Preprint manuscript No.
(will be inserted by the editor)

A general model of best response adaptation

Ulrich Berger

Vienna University of Economics, Department VW5
Augasse 2-6, A-1090 Vienna, Austria
e-mail: ulrich.berger@wu-wien.ac.at

November 2002 (first version: March 2002)

Abstract We develop a general model of best response adaptation in large populations for symmetric and asymmetric conflicts with role-switching. For special cases including the classical best response dynamics and the symmetrized best response dynamics we show that the set of Nash equilibria is attracting for zero-sum games. For asymmetric conflicts and equally large populations, convergence to a Nash equilibrium in the base game implies convergence to a Nash equilibrium on the Wright manifold in the role game.

Journal of Economic Literature classification numbers: C72, D83.

Key words Role Games, Best Response Adaptation, Learning, Evolution.

1 Evolutionary Game Dynamics

Solution concepts from non-cooperative game theory, i.e. Nash equilibrium and its refinements, form the essential basis of many economic models. It has been shown, however, that very strong rationality and consistency assumptions are necessary to justify Nash equilibrium behavior. In recent years another approach to behavior consistent with Nash equilibrium predictions has been advocated by evolutionary game theorists. The idea is that social or competition-based selection or learning processes guide the behavior of individuals. The question then is, if in the long run this kind of evolution leads to behavior corresponding to Nash equilibrium or a refinement thereof.

The most prominent deterministic dynamics in evolutionary game theory is the *replicator dynamics* (Taylor and Jonker, 1978), which stems from the biological sciences. The underlying assumption is that individuals are programmed to some pure strategy which their offspring inherits. The number of offsprings is determined by the success the strategy has when played against other individuals. The dynamic equations arising from these assumptions have been extensively studied, see Hofbauer and Sigmund (1998) or Weibull (1995). It can e.g. be shown that Nash equilibria are fixed points for the replicator dynamics, stable fixed points are Nash equilibria, and asymptotically stable fixed points are perfect Nash equilibria.

The replicator dynamics is undoubtedly a reasonable model in a biological context, but it is much less clear if it also appropriately expresses evolution in a social or economic environment. Although it can be shown to arise for special rules of imitation (Björnerstedt and Weibull, 1996, and Schlag, 1998) or learning (Börgers and Sarin, 1997), economists are interested in dynamic processes which allow for a greater variety of how strategies are adjusted in populations of interacting agents. Classes of such alternative dynamics have been studied since the early 1990s, among them the well known classes of *payoff positive*, *weakly payoff positive*, or *payoff monotonic* dynamics, see Nachbar (1990), Friedman (1991) and Samuelson and Zhang (1992). These are all so-called *selection dynamics*, which implies that initially unused strategies in a population will never appear. Dynamic processes for which the appearance of new strategies is not ruled out are called *innovative dynamics*, examples are the *Brown-von Neumann-Nash dynamics* (see Berger and Hofbauer, 2002) and the more prominent *best response dynamics* (Matsui, 1992). In this paper we concentrate on a generalization of the latter, which is particularly appealing for asymmetric conflicts.

2 Best Response Dynamics

The best response dynamics arises as a continuous approximation of the discrete fictitious play learning process, which is due to Brown (1951) and Robinson (1951). For a symmetric game with n pure strategies a sequence

$(x(t))_{t \geq 0}$ is a fictitious play process, if $x(0)$ is an element of the $(n-1)$-dimensional probability simplex S_n and

$$x(t+1) = \frac{tx(t) + b(x(t))}{t+1} \qquad \text{for } t \geq 0 \text{ and } b(x(t)) \in B(x(t)),$$

where $B(x)$ is the set of best responses to the empirical strategy distribution x.

Originally this learning process was proposed as an algorithm for calculating the value of a two-person zero-sum game. It is also known as the standard learning process for boundedly rational, myopic agents. The definition given above implies that each player plays a best response to the empirical distribution of his opponent's past play.

Robinson (1951) proved that under fictitious play the set of Nash equilibria is globally attractive for zero-sum games, and Miyasawa (1961) proved convergence for 2×2-games under a particular tie-breaking rule. For games with more than two strategies per player convergence need not occur, however (Shapley, 1964).

The cited results also hold for the continuous fictitious play process

$$\dot{x}(t) = \frac{b(x(t)) - x(t)}{t},$$

(Brown, 1951, see also Rosenmüller, 1971), and for the best response dynamics, which differs from this process only by a rescaling of time and (suppressing the time variable) reads

$$\dot{x} = b(x) - x, \qquad b(x) \in B(x). \tag{1}$$

The analogous best response dynamics for an $n\times m$ bimatrix game is given by

$$\dot{z} = b(y) - z, \quad \dot{y} = b(z) - y, \tag{2}$$

where $(z, y) \in S_n \times S_m$ and $b(.) \in B(.)$.

The best response dynamics was introduced by Gilboa and Matsui (1991) and Matsui (1992), and further analyzed by Hofbauer (1995). It is equivalent to continuous fictitious play, but its interpretation is rather different. The learning process interpretation of fictitious play assumes that the same two players play some game repeatedly. This interpretation of the fictitious play process has been criticized because of the inherent conflict between its rationality assumptions on the one hand (players calculate best responses) and the implied disability of players to detect patterns like cycles.[1] Moreover, convergence means convergence of the empirical distribution (beliefs), not convergence of actual play. The usual interpretation of the best response dynamics is immune to this kind of criticism, since it is based on the *population model* of game theory.

[1] See e.g. Fudenberg and Levine (1998) for a discussion.

3 The population model

The population model of game theory was already formulated by Nash (1950) in his Ph.D. thesis (see Weibull, 1994, for a quotation). For symmetric conflicts, this model assumes that there is a large population of agents. Pairs of agents are drawn randomly from this population and are matched to play some matrix game. Interactions may be anonymous, but also if they are not, the chance to meet an opponent again in the near future is so low that it does not pay to try to influence his behavior as it is common in repeated games of two fixed players as e.g. the repeated Prisoner's Dilemma game. Moreover, each agent is so small compared to the population that he can safely neglect the influence his strategy choice has on the average distribution of strategies in the population and hence on the future behavior of other agents. In this sense, myopic behavior is a plausible assumption. This is reminiscent of the price-taker assumption in the analysis of competitive equilibrium.

The population model of the best response dynamics assumes that each agent is bound to some fixed pure strategy, but every now and then a small fraction of agents review their strategy. Reviewing agents choose some pure best response to the average strategy in the population (in the opponent population, if the conflict is asymmetric). A slightly different, but equivalent, interpretation is that agents stick to their strategy for lifetime, and in each period some randomly selected agents die (leave the population) and some are born (enter the population), where each of the latter chooses some pure best response. This average strategy $x \in S_n$ is called the strategy profile or *state* of the population.

4 Asymmetric Conflicts and Role Games

For asymmetric conflicts, the standard population model supposes that there are two types of agents, corresponding to two large populations, and in each period one agent is drawn at random from each population. Then these agents are matched to play a bimatrix game with the agent from the first population as the row-player and the agent from the second population as the column-player. Common examples iclude games between males and females, buyers and sellers, or employers and employees.

The population models of symmetric and asymmetric conflicts do, however, not capture certain situations of interest. Imagine for example the "crossing game": Pairs of car drivers meet randomly at some unruled crossing. Both can either give way or drive on (with the well known consequences). The conflict seems to be symmetric, but the population model of symmetric conflicts is not appropriate for this situation, because the drivers can identify their player position in the game. One of the drivers comes from the right, the other one from the left. Hence the drivers can also condition their strategies on their player position, and this is obviously what they do in real life. In most countries the one coming from the right drives on and the other one gives way.[2] Another scenario could be one where car drivers randomly meet pedestrians wanting to cross the street. Car drivers can again give way or drive on, while pedestrians choose between stepping on the street and waiting until the car has passed. This looks like an asymmetric conflict, but note that in fact there is no such thing like a population of pedestrians or a population of car drivers. Instead, there is only one population of agents who are sometimes in the role of a pedestrian and sometimes in the role of a car driver. Of course these agents must condition their choice of action on their player position (their role) in the game.

In both these scenarios what is going on is that agents of the *same* type meet and play a bimatrix game (which can be symmetric, however), where they condition their strategies on their player position, or role, in the game. With this specification of the interaction, the agents are playing what is called a *role game*. A role game is a symmetric two player game based on a bimatrix game. The extensive form of a (static) role game is shown in Figure 1: With the first move, nature decides which player will play rows (role 1) and which will play columns (role 2). After that, the two players play the bimatrix game, which is called the *base game*, according to the roles they have been assigned. In order to distinguish the strategies of the players in the role game and the base game, we refer to the strategies in the latter game as *actions*. Then a *strategy* of a player is a pair of actions, one for each role he or she might be assigned. If there are n pure actions available in role 1, and m pure actions in role 2, then there are nm pure strategies in the role game.

The idea of constructing a role game from a bimatrix game already appears (for zero-sum games) in the classical work of von Neumann and Morgenstern (1944). Selten (1980), Maynard Smith (1982), Weibull (1995), and Binmore and Samuelson (2001) used it in an evolutionary context[3]. The replicator dynamics was studied for role games with 2×2 base games by Gaunersdorfer et al. (1991), see also Cressman et al. (2000) and Cressman (2000) for related work, introducing the notion of a *Wright equilibrium*, i.e. an equilibrium on the Wright manifold, which turns out to play an important role also in our analysis. The best response dynamics for role games with 2×2 base games has been studied by Berger (2001), see also Berger (2002).[4]

Formally, let the base game be given by an $n\times m$ bimatrix game with payoff matrices U^1 and U^2. Denote by z_i the probability of pure action i, then a mixed action of player 1 is given by the (column-) vector $z = (z_i)_{1\leq i\leq n}$, which is an element of the $n-1$ dimensional probability simplex S_n. Similarly, a mixed action of player 2 is a vector $y = (y_j)_{1\leq j\leq m}$ in S_m.

[2] At least this is what they are expected to do.

[3] Weibull (1995) uses the expression *role-conditioned game* for the role game, and Binmore and Samuelson (2001) call it the *role-completed game*.

[4] Unfortunately, the common method of local stability analysis via linearization of the vector field fails for the best response dynamics, so results are available only for special cases, and must be derived using iterated Poincaré maps or Ljapunov functions.

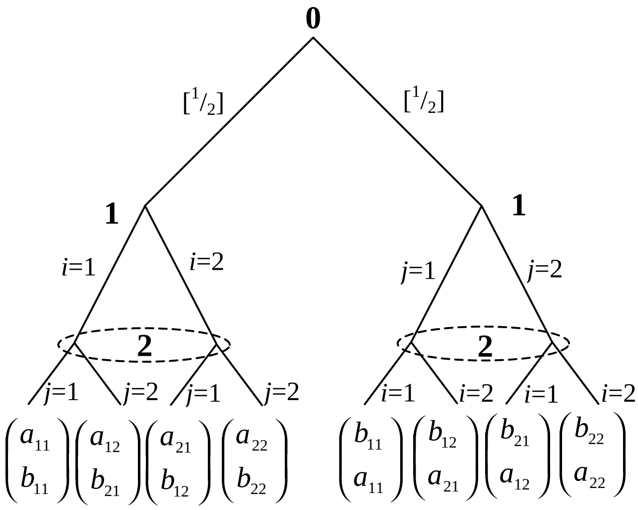

Fig. 1 The extensive form of the role game for a base game with two actions per player. In the base game, if player 1 chooses action i and player 2 chooses j, the former's payoff is a_{ij} and the latter gets b_{ji}. In the first move, player 1 is assigned the roles 1 and 2 with probability 1/2 each.

With a little abuse of notation we will not distinguish between a pure action i and the corresponding mixed action e^i placing probability 1 on this pure action. Let a dot denote the scalar product of two vectors, then the payoffs to player 1 and 2 can be written as $z \cdot U^1 y$ and $y \cdot U^2 z$, respectively.

The corresponding role game is a symmetric $nm \times nm$ game with payoff matrix U. Denoting a pure strategy by the pair ij of pure actions it consists of, the elements of this payoff matrix are given by $u_{ij,kl} = (u^1_{il} + u^2_{jk})/2$. A mixed strategy for the role game is represented by a vector $x = (x_{ij})$ in the simplex S_{nm}, and the payoff to strategy x against some strategy p is $x \cdot Up$.

Since by construction of the role game every pure strategy corresponds to a unique pair of pure base game actions, we get a natural correspondence between mixed strategies and pairs of mixed actions: Every mixed strategy x corresponds to the pair $(z(x), y(x))$ of actions with $z_i(x) = \sum_{l=1}^{m} x_{il}$ and $y_j(x) = \sum_{k=1}^{n} x_{kj}$. This means that $z(x)$ and $y(x)$ are just the marginal distributions of x. If $n \geq 2$ and $m \geq 2$, the space of mixed strategies has dimension $nm-1$, which is greater than $n+m-2$, the dimension of the space of pairs of mixed actions. Geometrically, the set of strategies corresponding to the same action pair (z, y) is the intersection of the simplex S_{nm} with a linear subspace of dimension $(nm-1) - (n+m-2) = (n-1)(m-1)$. If we denote this set by $X(z, y)$, then all the strategies $x \in X$ earn the same payoff $x \cdot Up = [z(x) \cdot U^1 y(p) + y(x) \cdot U^2 z(p)]/2$ against some strategy p. This yields (see also Berger, 2001, or Cressman et al., 2000)

Lemma 1 *If x' is a best response to x in the role game, then $(z(x'), y(x'))$ is a best response*[5] *to $(z(x), y(x))$ in the base game. If (z', y') is a best response to (z, y) in the base game, then x' is a best response to x in the role game, for any $x' \in X(z', y')$ and $x \in X(z, y)$.*

Since a (symmetric) Nash equilibrium of the role game is a strategy x which is a best response to itself, and a Nash equilibrium of the base game is a pair of actions (z, y) which is a best response to itself, it follows that

Lemma 2 *If x is a Nash equilibrium of the role game, then $(z(x), y(x))$ is a Nash equilibrium of the base game. If (z, y) is a Nash equilibrium of the base game, then x is a Nash equilibrium of the role game for any $x \in X(z, y)$.*

Note that the pre-image $X(z, y)$ has dimension $(n-1)(m-1)$ if and only if z and y are completely mixed actions. If e.g. both are pure actions (i, j), then the set $X(i, j)$ is the singleton consisting of the corresponding pure strategy ij of the role game.

The *Wright manifold* of the simplex S_{nm} is the set $W = \{x \in S_{nm} : x_{ij} = z_i y_j\}$. This manifold, originally stemming from the two-locus two-alleles model of population genetics, consists of all mixed strategies where the choices of actions for the two roles are independent. It connects the vertices of the simplex S_{nm} and intersects each pre-image $X(z, y)$ in exactly one point.

5 Role Switching and Best Response Adaptation

The usual interpretation of a static role game as in Figure 1 requires that the two roles are randomly assigned to the agents *after* a pair has been matched. This is natural for a one shot game, but it poses some difficulties for the population model we have in mind. In the crossing game depicted above the random assignment of roles might be acceptable, but in the driver-pedestrian conflict it is not. If you are a pedestrian, you stay a pedestrian, at least for the next few minutes. The next morning you might take the car, that is, you have switched your role. Likewise, an agent might on one day be in the role of a buyer (leader, donor, etc.) and on the next day in the role of a seller (follower, receiver, etc.). This kind of *role switching* has to be modeled if we want a realistic picture of the dynamic process involved. The easiest way to do this is to assume one type of agents, but two populations, corresponding to the two roles, where pairs are formed by randomly drawing one agent from each population. The idea that the role of an agent is not fixed can then be introduced by the assumption that every now and then an agent switches his role, i.e. he changes to the other population.

We could in principle directly apply the best response dynamics to this population model. However, in the usual formulation this dynamics seems

[5] To shorten the notation we call a pair (z', y') a best response to a pair (z, y), if z' is a best response to y and y' is a best response to z.

to be too inflexible to suit our model. Note that the best response dynamics demands that an agent with the possibility to review his strategy has to review both his actions at the same time. Recently, Cressman (2001, chapter 3.4.2) suggested an alternative version of best response dynamics for role games. In this version, which he calls the *symmetrized best response dynamics*, agents review their actions with probability 1/2 each.[6] One can imagine other specifications of strategy reviewing. If, for example, it is costly to review an action, and switching occurs only rarely, an agent might prefer only to review the action he is currently using (his *active* action), and to leave the currently unused (*passive*) action unchanged. He might also occasionally decide not to review at all. We would like to be able to incorporate these and other possibilities in our model. In the following we present a model of best response adaptation and derive the dynamic equations guiding the suggested evolutionary process. Let us state the exact assumptions of the model, starting with discrete time steps for convenience.

6 The Model

There are two large populations of agents, populations 1 and 2, corresponding to the two roles in the base game. The populations have N_1 and N_2 members, respectively. We assume $N_1 \leq N_2$ and denote the relative sizes by $w = N_1/(N_1 + N_2)$ and $1 - w = N_2/(N_1 + N_2)$. Each agent initially has some pure strategy ij, meaning "play action i if in role 1 and action j if in role 2". In period t, denote the fraction of agents in population 1 (called 1-agents) with strategy ij by $x^1_{ij}(t)$. For population 2, $x^2_{ij}(t)$ is defined analogously, and $x_{ij}(t) = wx^1_{ij}(t) + (1 - w)x^2_{ij}(t)$ is the total fraction of agents with strategy ij.

In each period t, $K(t) \leq N_1$ agents are randomly drawn from each population and switch to the other population, i.e. they switch their roles. Note that the population sizes remain constant.

After that, $k(t) \leq N_1$ agents are randomly selected from each population and get the possibility to review their strategies. Each selected agent reviews both his actions with probability p_1, he reviews only his active action with probability p_2, only his passive action with probability p_3, and he does not review at all with probability $p_4 = 1-p_1-p_2-p_3$. We assume that the four probabilities are the same for all agents and all periods. Furthermore, we can without loss of generality assume $p_4 = 0$, because a positive probability

[6] This interpretation is consistent with the dynamic equations he derives. Cressman's (2001) original formulation is that agents review only the action for their "current role". It is, however, not quite clear what an agent's current role is, if this role is assigned randomly during each matching. With a static role game, the agent does not know what his role will be in the next interaction, and it is hard to explain why he should review the action for the role he was assigned the last time. The idea is easier to formulate in our framework and appears as an important special case below.

of not reviewing if selected can be replicated by a lower selection probability (lower $k(t)$). So we have $p_1 + p_2 + p_3 = 1$.

An agent reviewing his active action chooses a pure best response to the average active action in the other population, while when reviewing his passive action, he chooses a pure best response to the average active action in his own population.[7] We assume that in the case of multiple best responses to both action distributions a reviewing agent makes his choices for the two roles independently. Furthermore, we assume that all agents use the same (but arbitrary) tie-breaking rule.

Finally, $M(t) \leq N_1$ agents are drawn randomly from each population and matched. Each pair of agents plays the base game (U^1, U^2), with the 1-agent in the role of the row-player and the 2-agent in the role of the column-player.

In the following notation we will sometimes suppress the time variable where this is not misleading. Let $z_i^1 = \sum_l x_{il}^1$ and $y_j^2 = \sum_k x_{kj}^2$ be the frequencies of the active actions and $z_i^2 = \sum_l x_{il}^2$ and $y_j^1 = \sum_k x_{kj}^1$ the frequencies of the inactive actions in the populations. The respective total frequencies are $z = wz^1 + (1-w)z^2$ and $y = wy^2 + (1-w)y^1$.

For convenience we use the same expression $B(.)$ to denote the (mixed) best response correspondences for action distributions z^1, z^2, y^1, and y^2.

Now consider the expected increase $\Delta x_{ij}^1(t) = x_{ij}^1(t+1) - x_{ij}^1(t)$ from period t to period $t+1$. First, there is an "inflow" caused by role-switching of Kx_{ij}^2 agents and a corresponding outflow of Kx_{ij}^1 agents. Then consider the change in x_{ij}^1 due to strategy-reviewing. In period t, each agent reviewing his 1-action chooses some pure best response from $B(y^2)$. Analogously, each agent reviewing his 2-action chooses a pure best response from $B(z^1)$. Now take a single agent from population 1. We want to calculate the probability that he has strategy ij after reviewing has taken place. Denote the probability that the agent chooses action i when reviewing his 1-action by b_i^1 and the probability that he chooses j when reviewing his 2-action by b_j^2. These values depend on the tie-breaking rule, if the best responses are not unique.

With probability $(1 - k/N_1)x_{ij}^1$, the agent already has the strategy ij and is not selected to review his strategy. With probability k/N_1 he is selected to review his strategy. In the latter case there are three ways to proceed. With probability p_1 the agent reviews both his actions and then (by the independence assumption) with probability $b_i^1 b_j^2$ he chooses ij as his new strategy. The second possibility is that he only reviews his 1-action (probability p_2) and chooses i (probability b_i^1), but has previously had the 2-action j (probability y_j^1). Finally, the agent may only review his 2-action (probability p_3) and choose j (probability b_j^2), but has already been using the 1-action i (probability z_i^1).

[7] He knows that his passive action will only become active if he switches to the other population, in which case he will play against a member of the population he currently belongs to.

Averaging over all agents in population 1 and subtracting $x^1_{ij}(t)$ yields an expected increase of

$$\begin{aligned}\Delta x^1_{ij} = {} & \frac{K}{N_1}(x^2_{ij} - x^1_{ij}) + \\ & + \frac{k}{N_1}[p_1 b^1_i b^2_j + p_2 b^1_i y^1_j + p_3 b^2_j z^1_i - x^1_{ij}].\end{aligned} \tag{3}$$

Analogously we get

$$\begin{aligned}\Delta x^2_{ij} = {} & \frac{K}{N_2}(x^1_{ij} - x^2_{ij}) + \\ & + \frac{k}{N_2}[p_1 b^1_i b^2_j + p_2 b^2_j z^2_i + p_3 b^1_i y^2_j - x^2_{ij}]\end{aligned} \tag{4}$$

for population 2. Summing up over all i and all j, respectively, in (3) and (4), it can be seen that the average distributions of actions in the two populations behave according to

$$\begin{aligned}\Delta z^1 &= \tfrac{K}{N_1}(z^2 - z^1) + \tfrac{k}{N_1}(1-p_3)(b^1 - z^1),\\ \Delta z^2 &= \tfrac{K}{N_2}(z^1 - z^2) + \tfrac{k}{N_2}(1-p_2)(b^1 - z^2),\\ \Delta y^2 &= \tfrac{K}{N_2}(y^1 - y^2) + \tfrac{k}{N_2}(1-p_3)(b^2 - y^2),\\ \Delta y^1 &= \tfrac{K}{N_1}(y^2 - y^1) + \tfrac{k}{N_1}(1-p_2)(b^2 - y^1).\end{aligned}$$

Going from discrete to continuous time, setting $R(t) = K(t)\frac{N_1+N_2}{N_1 N_2}$, $r(t) = k(t)\frac{N_1+N_2}{N_1 N_2}$ (we assume that $R(.)$ and $r(.)$ are continuous and positive), we therefore get the following system of differential inclusions.

$$\begin{aligned}\dot z^1 &\in (1-w)\ [R(z^2 - z^1) + r(1-p_3)(B(y^2) - z^1)],\\ \dot z^2 &\in \quad w\ [R(z^1 - z^2) + r(1-p_2)(B(y^2) - z^2)],\\ \\ \dot y^2 &\in \quad w\ [R(y^1 - y^2) + r(1-p_3)(B(z^1) - y^2)],\\ \dot y^1 &\in (1-w)\ [R(y^2 - y^1) + r(1-p_2)(B(z^1) - y^1)].\end{aligned} \tag{5}$$

The best response correspondences $B(.)$ are upper-semicontinuous with closed and convex values. Hence the existence of at least one solution through each initial value, which is Lipschitz continuous and defined for all positive times, is guaranteed, see e.g. Aubin and Cellina (1984). However, multiple solutions may exist if one of the involved best response sets is not a singleton. Every solution can be written as a function $t \mapsto (z^1(t), z^2(t), y^2(t), y^1(t))$, satisfying the system of differential equations

$$\begin{aligned}\dot z^1 &= (1-w)\ [R(z^2 - z^1) + r(1-p_3)(b(y^2) - z^1)],\\ \dot z^2 &= \quad w\ [R(z^1 - z^2) + r(1-p_2)(b(y^2) - z^2)],\\ \\ \dot y^2 &= \quad w\ [R(y^1 - y^2) + r(1-p_3)(b(z^1) - y^2)],\\ \dot y^1 &= (1-w)\ [R(y^2 - y^1) + r(1-p_2)(b(z^1) - y^1)].\end{aligned} \tag{6}$$

for almost all $t \geq 0$, where the functions $t \mapsto b(.(t)) \in B(.(t))$ are measurable.

In the space of strategies this gives rise to a solution $t \mapsto (x^1(t), x^2(t))$ with

$$\begin{aligned}\dot{x}^1_{ij} &= (1-w)[R(x^2_{ij} - x^1_{ij}) + \\ &\quad + r(p_1 b_i(y^2) b_j(z^1) + p_2 b_i(y^2) y^1_j + p_3 b_j(z^1) z^1_i - x^1_{ij})], \\ \dot{x}^2_{ij} &= w[R(x^1_{ij} - x^2_{ij}) + \\ &\quad + r(p_1 b_i(y^2) b_j(z^1) + p_2 b_j(z^1) z^2_i + p_3 b_i(y^2) y^2_j - x^2_{ij})]\end{aligned} \tag{7}$$

for almost all $t \geq 0$. Note that we can multiply the right hand side of (6) and (7) with any continuous and positive function of t without changing the shape of the orbits (this amounts to a rescaling of time). Allowing a little abuse of notation this implies that we can choose $r(.)$ arbitrarily (but positive and continuous).

The average frequency x_{ij} of strategy ij in the set of all agents then follows the dynamic equation

$$\begin{aligned}\dot{x}_{ij} &= w\dot{x}^1_{ij} + (1-w)\dot{x}^2_{ij} \\ &= w(1-w)r[2p_1 b_i(y^2) b_j(z^1) + b_i(y^2)(p_2 y^1_j + p_3 y^2_j) + \\ &\quad + b_j(z^1)(p_3 z^1_i + p_2 z^2_i) - x^1_{ij} - x^2_{ij}].\end{aligned} \tag{8}$$

Summing up over all i and j, respectively, yields

$$\begin{aligned}\dot{z}_i &= w(1-w)r[(1+p_1)b_i(y^2) - (1-p_2)z^2_i - (1-p_3)z^1_i], \\ \dot{y}_j &= w(1-w)r[(1+p_1)b_j(z^1) - (1-p_2)y^1_j - (1-p_3)y^2_j].\end{aligned} \tag{9}$$

The next Lemma states, loosely speaking, that constant solutions of (5) coincide with Nash equilibria of the base game.

Lemma 3 *The constant function $t \mapsto (z^1, z^2, y^2, y^1)$ is a solution of (5) if and only if $z^1 = z^2 = z^*$ and $y^1 = y^2 = y^*$, where (z^*, y^*) is a Nash equilibrium of the base game.*

Proof See appendix.

Note that a solution starting at a Nash equilibrium is not necessarily a constant one. For example, if there is a continuum of equilibria, we can induce arbitrary movements of the orbits in this continuum by choosing a proper selection $b(y^2(t))$ and $b(z^1(t))$ from the continuum. However, if the base game has a unique Nash equilibrium, the constant solution is also unique, and the equilibrium may be called a fixed point of (5).

7 Special Cases

The model described above is general enough to allow for a variety of behavior patterns based on best response adaptation. In the following we define two important special cases.

First we consider the case of *complete symmetry.* By complete symmetry (CSY) we mean symmetry in population size, in reviewing probability, and in initial conditions. More precisely,

$$\text{(CSY):} \qquad w = 1/2,\ p_2 = p_3,\ x^1(0) = x^2(0).$$

This is not meant to be a realistic description of initial conditions in a population model of asymmetric conflicts, it is defined in this way for purely notational purposes. To see why, note that since all agents use the same tie-breaking rule, $z^2(t) = z^1(t)$ implies that $b(z^2(t)) = b(z^1(t))$, and analogously for the agents' 2-actions. Hence the behavior of the two populations is identical (we suppress the superscripts 1 and 2 in such a case). Role-switching becomes meaningless in this scenario, and the model is equivalent to a model of two separate populations, where the passive actions never come into use and the active actions are reviewed with probability $p_1 + p_2$.

On the level of actions this is nothing but the best response dynamics for bimatrix games.[8] Indeed, setting $w = 1/2$, $p_2 = p_3 = (1 - p_1)/2$, choosing $r(t) \equiv 4/(1 + p_1)$, and assuming initial symmetry, (6) reduces to (2).

On the level of strategies we can identify two known subcases. First, with $p_1 = 1$ the above scenario replicates the best response dynamics for the role game. To see this, note that (CSY) with $p_1 = 1$ and $r(t) \equiv 2$ reduces (7) to two identical differential equations of the form $\dot{x} = b(x) - x$. [9] The specification $p_1 = 1$ is justified e.g. if the costs of strategy adjustment are extremely low. Second, for $p_1 = 0$ (and $r(t) \equiv 4$) we have $p_2 = p_3 = 1/2$, and while the orbits are the same as with $p_1 = 1$ at the level of actions, the orbits in the state space of strategies look different. They are given by the dynamic equations

$$\dot{x}_{ij} = (b_i(y)y_j - x_{ij}) + (b_j(z)z_i - x_{ij}),$$

which are Cressman's (2001) symmetrized best response dynamics. The reviewing probabilities of this dynamics may be justified if adjusting strategies is costly and the rate of role-switching R is very high. This would mean that agents prefer to review only one of their actions and choose each one with

[8] As usual, if the bimatrix game is symmetric and we assume symmetry in initial conditions, then also the case of the best response dynamics for matrix games (1) is included.

[9] The resulting dynamics is not exactly identical to the best response dynamics for the role game, because the assumption of independence implies that a completely mixed equilibrium of the role game is unstable as long as it is not a Wright equilibrium. This, however, does not influence orbits along which points of indifference between different best responses for both actions are isolated.

probability 1/2, since they do not know which one will be active in their next interaction.

As a second special case of our general model we define the case *review only active* (ROA).

$$\text{(ROA):} \qquad p_2 = 1.$$

As the name suggests, in this case all agents review only their active action. This makes sense e.g. if adjusting strategies is costly and the rate of role-switching is very low (agents can expect to stay in the same role for a long time).

8 Asymptotic Behavior for Zero-Sum Games

A special case is given, if $U^2 = -(U^1)^T$ (the negative transpose of U^1), i.e., if the base game is a zero-sum game. For such games we will show that in the cases (ROA) or (CSY), the set of Nash equilibria is globally asymptotically stable under (5).[10] Convergence behavior for zero-sum games is especially important for 2×2 games. While for generic 2×2 games convergence to Nash equilibrium is not difficult to show for the cases with one or two pure equilibria, respectively, this is a nontrivial task for the cyclic case (with a unique and completely mixed equilibrium). However, since in the cyclic case the game can be shown to be strategically equivalent to a zero-sum game (see Hofbauer and Sigmund (1998)), this case is covered by our result. As a consequence, for the class of generic 2×2 games, if (ROA) or (CSY), then every solution converges to a Nash equilibrium.

For the cyclic 2×2 case, a typical orbit of $(z^1(t), y^2(t))$ and $(z^2(t), y^1(t))$ is shown in Figure 2.

Theorem 1 *Let the base game be zero-sum. If (ROA) or (CSY), then both* $(z^1(t), y^2(t))$ *and* $(z^2(t), y^1(t))$ *– and hence* $(z(t), y(t))$ *– converge to the set of Nash equilibria.*

To prove the stability of the Nash equilibrium set we construct a Ljapunov function for solutions of (5). This function is a modification of the one used by Hofbauer (1995). So let $U^2 = -(U^1)^T$ and consider the continuous function

$$\begin{aligned} V(z^1, z^2, y^1, y^2) := (1-w)\ [\max_h (U^1 y^2)_h + \max_h (U^2 z^2)_h] + \\ + \ w\ [\max_h (U^1 y^1)_h + \max_h (U^2 z^1)_h]. \end{aligned}$$

The proof consists of combining the following four Lemmas.

Lemma 4 V *is nonnegative and vanishes if and only if* (z^1, y^2) *and* (z^2, y^1) *are Nash equilibria.*

[10] It follows easily that this result also holds for games which are strategically equivalent to a zero-sum game.

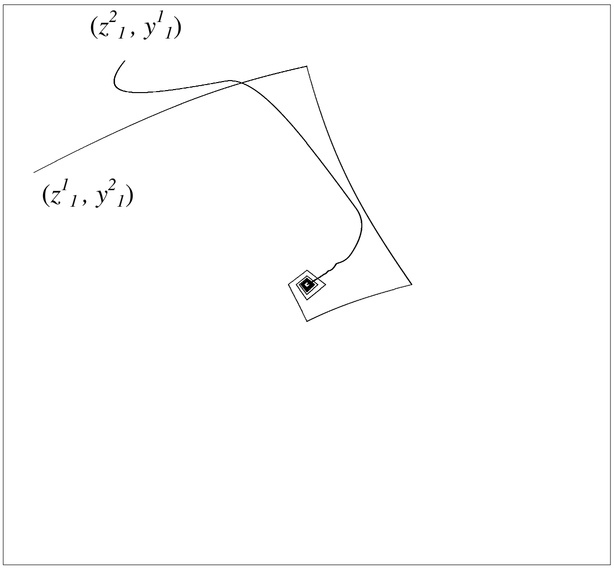

Fig. 2 A typical orbit for a game of the Matching Pennies type. Here, the parameters are $R = 2$, $r = 1$, $w = 1/2$, and $p_2 = 1$.

Proof See appendix.

Lemma 5 *If (ROA) or (CSY), then* V *is nonincreasing along any solution of (5).*

Proof See appendix.

Lemma 6 *If (ROA) or (CSY), then* $\dot{V} = 0$ *implies that* (z^1, y^2) *is a Nash equilibrium.*

Proof See appendix.

Combining these three Lemmas and applying Ljapunov's argument shows that $(z^1(t), y^2(t))$ converges to the set of Nash equilibria. It remains to prove that the same holds for $(z^2(t), y^1(t))$.

Lemma 7 *If (ROA) or (CSY),* $(z^2(t), y^1(t))$ *converges to the set of Nash equilibria.*

Proof See appendix.

This completes the proof of Theorem 1.

9 Convergence to the Wright manifold

Consider a base game which is zero-sum with a unique Nash equilibrium (z^*, y^*). For (ROA) or (CSY) we have shown that both the active part

(z^1, y^2) and the passive part (z^2, y^1) of any solution of (5) converge to (z^*, y^*). Since any solution of (5) corresponds to a solution of (7), it follows that x^1 and x^2 converge to the set $X(z^*, y^*)$ of Nash equilibria of the role game. If this set is a singleton, we know the distribution of role game strategies in the long run. However, if it is not – and this is the case if e.g. (z^*, y^*) is completely mixed, the question remains, which part of the equilibrium set is attractive. Subsequently we will show that if the populations are equally large, i.e. $w = 1/2$, and there is an orbit converging to a base game equilibrium, then the distribution of strategies along this orbit converges to the Wright manifold, and hence to a Wright equilibrium. In the situation described above, this means that all orbits converge to the (unique) Wright equilibrium x^* defined by $x^*_{ij} = z^*_i y^*_j$.

Theorem 2 *Let $w = 1/2$. If $z^1(t)$ and $z^2(t)$ converge to z^* and $y^2(t)$ and $y^1(t)$ converge to y^*, then $x_{ij}(t)$ converges to $z^*_i y^*_j$.*

Proof See appendix.

10 Discussion

From Theorem 1 with (CSY) we can derive the result of Robinson (1951) for continuous fictitious play, or the equivalent result of Hofbauer (1995) for the best response dynamics. Moreover, under the plausible assumption that agents only review their active strategy (ROA), we can see that convergence for zero-sum games continues to hold for different population sizes, varying rates of role-switching and strategy-reviewing, and for arbitrary initial distributions in both populations.

Theorem 2 is even more general. For the continuous fictitious play process, or, equivalently, the best response dynamics, which is covered by (CSY), there are other well known convergence results, e.g. for generic 2×2 games (Miyasawa (1961)), weighted potential games (Monderer and Shapley, 1996), or strongly dominance solvable games (Milgrom and Roberts, 1991). For all these classes of games, as a corollary of Theorem 2 we obtain convergence to a Wright equilibrium. This is trivial if the attracting base game equilibrium is in pure actions (every pure strategy equilibrium is contained in the Wright manifold), but if the base game equilibrium is completely mixed, there is a continuum of mixed Nash equilibria in the role game, of which only the Wright equilibrium is attracting. In this sense, our model with (ROA, $w = 1/2$) or (CSY) exhibits equilibrium selection. This happens e.g. for games of the Matching Pennies type,[11] and hence the result of Berger (2001) is obtained as a simple corollary. With a little additional

[11] The best response dynamics converges cyclically to the completely mixed equilibrium in these games. Note, however, that according to Krishna and Sjöström (1998), for the best response dynamics cyclic convergence to a completely mixed equilibrium is a nongeneric phenomenon if there are more than two pure actions in both roles.

effort it can also be shown that the result for 4×4 zero-sum matrix games of Berger (2002) follows from our analysis. Finally, for (CSY) and $p_1 = 0$ it can be seen that Theorem 2 recovers the result of Cressman (2001) for the symmetrized best response dynamics. In this case *all* the coefficients in (13) are zero and hence, as observed by Cressman (2001), the Wright manifold is attracting and invariant. From this we can infer even more. The behavior of orbits in the base game is completely reflected by the behavior of orbits on the Wright manifold in the role game. For example if there is an attracting Shapley polygon P in the base game, then there is a corresponding attracting set $X(P) \cap W$ in the role game, a Shapley polygon embedded in the Wright manifold.

11 Appendix

11.1 Proof of Lemma 3

The if-part is obvious, so let us assume there is a constant solution. Note that at least one of the factors $1 - p_3$ and $1 - p_2$ is positive. From the first line of (5) it can be seen that $\dot{z}^1$ is the sum of two vectors, one pointing in the direction of z^2 and the other one pointing at $b(y^2)$. Since this sum is zero, z^1 must lie (weakly) between z^2 and $b(y^2)$. However, from the second line of (5), the same is true vice versa: z^2 must lie (weakly) between z^1 and $b(y^2)$. It follows that $z^1 = z^2 = b(y^2)$. Analogously, $y^1 = y^2 = b(z^1)$. Calling these variables z^* and y^* we have $z^* = b(y^*)$ and $y^* = b(z^*)$, i.e., (z^*, y^*) is a Nash equilibrium. □

11.2 Proof of Lemma 4

Note that

$$V \geq (1 - w)(z^2 \cdot U^1 y^2 + y^2 \cdot U^2 z^2) + w(z^1 \cdot U^1 y^1 + y^1 \cdot U^2 z^1) = 0,$$

since $U^2 = -(U^1)^T$, and equality holds if and only if

$$z^2 \in B(y^2) \text{ and } y^2 \in B(z^2) \text{ and } z^1 \in B(y^1) \text{ and } y^1 \in B(z^1),$$

i.e., if and only if (z^2, y^2) and (z^1, y^1) are Nash equilibria. In a zero-sum game, Nash equilibria are interchangeable, and thus the above condition is equivalent to the condition that (z^1, y^2) and (z^2, y^1) are Nash equilibria. □

11.3 Proof of Lemma 5

Let us first write $V = (1 - w)(V_1 + V_2) + w(V_3 + V_4)$, where $V_1 = \max_h (U^1 y^2)_h$, $\ldots, V_4 = \max_h (U^2 z^1)_h$. Note that along any solution of (5), $V_1(t)$ is absolutely continuous as the maximum of absolutely continuous functions. Thus $V_1(t)$ is differentiable almost everywhere. The same holds for V_2, V_3, V_4, and so V itself is differentiable almost everywhere. We know that this also holds for the functions $z^1(t), y^2(t), z^2(t), y^1(t)$. Let t_0 be a point of differentiability of these four functions as well as of $V(t)$. Now let us take a closer look at the derivative of e.g. $V_1(t)$. Our goal for the moment is to show that the derivative with respect to t of V_1 at time t_0 can be written as $\dot{V}_1(t_0) = b(y^2(t_0)) \cdot U^1 \dot{y}^2(t_0)$. To this end, we have to show that $\dot{V}_1(t_0) = e^i \cdot U^1 \dot{y}^2(t_0)$ for *all* pure actions $i \in B(y^2(t_0))$. The next paragraph covers this part of the proof.

For any sequence $t_k \to t_0$, $t_k \neq t_0$ there is a sequence of pure actions $i_k \in B(y^2(t_k))$. This sequence always has a constant subsequence, since the number of available indices i_k is finite. Assume the constant subsequence is $i_c, i_c, \ldots$, then we know that $i_c \in B(y^2(t_0))$. It follows that

$$\begin{aligned}\dot{V}_1(t_0) &= \lim_{k\to\infty} \frac{V_1(t_k) - V_1(t_0)}{t_k - t_0} \\ &= \lim_{k\to\infty} \frac{e^{i_c} \cdot U^1 y^2(t_k) - e^{i_c} \cdot U^1 y^2(t_0)}{t_k - t_0} \\ &= e^{i_c} \cdot U^1 \lim_{k\to\infty} \frac{y^2(t_k) - y^2(t_0)}{t_k - t_0} \\ &= e^{i_c} \cdot U^1 \dot{y}^2(t_0)\,.\end{aligned}$$

If $B(y^2(t_0))$ is not a singleton, and there is another subsequence with a different limit $i_d \in B(y^2(t_0))$, then the existence of $\dot{V}_1(t_0)$ implies $e^{i_d} \cdot U^1 \dot{y}^2(t_0) = e^{i_c} \cdot U^1 \dot{y}^2(t_0)$. Now suppose that for some index i_d with $i_d \in B(y^2(t_0))$ there is no sequence $t_k \to t_0$ with $i_d \in B(y^2(t_k))$ for all k. Then there is a neighborhood N of t_0, such that i_d is not a best response to $y^2(t)$ for $t \in N - \{t_0\}$. This in turn implies that

$$(e^{i_c} - e^{i_d}) \cdot U^1 \dot{y}^2(t_0) = 0,$$

since otherwise there are always points t_k arbitrarily close to t_0, with

$$(e^{i_d} - e^{i_c}) \cdot U^1 y^2(t_k) > (e^{i_d} - e^{i_c}) \cdot U^1 y^2(t_0) = 0,$$

contradicting the fact that i_d is not a best response for such t_k.

We have now proved that $\dot{V}_1(t_0) = e^i \cdot U^1 \dot{y}^2(t_0)$ for all $i \in B(y^2(t_0))$. Hence also $\dot{V}_1(t_0) = b(y^2(t_0)) \cdot U^1 \dot{y}^2(t_0)$. Analogous considerations for $\dot{V}_2$, $\dot{V}_3$, and $\dot{V}_4$ yield

$$\begin{aligned}&\dot{V}_1(t_0) = b(y^2(t_0)) \cdot U^1 \dot{y}^2(t_0),\ \dot{V}_2(t_0) = b(z^2(t_0)) \cdot U^2 \dot{z}^2(t_0), \\ &\dot{V}_3(t_0) = b(y^1(t_0)) \cdot U^1 \dot{y}^1(t_0),\ \dot{V}_4(t_0) = b(z^1(t_0)) \cdot U^2 \dot{z}^1(t_0),\end{aligned}$$

and we can finally calculate

$$\begin{aligned}&[r\ w(1-w)]^{-1}\,\dot{V} = w^{-1}(\dot{V}_1 + \dot{V}_2) + (1-w)^{-1}(\dot{V}_3 + \dot{V}_4) \\ &= (1-p_3)[b(y^2) \cdot U^1 b(z^1) + b(z^1) \cdot U^2 b(y^2)] - \\ &- (1-p_3)[b(y^2) \cdot U^1 y^2 + b(z^1) \cdot U^2 z^1] + \\ &+ (1-p_2)[b(y^1) \cdot U^1 b(z^1) + b(z^2) \cdot U^2 b(y^2)] - \\ &- (1-p_2)[b(y^1) \cdot U^1 y^1 + b(z^2) \cdot U^2 z^2] + \\ &+ (R/r)[(b(y^2) - b(y^1)) \cdot U^1 (y^1 - y^2) + (b(z^1) - b(z^2)) \cdot U^2 (z^2 - z^1)].\end{aligned} \tag{10}$$

The first term in this sum is zero since the game is zero-sum, and the second term is nonpositive, because

$$b(y^2) \cdot U^1 y^2 + b(z^1) \cdot U^2 z^1 \geq z^1 \cdot U^1 y^2 + y^2 \cdot U^2 z^1 = 0. \tag{11}$$

By analogous reasoning, the fourth term is nonpositive. For the fifth term, note that $b(y^2)\cdot U^1 y^2 \geq b(y^1)\cdot U^1 y^2$ and $b(y^2)\cdot U^1 y^1 \leq b(y^1)\cdot U^1 y^1$. This yields $(b(y^2) - b(y^1))\cdot U^1(y^1 - y^2) \leq 0$. Analogously, we get $(b(z^1) - b(z^2))\cdot U^2(z^2 - z^1) \leq 0$, so also the fifth term is nonpositive. Considering the third term, we can see that this term vanishes if either $p_2 = 1$, i.e. (ROA), or (CSY).

Hence for (ROA) or (CSY) we have $\dot{V} \leq 0$ for almost all $t > 0$ along any solution of (5). □

11.4 Proof of Lemma 6

For $\dot{V} = 0$, the second and the fourth term of the sum in (10) must vanish. If (ROA) then $1 - p_3 = 1$ and a necessary condition for the second term to vanish is equality in (11). Equality holds if and only if $z^1 \in B(y^2)$ and $y^2 \in B(z^1)$, i.e. if and only if (z^1, y^2) is a Nash equilibrium. If (CSY), then $1 - p_3 = 1 - p_2 > 0$. By the same reasoning as above, $(z^1, y^2) = (z^2, y^1)$ is a Nash equilibrium. □

11.5 Proof of Lemma 7

In case (CSY) this follows by definition, so assume (ROA). If $p_2 = 1$, the second line of (5) reads $\dot{z}^2 = wR(z^1 - z^2)$, that means $\dot{z}^2$ points into the direction of z^1. We know that $z^1(t)$ converges to the 1-component of the Nash equilibrium set. Let NE denote the set of Nash equilibria, then this 1-component is the set $NE_1 = \{z^* : \exists y^* : (z^*, y^*) \in NE\}$. Analogously we define NE_2. From the interchangeability of Nash equilibria for zero-sum games it follows that $NE = NE_1 \times NE_2$. We also know that NE is convex for zero-sum games, and so are the component sets of NE. Assume z^2 is not in NE_1 and let N be an open, convex neighborhood of NE_1. From convergence it follows that $z^1(t)$ stays in N for large t. Since $\dot{z}^2$ points into the direction of z^1, it points to N for any large t. So let us assume that z^2 is not in N and $\dot{z}^2$ points to N, and let $\hat{z}$ be the point in the closure of N which has minimal distance to z^2 (this point exists, since N is convex). Denote by H the hyperplane which contains $\hat{z}$ and is perpendicular to $\hat{z} - z^2$. This is a separating hyperplane for the convex sets N and $\{z^2\}$. Since $\dot{z}^2$ points to N, it also points to H. Hence the distance between z^2 and N strictly decreases until z^2 reaches N. Letting N shrink to NE_1 yields the convergence of $z^2(t)$ to NE_1. Analogously, $y^1(t)$ converges to NE_2, and hence $(z^2(t), y^1(t))$ converges to $NE_1 \times NE_2 = NE$. □

11.6 Proof of Theorem 2

Convergence to the Wright manifold is shown by invoking another Ljapunov function, which is based on the one used by Cressman (2001). Remember we denoted the frequencies of the actions and strategies within the whole set of agents by $z = wz^1 + (1 - w)z^2$, $y = (1 - w)y^2 + wy^1$, and $x = wx^1 + (1 - w)x^2$. Now we define the functions

$$L_{ij}(x^1, x^2) = z_i y_j - x_{ij} + C(z_i - z_i^*)\dot{y}_j, \tag{12}$$

with $C = 4p_1(1+p_1)^{-2}$, and subsequently we show that $L_{ij}(t)$ goes to zero for every strategy ij. This means that $x(t)$ converges to the Wright manifold. We know that $x(t)$ converges to $X(z^*, y^*)$, hence it follows that $x(t)$ converges to a Wright equilibrium.

It suffices to show that $\dot{L}_{ij}(t) = -L_{ij}(t) + G_{ij}(t)$ for almost all $t > 0$, with $G_{ij}(t) \to 0$ for $t \to \infty$. Thus we calculate

$$\begin{aligned} G_{ij} &= \dot{L}_{ij} + L_{ij} = \dot{z}_i y_j + z_i \dot{y}_j - \dot{x}_{ij} + \\ &\quad + C\dot{z}_i \dot{y}_j + C(z_i - z_i^*)\ddot{y}_j + z_i y_j - x_{ij} + C(z_i - z_i^*)\dot{y}_j. \end{aligned}$$

Inserting from (8) and (9), G_{ij} can be written as

$$\begin{aligned} G_{ij} &= \alpha_1 b_i(y^2) y_j^2 + \alpha_2 b_i(y^2) y_j^1 + \\ &+ \beta_1 b_j(z^1) z_i^1 + \beta_2 b_j(z^1) z_i^2 + \\ &+ \delta_1 z_i^1 y_j^2 + \delta_2 z_i^1 y_j^1 + \delta_3 z_i^2 y_j^2 + \delta_4 z_i^2 y_j^1 + \\ &+ \gamma b_i(y^2) b_j(z^1) + \lambda_1 x_{ij}^1 + \lambda_2 x_{ij}^2 + \\ &+ C\dot{z}_i \dot{y}_j + C(z_i - z_i^*)(\ddot{y}_j - \dot{y}_j). \end{aligned} \tag{13}$$

Note that the last term of this sum is uniformly bounded for all t where the derivatives exist. Setting $r(t) \equiv [2w(1-w)]^{-1}$, computing the coefficients in (13) for $w = 1/2$ yields

$$\begin{aligned} &\alpha_1 = [(1-w)(1+p_1) - p_3]/2, \quad \alpha_2 = [w(1+p_1) - p_2]/2, \\ &\beta_1 = [w(1+p_1) - p_3]/2, \quad \beta_2 = [(1-w)(1+p_1) - p_2]/2, \\ &\delta_1 = w(1-w) - (p_1+p_2)/2, \quad \delta_2 = w^2 - w(1+p_1)/2, \\ &\delta_3 = (1-w)^2 - (1-w)(1+p_1)/2, \quad \delta_4 = w(1-w) - (p_1+p_3)/2, \\ &\gamma = -p_1, \quad \lambda_1 = 1/2 - w, \quad \lambda_2 = w - 1/2. \end{aligned}$$

Note that

$$\alpha_1 + \alpha_2 = \beta_1 + \beta_2 = p_1$$

and

$$\gamma = \delta_1 + \delta_2 + \delta_3 + \delta_4 = -p_1.$$

Moreover, λ_1 and λ_2 vanish for $w = 1/2$.

If z_i^1 and z_i^2 converge to z_i^*, then so does z_i, and analogously for y_j. With a little abuse of notation writing $f \to g$ instead of $f(t) - g(t) \to 0\,(t \to \infty)$, we can therefore see that

$$\begin{aligned} G_{ij} &\to (\alpha_1 + \alpha_2) b_i(y^2) y_j^* + (\beta_1 + \beta_2) b_j(z^1) z_i^* + \\ &\quad + (\delta_1 + \delta_2 + \delta_3 + \delta_4) z_i^* y_j^* + \gamma b_i(y^2) b_j(z^1) + C\dot{z}_i \dot{y}_j \\ &= p_1[b_i(y^2) y_j^* + b_j(z^1) z_i^* - z_i^* y_j^* - b_i(y^2) b_j(z^1)] + C\dot{z}_i \dot{y}_j \\ &= -p_1[b_i(y^2) - z_i^*][b_j(z^1) - y_j^*] + C\dot{z}_i \dot{y}_j \\ &\qquad \text{(inserting from (9) and for } C) \\ &\to -p_1[b_i(y^2) - z_i^*][b_j(z^1) - y_j^*] + p_1[b_i(y^2) - z_i^*][b_j(z^1) - y_j^*] = 0. \end{aligned}$$

So $G_{ij} \to 0$, and hence also L_{ij} vanishes for $t \to \infty$. This implies $x_{ij}(t) \to z_i^* y_j^*$. □

References

1. Aubin, Jean-Pierre and Arrigo Cellina, 1984, Differential inclusions (Springer, Berlin).
2. Berger, Ulrich, 2001. Best response dynamics for role games. International Journal of Game Theory **30,** 527-538.
3. Berger, Ulrich, 2002. Continuous fictitious play and projective geometry. Mimeo, Vienna University of Economics.
4. Berger, Ulrich and Josef Hofbauer, 2002. Irrational behavior in the Brown-von Neumann-Nash dynamics. Mimeo, University of Vienna.
5. Börgers, Tilman and Rajiv Sarin, 1997. Learning through reinforcement and replicator dynamics. Journal of Economic Theory **77,** 1-14.
6. Binmore, Ken and Larry Samuelson, 2001. Evolution and mixed strategies. Games and Economic Behavior, **34,** 200-226.
7. Björnerstedt, Jonas and Jörgen W. Weibull, 1996, Nash equilibrium and evolution by imitation, In: K. Arrow and E. Colombatto (Eds.), The rational foundations of economic behaviour. New York, MacMillan, pp. 155-171.
8. Brown, George W., 1951, Iterative solution of games by fictitious play, In: T. C. Koopmans, (Ed.), Activity analysis of production and allocation. New York, Wiley, pp. 347-376.
9. Cressman, Ross, 2000. Subgame monotonicity in extensive form evolutionary games. Games and Economic Behavior, **32,** 183-205.
10. Cressman, Ross, 2001, Evolutionary Dynamics and Extensive Form Games (mimeo).
11. Cressman, Ross, Gaunersdorfer, Andrea and Jean-Francois Wen, 2000. Evolutionary and dynamic stability in symmetric evolutionary games with two independent decisions. International Game Theory Review, **2,** 67-81.
12. Friedman, Daniel, 1991. Evolutionary games in economics. Econometrica **59,** 637-666.
13. Fudenberg, Drew and David K. Levine, 1998, The Theory of Learning in Games (MIT Press, Cambridge, MA).
14. Gaunersdorfer, Andrea, Hofbauer, Josef and Karl Sigmund, 1991. On the dynamics of asymmetric games. Theoretical Population Biology, **39,** 345-357.
15. Gilboa, Itzhak and Akihiko Matsui, 1991. Social stability and equilibrium. Econometrica, **59,** 859-867.
16. Hofbauer, Josef, 1995. Stability for the best response dynamics. Mimeo, University of Vienna.
17. Hofbauer, Josef and Karl Sigmund, 1998, Evolutionary games and population dynamics (Cambridge Univ. Press, Cambridge, UK).
18. Krishna, Vija and Tomas Sjöström, 1998. On the convergence of fictitious play. Mathematics of Operations Research, **23,** 479-511.
19. Matsui, Akihiko, 1992. Best response dynamics and socially stable strategies. Journal of Economic Theory, **57,** 343-362.
20. Maynard Smith, John, 1982, Evolution and the theory of games (Cambridge University Press, Cambridge, UK).
21. Milgrom, Paul and John Roberts, 1991. Adaptive and sophisticated learning in repeated normal form games. Games and Economic Behavior **3,** 82-100.
22. Miyasawa, Koichi, 1961. On the convergence of the learning process in a 2×2 non-zero-sum two person game. Economic Research Program, Princeton University, Research Memorandum No. 33.

23. Monderer, Dov and Lloyd S. Shapley, 1996. Potential Games. Games and Economic Behavior, **14,** 124-143.
24. Nachbar, John, 1990. "Evolutionary" selection dynamics in games: Convergence and limit properties. International Journal of Game Theory **19,** 59-89.
25. Nash, John, 1950, Non-cooperative games, unpublished Ph.D. thesis (Princeton University).
26. von Neumann, John and Oskar Morgenstern, 1944, Theory of games and economic behavior (Princeton University Press).
27. Robinson, Julia, 1951. An iterative method of solving a game. Annals of Mathematics, **54,** 296-301.
28. Rosenmüller, Joachim, 1971. Über Periodizitätseigenschaften spieltheoretischer Lernprozesse. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, **17,** 259-308.
29. Samuelson, Larry and Jianbo Zhang, 1992. Evolutionary stability in asymmetric games. Journal of Economic Theory **57,** 363-391.
30. Schlag, Karl, 1998. Why imitate, and if so, how? A bounded rational approach to multi-armed bandits. Journal of Economic Theory **78,** 130-156.
31. Selten, Reinhard, 1980. A note on evolutionarily stable strategies in asymmetric animal conflicts. Journal of Theoretical Biology, **84,** 93-101.
32. Shapley, Lloyd S., 1964. Some topics in two-person games. Annals of Mathematical Studies, **5,** 1-28.
33. Taylor, Peter D. and Leo B. Jonker, 1978. Evolutionarily stable strategies and game dynamics. Mathematical Biosciences, **40,** 145-156.
34. Weibull, Jörgen W., 1994. The "as if" approach to game theory: Three positive results and four obstacles. European Economic Review **38,** 868-881.
35. Weibull, Jörgen W., 1995, Evolutionary game theory (MIT Press, Cambridge, MA).