

Extensive Form Games, Asymmetric Games and Games with Continuous Strategy Spaces

Ross Cressman, Wilfrid Laurier University

Introduction

The initial development of evolutionary game theory and evolutionary stability typically assumed

1. that an individual's payoff depends either on his strategy and that of his opponent used during a single interaction with another player (normal form game) or on his strategy and the current behavioral distribution of the population through a single random interaction (population game; playing-the-field model),
2. that the (pure) strategy set S available to an individual is finite and the same for each player (symmetric game).

Today, I will relax these assumptions in three different ways and consider the consequences for evolutionary dynamics, especially the replicator equation.

First, suppose that pairs of individuals have a series of interactions with each other and that the set of actions available at later interactions may depend on what choices were made in earlier ones. Many parlour games (e.g. tic-tac-toe, chess) are of this sort and it is my contention that most important "real-life" games involving humans or other species include a series of interactions among the same individuals. It is often more appropriate to represent such games in extensive form rather than normal form.

Second, in many cases, it is more reasonable to assume that strategies available to one player are different than those available to another. For instance, choices available when Black moves in chess are not usually the same as for White. Similarly, if players are from two different species, their strategy sets will almost surely be different (e.g. predator and prey). Suppose that there are two (or more) types of players and a finite set of strategies for each type. If there are exactly two types and the only interactions are single ones between a player of each type, we have a bimatrix game. Otherwise, it is a more general asymmetric game in either extensive or normal form.

Finally, I will consider briefly symmetric (asymmetric) games where the pure strategy set for each (type of) player is a continuum such as a subinterval of real numbers. Now the replicator equation is an infinite dimensional dynamical system on the space(s) of probability measures over the subinterval(s) that correspond to the distribution(s) of individual behaviors. Generalizations of the ESS (evolutionarily stable strategy) concept can be defined that characterize stability of single strategies (i.e. Dirac delta distributions) under the replicator equation as well as under the simpler canonical equation of adaptive dynamics that approximates the evolution of the mean distribution(s).

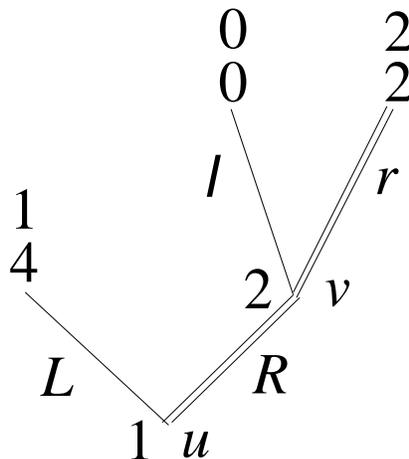


Figure 1: Extensive form for Example 1

1 Extensive Form Games

Although (finite, two-player) extensive form games are most helpful when used to represent a game with long (but finite) series of interactions between the same two players, differences with normal form intuition already emerge for short games with perfect information.

1.1 Perfect information games

A (finite, two-player) *perfect information game* is given by a rooted game tree Γ where each non-terminal node is a decision point of one of the players or of nature. The edges leading away from the root at each player decision node are this player's choices (or actions) at this node. There must be at least two choices at each player decision node. A pure (behavior) strategy for a player specifies a choice at all of his decision nodes. A mixed behavior strategy for a player specifies a probability distribution over the set of actions at each of his decision nodes. Payoffs to both players are specified at each terminal node $z \in Z$. A probability distribution over Z is called an outcome.

Example 1. (Weibull, 1995) Figure 1 is an elementary perfect information game with no moves by nature (i.e. at each non-terminal node, either player 1 or player 2 has a decision point). Player 1 has one decision node u where he chooses between the actions L and R . If he takes action L , player 1 gets payoff 1 and player 2 gets 4. If he takes action R , then we reach the decision point v of player 2 who then chooses between ℓ and r leading to both players receiving payoff 0 or both payoff 2 respectively.

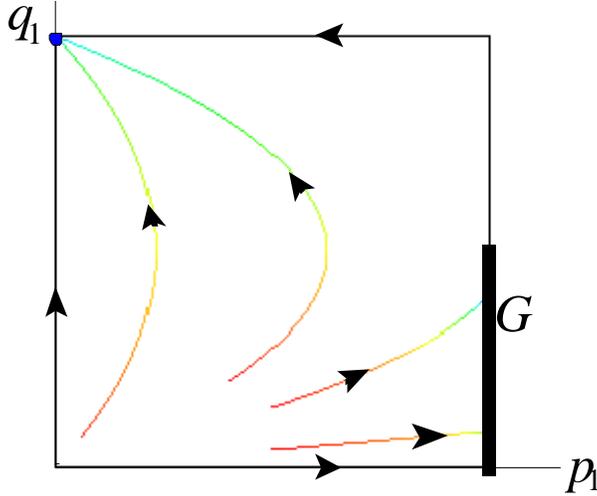


Figure 2: Trajectories for the replicator equation for Example 1

What are the Nash equilibria (NE) for this example? If players 1 and 2 choose R and r respectively with payoff 2 for both, then

1. player 2 does worse through unilaterally changing his strategy by playing r with probability q less than 1 (since $0(1 - q) + 2q < 2$) and
2. player 1 does worse through unilaterally changing his strategy by playing L with positive probability p (since $1p + 2(1 - p) < 2$).

Thus, the strategy pair (R, r) is a strict NE corresponding to the outcome $(2, 2)$.

In fact, if player 1 plays R with positive probability at a NE, then player 2 must play r . From this it follows that player 1 must play R with certainty (i.e. $p = 0$) (since his payoff of 2 is better than 1 obtained by switching to L). Thus any NE with $p < 1$ must be (R, r) . On the other hand, if $p = 1$ (i.e. player 1 chooses L), then player 2 is indifferent to what strategy he uses since his payoff is 4 for any (mixed) behavior. Furthermore, player 1 is no better off by playing R with positive probability if and only if player 2 plays ℓ at least half the time (i.e. $0 \leq q \leq \frac{1}{2}$). Thus

$$G \equiv \{(L, (1 - q)\ell + qr \mid 0 \leq q \leq \frac{1}{2}\}$$

is a set of NE, all corresponding to the outcome $(1, 4)$. G is called a NE component since it is a connected set of NE that is not contained in any larger connected set of NE. The NE structure of Example 1 consists of the single strategy pair $G^* = \{(R, r)\}$ and the set G . These are indicated as a solid point and line segment respectively in Figure 2 where $G^* = \{(p, q) \mid p = 0, q = 1\} = \{(0, 1)\}$.

What about dynamics? Suppose players 1 and 2 use mixed strategies p and q respectively. The payoffs of pure strategies L and R are 1 and $(1 - q)0 + 2q$

respectively and the payoffs of pure strategies ℓ and r are $4p + (1 - p)0$ and $4p + (1 - p)2$ respectively. Thus, the expected payoffs are $p + (1 - p)2q$ and $(1 - q)4p + q(4p + (1 - p)2)$ for players 1 and 2 respectively. From this, the replicator equation is

$$\begin{aligned}\dot{p} &= p(1 - (p + (1 - p)2q)) = p(1 - p)(1 - 2q) \\ \dot{q} &= q(4p + (1 - p)2 - [(1 - q)4p + q(4p + (1 - p)2)]) = q(1 - q)2(1 - p).\end{aligned}$$

The rest points are the two vertices $\{(0, 0), (0, 1)\}$ and the edge $\{(1, q) \mid 0 \leq q \leq 1\}$. Notice that, for any interior trajectory, q is strictly increasing and that p is strictly increasing (decreasing) if and only $q < \frac{1}{2}$ ($q > \frac{1}{2}$).

The following results for Example 1 are straightforward to prove.

1. Every NE outcome is a single terminal node (either $(2, 2)$ or $(1, 4)$).
2. Every NE component G includes a pure strategy pair.
3. The outcomes of all elements of G are the same.
4. Every interior trajectory of the replicator equation converges to a NE.
5. Every pure strategy NE is stable but not necessarily asymptotically stable.
6. Every NE that has a neighborhood whose only rest points are NE is stable.
7. If a NE component is interior attracting, it includes the SPNE defined below.
8. A NE is asymptotically stable if and only if it is strict if and only if it is pervasive (i.e. it reaches every player decision point).

Theorem 1. (Cressman, 2003) Results 2 to 8 are true for all generic perfect information games. Result 1 holds for generic perfect information games without moves by nature.

Note: Some of these results are true for general perfect information games and some are not. For instance, Result 1 is not true for some non-generic games or for generic games with moves by nature. Result 4, which provides the basis to connect dynamics with NE in Results 5 to 8, remains an open problem for non-generic perfect information games.

Definition. An extensive form game Γ is *generic* if no two pure strategy pairs that yield different outcomes have the same payoff for one of the players. If Γ is a perfect information game and there are no moves by nature, this is equivalent to the property that no two terminal nodes have the same payoff for one of the players.

Remark 1. Example 1 is a famous game known as the Entry Deterrence Game or the Chain Store Game introduced by the Nobel laureate Reinhard Selten (Selten, 1978). Player 2 is a monopolist who wants to keep the potential entrant (player 1) from entering the market that has a total value of 4. He does this by threatening to ruin the market (play ℓ with both payoffs 0) if player 1 enters (plays R) rather than accepting the entrant (play r and split the total value of 4 to yield payoff 2 for each player). However, this is often viewed as an incredible

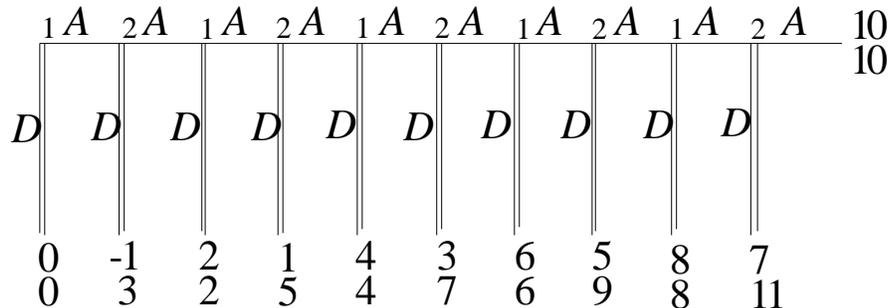


Figure 3: Centipede game of length ten.

(i.e. unbelievable) threat since the monopolist should accept the entrant if his decision point is reached (i.e. if player 1 enters) since this gives the higher payoff to him (i.e. $2 > 0$). Some game theorists argue that a generic perfect information game has only one rational NE equilibrium outcome and this can be found by backward induction. This procedure starts at a final player decision point (i.e. a player decision point that has no player decision points following it) and decides which unique action this player chooses there to maximize his payoff in the subgame with this as its root. The original game tree is then truncated at this node by creating a terminal node there with payoffs to the two players given by this action. The process is continued until the game tree has no player decision nodes left and yields the subgame perfect NE (SPNE). This is a pure strategy pair and is indicated by the double lines in the game tree. If a NE is not subgame perfect, then this perspective argues that there is some player decision node where an incredible threat has been used.

Remark 2. The above theory continues to hold for arbitrary generic perfect information games such as in Figures 3 and 4. For the centipede game of Figure 3 (Rosenthal, 1981), the SPNE is for both players to play D (down) at each of their decision points. In fact, the only NE outcome is $(0, 0)$ (i.e. player 1 plays D immediately) and so every interior trajectory converges to a NE with this outcome (the dynamics here is in a $2(32 - 1) = 62$ dimensional space). Note that, at each player decision point besides the last, both players are better off if this player plays A (across) there and his opponent plays A at the next decision point. From this it follows that, if any choice D is eliminated, then the SPNE of the new perfect information game is the terminal node that immediately follows the last D eliminated.

Since no pure strategy pair in Figure 4 can reach both the left-side subgame and the right-side subgame, none are pervasive. Thus no NE can be asymptotically stable (Theorem 1, Result 8). This is a more elementary example than Figure 3 (which also has no asymptotically stable NE). The dynamics here is only 6 dimensional. In fact, Figure 4 is probably the easiest example (Cress-

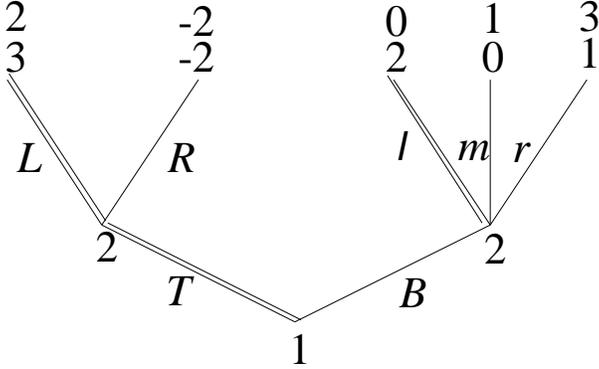


Figure 4: Perfect information game with unstable subgame perfect NE component.

man, 2003) of a perfect information game where the NE component of the SPNE outcome $(2, 3)$ is not interior attracting since there are trajectories that start near the boundary of this component that converge to the NE component with outcome $(0, 2)$.

Remark 3. Extensive form games can always be represented in normal form. The bimatrix normal form of Example 1 is

$$\begin{array}{cc} & \begin{array}{cc} \text{Ruin} & \text{Accept} \end{array} \\ \begin{array}{c} \text{Not Enter} \\ \text{Enter} \end{array} & \begin{bmatrix} 1, 4 & 1, 4 \\ 0, 0 & 2, 2 \end{bmatrix} . \end{array}$$

This elementary example already shows a common feature of the normal form approach for such games; namely, that some payoff entries are repeated in the bimatrix. As a normal form, this means the game is non-generic even though it arose from a generic perfect information game. For this reason, most normal form games cannot be represented as perfect information games. However, they can all be represented as a simultaneity game.

1.2 Simultaneity games

A (finite, two-player) *simultaneity game* is an extensive form game that involves n stages such that, at each stage, both players know all actions that have already occurred in previous stages but not the opponent's action in the current stage. If there are moves by nature, each of these nodes occur at the beginning of some stage and both players know what action nature takes at these nodes. Thus, the first player decision point x at any stage is the root of a subgame (and so an information set u for this player). Any decision node y following x in the same stage must be a decision point of the other player. The set v of all such y forms

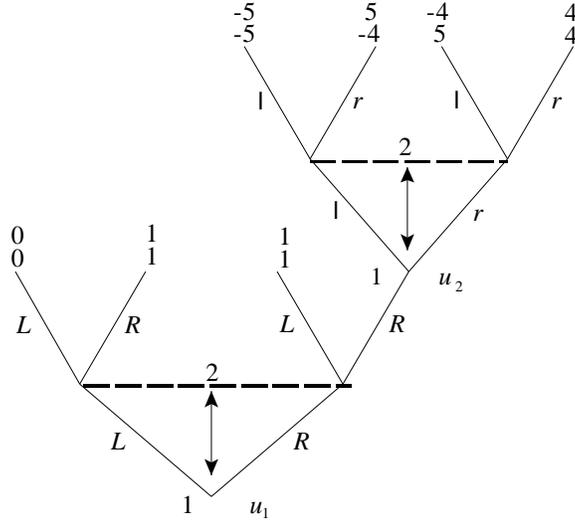


Figure 5: Extensive form of Example 2.

an information set for this other player and so each y in v has the same set of actions. The simultaneity game is symmetric if there is a bijection between the information sets and actions of player 1 and those of player 2 that makes its normal form representation a symmetric game (as explained in the following example).

Example 2. (van Damme, 1991) Figure 5 is an example of an elementary two-stage symmetric simultaneity game. The two information sets of player 2 both include two decision points and are indicated by dashed horizontal lines. At each of these information sets, player 2 must make the same choice at both decision points. The (reduced strategy) normal form is

$$\begin{array}{c}
 \\
 L \\
 R\ell \\
 Rr
 \end{array}
 \begin{array}{c}
 L \\
 R\ell \\
 Rr
 \end{array}
 \begin{array}{c}
 R\ell \\
 Rr \\
 Rr
 \end{array}
 \left[\begin{array}{ccc}
 0,0 & 1,1 & 1,1 \\
 1,1 & -5,-5 & 5,-4 \\
 1,1 & -4,5 & 4,4
 \end{array} \right].$$

This is a symmetric game since the column player (player 2) has payoff matrix $\begin{bmatrix} 0 & 1 & 1 \\ 1 & -5 & -4 \\ 1 & 5 & 4 \end{bmatrix}$ which is the transpose of the payoff matrix $A \equiv \begin{bmatrix} 0 & 1 & 1 \\ 1 & -5 & 5 \\ 1 & -4 & 4 \end{bmatrix}$ of the row player (player 1). Thus, a player's payoff depends only on the strategy pair used and not on his designation as a row or column player.

To apply backward induction to this example, the only proper subgame Γ_{u_2} has root at u_2 and payoff matrix $\begin{array}{c} \ell \\ r \end{array} \begin{bmatrix} -5 & 5 \\ -4 & 4 \end{bmatrix}$. This is equivalent to

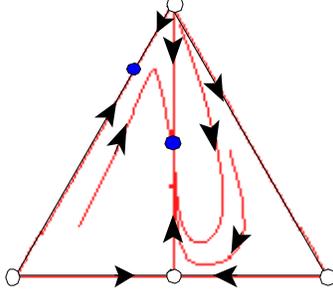


Figure 6: Trajectories of the replicator equation for Example 2.

a Hawk-Dove Game with a unique symmetric NE $\frac{1}{2}\ell + \frac{1}{2}r$ (which is also an ESS) and corresponding payoff 0. The truncated game with root at u_2 has payoff matrix $\begin{matrix} L & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ R & \end{matrix}$ and it also has unique ESS at $\frac{1}{2}L + \frac{1}{2}R$ and corresponding payoff $\frac{1}{2}$. Thus, $\frac{1}{2}L + \frac{1}{2}(\frac{1}{2}R\ell + \frac{1}{2}Rr) = \frac{1}{2}L + \frac{1}{4}R\ell + \frac{1}{4}Rr$ is a symmetric NE of Example 2 which can be easily confirmed since

$$Ap^* = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \text{ where } p^* = \begin{bmatrix} 1/2 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

Somewhat surprisingly, p^* is not an ESS of A since, for example, $p^* \cdot Ae_3 < e_3 \cdot Ae_3$ (i.e. $\frac{1}{2} + \frac{5}{4} + \frac{4}{4} < 4$). However, p^* is globally asymptotically stable under the replicator equation (Figure 6) by the following theorem.

Recall that the replicator equation for a symmetric game with $n \times n$ payoff matrix A is

$$\dot{p}_i = (e_i - p) \cdot Ap$$

for $i = 1, \dots, n$ where e_i is the unit vector (corresponding with the i^{th} pure strategy) that has 1 in its i^{th} component and 0 everywhere else and p_i is the proportion of the population using strategy e_i .

Theorem 2. (Cressman, 2003) Suppose that Γ is a symmetric simultaneity game. If p^* is an asymptotically stable NE of the standard normal form of Γ under the replicator equation, then p^* is pervasive and subgame perfect. If Γ has no moves by nature, then a pervasive NE p^* of the reduced-strategy normal form of Γ is asymptotically stable under the replicator equation if and only if p^* is given by backward induction applied to the asymptotically stable pervasive NE of the subgames of Γ and their truncations.

Remark 4. Selten mistakenly asserted in 1983 (correcting himself in 1988) that the backward induction procedure applied to the ESSs of subgames and

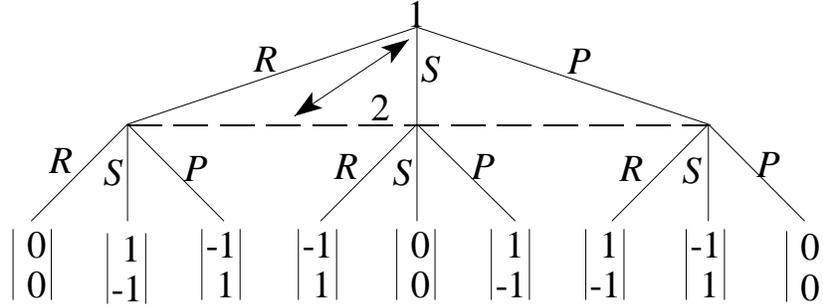


Figure 7: Extensive form of the standard RSP Game.

their truncations yields a direct ESS (i.e. an ESS in behavior strategies) for symmetric extensive form games. Example 2 is the well-known counterexample due to Eric van Damme (1991). By Theorem 2, Selten's assertion is true when "ESS" is replaced by "asymptotic stability under the replicator equation" and there are no moves by nature. One must be careful extending this result when there are moves by nature as illustrated by Example 3 below.

Every (symmetric) normal form game can be represented as a single-stage (symmetric) simultaneity game. Thus, unlike perfect information games, generic symmetric simultaneity games can have a NE outcome (and ESS) whose NE component does not include a pure strategy. For instance, the standard zero-

sum Rock-Scissors-Paper (RSP) Game with payoff matrix

$$\begin{matrix} R \\ S \\ P \end{matrix} \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

and unique NE $(1/3, 1/3, 1/3)$ has extensive form given in Figure 7. The following example uses the generalized RSP game with payoff matrix

$$\begin{bmatrix} 0 & b_2 & -a_3 \\ -a_1 & 0 & b_3 \\ b_1 & -a_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 6 & -4 \\ -4 & 0 & 4 \\ 2 & -2 & 0 \end{bmatrix} \quad (1)$$

and unique NE $(10/29, 8/29, 11/29)$. From Hofbauer and Sigmund (1998, Section 7.7), this is not an ESS since $b_1 < a_3$ (i.e. $2 < 4$) but it is globally asymptotically stable under the replicator equation since $a_1 a_2 a_3 < b_1 b_2 b_3$ (i.e. $2 \cdot 4 \cdot 4 < 2 \cdot 4 \cdot 6$).

Example 3. (Chamberland and Cressman, 2000) Suppose that, on even numbered days, players play the left-hand subgame of Figure 8 and on odd numbered days the right-hand subgame (an alternative interpretation is that nature flips a fair coin at the root of Figure 8 that determines which subgame is played). However, for both types of days, the same RSP game given by (1) is played.

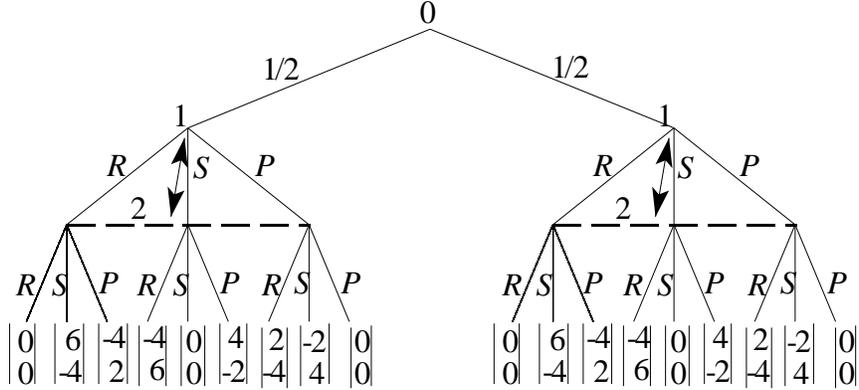


Figure 8: Extensive form of a single-stage simultaneity game with a move by nature and identical generalized RSP subgames.

In this single-stage symmetric simultaneity game, both players have 9 pure (behavior) strategies RR, RS, \dots, PP that specify a choice of R, S or P in each of the subgames. The unique symmetric NE outcome is for both players to play $p^* = (10/29, 8/29, 11/29)$ in both subgames. The corresponding NE component is a four-dimensional set E of points $p = (p_{11}, \dots, p_{33})$ formed by intersecting a four-dimensional hyperplane with the eight-dimensional strategy simplex Δ^9 . These are the points whose induced marginal strategies in the two subgames (i.e. p^1 and p^2 where $p_i^1 \equiv \sum_{j=1}^3 p_{ij}$ and $p_j^2 \equiv \sum_{i=1}^3 p_{ij}$) are both equal to p^* . It can be shown (Chamberland and Cressman, 2000) that some points in E near the boundary of Δ^9 are unstable since the linearization of the replicator equation there yields an eigenvalue with positive real part.

The reason this can occur is that, at a general $p \in \Delta^9$, the evolution of strategy frequencies in one subgame can be influenced by payoffs received in the other subgame. In particular, the frequency of R use in the left-hand subgame can be increasing even if the population state there is mostly P users. To avoid this type of unintuitive situation, the replicator equation can be restricted to the four-dimensional invariant Wright manifold

$$W \equiv \{p \in \Delta^9 \mid p_{ij} = p_i^1 p_j^2\}.$$

On W , the dynamics for the induced strategy in each subgame is the same as the replicator equation for the payoff matrix (1). Thus, each interior trajectory that starts on W converges to the single point p^* with $p_{ij}^* = p_i^* p_j^*$.

Remark 5. The Wright manifold W can be defined for all simultaneity games (in fact, all extensive form games) and it is invariant under the replicator equation. On W , Theorem 2 is true for all symmetric simultaneity games whether or not there are moves by nature.

Remark 6. Extensive form games, with an explicit description of the sequential feature of the players' possible actions, played a central role in the initial development of classical game theory by von Neumann and Morgenstern (1944). On the other hand, most dynamic analyses of evolutionary games are based on their normal forms. One consequence of this is that typical normal form examples considered in evolutionary game theory have a small number of pure strategies since it is well-known that the high-dimensional systems of evolutionary dynamics associated to a large number of pure strategies can exhibit all the complexities of arbitrary dynamical systems such as periodic orbits, limit cycles, bifurcations and chaos. The above discussion was meant to convince you that the extensive form structure (which is usually associated with a large number of pure strategies) imparts special properties on the evolutionary dynamics that makes its analysis more tractable than would otherwise be expected.

2 Asymmetric Games

A (finite, two-player) *asymmetric game* has a set $\{u_1, u_2, \dots, u_N\}$ of N roles. Players 1 and 2 are assigned roles u_k and u_ℓ respectively with probability $\rho(u_k, u_\ell)$. We assume that role assignment is independent of player designation (i.e. $\rho(u_k, u_\ell) = \rho(u_\ell, u_k)$). If players are assigned the same role (i.e. $k = \ell$), then they play a symmetric (normal form) game with payoff matrix A_{kk} . When they are assigned different roles (i.e. $k \neq \ell$), they play a bimatrix (normal form) game with payoff matrices $A_{k\ell}$ and $A_{\ell k}$,

Figure 9 is the extensive form of a two role game with two pure strategies in role u_1 and three in role u_2 . Here, the initial move by nature indicates $\rho(u_k, u_\ell) = \frac{1}{4}$ for all $1 \leq k, \ell \leq 2$. On the other hand, if $N = 1$, then $\rho(u_1, u_1) = 1$ and we have a symmetric game (e.g. only the left-hand subtree of Figure 9 formed by nature following the left-most direction at the root with probability 1). Similarly, if $N = 2$, then $\rho(u_1, u_2) = \rho(u_2, u_1) = \frac{1}{2}$ and $\rho(u_1, u_1) = \rho(u_2, u_2) = 0$ and so we have a bimatrix game (e.g. only the middle two subtrees of Figure 9 formed by nature following these two directions at the root with probability $\frac{1}{2}$). Thus, asymmetric games include both symmetric and bimatrix normal form games as special cases.

All asymmetric games have a single-stage extensive form representation with an initial move by nature and information sets u_1, u_2, \dots, u_N for both players. A pure strategy for player 1 specifies a choice at each of his information sets. It has the form e_i where $i = (i_1, \dots, i_N)$ is a multi-index with i_k giving the choice of e_i at u_k . Each mixed strategy p is a discrete probability distribution over the finite set $\{e_i\}$ with weight p_i on e_i . This p induces a local behavior strategy p^k at each information set u_k given by

$$p_r^k = \sum \{p_i \mid i_k = r\}$$

and the Wright manifold is $W \equiv \{p \mid p_i = p_{(i_1, \dots, i_N)} = p_{i_1}^1 p_{i_2}^2 \dots p_{i_N}^N\}$. W is invariant under the replicator equation.

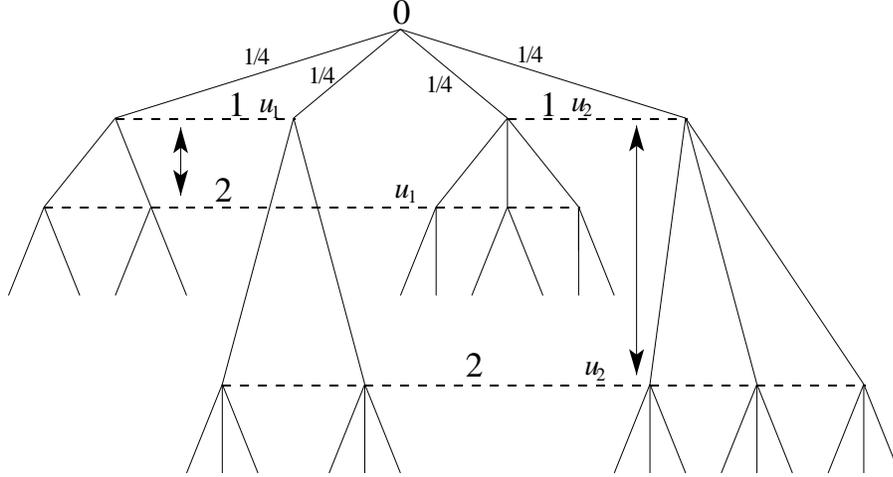


Figure 9: Extensive form of an asymmetric two-role game.

2.1 Bimatrix games

Here, $N = 2$, $\rho(u_1, u_2) = \rho(u_2, u_1) = \frac{1}{2}$ and $\rho(u_1, u_1) = \rho(u_2, u_2) = 0$. By abuse of notation, let $A_{12} = A, A_{21} = B, p^1 = p, p^2 = q, e_i$ be the pure strategies in role u_1 , and f_j be the pure strategies in role u_2 . Then, on W ,

$$\begin{aligned}\dot{p}_i &= p_i(e_i - p) \cdot Aq \\ \dot{q}_j &= q_j(f_j - q) \cdot Bp\end{aligned}$$

is the replicator equation restricted to W when the pure strategies are given as the appropriate unit vectors. This is the bimatrix replicator dynamics which we illustrate in Example 4.

Example 4. (Cressman, 2003) Consider the game with bimatrix $\begin{array}{cc} & \begin{array}{cc} H & C \end{array} \\ \begin{array}{c} T \\ I \end{array} & \begin{bmatrix} 5, 4 & 1, 6 \\ 4, 0 & 3, -2 \end{bmatrix} \end{array}$.

Unlike perfect information games, this has no NE given by a pure strategy pair (e.g. at (T, H) , player 2 does better by switching to C since $6 > 4$). In fact, no strategy pair is a NE if either player uses a pure strategy. Thus, any NE (p^*, q^*) must be a completely mixed strategy for each player. In particular, there is a unique NE given by $(p_1^*, q_1^*) = (\frac{1}{2}, \frac{2}{3})$ since

$$Aq^* = \begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 11/3 \\ 11/3 \end{bmatrix} \text{ and } Bp^* = \begin{bmatrix} 4 & 0 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

However, this NE is not asymptotically stable since $H(p_1, q_1) \equiv p_1^2(1-p_1)^2q_1^2(1-$

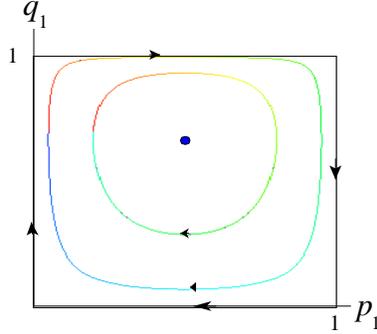


Figure 10: Trajectories of the bimatrix replicator equation for Example 3.

q_1) is a constant of motion under the replicator equation (i.e. $\frac{dH}{dt} = 0$) whose level curves are given in Figure 10. Trajectories of the replicator equation

$$\begin{aligned}\dot{p}_1 &= p_1(1-p_1)(3q_1-2) \\ \dot{q}_1 &= q_1(1-q_1)(2-4p_1)\end{aligned}\tag{2}$$

evolve clockwise around the interior rest point (p_1^*, q_1^*) along these curves.

Another way to see that $(p_1^*, q_1^*) = (\frac{1}{2}, \frac{2}{3})$ is not asymptotically stable is to consider the time-adjusted replicator equation in the interior of the unit square that divides the vector field in (2) by the Dulac function $p_1(1-p_1)q_1(1-q_1)$ to obtain

$$\begin{aligned}\dot{p}_1 &= \frac{(3q_1-2)}{q_1(1-q_1)} \\ \dot{q}_1 &= \frac{(2-4p_1)}{p_1(1-p_1)}.\end{aligned}$$

A rectangle $\Delta p \Delta q$ in the interior does not change area as it evolves since its horizontal and vertical cross-sections maintain the same lengths under this dynamics. (This invariance of area also follows from Liouville's result that "volumes" remain constant when the vector field is divergence free.) Thus no interior point can be asymptotically stable since no small rectangle containing it evolves to this point (since it is a region with zero area).

Example 4 is the well-known Buyer-Seller Game where a Buyer of some merchandise can either Trust the Seller to give an accurate value of this item or the Buyer can have the item inspected (i.e. play Inspect) to determine its true value. The Seller has a choice between Honest (give an accurate value) or Cheat (misrepresent its true value). The clockwise rotation of the trajectories in Figure 10 is not surprising given the cycling of the best responses to pure strategy pairs. What is not *a priori* clear is why trajectories cannot spiral in

to the interior rest point making it asymptotically stable. This follows from the analysis above for the two-dimensional dynamics of Figure 10. It is also a consequence of part (a) of the following theorem for general bimatrix games (which can be proved using Liouville's result in higher dimensions).

Theorem 3. (Hofbauer and Sigmund, 1998) (a) A strategy pair (p^*, q^*) is an asymptotically stable rest point of the bimatrix replicator equation if and only if it is a strict NE. In particular, (p^*, q^*) is a pure strategy pair.
 (b) If there is no interior NE, then all trajectories of the bimatrix replicator equation converge to the boundary.

2.2 Two-species ESS

Asymmetric games with two roles (i.e. $N = 2$) can be interpreted as games between two species by equating intraspecific interactions as between individuals playing a symmetric game in the same roles and interspecific interactions between individuals playing a bimatrix game in opposite roles. From this perspective, Figure 9 is then an example where there are both intra and inter specific interactions. On the other hand, bimatrix games such as the Buyer-Seller Game of Example 4 are then ones where all interactions are interspecific.

Suppose we extend Maynard Smith's original idea by saying that a (two-species) ESS is a monomorphic system with strategy pair (p^*, q^*) that cannot be successfully invaded by a rare (mutant) subsystem using a different strategy pair (p, q) . That is, define (p^*, q^*) as a two-species ESS if it is asymptotically stable under the two-dimensional replicator equation based on the strategy pairs (p^*, q^*) and (p, q) whenever $(p, q) \neq (p^*, q^*)$. Then

Theorem 4. (Cressman, 2003) (a) (p^*, q^*) is a two-species ESS if and only if
 either $p^* \cdot (Ap + Bq) > p \cdot (Ap + Bq)$
 or $q^* \cdot (Cp + Dq) > q \cdot (Cp + Dq)$
 for all strategy pairs (p, q) that are sufficiently close (but not equal) to (p^*, q^*) .
 (b) If (p^*, q^*) is a two-species ESS, then it is asymptotically stable for the two-species replicator equation (i.e. based on all pure strategies of the asymmetric game).

Here A and D are the payoff matrices for intraspecific interactions (i.e. symmetric games) of species one and two respectively whereas B and C form the bimatrix game corresponding to interspecific interactions.

If there are no intraspecific interactions (take A and D as the zero matrices 0 of the appropriate size), then (p^*, q^*) is a two-species ESS if and only if it is a strict NE (e.g. by taking $q = q^*$, we have that $p^* \cdot Bq^* > p \cdot Bq^*$ if $p \neq p^*$ since $q^* \cdot Cp = q \cdot Cp$). That is, by Theorems 3 and 4, (p^*, q^*) is a two-species ESS for a bimatrix game if and only if it is a strict NE if and only if it is asymptotically stable under the bimatrix replicator equation.

At the other extreme, suppose that there are no interspecific interactions (take B and C as zero matrices). Then, (p^*, q^*) is a two-species ESS if and

only p^* is a single-species ESS for species one and q^* is a single-species ESS for species two. For example, when $q = q^*$, we need $p^* \cdot Ap > p \cdot Ap$ for all p that are sufficiently close (but not equal) to p^* . Recall that this inequality condition, called local superiority by Weibull (1995), characterizes the single-species ESS (of species one). From this result, there can be two-species ESSs that are not strict NE (see Example 5 below). In particular, there can be completely mixed ESSs.

From these two extremes, we see that the concept of a two-species ESS combines and generalizes the concepts of single-species ESS of symmetric games and the strict NE of bimatrix games.

Example 5. (Krivan et al., 2008) Suppose that there are two species competing in two different habitats (or patches) and that the overall population size (i.e. density) of each species is fixed. Also assume that the fitness of an individual depends only on its species, the patch it is in and the density of both species in this patch. Then strategies of species one and two can be parameterized by the proportions p_1 and q_1 respectively of these species that are in patch one. If individual fitness (i.e. payoff) is positive when a patch is unoccupied and linearly decreasing in patch densities, it is of the form

$$\begin{aligned} F_i &= r_i \left(1 - \frac{p_i M}{K_i} - \frac{\alpha_i q_i N}{K_i} \right) \\ G_i &= s_i \left(1 - \frac{q_i N}{L_i} - \frac{\beta_i p_i M}{L_i} \right). \end{aligned}$$

Here, F_i is the fitness of a species one individual in patch i , G_i is the fitness of a species two individual in patch i , $p_2 = 1 - p_1$ and $q_2 = 1 - q_1$. All other parameters are fixed and positive (see Remark 7 below).

By linearity, these fitnesses can be represented by a two-species asymmetric game with payoff matrices

$$\begin{aligned} A &= \begin{bmatrix} r_1 - \frac{r_1 M}{K_1} & r_1 \\ r_2 & r_2 - \frac{r_2 M}{K_2} \end{bmatrix} & B &= \begin{bmatrix} -\frac{\alpha_1 r_1 N}{K_1} & 0 \\ 0 & -\frac{\alpha_2 r_2 N}{K_2} \end{bmatrix} \\ C &= \begin{bmatrix} -\frac{\beta_1 s_1 M}{L_1} & 0 \\ 0 & -\frac{\beta_2 s_2 M}{L_2} \end{bmatrix} & D &= \begin{bmatrix} s_1 - \frac{s_1 N}{L_1} & s_1 \\ s_2 & s_2 - \frac{s_2 N}{L_2} \end{bmatrix}. \end{aligned}$$

For example, $F_i = e_i \cdot (Ap + Bq)$. At an equilibrium (p, q) , all individuals present in species one must have the same fitness as do all individuals present in species two.

Suppose that both patches are occupied at the equilibrium (p, q) . Then (p, q) is a NE and (p_1, q_1) is a point in the interior of the unit square that satisfies

$$\begin{aligned} r_1 \left(1 - \frac{p_1 M}{K_1} - \frac{\alpha_1 q_1 N}{K_1} \right) &= r_2 \left(1 - \frac{(1 - p_1) M}{K_2} - \frac{\alpha_2 (1 - q_1) N}{K_2} \right) \\ s_1 \left(1 - \frac{q_1 N}{L_1} - \frac{\beta_1 p_1 M}{L_1} \right) &= s_2 \left(1 - \frac{(1 - q_1) N}{L_2} - \frac{\beta_2 (1 - p_1) M}{L_2} \right). \end{aligned}$$

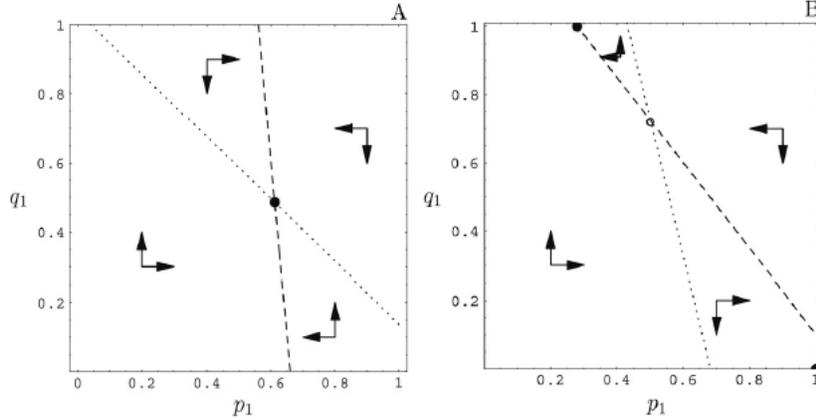


Figure 11: The two-patch Habitat Selection Game. Solid dots are ESSs.

That is, these two "equal fitness" lines (which have negative slopes) intersect at (p_1, q_1) as in Figure 11.

The interior NE (p, q) is a two-species ESS if and only if the equal fitness line of species one is steeper than that of species two. That is, (p, q) is an interior two-species ESS in Figure 11A but not in Figure 11B. The interior two-species ESS in Figure 11a is globally asymptotically stable under the replicator equation.

Figure 11B has two two-species ESSs, both on the boundary of the unit square. One is a pure strategy pair strict NE with species one and two occupying separate NE with species one and two occupying separate patches $(p_1 = 1, q_1 = 0)$ and the other has species two in patch one and species one split between the two patches $(0 < p_1 < 1, q_1 = 1)$. Both are locally asymptotically stable under the replicator equation with basins of attraction formed by an invariant separatrix joining the two vertices corresponding to both species in the same patch on which trajectories evolve to the interior NE.

If the equal fitness lines do not intersect in the interior of the unit square, then there is exactly one two-species ESS. This is on the boundary (either a vertex or on an edge) and is globally asymptotically stable under the replicator equation.

Remark 7. Example 5 is called a (two-patch) Habitat Selection Game for two competitive species. The fixed parameters then have biological interpretations. Specifically, for species one, M is the total population size; r_i, K_i and α_i are its intrinsic growth rate, carrying capacity and interspecific competition coefficient (modelling the the effect of the second species on the first) in patch i respectively. The analogous parameters for species two are $N; s_i, L_i$ and β_i . Linearity of the fitness functions corresponds to Lotka-Volterra interactions.

Habitat selection games for a single species were already introduced before

evolutionary game theory was developed when Fretwell and Lucas (1969) defined an ideal free distribution (IFD) to be a patch distribution whereby the fitness of any individual in an occupied patch was the same as the fitness of any other individual and at least as high as what would be the fitness in any unoccupied patch. If patch fitness is decreasing in patch density, then the IFD and ESS concepts are identical for a single species. In fact, there is a unique IFD and it is globally asymptotically stable under the replicator equation.

For two species, some authors consider an interior NE to be a (two-species) IFD. Example 5 shows such NE may be unstable (Figure 11B) and so justifies the perspective of others who restrict the IFD concept to two-species ESSs.

Remark 8. The generalization of Theorem 4 to three (or more) species is a difficult problem (Cressman et al., 2001). It is possible to characterize a monomorphic three-species ESS as one where, at all nearby strategy distributions, at least one species does better using its ESS strategy. However, such an ESS concept does not always imply stability of the three-species replicator equation that is based on the entire set of pure strategies for each species.

3 Games with Continuous Strategy Spaces

When players can choose from a continuum of pure strategies, the connections between NE and dynamic stability become more complicated. For example, the standard result forming one part of the Folk Theorem of Evolutionary Game Theory (Hofbauer and Sigmund, 1998) that a strict NE is asymptotically stable under the replicator equation (as well as most other deterministic evolutionary dynamics) is true for all games that have a finite set of pure strategies including the symmetric and asymmetric games of extensive or normal form in Sections 1 and 2. However, a strict NE is not always dynamically stable for games with continuous strategy spaces as seen in Example 6 below. Now, dynamic stability requires additional conditions such as that of a continuously stable strategy (CSS) or a neighborhood invader strategy (NIS) introduced by Eshel (1983) and Apaloo (1997) respectively. In fact, the exact requirement depends on the form of the evolutionary dynamics.

3.1 Symmetric games with a continuous strategy space

In this section, static conditions are developed for stability under the two standard dynamics for games with a continuous strategy space S ; namely, adaptive dynamics and the replicator equation. The canonical equation of adaptive dynamics (Dieckmann and Law, 1996; Dercole and Rinaldi, 2008) models the evolution of the population mean strategy $x \in S$ by assuming that the population is always monomorphic at its mean. Then x evolves through trait substitution in a direction y of nearby mutants that can invade due to their higher payoff than x when playing against this monomorphism.

The replicator equation is now a dynamic on the space $\Delta(S)$ of Borel probability measures over the strategy space S (Bomze, 1991). This infinite-dimensional dynamical system restricts to the replicator equation of a symmetric normal form game when a finite subset of S is taken as the strategy set. From the perspective of the replicator equation that describes the evolution of the population strategy distribution $P \in \Delta(S)$ rather than the evolution of the population mean, the canonical equation becomes a heuristic tool that approximates how the mean evolves by ignoring effects due to the diversity of strategies in the population.

3.1.1 One dimensional strategy space

Suppose that S is a convex compact subset of \mathbf{R} (i.e. a closed and bounded interval). Following Hofbauer and Sigmund (1990), if $\pi(y, x)$ is continuously differentiable, then adaptive dynamics has the form (up to a change in time scale)

$$\dot{x} = \frac{\partial \pi(y, x)}{\partial y} \Big|_{y=x} \quad (3)$$

at interior points of S (i.e. for $x \in \text{int}(S)$). That is, x increases if $\pi(y, x)$ is an increasing function of y for y close to x .

An $x^* \in \text{int}(S)$ is an equilibrium if $\pi_1(x^*, x^*) = 0$ (here π_1 is the partial derivative of π with respect to the first variable). Such an equilibrium is called convergence stable (Christiansen, 1991) if it is asymptotically stable under (3). If $\pi(y, x)$ has continuous partial derivatives up to second order, then x^* is convergence stable if $\frac{d}{dx} \left[\frac{\partial \pi(y, x)}{\partial y} \Big|_{y=x^*} \right] \Big|_{x=x^*} < 0$. If π_{ij} is the second order partial derivative of π with respect to i and j , then this inequality condition is $\pi_{11}(x^*, x^*) + \pi_{12}(x^*, x^*) \equiv \pi_{11} + \pi_{12} < 0$. Conversely, if x^* is convergence stable, then $\pi_{11} + \pi_{12} \leq 0$. The following example illustrates the NE and convergence stability concepts for quadratic payoff functions.

Example 6. Let $S = [-1, 1]$ be the set of pure strategies and $\pi(x, y) = ax^2 + bxy$ be the payoff to x playing against y for all $x, y \in S$ where a and b are fixed real numbers. Then x^* in the interior of S is a NE (i.e. $\pi(x, x^*) \leq \pi(x^*, x^*)$ for all $x \in S$) if and only if $x^* = 0$ and $a \leq 0$. Also, $x^* = 0$ is a strict NE if and only if $a < 0$. These results exclude the degenerate case where $2a + b = 0$ and $\pi(x, y) = a(x - y)^2 - ay^2$. In this case (which we ignore from now on), every $x \in S$ is a strict NE if $a < 0$, a NE if $a = 0$, and there are no pure strategy NE if $a > 0$.

From (3), adaptive dynamics is now

$$\dot{x} = \frac{\partial \pi(y, x)}{\partial y} \Big|_{y=x} = (2a + b)x.$$

The only equilibrium is $x^* = 0$ and it is convergence stable if and only if $2a + b < 0$. In particular, $x^* = 0$ may be a strict NE but not convergence stable (e.g. $a < 0$ and $2a + b > 0$) or may be a convergence stable rest point that is not a

NE (e.g. $a > 0$ and $2a + b < 0$). In the first case, the strict NE is a rest point of adaptive dynamics that is unattainable from nearby monomorphic populations. The population evolves to the endpoint of S closest to the initial value of x (e.g. x evolves to 1 if x is positive initially).

In the latter case, once $x^* = 0$ becomes established as the population monomorphism, it is vulnerable to invasion by mutants since $\pi(x, x^*) > \pi(x^*, x^*)$ for all nonzero $x \in S$. Alternatively, an $x \in S$ closer to x^* than y cannot invade a dimorphic population that is evenly split between the strategies $\pm y$ for $|y| > |x|$ since the expected payoff to x (namely, $\frac{1}{2}\pi(x, y) + \frac{1}{2}\pi(x, -y)$) is less than the expected payoff to either y or $-y$. In fact, x can invade this dimorphism if and only if $|x| > |y|$. For either of these reasons, it is usually assumed that evolutionary stability of an $x^* \in S$ requires a convergence stable equilibrium that is also a strict NE (see the following Definition 1 and Theorem 5). A convergence stable rest point that is not a strict NE forms the basis of an initial evolutionary branching (Doebeli and Dieckmann, 2000) into a dimorphic system.

There is some disagreement whether the strict NE condition for x^* should hold for all $x \in S$ or be restricted to those strategies close to x^* . In the following, we take the second approach and call these neighborhood strict NE to make this choice clear. Such NE are also called ESS (Marrow et al., 1996) or are said to satisfy the ESS Maximum Principle (Vincent and Brown, 2005). We will not use the ESS terminology in Section 3 since the meaning of ESS is not universally accepted for games with a continuous strategy space (Apaloo et al, 2009).

Definition 1. (Eshel, 1983) Suppose the strategy space S is a subinterval of real numbers. An $x^* \in S$ is a *neighborhood continuously stable strategy (CSS)* if there exists an $\varepsilon > 0$ such that, for all $x \in S$ with $0 < |x - x^*| < \varepsilon$, the following two conditions hold.

- (i) $\pi(x, x^*) < \pi(x^*, x^*)$ (neighborhood strict NE condition).
- (ii) There exists $\eta > 0$ (which depends on x) such that, for all $x' \in S$ with $0 < |x' - x| < \eta$, $\pi(x', x) > \pi(x, x)$ if and only if $|x' - x^*| < |x - x^*|$ (convergence stability condition).

The CSS concept, like the ESS definition of Maynard Smith (1982), is defined in terms of static payoff comparisons that are meant to predict the outcome of evolutionary dynamics. For the CSS (as well as the related notion of neighborhood half superiority given in Definition 2), the connection to dynamic stability is summarized in Theorem 5. Neighborhood superiority for a pure strategy requires the extension of the payoff functions to distributions $P \in \Delta(S)$. Let δ_x be the Dirac delta distribution with full weight on the point $x \in S$ and assume that individuals interact in random pairwise contests. Then $\pi(\delta_x, P) \equiv \int_S \pi(x, y)P(dy)$ (also denoted by $\pi(x, P)$) is the expected payoff to an individual who uses strategy x if the population mean behavior is P and $\pi(P, P) \equiv \int_S \pi(x, P)P(dx)$ is the mean payoff of the population (Bomze and Potscher, 1989).

The following definition is given for general multi-dimensional strategy spaces. The support of P is the closed set given by $\{x \in S \mid P(\{y \mid |y - x| > \varepsilon\}) > 0 \text{ for all } \varepsilon > 0\}$.

Definition 2. (Cressman, 2009) Suppose the strategy space S of a symmetric game is a subset of \mathbf{R}^n and $0 \leq p^* < 1$ is fixed. Strategy $x^* \in S$ is *neighborhood p^* -superior* if $\pi(x^*, P) > \pi(P, P)$ for all $P \in \Delta(S)$ with $1 > P(\{x^*\}) \geq p^*$ and the support of P sufficiently close to x^* . It is *neighborhood superior* (respectively, *neighborhood half-superior*) if $p^* = 0$ (respectively, $p^* = \frac{1}{2}$). Strategy $x^* \in S$ is *globally p^* -superior* if $\pi(x^*, P) > \pi(P, P)$ for all $P \in \Delta(S)$ with $1 > P(\{x^*\}) \geq p^*$.

Theorem 5. Suppose that S is one dimensional and $x^* \in \text{int}(S)$ is a rest point of adaptive dynamics (3) (i.e. $\pi_1(x^*, x^*) = 0$).

(a) If $\pi_{11} < 0$, then x^* is a neighborhood strict NE. Conversely, if x^* is a neighborhood strict NE, then $\pi_{11} \leq 0$.

(b) If $\pi_{11} + \pi_{12} < 0$, then x^* is convergence stable. Conversely, if x^* is convergence stable, then $\pi_{11} \leq 0$ and $\pi_{11} + \pi_{12} \leq 0$.

(c) If $\pi_{11} < 0$ and $\pi_{11} + \pi_{12} < 0$, then x^* is a neighborhood CSS and neighborhood half-superior. Conversely, if x^* is a neighborhood CSS or neighborhood half-superior, then $\pi_{11} \leq 0$ and $\pi_{11} + \pi_{12} \leq 0$.

Except in borderline cases when $\pi_{11} = 0$ or $\pi_{11} + \pi_{12} = 0$, Theorem 5 characterizes both the stability of interior neighborhood strict NE under adaptive dynamics and those $x^* \in \text{int}(S)$ that are neighborhood half-superior (i.e. $p^* = \frac{1}{2}$). On the other hand, a neighborhood CSS need not be stable under the replicator equation. This is shown in the continuation of Example 6 below that follows a brief development of the essential properties of this latter dynamics.

A trajectory P_t for $t \geq 0$ is a solution of the replicator equation if the weight $P_t(B)$ assigned to any Borel subset B of S satisfies

$$\frac{dP_t}{dt}(B) = \int_B (\pi(\delta_x, P_t) - \pi(P_t, P_t)) P(dx). \quad (4)$$

If π is continuous, there is a unique solution given any initial $P_0 \in \Delta(S)$ (Oechssler and Riedel, 2001). Stability under this replicator equation is typically analyzed with respect to the weak topology for $\Delta(S)$. We are most interested in the dynamic stability of a monomorphic population (i.e. of δ_{x^*} for some $x^* \in S$) when x^* is in the support of P_0 . Every neighborhood of δ_{x^*} in the weak topology contains the set of all distributions P whose support is within ε of x^* (i.e. $P(\{x \mid |x - x^*| > \varepsilon\}) = 0$) for some $\varepsilon > 0$ and so δ_x is in this neighborhood if $|x - x^*| < \varepsilon$.

Example 6. (Continued) Consider Example 6 again where now all individuals play either $x^* = 0$ or a nearby strategy x (which is fixed). For this restricted two-strategy game, the replicator equation becomes the one-dimensional dynamics

$\dot{p} = p(1-p)(a+bp)x^2$ (here p is the frequency of strategy x) corresponding to the symmetric game whose normal form is

$$\begin{array}{c} x \\ x^* \end{array} \quad \begin{bmatrix} x & x^* \\ (a+b)x^2 & ax^2 \\ 0 & 0 \end{bmatrix} .$$

Take $a = -2$ and $b = 3$ so that $x^* = 0$ is a CSS and also half-superior (i.e. $a < 0$ and $2a + b < 0$). Since $a + b > 0$ and $a < 0$, both pure strategies are strict NE and so locally asymptotically stable for the replicator equation applied to this two-strategy game. However, neither is asymptotically stable for the infinite-dimensional replicator equation for the full game since, in the weak topology on $\Delta(S)$, any neighborhood of x^* includes all probability measures whose support is sufficiently close to x^* . Thus, asymptotic stability of x^* requires that x^* is globally stable for the two-strategy game and so $a + b \leq 0$. Instability in this example results from $a + b = 1 > 0$.

In the following section, we examine stability with respect to the replicator equation for multi-dimensional strategy spaces.

3.1.2 Multi-dimensional strategy space

The above one-dimensional model and theory can be extended to multi-dimensional strategy spaces S that are compact convex subsets of \mathbf{R}^n with $x^* \in S$ in its interior. Following the static approach of Lessard (1990), x^* is a neighborhood CSS if it is a neighborhood strict NE that satisfies condition (ii) of Definition 1 along each line through x^* . Theorem 6 then generalizes Theorem 5 in terms of the Taylor expansion about (x^*, x^*) of the payoff function

$$\begin{aligned} \pi(x, y) &= \pi(x^*, x^*) + \nabla_1 \pi(x^*, x^*)(x - x^*) + \nabla_2 \pi(x^*, x^*)(y - x^*) \\ &\quad + \frac{1}{2} [(x - x^*) \cdot A(x - x^*) + 2(x - x^*) \cdot B(y - x^*) + (y - x^*) \cdot C(y - x^*)] \\ &\quad + \text{higher order terms.} \end{aligned}$$

Here, ∇_1 and ∇_2 are gradient vectors with respect to x and y respectively (e.g. the i^{th} component of $\nabla_1 \pi(x^*, x^*)$ is $\frac{\partial \pi(x', x^*)}{\partial x'_i} |_{x'=x^*}$) and A, B, C are the $n \times n$ matrices with ij^{th} entries (all partial derivatives are evaluated at x^*)

$$A_{ij} \equiv \left[\frac{\partial^2}{\partial x'_j \partial x'_i} \pi(x', x^*) \right]; B_{ij} \equiv \left[\frac{\partial}{\partial x'_i} \frac{\partial}{\partial x_j} \pi(x', x) \right]; C_{ij} \equiv \left[\frac{\partial}{\partial x'_j} \frac{\partial}{\partial x'_i} \pi(x^*, x') \right].$$

An $n \times n$ matrix M is negative definite (respectively, negative semi-definite) if, for all nonzero $x \in \mathbf{R}^n$, $x \cdot Mx < 0$ (respectively, $x \cdot Mx \leq 0$).

Adaptive dynamics for multi-dimensional strategy spaces generalizing (3) now has the form

$$\frac{dx}{dt} = C_1(x) \nabla_1 \pi(y, x) |_{y=x} \quad (5)$$

where $C_1(x)$ is an $n \times n$ covariance matrix modeling the mutation process (and its rate) in different directions. We will assume that $C_1(x)$ is a positive-definite symmetric matrix for $x \in \text{int}(S)$ that depends continuously on x . System (5) is called the canonical equation of adaptive dynamics.

Theorem 6. (Cressman, 2009; Leimar, 2009) Suppose $x^* \in \text{int}(S)$ is a rest point of (5) (i.e. $\nabla_1 \pi(x^*, x^*) = 0$).

(a) If A is negative definite, then x^* is a neighborhood strict NE. Conversely, if x^* is a neighborhood strict NE, then A is negative semi-definite.

(b) If $A + B$ is negative definite, then x^* is convergence stable for any choice of covariance matrix $C_1(x)$ (i.e. x^* is an asymptotically stable rest point of (5)). Conversely, if x^* is convergence stable for any choice of covariance matrix $C_1(x)$, then $A + B$ is negative semi-definite.

(c) If A and $A + B$ are negative definite, then x^* is a neighborhood CSS and neighborhood half-superior. Conversely, if x^* is a neighborhood CSS or neighborhood half-superior, then A and $A + B$ are negative semi-definite.

In the continuation of Example 6, we saw that dynamic stability with respect to the replicator equation requires more than the CSS concept. In Theorem 7, it is NIS that takes the place of the convergence stability CSS condition and attractivity replaces dynamics stability for the replicator equation. These are defined as follows.

Definition 3. (a) (Apaloo, 1997) Suppose the strategy space S of a symmetric game is a subset of \mathbf{R}^n . Strategy $x^* \in S$ is a *neighborhood invader strategy* (NIS) if $\pi(x^*, x) > \pi(x, x)$ for all $x \in S$ sufficiently close (but not equal) to x^* .

(b) (Cressman et al., 2006) δ_{x^*} is *neighborhood attracting* under the replicator equation if, for any initial distribution P_0 with support sufficiently close to x^* and $P_0(\{x^*\}) > 0$, P_t converges to δ_{x^*} in the weak topology.

Theorem 7. (Cressman et al., 2006) Suppose x^* is in the interior of S and satisfies $\nabla_1 \pi(x^*, x^*) = 0$.

(a) If $A + 2B$ is negative definite, then x^* is a NIS. Conversely, if x^* is a NIS, then $A + 2B$ is negative semi-definite.

(b) If A and $A + 2B$ are negative definite, then x^* is neighborhood superior and neighborhood attracting under the replicator equation. Conversely, if x^* is neighborhood superior or neighborhood attracting, then A and $A + 2B$ are negative semi-definite.

Remark 9. The analysis in Cressman et al. (2006) shows that one must be careful in extending the statements of Theorem 7, part b, from neighborhood attractivity to asymptotic stability or from $P_0(\{x^*\}) > 0$ to x^* in the support of P_0 , especially if the payoff function is not symmetric (i.e. $\pi(x, y) \neq \pi(y, x)$ for some $x, y \in S$). In fact, there remain open problems in these cases. On the other hand, there are examples that show neither negative definiteness nor negative semi-definiteness provide complete characterizations in any part of Theorems 6

or 7. For example, there are borderline cases with x^* a neighborhood strict NE and $A + 2B$ negative semi-definite for which x^* is a NIS in one case but not in the other (Cressman et al., 2006).

3.2 Asymmetric games with a continuous strategy spaces

The above theory of multi-dimensional CSS and NIS as well as their connections to evolutionary dynamics have been extended to asymmetric games with continuous strategy spaces (Cressman, 2009, 2010). When there are two roles, it is shown there that the CSS and NIS can be characterized (excluding borderline cases) by payoff comparisons similar to those found for the two-species ESS when both roles have a finite number of strategies (see Theorem 4 and Definition 4). In this section, we will assume that the continuous strategy sets S and T for the two roles are both one-dimensional compact intervals and that payoff functions have continuous partial derivatives up to second order in order to avoid technical and/or notational complications.

For $(x, y) \in S \times T$, let $\pi_1(x'; x, y)$ (respectively, $\pi_2(y'; x, y)$) be the payoff to a player in role one (respectively, in role 2) using strategy $x' \in S$ (respectively $y' \in T$) when the population is monomorphic at (x, y) . Note that π_1 has a different meaning here than in Section 3.1 where it was used to denote a partial derivative. With this terminology, the canonical equation of adaptive dynamics (c.f. (3)) becomes

$$\begin{aligned} \frac{dx}{dt} &= k_1(x, y) \frac{\partial}{\partial x'} \pi_1(x'; x, y) \Big|_{x'=x} \\ \frac{dy}{dt} &= k_2(x, y) \frac{\partial}{\partial y'} \pi_2(y'; x, y) \Big|_{y'=y} \end{aligned} \quad (6)$$

where $k_i(x, y)$ for $i = 1, 2$ are positive continuous functions of (x, y) . At an interior rest point (x^*, y^*) of (6), $\frac{\partial \pi_1}{\partial x'} = \frac{\partial \pi_2}{\partial y'} = 0$. In particular, if (x^*, y^*) is a neighborhood strict NE (i.e. if $\pi_1(x; x^*, y^*) < \pi_1(x^*; x^*, y^*)$ and $\pi_2(y; x^*, y^*) < \pi_2(y^*; x^*, y^*)$ for all x and y sufficiently close but not equal to x and y respectively) in the interior of $S \times T$, then it is a rest point (x^*, y^*) of (6). (x^*, y^*) is called convergence stable (or strongly convergence stable as in Leimar, 2009) if it is asymptotically stable under (6) for any choice of k_1 and k_2 .

The characterizations of these concepts in the following theorem are given in terms of the linearization of (6) about (x^*, y^*) ; namely,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} k_1(x^*, y^*) & 0 \\ 0 & k_2(x^*, y^*) \end{bmatrix} \begin{bmatrix} A + B & C \\ D & E + F \end{bmatrix} \begin{bmatrix} x - x^* \\ y - y^* \end{bmatrix} \quad (7)$$

where

$$\begin{aligned} A &\equiv \frac{\partial^2}{\partial x' \partial x'} \pi_1(x'; x^*, y^*); B \equiv \frac{\partial}{\partial x'} \frac{\partial}{\partial x} \pi_1(x'; x, y^*); C \equiv \frac{\partial}{\partial x'} \frac{\partial}{\partial y} \pi_1(x'; x^*, y) \\ D &\equiv \frac{\partial}{\partial y'} \frac{\partial}{\partial x} \pi_2(y'; x, y^*); E \equiv \frac{\partial}{\partial y'} \frac{\partial}{\partial y} \pi_2(y'; x^*, y); F \equiv \frac{\partial^2}{\partial y' \partial y'} \pi_2(y'; x^*, y^*) \end{aligned}$$

and all partial derivatives are evaluated at the equilibrium.

Theorem 8. (Cressman, 2010) Suppose (x^*, y^*) is a rest point (x^*, y^*) of (6) in the interior of $S \times T$.

- (a) (x^*, y^*) is a neighborhood strict NE if A and F are negative. Conversely, if (x^*, y^*) is a neighborhood NE, then A and F are non-positive.
- (b) (x^*, y^*) is convergence stable if, for all nonzero $(x, y) \in \mathbf{R}^2$, either $x((A+B)x + Cy) < 0$ or $y(Dx + (E+F)y) < 0$. Conversely, if (x^*, y^*) is convergence stable, then either $x((A+B)x + Cy) \leq 0$ or $y(Dx + (E+F)y) \leq 0$ for all $(x, y) \in \mathbf{R}^2$.
- (c) (x^*, y^*) is convergence stable if $A+B < 0, E+F < 0$ and $(A+B)(E+F) > CD$. Conversely, if (x^*, y^*) is convergence stable, then $A+B \leq 0, E+F \leq 0$ and $(A+B)(E+F) \geq CD$.

Proof. (a) These statements are straightforward consequences of the Taylor expansion of the payoff functions $\pi_1(x'; x, y)$ and $\pi_2(y'; x, y)$ about (x^*, y^*) .

(b) (x^*, y^*) is convergence stable if both eigenvalues of the linearization (7) have negative real parts for any choice of positive $k_1(x^*, y^*)$ and $k_2(x^*, y^*)$. This latter condition holds if and only if the trace is negative (i.e. $k_1(x^*, y^*)(A+B) + k_2(x^*, y^*)(E+F) < 0$) and the determinant is positive (i.e. $k_1(x^*, y^*)k_2(x^*, y^*)[(A+B)(E+F) - DC] > 0$).

Assume that either $x((A+B)x + Cy) < 0$ or $y(Dx + (E+F)y) < 0$ for all nonzero $(x, y) \in \mathbf{R}^2$. In particular, with $(x, y) = (x, 0)$, we have $A+B < 0$. Analogously $E+F < 0$ and so the trace is negative. For a fixed nonzero y , let $x \equiv -\frac{C}{A+B}y$. Then $(A+B)x + Cy = 0$ and so $y(Dx + (E+F)y) < 0$. That is, $y\left(-\frac{CD}{A+B}y + (E+F)y\right) = \frac{(A+B)(E+F) - CD}{A+B}y^2$ is negative and this implies the determinant is positive. Thus, (x^*, y^*) is convergence stable.

Conversely, assume that (x^*, y^*) is convergence stable. Then, the trace must be non-positive and the determinant non-negative for any choice of positive $k_1(x^*, y^*)$ and $k_2(x^*, y^*)$ (otherwise, there is an eigenvalue with positive real part). In particular, $A+B \leq 0$ and $E+F \leq 0$. Case 1. If $CD \leq 0$, then either $xCy \leq 0$ or $yDx \leq 0$. Thus, either $x((A+B)x + Cy) \leq 0$ or $y(Dx + (E+F)y) \leq 0$ for all $(x, y) \in \mathbf{R}^2$. Case 2. If $CD > 0$, we may assume without loss of generality that $C > 0$ and $D > 0$. Suppose that $x((A+B)x + Cy) > 0$. Then $xy > -\frac{(A+B)x^2}{C} > 0$. Thus $y(Dx + (E+F)y) = \frac{y}{x}(Dx^2 + (E+F)xy) < \frac{y}{x}(Dx^2 + -\frac{(A+B)(E+F)x^2}{C}xy) \leq 0$.

(c) These statements follow from the arguments used to prove part b. ■

As in Section 3.1 for symmetric games, a neighborhood CSS is a neighborhood strict NE that is convergence stable (when borderline cases are excluded). For one-dimensional strategy spaces, S and T , parts b and c of Theorem 8 give equivalent conditions for convergence stability. Although the inequalities in part c are the easiest to use in practical examples, it is the approach in part b that is most directly tied to the theory of CSS, NIS and neighborhood superiority as well as their connections to evolutionary dynamics, especially as the strategy spaces become multi-dimensional. It is again neighborhood superiority accord-

ing to part a of the following definition that unifies this theory (see Theorem 9 that assumes borderline cases are excluded).

Definition 4. (Cressman, 2010) Suppose (x^*, y^*) is in the interior of $S \times T$.

(a) Fix $0 \leq p^* < 1$. Strategy pair (x^*, y^*) is *neighborhood p^* -superior* if

$$\text{either } \pi_1(x^*; P, Q) > \pi_1(P; P, Q) \text{ or } \pi_2(y^*; P, Q) > \pi_2(Q; P, Q) \quad (8)$$

for all $(P, Q) \in \Delta(S) \times \Delta(T)$ with $1 \geq P(\{x^*\}) \geq p^*$, $1 \geq Q(\{y^*\}) \geq p^*$ and the support of (P, Q) sufficiently close (but not equal) to (x^*, y^*) . (x^*, y^*) is *neighborhood half-superior* if $p^* = \frac{1}{2}$. (x^*, y^*) is *neighborhood superior* if $p^* = 0$. (x^*, y^*) is *(globally) p^* -superior* if the support of (P, Q) in (8) is an arbitrary subset of $S \times T$ (other than $\{(x^*, y^*)\}$).

(b) Strategy pair (x^*, y^*) is a *neighborhood invader strategy (NIS)* if, for all (x, y) sufficiently close (but not equal) to (x^*, y^*) , either $\pi_1(x^*; x, y) > \pi_1(x; x, y)$ or $\pi_2(y^*; x, y) > \pi_2(y; x, y)$.

Theorem 9. (Cressman, 2010) Suppose that (x^*, y^*) is in the interior of $S \times T$.

(a) (x^*, y^*) is a neighborhood CSS if and only if it is neighborhood half-superior.

(b) (x^*, y^*) is a neighborhood strict NE and NIS if and only if it is neighborhood superior.

(c) Consider evolution under the replicator equation generalizing (4) to asymmetric games and initial population distributions $(P_0, Q_0) \in \Delta(S) \times \Delta(T)$ that satisfy $P_0(\{x^*\})Q_0(\{y^*\}) > 0$. If (x^*, y^*) is a strict neighborhood NE and a NIS, then (P_t, Q_t) converges to $(\delta_{x^*}, \delta_{y^*})$ in the weak topology whenever the support of (P_0, Q_0) is sufficiently close to (x^*, y^*) . Conversely, if (P_t, Q_t) converges to $(\delta_{x^*}, \delta_{y^*})$ in the weak topology for every (P_0, Q_0) with support sufficiently close to (x^*, y^*) , then (x^*, y^*) is a neighborhood strict NE and NIS.

Theorem 8 is the (two-role) asymmetric counterpart of Theorem 5 for symmetric games when the continuous strategy spaces are one dimensional. Definition 4 and Theorem 9 generalize Definition 3 and Theorems 6 and 7 of Section 3.1 to these asymmetric games. Based on a thorough analysis of the Taylor expansions of the two payoff functions, their statements remain correct when S and T are multi-dimensional.

4 Conclusion

The static payoff comparisons (e.g. the ESS conditions) introduced by Maynard Smith (1982) to predict the behavioral outcome of evolution in symmetric games with finitely many strategies have been extended in many directions during the intervening years. These include biological extensions to multiple species and to population games as well as the equally important extensions to predict rational individual behavior in human conflict situations. As is apparent from this article, there is a complex relationship between these static conditions and evolutionary stability of the underlying dynamical system.

This article has emphasized evolutionary stability in (symmetric or asymmetric) extensive form games and games with continuous strategy spaces under the deterministic replicator equation that is based on random pairwise interactions. Evolutionary stability is also of much current interest for other game-theoretic models such as those that incorporate stochastic effects due to finite populations; models with assortative (i.e. non-random) interactions (e.g. games on graphs); models with multi-player interactions (e.g. public goods games). As the evolutionary theory behind these (and other) models is a rapidly expanding area of current research, it is impossible to know in what guise the evolutionary stability conditions will emerge in future applications. On the other hand, it is certain that Maynard Smith's original idea will continue to play a central role.

References

- [1] Apaloo, J. (1997) Revisiting strategic models of evolution: The concept of neighborhood invader strategies. *Theor. Pop. Biol.* **52**, 71-77.
- [2] Apaloo, J., J.S. Brown and T.L. Vincent. (2009) Evolutionary game theory: ESS, convergence stability, and NIS. *Evol. Ecol. Res.* **11**, 489-515.
- [3] Bomze, I.M. (1991) Cross entropy minimization in uninvadable states of complex populations. *J. Math. Biol.* **30**, 73-87.
- [4] Bomze, I.M. and B.M. Potscher (1989) *Game Theoretical Foundations of Evolutionary Stability*. Springer, Berlin.
- [5] Chamberland, M. and R. Cressman. (2000) An example of dynamic (in)consistency in symmetric extensive form evolutionary games. *Games and Econ. Behav.* **30** 319-326.
- [6] Christiansen, F.B. (1991) On conditions for evolutionary stability for a continuously varying character. *Amer. Nat.* **138**, 37-50.
- [7] Cressman, R. (2003) *Evolutionary Dynamics and Extensive Form Games*, MIT Press, Cambridge, MA.
- [8] Cressman, R. (2009) Continuously stable strategies, neighborhood superiority and two-player games with continuous strategy spaces. *Int. J. Game Theory*, **38**, 221-247.
- [9] Cressman, R. (2010) CSS, NIS and dynamic stability for two-species behavioral models with continuous trait spaces. *J. Theor. Biol.*, **262**, 80-89.
- [10] Cressman, R., J. Garay and J. Hofbauer (2001) Evolutionary stability concepts for N-species frequency-dependent interactions. *J. Theor. Biol.*, **211**, 1-10.

- [11] Cressman, R., J. Hofbauer and F. Riedel (2006) Stability of the replicator equation for a single-species with a multi-dimensional continuous trait space. *J. Theor. Biol.*, **239**, 273-288.
- [12] Dercole, F. and S. Rinaldi (2008) *Analysis of Evolutionary Processes. The Adaptive Dynamics Approach and its Applications*. Princeton: Princeton University Press.
- [13] Dieckmann, U. and R. Law (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579-612.
- [14] Doebeli, M. and U. Dieckmann (2000) Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Am. Nat.* **156**, S77-S101.
- [15] Eshel I. (1983) Evolutionary and continuous stability. *J. Theor. Biol.* **103**, 99-111.
- [16] Fretwell, D.S. and H.L. Lucas (1969) On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* **19**, 16-32.
- [17] Hofbauer, J. and K. Sigmund (1998) *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge.
- [18] Krivan, V., R. Cressman and C. Schneider (2008) The ideal free distribution: A review and synthesis of the game theoretic perspective. *Theor. Pop. Biol.* **73**, 403-425.
- [19] Leimar, O. (2009). Multidimensional convergence stability. *Evol. Ecol. Res.* **11**, 191-208..
- [20] Lessard, S. (1990) Evolutionary stability: one concept, several meanings. *Theoret. Pop. Biol.* **37**, 159-170.
- [21] Marrow P., U. Dieckmann and R. Law (1996) Evolutionary dynamics of predator-prey systems: an ecological perspective. *J. Math. Biol.* **34**, 556-578.
- [22] Maynard Smith, J. (1982) *Evolution and the Theory of Games*, Cambridge University Press.
- [23] Oechssler, J. and F. Riedel (2001) Evolutionary dynamics on infinite strategy spaces. *Econ. Theory* **17**, 141-162.
- [24] Rosenthal, R. (1981) Games of perfect information, predatory pricing and the chain-store paradox. *J. Econ. Theor.* **25**, 92-100.
- [25] Selten, R. (1978) The chain-store paradox. *Theory and Decision* **9**, 127-159.

- [26] Selten, R. (1983) Evolutionary stability in extensive two-person games. *Math. Soc. Sci.* **5** 269-363.
- [27] Selten, R. (1988) Evolutionary stability in extensive two-person games - correction and further development. *Math. Soc. Sci.* **16** 223-266.
- [28] van Damme, E. (1991) *Stability and Perfection of Nash Equilibria* (2nd Edition). Berlin: Springer-Verlag.
- [29] Vincent, T.L. and J.S. Brown (2005) *Evolutionary Game Theory, Natural Selection and Darwinian Dynamics*, Cambridge University Press.
- [30] von Neumann, J. and O. Morgenstern (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
- [31] Weibull, J. (1995) *Evolutionary Game Theory*, MIT Press, Cambridge. MA.