Fisher model, however, selection favors A replacing B for $N = 2$ and $N \geq 8$, but opposes A replacing B for $N = 3, \dots, 7$.

We next present a dual result to Theorem 4 for games with an interior equilibrium point x^* . If $x^* < 1/3$, selection now favors B ; if $x^* > 2/3$, selection favors A . Otherwise selection favors change, and whether selection favors change from A to B more strongly than change from B to A depends simply on whether $x^* < 1/2$.

Theorem 5. *Suppose $a < c$ and $d < b$. Let i_0 be a fixed group size. Then there exists $N_0 \geq 2$ such that for every population size $N \geq N_0$ the following holds, provided selection is sufficiently weak, that is, $0 < w \leq w_0(N)$.*

a) *If $x^* < 1/3$, then*

$$\rho_{AB}(i) < \frac{i}{N} < \rho_{BA}(i), \quad i = 1, \dots, N - 1.$$

b) *If $1/3 < x^* < 1/2$, then*

$$\frac{i}{N} < \rho_{AB}(i) < \rho_{BA}(i), \quad i = 1, \dots, i_0.$$

c) *If $1/2 < x^* < 2/3$, then*

$$\frac{i}{N} < \rho_{BA}(i) < \rho_{AB}(i), \quad i = 1, \dots, i_0.$$

d) *If $2/3 < x^*$, then*

$$\rho_{BA}(i) < \frac{i}{N} < \rho_{AB}(i), \quad i = 1, \dots, N - 1.$$

The proof of Theorem 5 is similar to that of Theorem 4 and is therefore omitted. Note that it is again straightforward to compute explicitly a threshold N_0 such that the claimed inequalities hold for every $N \geq N_0$.

4. A Different Playing Scheme

In calculating the fitness of A - and B -players we have so far used the expected payoffs for random mating. That is, we assumed that between two consecutive time steps of the Markov chain, either the game is infinitely often played or everyone plays everyone else exactly once. A more realistic approach would be to assume that the game is played by finitely many randomly chosen pairs. Then the realized payoffs are stochastic, since they depend on what pairs have been chosen, and so the transition probabilities of the Wright-Fisher process become more involved.

In this section we study the simple case where at each time point n only one randomly chosen pair plays the game. Every member of the population has fitness equal to 1, except for the two players chosen. Their fitness is determined by the payoffs of the game. More generally, the fitness of the two players is a convex combination of their initial fitness, also equal to 1, and the payoffs. The degree to which the payoffs contribute to fitness is measured by $w \in [0, 1]$. To calculate the transition probabilities for a frequency-dependent Wright-Fisher process based on this mating scheme, suppose that in the present population i players use strategy A and $N - i$ players use B . Then the probabilities that a pair of A -players, a mixed pair, or a pair of B -players is chosen are, respectively,

$$\pi_{AA}(i) = \frac{i(i-1)}{N(N-1)}, \quad \pi_{AB}(i) = 2 \frac{i(N-i)}{N(N-1)}, \quad \pi_{BB}(i) = \frac{(N-i)(N-i-1)}{N(N-1)}.$$

If two A -players are chosen, the total fitness of all the A -players is $i - 2 + 2(1 - w + wa) = i + 2w(a - 1)$ and the total fitness of the population is $N + 2w(a - 1)$. In this case, the next generation has a binomial distribution with parameters N and

$$\beta(i|AA) = \frac{i + 2w(a - 1)}{N + 2w(a - 1)}.$$

The corresponding parameters for the two other possible pairs are

$$\beta(i|AB) = \frac{i + w(b - 1)}{N + w(b + c - 2)}, \quad \beta(i|BB) = \frac{i}{N + 2w(d - 1)}.$$

The transition probabilities are mixtures of three binomial distributions:

$$\begin{aligned} p_{ij} = & \binom{N}{j} \{ \pi_{AA}(i) \beta(i|AA)^j [1 - \beta(i|AA)]^{N-j} \\ & + \pi_{AB}(i) \beta(i|AB)^j [1 - \beta(i|AB)]^{N-j} \\ & + \pi_{BB}(i) \beta(i|BB)^j [1 - \beta(i|BB)]^{N-j} \} \end{aligned}$$

Let

$$\begin{aligned} \alpha &= b - d + \frac{d - c + w(b - c)(d - 1)}{N - 1}, \\ \gamma &= a - c + \frac{b - a + w(a - 1)(b - c)}{N - 1}. \end{aligned}$$

Theorem 6. *a) If $\alpha > 0$ and $\gamma > 0$, then $x_i > i/N$ for all $i = 1, \dots, N - 1$, in particular $\rho_{AB} > 1/N$ and $\rho_{BA} < 1/N$.*

b) If $\alpha < 0$ and $\gamma < 0$, then $x_i < i/N$ for all $i = 1, \dots, N - 1$, in particular $\rho_{AB} < 1/N$ and $\rho_{BA} > 1/N$.

Proof. Set $h(i) = E[X_1 | X_0 = i] - i$. As there are no mutations, $h(0) = h(N) = 0$. Furthermore,

$$h(i) = N \{ \pi_{AA}(i) \beta(i|AA) + \pi_{AB}(i) \beta(i|AB) + \pi_{BB}(i) \beta(i|BB) \} - i,$$

which shows that h is polynomial of degree 3. Thus h can have at most one sign change in $(0, N)$. We have

$$h(1) = \frac{2(N-1)w\alpha}{(N+w(b+c-2))(N+2w(d-1))},$$

$$h(N-1) = \frac{2(N-1)w\gamma}{(N+w(b+c-2))(N+2w(a-1))}.$$

Hence if $\alpha > 0$ and $\gamma > 0$, then $h(1) > 0$ and $h(N-1) > 0$. This in turn implies that $h(i) > 0$ for all $i = 1, \dots, N-1$. That is, $\{X_n\}$ is a submartingale, and it follows from the optional stopping theorem (Karlin & Taylor 1975) that

$$i < E[X_1|X_0 = i] \leq E[X_\tau|X_0 = i] = NP\{X_\tau = N|X_0 = i\} = Nx_i,$$

where $\tau = \inf \{n : X_n \in \{0, N\}\}$. This proves a). The proof of b) is similar. \square

All the results for weak selection presented in Section 3 hold unchanged for the mating scheme of the present section. To see this set $y_i = x'_i(0)$. Then, for $i = 1, \dots, N-1$,

$$\begin{aligned} y_i - \sum_{j=1}^{N-1} p_{ij}(0)y_j &= \frac{1}{N} \sum_{j=1}^N j p'_{ij}(0) \\ &= \frac{d}{dw} \{ \pi_{AA}(i)\beta(i|AA) + \pi_{AB}(i)\beta(i|AB) \\ &\quad + \pi_{BB}(i)\beta(i|BB) \}_{w=0} \\ &= \frac{2i(N-i)}{N^3(N-1)} \{ d-a + (a-b-c+d)i + (b-d)N \} \\ &= \frac{2}{N} \frac{i}{N} \left(1 - \frac{i}{N} \right) (f'_i - g'_i). \end{aligned}$$

Thus the derivatives of the current fixation probabilities satisfy exactly the same system of equations as those in Section 3, apart from an unimportant constant factor $2/N$. It is therefore obvious that Theorems 3 to 5 hold in the present situation as well.

5. Discussion

In this paper, we have studied a Wright-Fisher process with frequency-dependent selection in order to investigate game dynamics in finite populations with discrete generations. We have compared the probability ρ_{AB} that a single individual using strategy A takes over a population of B -players with the corresponding probability under neutral drift, which is $1/N$. We say that selection favors A replacing B if $\rho_{AB} > 1/N$, selection favors change if $\rho_{AB} > 1/N$ and $\rho_{BA} > 1/N$, and selection opposes change if $\rho_{AB} < 1/N$ and $\rho_{BA} < 1/N$. In the case of strong selection, we have derived simple sufficient conditions for selection to favor one strategy over the other and necessary conditions for selection to favor or oppose change, respectively.

For weak selection, we have obtained a complete characterization for selection to favor one strategy, or to favor or to oppose change. The characterization involves only two simple quadratic polynomials in N . We have shown in an example that the frequency dependent Wright-Fisher process can behave very differently from the frequency dependent Moran process studied by Nowak et al. (2004) when the population is small. However, the characterization yields that for sufficiently large (but finite) populations, the ‘1/3 rule’ of the Moran process also holds for the Wright-Fisher process: in the coordination case, selection favors A replacing B if the unstable equilibrium point is less than $1/3$. Our result also shows that in a finite population model, the standard ESS condition does not imply protection against invasion and replacement. This is in stark contrast to the standard model for infinite populations, the replicator dynamics, where every ESS is proof against invasion.

6. Appendix

Proof of Lemma 1. The i th row of the transition matrix describes a binomial distribution with parameters N and $\pi_i = if_i/(if_i + (N - i)g_i)$. We will show that $\pi_0 < \pi_1 < \dots < \pi_N$. This implies that the matrix is totally positive, see Karlin (1968), page 19. Note that $\pi_i < \pi_{i+1}$ is equivalent to $E[X_1|X_0 = i] < E[X_1|X_0 = i + 1]$, which seems reasonable but needs to be verified. Regarding i as a continuous variable, one has

$$\frac{d}{di}\pi_i = \frac{(N - i)^2 f_1 g_0 + [(N - 1)ig_i + (N - i)(i - 1)g_0]f_N}{(N - 1)[if_i + (N - i)g_i]^2} > 0. \quad \square$$

Proof of (10). To solve (9), consider candidates of the form

$$y_i^{(1)} = i(N - i), \quad y_i^{(2)} = i(N^2 - i^2), \quad i = 1, \dots, N - 1.$$

Using the moment formulas for the binomial distribution, in particular

$$\sum_{j=0}^N j^3 \binom{N}{j} \left(\frac{i}{N}\right)^j \left(\frac{N - i}{N}\right)^{N - j} = \frac{(N - 1)(N - 2)}{N^2} i^3 + 3 \frac{N - 1}{N} i^2 + i,$$

one may verify that

$$y_i^{(1)} - \sum_{j=1}^{N-1} p_{ij}(0) y_j^{(1)} = \frac{i(N - i)}{N},$$

$$y_i^{(2)} - \sum_{j=1}^{N-1} p_{ij}(0) y_j^{(2)} = \frac{i(N - i)}{N} \left\{ 1 + \left(3 - \frac{2}{N} \right) i \right\}.$$

As the right-hand side of (9) coincides with a third degree polynomial in i that vanishes at $i = 0$ and $i = N$, it follows that the solution can be written in the form $y_i = c_1 y_i^{(1)} + c_2 y_i^{(2)}$. A somewhat tedious calculation shows that

$$c_1 = -\frac{1}{N - 1} \left\{ \frac{a - b - c + d}{3N - 2} + d - b + \frac{a - d}{N} \right\}, \quad c_2 = \frac{a - b - c + d}{(N - 1)(3N - 2)}. \quad \square$$

Acknowledgements. We wish to thank the referee for several comments that helped to improve the presentation of the manuscript. Parts of the paper were written during a visit of the first author at Harvard University. This author would like to thank the Program for Evolutionary Dynamics for its hospitality. The Program for Evolutionary Dynamics at Harvard University is sponsored by Jeffrey Epstein.

References

- Antal, T., Scheuring, I.: Fixation of strategies for an evolutionary game in finite populations. Mimeo, Center for Polymer Studies and Department of Physics, Boston University, 2005
- Binmore, K.: *Game Theory and the Social Contract: Playing Fair*. MIT Press, Cambridge, MA, 1993
- Binmore, K.: *Game Theory and the Social Contract: Just Playing*. MIT Press, Cambridge, MA, 1998
- Bürger, R.: *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley, Chichester, 2000
- Corradi, V., Sarin, R.: Continuous approximations of stochastic evolutionary game dynamics. *J. Econom. Theory* **94**, 163–191 (2000)
- Dostálková, I., Kindlmann, P.: Evolutionarily stable strategies for stochastic processes. *Theoret. Population Biol.* **65**, 205–210 (2004)
- Ewens, W.J.: *Mathematical Population Genetics, I. Theoretical Introduction*. Second edition. Springer, New York, 2004
- Ficici, S., Pollack, J.: Effects of Finite Populations on Evolutionary Stable Strategies. Proceedings of the 2000 Genetic and Evolutionary Computation Conference, L. Darrell Whitley (ed.), Morgan-Kaufmann (2000)
- Fogel, G., Andrews, P., Fogel, D.: On the instability of evolutionary stable strategies in small populations. *Ecological Modeling* **109**, 283–294 (1998)
- Foster, D., Young, P.: Stochastic evolutionary game dynamics. *Theoret. Population Biol.* **38**, 219–232 (1990)
- Fudenberg, D., Harris, C.: Evolutionary dynamics with aggregate shocks. *J. Econom. Theory* **57**, 420–441 (1992)
- Fudenberg, D., Imhof, L.A., Nowak, M.A., Taylor, C.: Evolutionary game dynamics in finite populations with strong selection and weak mutation. Mimeo, Program for Evolutionary Dynamics, Harvard University, 2004.
- Fudenberg, D., Levine, D.K.: *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998
- Fudenberg, D., Tirole, J.: *Game Theory*. MIT Press, Cambridge, MA, 1991
- Gintis, H.: *Game Theory Evolving*. Princeton University Press, 2000
- Hofbauer, J., Schuster, P., Sigmund, K.: A note on evolutionarily stable strategies and game dynamics. *J. Theoret. Biol.* **81**, 609–612 (1979)
- Hofbauer, J., Sigmund, K.: *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998
- Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. *Bull. Amer. Math. Soc.* **40**, 479–519 (2003)
- Imhof, L.A.: The long-run behavior of the stochastic replicator dynamics. *Ann. Appl. Probab.* **15**, 1019–1045 (2005a)
- Imhof, L.A.: Evolutionary game dynamics under stochastic influences. *Jahresber. Deutsch. math.-verein.* **107**, 197–213 (2005b)
- Imhof, L.A., Fudenberg D., Nowak, M.A.: Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA* **102**, 10797–10800 (2005)
- Kandori, M., Mailath, G., Rob, R.: Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56 (1993)
- Karlin, S.: *Total Positivity*. Stanford University Press, 1968
- Karlin, S., Taylor, H.M.: *A First Course in Stochastic Processes*, Second edition. Academic Press, London, 1975

- Maynard Smith, J.: *Evolution and the Theory of Games*, Cambridge University Press, 1982
- Moran, P.A.P.: *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford, 1962
- Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D.: Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004)
- Nowak, M.A., Sigmund, K.: Evolutionary dynamics of biological games. *Science* **303**, 793–799 (2004)
- Samuelson, L.: *Evolutionary Games and Equilibrium Selection*. MIT Press, Cambridge, MA, 1997
- Schaffer, M.: Evolutionarily stable strategies for a finite population and a variable contest size. *J. Theoret. Biol.* **132**, 469–478 (1988)
- Schreiber, S.J.: Urn models, replicator processes, and random genetic drift. *SIAM J. Appl. Math.* **61**, 2148–2167 (2001)
- Seneta, E.: *Non-negative Matrices and Markov Chains*. Second edition. Springer, New York, 1981
- Taylor, C., Fudenberg, D., Sasaki, A., Nowak, M.A.: Evolutionary game dynamics in finite populations. *Bull. Math. Biol.* **66**, 1621–1644 (2004)
- Taylor, P.D., Jonker, L.B.: Evolutionarily stable strategies and game dynamics. *Math. Biosciences* **40**, 145–156 (1978)
- Weibull, J.: *Evolutionary Game Theory*. MIT Press, Cambridge, MA, 1995
- Wild, G., Taylor, P.D.: Fitness and evolutionary stability in game theoretic models of finite populations. *Proc. R. Soc. Lond. B* **271**, 2345–2349 (2004)