

# Learning and Evolution in a Heterogeneous Population

Ed Hopkins\*  
Department of Economics  
University of Edinburgh  
Edinburgh EH8 9JY, UK  
E.Hopkins@ed.ac.uk

April, 1997

## Abstract

A framework is proposed for investigating the effect of evolutionary selection on a population where some agents learn. It is shown that learning behaviour when aggregated has different properties than when considered at the level of the individual and that a combination of learning and evolution has different properties in terms of stability than when considered separately. Convergence is shown for all  $2 \times 2$  games and a famous  $3 \times 3$  example.

---

\*I would like to thank Alan Kirman, Mark Salmon, Ken Binmore, Debora di Gioacchino and an anonymous referee for comments and suggestions.

# 1 Introduction

Game theorists have recently shown an increasing interest in modelling both learning and evolution. Nash equilibrium (and its refinements) place strong requirements on the rationality and the computational ability of players and on the information they must possess. In switching to models with boundedly-rational agents the hope has been not only to weaken those demands but also to select between equilibria in a manner which is more intuitive. Unfortunately, the dynamics considered do not necessarily converge and thus fail to give clear predictions. The results here indicate that in part this failure arises from too narrow a focus. Most research has concentrated on properties of individual algorithms. We examine a model where there is both learning and evolution and find quite different results from when they are considered separately. In particular, there is convergence for a wider class of games.

There are obvious similarities between the properties of adaptive learning and evolutionary dynamics. Typically, both are concerned with the development of the distribution of strategies within some large population<sup>1</sup>. As Cabrales and Sobel (1992) show, evolutionary dynamics under certain conditions can be “consistent with adaptive learning” in the sense of Milgrom and Roberts (1991). But this is only a condition on the asymptotic behaviour of a selection or learning process. In the short run, although “consistent”, different processes may behave quite differently. In particular, while selection dynamics are typically smooth functions of current strategy distributions, under fictitious play or Cournotian dynamics, where players make best responses to previous play(s) of opponents, there can be discontinuous jumps in play. Convergence to mixed strategies is in particular troublesome (for example, see Fudenberg and Kreps 1993; Jordan, 1993). Here it is shown that if one aggregates such behaviour across a large population, smoothness is obtained.

The standard evolutionary dynamic framework assumes that agents compete in some game and then reproduce according to the success they obtain. Here I make the (strong) assumption that the population is randomly matched an infinite number of times in each “generation” to play the game. The population is heterogeneous in that some agents learn. At the end of their “lifespan” agents reproduce according to the success of the strategies they develop, or, to be precise, according to the limit of this learning process. Thus, there are two mechanisms that can change the mix of strategies in the population. Agents can change their own strategies, a “learning” process, and an evolutionary mechanism also chooses between different agents, the “selection” process.

The combination of the two has quite different implications for the stability of equilibrium than each considered in isolation. We show that the distribution of strategies in the population converges to Nash frequencies for all  $2 \times 2$  asymmetric games and also for a famous  $3 \times 3$  game first proposed by Shapley in 1964. Shapley’s original pessimistic result has been confirmed and generalised by more recent research, (Jordan, 1993). It is

---

<sup>1</sup>Some papers in the first camp include Milgrom and Roberts (1991), Kandori et al. (1993), Young (1993); in the second, Nachbar (1990), Samuelson and Zhang (1992).

therefore particularly striking that, even given the particular assumptions of this model, that a population can converge to the Nash equilibrium of such a game.

## 2 Learning and Selection

In this section, we first set out a standard model of evolutionary dynamics. We then explain why mixed strategies of asymmetric games are typically unstable in this setup. We go on to modify the model by the introduction of a simple learning process.

An infinite population is repeatedly, randomly matched to play a two-player normal-form game,  $G = (\{1, 2\}, I, J, A, B)$ . We develop the model and notation on the basis that the game is asymmetric (in the evolutionary sense), in which case the players labelled 1 are drawn from a different “population” from the players labelled 2. For example, in the “Battle of the Sexes” game, players are matched so that a female always plays against a male.  $I$  is a set of  $n$  strategies, available to the first population,  $J$ , the set of  $m$  strategies of the second population. Payoffs for the first population are determined by  $A$ , a  $n \times m$  matrix of payoffs, with typical element  $a_{ij}$ , which is the payoff a member of the first population receives when playing strategy  $i$  against a member of the second population playing strategy  $j$ .  $B$ , with typical element  $b_{ji}$ , is the  $m \times n$  equivalent for the second population. There are  $n + m$  “types” of agent, each associated with one strategy. The state of the system can thus be summarised by the proportions of the population playing each strategy  $\mathbf{x} = (x_1, \dots, x_n)$ ,  $\mathbf{y} = (y_1, \dots, y_m)$ . That is, the state space is the Cartesian product of the simplexes,  $S_n \times S_m$  where  $S_n = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^n : \sum x_i = 1, x_i \geq 0, \text{for } i = 1, \dots, n\}$ . Define the interior (or,  $\text{int } S_n \times S_m$ ), as all states where all types have strictly positive representation, and define the boundary as all states where at least one type has zero representation. The symbol “ $\cdot$ ” indicates multiplication by a transpose, and the notation  $(A\mathbf{y})_i$  indicates the  $i$ th element of the vector in parentheses.

The problem with which we are really concerned with here is the generic instability of mixed strategy equilibria in asymmetric games under adaptive dynamics. Hofbauer and Sigmund (1988) set out the reasons for this in the case of evolutionary dynamics. In an environment where each member of the first population is randomly matched with a member of the second, the expected payoffs for the first population are  $A\mathbf{y}$  and  $B\mathbf{x}$  for the second. We assume that

$$\dot{\mathbf{x}} = Q(\mathbf{x})A\mathbf{y} \text{ and } \dot{\mathbf{y}} = P(\mathbf{y})B\mathbf{x} \quad (1)$$

where  $Q, P$  are symmetric positive semidefinite matrices. This is a very general formulation for adaptive processes, including the evolutionary replicator dynamics and some learning processes as special cases (see Hopkins, 1995; Hofbauer and Sigmund, 1990). If we linearise the dynamics at a fully mixed fixed point  $\xi$ , we obtain

$$R = \begin{pmatrix} 0 & Q(\xi)A \\ P(\xi)B & 0 \end{pmatrix}. \quad (2)$$

Because the payoffs of the first population depend only on  $\mathbf{y}$  and not  $\mathbf{x}$  and conversely for the second population, the trace of this matrix (2) is zero. Consequently the eigenvalues are either a mixture of negative and positive or they have zero real part. In discrete time all such equilibria are always unstable. If we replace  $\dot{\mathbf{x}}$  and  $\dot{\mathbf{y}}$  in (1) by  $\mathbf{x}(t+1) - \mathbf{x}(t)$  and  $\mathbf{y}(t+1) - \mathbf{y}(t)$  respectively, then the linearisation at a mixed equilibrium is  $I + R$ , and has eigenvalues  $1 + \mathbf{r}$ , where  $\mathbf{r}$  is the vector of eigenvalues of  $R$ . Given that, as we have seen,  $R$  possesses either a mixture of positive and negative eigenvalues or eigenvalues with zero real part, it is easy to show that the matrix  $I + R$  always has at least one eigenvalue of absolute value greater than one. However, in continuous time, in the case of eigenvalues with zero real part, the linearisation does not determine stability, this will be determined by the equations' higher order terms. Such equilibria are not structurally stable in that small changes in the structure of the game or the dynamics will affect stability. For example, equilibria can be stable in continuous time even though unstable in discrete time.

However, many mixed equilibria in asymmetric games are saddlepoints. Saddlepoints are of course unstable and this property is structurally stable. In other words, small variations in the specifications of the dynamic cannot make the equilibrium stable. Instability of these mixed equilibria can often make intuitive sense in that there are games which also possess stable pure equilibria which seem more plausible outcomes (see the discussion of asymmetric games in Maynard Smith, 1982; or in the context of human society, Sugden, 1989). Or to put it another way, the instability allows us to select between equilibria. However, there are many games which possess a unique mixed equilibrium. A famous example is the following game first discussed by Shapley (1964). This possesses a unique mixed equilibrium which is a saddle, with convergence only occurring if the first population starts in its equilibrium state, that is, with each of the three strategies with equal representation. From all other initial conditions, any dynamic satisfying (1) will diverge from equilibrium.

$$A = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} \quad B = \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array} \quad (3)$$

Could heterogeneity help with this problem? That is, if there were a diversity in the types of behaviour present in either one or both populations, could this change the stability properties of mixed equilibria? In the framework we have sketched up to now, it will not. If we require learning rules to be expressible in the manner of (1) as a positive definite transformation of the payoffs, it does not matter how many such rules are present in the population. It is easy to verify that the aggregation of any such rules would itself be a positive definite function of the payoffs. To produce real qualitative change, it is necessary to consider a wider deviation.

It is tempting to look in the direction of best response dynamics or fictitious play, because they offer behaviour which is qualitatively different. Rather than offering a smooth reaction to payoffs, there can be discrete jumps in play. It is not that this in

itself makes convergence properties any better. For example, Krishna and Sjöström (1995) have recently found that mixed strategy equilibria in non-zero sum games with more than two strategies are unstable for fictitious play. Rather it is the possibilities offered by the combination of different processes. Banerjee and Weibull (1995) consider the case where a proportion of the population is “rational”, that is, they play a strategy which is a best reply to the current state of the population. The result is quite striking. Every Nash equilibrium becomes stable. It is very interesting to see rationality only on a part of the population is enough to give results very similar to those of standard game theory. However, given that one of the main reasons for investigating adaptive dynamics is to select between equilibria, another approach is required.

The basic model is modified here by the addition of a type capable of inheriting rules more complex than simply to play a fixed strategy. As noted above, Banerjee and Weibull (1995) use a similar approach in the context of symmetric games, as does Stahl (1993). Thus, we now have  $n + 1, m + 1$  types, and we work in  $S_{n+1} \times S_{m+1}$ . We can think of each generation being divided into an infinite number of subperiods  $(0, 1, \dots, s, \dots)$ . As a reminder, the selection process operates between generations, the learning process within generations. We assume that the  $n+1$ th and  $m+1$ th type adjust their strategies so that they play what is an optimum response to the strategy of their previous opponent: the “best-response” or Cournotian dynamic. Similar behavioural hypotheses have been employed in recent learning literature (for example, Milgrom and Roberts, 1991; Kandori et al., 1993; Young, 1993), but here the implementation is particularly simple. Agents do not need to know anything about the overall distribution of strategies in the population or to have a memory longer than one subperiod. Yet, as we will see, this is enough to ensure convergence to Nash equilibrium in a large class of games.

Thus, at any given time, different members of the additional type may be playing different strategies. Let  $\mathbf{p}(s) = (p_1(s), \dots, p_n(s))$  and  $\mathbf{q}(s) = (q_1(s), \dots, q_m(s))$  where  $p_i(s)$  and  $q_j(s)$  denote respectively the proportion of this  $n + 1$ th type of population 1 playing the  $i$ th strategy, and the proportion of the  $m + 1$ th type playing the  $j$ th strategy at a given subperiod  $s$ . As  $I, J$  are finite, it is a standard result that for any pure strategy in  $I$ , there exists at least one element of  $J$  which is a best response to that strategy. Or  $\forall i \in I \quad \exists b_{j^*i} \geq b_{ji}$ . First, define  $I', J'$  as those subsets of  $I$  and  $J$  respectively of strategies which have current positive representation in the two populations. Second, let  $\mu_j$  represent the number of strategies in  $I$  which are equal best responses to strategy  $j$ . Third, let  $J_i = \{j \in J' : i = \operatorname{argmax}_{i \in I} a_{ij}\}$  be the set of strategies to which  $i$  is the best reply, and, equivalently, let  $I_j = \{i \in I' : j = \operatorname{argmax}_{j \in J} b_{ji}\}$ .

The probability that an individual of type  $n+1$ , in population 1, meets an individual of type  $j$  in the second population is  $y_j$ . There is also a probability  $q_j y_{m+1}$  of meeting an individual of type  $m + 1$  currently playing strategy  $j$ . In either case, faced with an opponent playing strategy  $j$ , the individual will play in the next subperiod a strategy which is a best reply to  $j$ . Thus, *within* each generation, each  $p_i$  evolves according to a

mapping  $S_m \rightarrow [0, 1]$

$$p_i(s+1) = \sum_{j \in J_i} \frac{y_j}{\mu_j} + \sum_{j \in J_i} \frac{y_{m+1} q_j(s)}{\mu_j} \quad (4)$$

Thus, although individual choices are made according to the best-reply dynamic, the distribution of strategies in the population is a continuous function of the previous subperiod's distribution. We make the assumption that when there are alternative best responses each agent chooses independently. Then by the law of large numbers each alternative response is chosen by an equal number of agents. This explains the presence of  $\mu_j$ , denoting the number of alternative best replies. Naturally if  $J_i = \emptyset$ ,  $p_i = 0$ , and if  $J_i = J'$ ,  $p_i = 1$ . These represent respectively the cases where  $i$  is not a best reply to any strategy, and where it is the dominant strategy. Similarly, for the second population,

$$q_j(s+1) = \sum_{i \in I_j} \frac{x_i}{\mu_i} + \sum_{i \in I_j} \frac{x_{n+1} p_i(s)}{\mu_i}. \quad (5)$$

**Lemma 1** *If, at time  $t$ ,  $1 > x_{n+1}(t), y_{m+1}(t)$ , the learning process represented by equations (4), (5) has an unique fixed point  $\mathbf{p}^*, \mathbf{q}^* \in S_n \times S_m$ .*

**Proof:** Though they change between generations, within each generation the population proportions  $\mathbf{x}, \mathbf{y}$  are fixed and are therefore constants for (4), (5). Consequently, the equations are simple, linear difference equations. Written in matrix form, they become

$$\mathbf{P}(s+1) = \mathbf{x}_0 + X_1 \mathbf{P}(s)$$

where  $\mathbf{P} = (p_1, \dots, p_n, q_1, \dots, q_m)$ ,  $\mathbf{x}_0$  is the vector of terms in  $x_i, y_j$ , and  $X_1$  is the matrix of terms in  $x_{n+1}$  or  $y_{m+1}$ . In equilibrium we have,  $\mathbf{P}^* = (I - X_1)^{-1} \mathbf{x}_0$ . By inspection of (4) and (5), it is possible to see that the coefficients on the  $p_i$  in the first  $n$  equations are all zero, as are the coefficients on the  $q_j$  in the next  $m$  equations. Hence,  $(I - X_1)$  can be partitioned in the following manner:

$$I - X_1 = \begin{pmatrix} I & -X_{12} \\ -X_{21} & I \end{pmatrix}$$

Each column of  $X_{12}$  and  $X_{21}$  sums to  $y_{m+1}$  and  $x_{n+1}$  respectively. Thus  $(I - X_1)$  is singular if and only if neither  $X_{12}$  and  $X_{21}$  are linearly independent of  $I$ , which can only be the case if  $x_{n+1} = y_{m+1} = 1$ . Otherwise, there is a unique fixed point,  $\mathbf{P}^*(\mathbf{x}, \mathbf{y}) = (\mathbf{p}^*, \mathbf{q}^*)$ . Because of the linearity of these equations, this solution will be a function of  $(\mathbf{x}, \mathbf{y})$ , Lipschitz continuous on the interior of  $S_{n+1} \times S_{m+1}$ .  $\square$

The exact value of this solution depend entirely on the value of the  $x_i, y_i$  and not on the value of  $\mathbf{p}, \mathbf{q}$  at the beginning of the learning process. Furthermore, the sufficient condition for the existence of an unique fixed point is also a sufficient condition for convergence.

**Lemma 2** *If, at time  $t$ ,  $1 > x_{n+1}(t), y_{m+1}(t)$ , the learning process converges to its unique fixed point.*

**Proof:** (4), (5) represent a system of  $n + m$  linear first order difference equations. The  $x_i, y_i$  are constant within each generation, and therefore are constants for (4), (5). In particular, the coefficients on the variables  $\mathbf{p}, \mathbf{q}$  on the right hand side are the  $x_{n+1}/\mu_i, y_{m+1}/\mu_j$ , the sum of which in each equation have an upper bound in value of either  $x_{n+1}$  or  $y_{m+1}$ . By the elementary theory of difference equations if this sum is less than unity for all equations, so are all the roots of the dynamic system.  $\square$

It is worth remarking that here convergence is not convergence in empirical frequencies, a notion of convergence that has been forcefully criticised in the recent literature (Young, 1993; Fudenberg and Kreps, 1993; Jordan, 1993). In this case, one does not have to take a time average. As the limit approaches, strategies are actually played at limiting frequencies.

I make the assumption that payoffs during the learning process do not affect the rate of reproduction. Rather it is the limit of the learning process, denoted  $(\mathbf{p}^*, \mathbf{q}^*)$  which determines reproductive fitness. (Compare Harley's assumption (e): "The learning period is short compared to the subsequent period of stable behaviours"; 1981, p613). This construction has some analytic convenience: if one assumes only a finite number of plays each period, the values of  $\mathbf{p}, \mathbf{q}$  will be dependent on their (arbitrary) initial values. We would have to make further assumptions about how much of the behaviour learnt within a period is transmitted between the generations. For example, we could assume that each generation starts from scratch: at the beginning of each period  $\mathbf{p}(0), \mathbf{q}(0)$  are randomly determined. That is, "children" learn nothing from their "parents". Or we can assume that the initial values are some function of play by the previous generation. However, using the limit, the value of  $(\mathbf{p}^*, \mathbf{q}^*)$  will be the same in either case.

As stated we use these limiting values to determine fitness. At the end of the learning process the total proportion of the first population adopting the  $i$ th strategy will be given by  $z_i = x_i + x_{n+1}p_i^*$ , and, the proportion of the second population adopting the  $j$ th strategy by  $w_j = y_j + y_{m+1}q_j^*$ . Given the assumption of random matching it is these overall distributions which decide fitness. For the first  $n, m$  types this will be, given the normal form game  $G$ ,

$$\pi_{xi} = (A\mathbf{w})_i, \quad \pi_{yj} = (B\mathbf{z})_j, \quad (6)$$

and for the learners,

$$\pi_{xn+1} = \mathbf{p}^* \cdot A\mathbf{w}, \quad \pi_{ym+1} = \mathbf{q}^* \cdot B\mathbf{z} \quad (7)$$

With fitnesses defined, we can propose as a selection mechanism the following replicator dynamics:

$$x_i(t+1) = f_{xi}(\mathbf{x}, \mathbf{y}) = x_i(t) \frac{\pi_{xi} + C}{\mathbf{z} \cdot A\mathbf{w} + C}, \quad y_j(t+1) = f_{yj}(\mathbf{x}, \mathbf{y}) = y_j(t) \frac{\pi_{yj} + C}{\mathbf{w} \cdot B\mathbf{z} + C}, \quad (8)$$

where  $C$  is an arbitrary constant. Alternatively, taking the limit, as generations become

arbitrarily short:

$$\dot{x}_i = F_{xi}(\mathbf{x}, \mathbf{y}) = x_i(t)(\pi_{xi} - \mathbf{z} \cdot A\mathbf{w}), \quad \dot{y}_j = F_{yj}(\mathbf{x}, \mathbf{y}) = y_j(t)(\pi_{yj} - \mathbf{w} \cdot B\mathbf{z}) \quad (9)$$

where  $\mathbf{z} \cdot A\mathbf{w}$ ,  $\mathbf{w} \cdot B\mathbf{z}$  are the average payoffs for the two populations. Inspection of (9) shows that this continuous selection mechanism has the following important property:

*Invariance.* As  $\sum_{i=1}^{n+1} F_{xi} = \sum_{j=1}^{m+1} F_{yj} = 0$ , the interior of the simplex is invariant under  $F$ . Starting from any interior point, the boundary is never reached in finite time. That is, if  $(\mathbf{x}(0), \mathbf{y}(0)) \in \text{int } S_{n+1} \times S_{m+1}$ , then  $(\mathbf{x}(t), \mathbf{y}(t)) \in \text{int } S_{n+1} \times S_{m+1}$  for all  $t \in \mathbf{R}$ .

If we impose the condition that  $C$  is sufficiently large such that both denominator and numerator in (8) are strictly positive<sup>2</sup> for all  $i, j$ , invariance will also hold for the discrete dynamic  $f$ . Given that  $\mathbf{p}^*, \mathbf{q}^*$  are themselves functions of the frequencies of types in the population, fitnesses will not be linear in  $\mathbf{x}, \mathbf{y}$  - a usual assumption of the replicator dynamics - and perhaps not even be defined when  $x_{n+1}$  and  $y_{m+1}$  are equal to one. However, by Lemmas 1 and 2, fitnesses are continuous functions of  $\mathbf{x}, \mathbf{y}$  elsewhere. This, combined with invariance implies that from any fully-mixed initial conditions, (that is,  $x_i > 0, i = 1 \dots n + 1$  and  $y_j > 0, j = 1 \dots m + 1$ ), the learning process converges, and fitnesses are defined, for all  $t \in \mathbf{R}$ . Thus while both  $f$  and  $F$  are not continuous on all of  $S_{n+1} \times S_{m+1}$  they are continuous on its interior. In other words, both  $f$  and  $F$  possess a limit even along a dynamic path with an accumulation point on the boundary of  $S_{n+1} \times S_{m+1}$ , even if that limit may be path-dependent.

What is important about this definition of fitness is that there is a fundamental difference from the standard evolutionary model. Fitnesses for the first population, for example, depend on  $\mathbf{w}$  which through  $\mathbf{q}^*$  depends on  $\mathbf{x}$ . Consequently  $d\pi_{xi}/d\mathbf{x} \neq 0$  and any linearisation at a fully mixed equilibrium does not have the same structure as (2). That is, mixed strategy equilibria of asymmetric games could be asymptotically stable in this framework. This possibility we now investigate.

### 3 Equilibrium

Equilibrium in this model consists of a population distribution which is a rest point for both selection and learning processes. That is, a state of the system where the limit of the learning process is such that all types present in the population earn the same average payoff. In the standard evolutionary model, that is, in the absence of the learners, under the selection dynamics defined by (8) or (9), denote the rest points for the game  $G$  in the interior of  $S_n \times S_m$ ,  $(\mathbf{z}^*, \mathbf{w}^*)$ . It is well known that such rest points

---

<sup>2</sup>An increase in the value of  $C$  is equivalent to the addition of an equal amount to the game matrices  $A, B$ . This will not change the best response structure or Nash equilibria but may change the qualitative behaviour of the discrete replicator dynamics. See Cabrales and Sobel (1992) for a discussion of the issues involved.



are Nash equilibria (Hofbauer and Sigmund, 1988; Nachbar, 1990)<sup>3</sup>. For the extended game, the conditions for an interior rest point under the selection dynamics are

$$\pi_{x_1} = \dots = \pi_{x_{n+1}}, \quad \pi_{y_1} = \dots = \pi_{y_{m+1}} \quad (10)$$

Furthermore, as  $(\mathbf{x}, \mathbf{y})$  are both constant if (10) holds, the limit for the learning process is also unchanging across all subsequent generations. The consequent distribution of strategies is a Nash equilibrium. Comparison of equations (6), (7), reveal that any values of  $(\mathbf{x}, \mathbf{y})$  that satisfy the above condition (10), also satisfy  $x_i + p_i^* x_{n+1} = z_i^*$ , and  $y_j + q_j^* y_{m+1} = w_j^*$ . That is, it is a Nash equilibrium for the original game  $G$  in the sense that an outside observer would see, as the learning process reached its limit, strategies being played with the Nash equilibrium frequencies,  $(\mathbf{z}^*, \mathbf{w}^*)$ . Note that for each population there is now one less independent equation than there are independent variables. This means that any isolated equilibrium of the original game in the interior of  $S_n \times S_m$  will be represented by a continuum of fixed points in the interior of  $S_{n+1} \times S_{m+1}$ .

Furthermore, we can show that for all  $2 \times 2$  games the system will converge to a Nash equilibrium. (11) gives a general  $2 \times 2$  game.

$$A = \begin{array}{|c|c|} \hline 0 & a \\ \hline 1-a & 0 \\ \hline \end{array} \quad B = \begin{array}{|c|c|} \hline 0 & b \\ \hline 1-b & 0 \\ \hline \end{array} \quad (11)$$

where  $-1 \leq a, b \leq 1$ . If  $1 > |a| > 0$  and  $1 > |b| > 0$  then there is a mixed Nash equilibrium where the first strategy of each population are represented with frequencies  $(b, a)$  respectively. The interesting case is when additionally  $ab < 0$ , as in this case, the mixed Nash equilibrium is unique, yet the replicator dynamics do not converge. However, the addition of an arbitrarily small initial population of learners is enough to stabilise the dynamics.

We start by assuming (without loss of generality) that  $a < 0$ . We can then solve for  $\mathbf{p}^*$  and  $\mathbf{q}^*$  and we find that  $(\mathbf{p}^*, \mathbf{q}^*) = (y_1 + y_3(1 - x_1), 1 - x_1 - x_3 y_1) / (1 + x_3 y_3)$ . Given this and the condition that in equilibrium  $(x_1 + p_1 x_3, y_1 + q_1 y_3) = (b, a)$ , we have a continuum of equilibrium points defined by the two equations

$$x_1 = b - a x_3, \quad y_1 = a - (1 - b) y_3.$$

Using  $x_1, x_3, y_1, y_3$  as our variables, if we take the linearisation of the continuous time dynamics  $F(\mathbf{x}, \mathbf{y})$  at any point in this continuum, it gives us a Jacobian:

$$J = \frac{1}{1 + x_3 y_3} \begin{pmatrix} -a y_3 & -a^2 y_3 & a & a(1 - b) \\ (1 - 2a) y_3 & (1 - 2a) a y_3 & 2a - 1 & (2a - 1)(1 - b) \\ -b & -ab & -b x_3 & -b(1 - b) x_3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Some further calculation reveals that  $J$  possesses two zero eigenvalues, which reflect the continuum of equilibrium points, and two with real part negative for  $0 < x_3, y_3 < 1$ .

<sup>3</sup>All states that consists of just one type are also rest points, but not all are Nash equilibria.

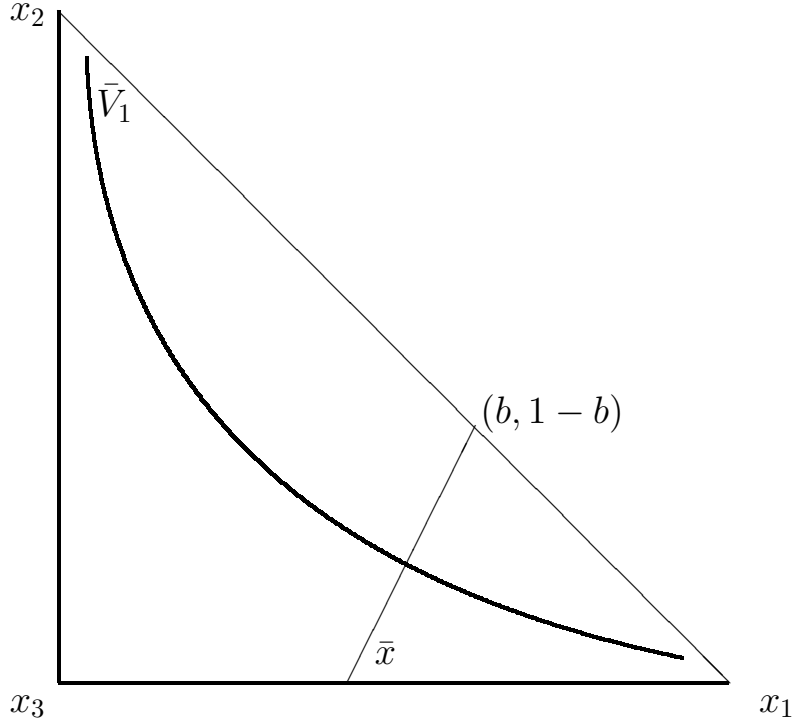


Figure 1: Continuum of equilibria  $\bar{x}$  and a level curve of the function  $V_1$ .

Given the presence of the zero eigenvalues, the linearisation is not in itself sufficient to determine stability. However, we can establish this by other means. Basically, the proof follows from the example given by Hofbauer and Sigmund (1988, p131-3) who introduce an additional type to the standard (symmetric) evolutionary model which plays a fixed mixed strategy. In the model considered here the behaviour of the additional type changes over time and hence the orbits of the evolutionary dynamic will be somewhat different. However, they all approach the continuum of mixed strategy equilibria.

The structure of the proof is illustrated in Figure 1. To draw this diagram we fix the value of  $\mathbf{y}$  at some arbitrary value but with  $1 > y_3 > 0$ . We can then draw in the continuum of equilibria  $\bar{x}$ . We define a function  $V_1$  first demonstrate that all orbits stay inside the set.

**Proposition 1** *If  $ab < 0$ , then the mixed equilibrium is asymptotically stable, for any fully mixed initial conditions.*

**Proof:** Define  $V_1 = x_1^b x_2^{1-b} y_1^a y_2^{1-a}$ .

$$\dot{V}_1 = V_1[(b - z_1)((A\mathbf{w})_1 - (A\mathbf{w})_2) + (a - w_1)((B\mathbf{z})_1 - (B\mathbf{z})_2)] = 0$$

or in other words  $V_1$  is a constant of motion. All orbits on the interior of  $S_{n+1} \times S_{m+1}$  will be within the level sets of  $V_1$ . This implies that all orbits either flow toward the

interior equilibrium, or flow toward a boundary where  $x_3$  and/or  $y_3$  are zero. However, define  $V_2 = x_1^a x_2^{1-a} x_3^{-1} y_1^{1-b} y_2^b y_3^{-1}$ , which approaches infinity as it approaches such a boundary.

$$\dot{V}_2 = V_2[(a - p_1)((A\mathbf{w})_1 - (A\mathbf{w})_2) + (q_1 - b)((B\mathbf{z})_1 - (B\mathbf{z})_2)] \leq 0$$

Evidently, orbits flow away from the boundary to the equilibrium, which must attract all the interior.  $\square$

In the case of discrete time, it is not possible to construct similar Liapunov functions, but we can note the following. The equivalent discrete system has linearisation  $I + J'$ , where  $J'$  is identical to  $J$  except that the first two rows are divided by  $\mathbf{z} \cdot A\mathbf{w} + C$ , and the third and fourth by  $\mathbf{w} \cdot B\mathbf{z} + C$ . Hence,  $J'$  also has eigenvalues with real part nonpositive, whose absolute value decreases to zero as  $C \rightarrow \infty$ . Thus, there is a  $C$  for which the eigenvalues of  $I + J'$  are less than or equal to one in absolute value. We can therefore be reasonably confident that also in discrete time the mixed equilibrium is asymptotically stable.

Note that not all mixed strategies are stable. That is, the dynamics can still be used to select between equilibria. If  $ab > 0$  then this interior equilibrium is a saddle. Similar arguments to those employed in Proposition 1, can be used to show that in this case, the system behaves in much the same way as standard evolutionary dynamics and flows toward the Nash equilibria located on the boundaries of the simplex.

The mixed equilibrium when  $ab < 0$  is non-hyperbolic (that is, the linearisation has eigenvalues with zero real part) for the continuous time replicator dynamics and hence not structurally stable. In this sense, a modification of the dynamics would be expected to change their qualitative behaviour. However, for the discrete time dynamics, the equilibrium is hyperbolic and hence robustly unstable. We have seen that even in this case, the addition of learning can stabilise the equilibrium. We go on to show that it can drive convergence to an equilibrium which seems to be unstable under every form of adaptive dynamic.

## 4 A $3 \times 3$ Example

The famous example given by Shapley (1964) to demonstrate non-convergence of fictitious play is shown in (3). The only Nash equilibrium of this game is interior, where both row and column play each of their strategies with equal probability. As we have seen, interior (mixed) equilibria of asymmetric games are never asymptotically stable under the replicator dynamics. Thus this game does not converge for the replicator dynamics, just as it does not for fictitious play. Recent research on learning and evolution has only served to confirm the robustness of this result (see for example, Jordan, 1993). However, under this modified system this game converges to the unique Nash equilibrium.

Starting from a fully-mixed initial state, the proportions of type 4 playing each strategy evolve according to:

$$\mathbf{p}(s+1) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + y_4 \mathbf{q}(s), \quad \mathbf{q}(s+1) = \begin{pmatrix} x_3 \\ x_1 \\ x_2 \end{pmatrix} + x_4 \begin{pmatrix} p_3(s) \\ p_1(s) \\ p_2(s) \end{pmatrix} \quad (12)$$

This is a system of six linear difference equations. By Lemma 2 we know that the fixed point of this system is the limit of the learning process. This can be calculated using standard methods. It would be possible to eliminate  $\mathbf{p}, \mathbf{q}$  by substitution using these results. However, it is easier to work in the other direction. We construct  $z_i(t) = x_i(t) + x_4(t)p_i^*(t)$ ,  $w_j(t) = y_j(t) + y_4(t)q_j^*(t)$ ;  $i, j = 1, 2, 3$ , where  $z_i$  is the total number of the first population playing strategy  $i$ , and  $w_j$  is the total number of the second playing strategy  $j$ . Note that (12) here implies that  $\mathbf{z}(t) = (z_2^*(t), z_3^*(t), z_1^*(t))$ , that  $\mathbf{w} = \mathbf{p}$  and that  $\mathbf{q} \cdot B = \mathbf{z} \cdot A$ . There is an interior equilibrium for this system: the plane such that  $x_1 = x_2 = x_3$ ,  $y_1 = y_2 = y_3$ , which we denote  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ . In such an equilibrium, (12) in turn implies that  $\mathbf{p} = \mathbf{q} = (1/3, 1/3, 1/3)$ . I now prove that the limit point of all solutions under  $f$ , given fully-mixed initial conditions, is on this plane (normally for the discrete dynamics the interior equilibrium is a repellor).

**Proposition 2** *The plane of equilibria  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  under  $f$  attracts all other points on the interior of  $S_4 \times S_4$ .*

**Proof:** Define  $V(\mathbf{x}, \mathbf{y}) = x_4 y_4$ . Given that  $x_4(t+1) = x_4(t) \frac{\mathbf{p} \cdot A \mathbf{w}}{\mathbf{z} \cdot A \mathbf{w}}$ , and that  $y_4(t+1) = y_4(t) \frac{\mathbf{q} \cdot B \mathbf{z}}{\mathbf{w} \cdot B \mathbf{z}}$ , it follows that  $V(t+1) - V(t) > 0$  if and only if

$$\mathbf{p} \cdot A \mathbf{w} \mathbf{q} \cdot B \mathbf{z} - \mathbf{z} \cdot A \mathbf{w} \mathbf{w} \cdot B \mathbf{z} = \mathbf{w} \cdot \mathbf{w} \mathbf{z} \cdot \mathbf{z} - \mathbf{z} \cdot \mathbf{w} \mathbf{w} \cdot \mathbf{z}^* > 0 \quad (13)$$

where  $\mathbf{z}^* = (z_3, z_1, z_2)$ . Divide through by  $\mathbf{w} \cdot \mathbf{w} \mathbf{z} \cdot \mathbf{z}$  to obtain:

$$1 - \cos \theta_{zw} \cos \theta_{wz^*} \geq 0$$

It follows that  $V(t+1) \geq V(t)$  with equality only at  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ .  $V(\mathbf{x}, \mathbf{y})$  is therefore a strict Liapunov function on all of the interior of  $S_4 \times S_4$  less  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ . It is therefore unclear whether the system will have its limit at  $V = 1$ , or whether it will come to rest at another point on  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ . However, in either case  $\lim_{t \rightarrow \infty} \mathbf{p} = \lim_{t \rightarrow \infty} \mathbf{q} = (1/3, 1/3, 1/3)$ .  $\square$

Note that a similar result holds in continuous time. Again defining  $V(\mathbf{x}, \mathbf{y}) = x_4 y_4$ , we have  $\dot{V} > 0$  if and only if  $\mathbf{w} \cdot \mathbf{w} + \mathbf{z} \cdot \mathbf{z} - \mathbf{z} \cdot \mathbf{w} - \mathbf{w} \cdot \mathbf{z}^* > 0$ . It is easy to show given the above proof that this inequality holds on all of the interior of  $S_4 \times S_4$  less  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ .

One other question needs to be answered. Obviously the fact that the Liapunov function  $V$  is increasing implies that the population share of the learners is increasing also. But as  $x_4$  and  $y_4$  approach 1, the convergence of the learning process slows,

and further convergence really does become dependent on our assumption that learners play infinitely often. However, simulation reveals that in fact the limit of  $V$  is not, in general, 1, and equilibrium can be reached with the proportion of learners in the population remaining quite low.

We consider briefly two other examples. The first (14) is the familiar ROCK-SCISSORS-PAPER game, the second (15) is a game proposed by Dekel and Scotchmer (1992). They show that the DUMB strategy survives in the limit under the discrete replicator dynamics although it is never a best response and therefore not rationalizable.

$$A = B = \begin{array}{|c|c|c|c|} \hline \text{ROCK} & b & a & c \\ \hline \text{SCISSORS} & c & b & a \\ \hline \text{PAPER} & a & c & b \\ \hline \end{array} \quad a > b > c \quad (14)$$

ROCK-SCISSORS-PAPER is well-known as a problem game. While it does converge for fictitious play, it does so only in empirical frequencies. It only converges for the discrete replicator dynamics if  $ac > b^2$ . As for the first example, the limit of these games when learners are also present is the unique Nash equilibrium. As both these two additional examples have a similar structure, it is not surprising that they elicit similar behaviour. The function,  $x_{n+1}y_{m+1}$ , will again work as a Liapunov function and shows that in both cases there is convergence in population frequencies to the unique Nash equilibrium.

$$A = B = \begin{array}{|c|c|c|c|c|} \hline \text{ROCK} & 1 & 2.35 & 0 & 0.1 \\ \hline \text{SCISSORS} & 0 & 1 & 2.35 & 0.1 \\ \hline \text{PAPER} & 2.35 & 0 & 1 & 0.1 \\ \hline \text{DUMB} & 1.1 & 1.1 & 1.1 & 0 \\ \hline \end{array} \quad (15)$$

## 5 Discussion

Games such as (3) cause problems for conventional models because they possess cycles of best responses. For example, in Young (1993), cyclic games are excluded from the results on the convergence of a stochastic learning process. The fundamental reason that this model gives qualitatively different behaviour is that there are two distinct processes determining the change in the distribution of strategies, working at different speeds. By changing strategies, the learners anticipate the next stage of the cycle and “damp” the non-convergent tendencies of the original model. The dependence is two-way. Without the non-learners, the best-response process would not converge for this game.

One might argue that the simple learning rule considered here would be displaced by more sophisticated behaviour. However, this is an argument that does not find support in more recent work (Banerjee and Weibull, 1995; Blume and Easley, 1992; Stahl, 1993), where rational agents do not necessarily displace less rational ones. There is no claim that the learning rule considered here is the “correct” one. However, there is also no

strong evidence that evolution will select for more complex or sophisticated behaviour in a strategic environment.

Learning and evolution are ostensibly similar processes. However, while evolution is defined at the level of a population, learning is carried out by individuals. Crawford (1989) demonstrates that even when agents' learning is modelled in a similar manner to the replicator dynamics, an aggregation of their behaviour does not have the same properties in terms of stability as evolutionary dynamics. Similarly, in this paper even the most elementary learning behaviour gives increased stability when considered at the level of the population. This opens up the possibility of further research about the aggregate properties of populations where a number of different classes of behaviour are present.

## References

- A. Banerjee and J. Weibull** Evolutionary selection and rational behavior, in A. Kirman and M. Salmon eds. "Rationality and Learning in Economics," Blackwell, Oxford, 1995.
- L. Blume and D. Easley** Evolution and market behavior, *J. Econ. Theory*, **58** (1992), 9-40.
- A. Cabrales and J. Sobel** On the limit points of discrete selection dynamics, *J. Econ. Theory*, **57** (1992), 407-419.
- V. Crawford** Learning and mixed strategy equilibria in evolutionary games, *J. Theor. Biol.*, **140** (1989), 537-550.
- E. Dekel and S. Scotchmer** On the evolution of optimizing behavior, *J. Econ. Theory*, **57** (1992), 392-406.
- D. Fudenberg and D. Kreps** Learning mixed equilibria, *Games Econ. Behav.*, **5** (1993), 320-367.
- C. Harley** Learning the evolutionarily stable strategy, *J. Theor. Biol.*, **89** (1981), 611-633.
- J. Hofbauer and K. Sigmund** "The Theory of Evolution and Dynamical Systems," Cambridge University Press, Cambridge, 1988.
- J. Hofbauer and K. Sigmund** Adaptive dynamics and evolutionary stability, *Appl. Math. Lett.*, **3** (1990), 75-79.
- E. Hopkins** Learning, matching and aggregation, Working Paper, European University Institute, Florence, 1995.

- J. Jordan** Three problems in learning mixed strategy equilibria, *Games Econ. Behav.*, **5** (1993), 368-386.
- V. Krishna and T. Sjöström** On the convergence of fictitious play, Working Paper, 1995.
- M. Kandori, G. Mailath and R. Rob** Learning, mutation and long run equilibrium in games, *Econometrica*, **61** (1993), 29-56.
- P. Milgrom and J. Roberts** Adaptive and sophisticated learning in normal form games, *Games Econ. Behav.*, **3** (1991), 82-100.
- J.H. Nachbar** Evolutionary selection dynamics in games: convergence and limit properties, *Int. J. Game Theory*, **19** (1990), 59-89.
- L. Samuelson and J. Zhang** Evolutionary stability in asymmetric games, *J. Econ. Theory*, **57** (1992), 363-392.
- L. Shapley** Some topics in two person games, in eds. M. Dresher et al., "Advances in Game Theory," Princeton University Press, Princeton, 1964.
- D.O. Stahl** Evolution of smart-n players. *Games Econ. Behav.*, **5** (1993), 604-617.
- H.P. Young** The evolution of conventions, *Econometrica*, **61** (1993), 57-84.