ROSS CRESSMAN, JÓZSEF GARAY, ANTONINO SCARELLI and ZOLTÁN VARGA

THE DYNAMIC STABILITY OF COALITIONIST BEHAVIOUR FOR TWO-STRATEGY BIMATRIX GAMES*

ABSTRACT. In this paper replicator dynamics are introduced to describe the propagation of coalitionist behaviour in conflicts given by a two-strategy bimatrix games. In the proposed approach non-coalitionists play either Nash strategies or choose one of two pure strategies. In the first case it is proved that non-coalitionists are asymptotically eliminated. In the second case coalitionists can propagate without eliminating all non-coalitionists.

KEY WORDS: bimatrix game, replicator dynamics, imitation, coalition

1. INTRODUCTION

Evolutionary game theory applied to bimatrix games has traditionally been restricted to predicting the players' behaviour in the context of a non-cooperative game. The purpose of this paper is to investigate continuous dynamic models that describe the evolution of the frequency for cooperative coalitionist behaviour relative to that of non-cooperative behaviour in pair-wise conflicts between agents of two populations that are given as a bimatrix game.

The basic idea is to adapt the so-called replicator dynamics describing the time variation of certain behaviour types in a population (Hofbauer and Sigmund, 1998; Cressman, 1992). In the biological evolutionary game context, where coalition formation has received little attention, replication means that "like begets like": a behaviour type is passed down from parent to offspring. Adapting this idea sketched in Weibull (1995), see also Hofbauer and Sigmund (1998), our model can also be applied to

* The research was supported by the Hungarian National Research Fund (OTKA) No. T037271. The final version was completed while one of the authors (R.C.) was a Fellow at the Collegium Budapest.

Theory and Decision **56:** 141–152, 2004.

the social evolution of behaviour types (strategies) in a population of pair-wise interacting agents, such as when individuals represent firms or some other social or economic units. The replication of carriers of a behaviour type can then be thought of as resulting from imitation. An agent is ready to review his strategy at a certain rate depending on the success of another strategy. In the corresponding replicator dynamics, the time rate of a strategy's relative frequency change is positive if this strategy performs better than the average, and negative in the opposite case.

Originally the replicator dynamics was introduced for the study of game-theoretical conflicts in biological evolutionary processes. This is based on the Darwinian theory where the competitiveness of behaviour types is the key issue since these types will increase in relative frequency if their gain (fitness) is greater than the mean gain (fitness) of the whole population. The aim of the present paper is then to study whether replicator dynamics corresponding to the theory of cooperative games can describe the propagation of coalitionist behaviour.

We propose dynamic models where in each population of agents the basic behaviour types are: coalitionist and non-coalitionist. In the *monomorphic dynamic model*, based on a two-strategy two-person (bimatrix) game, in the pair-wise conflict of coalitionists a two-person coalition is formed, and always the Nash equilibrium strategies are played if at least one of the agents is non-coalitionist (Owen, 1968, 1995). In the *polymorphic dynamic model*, the pair-wise conflict of two coalitionists is again modelled by a two-person coalition game, but if at least one of the agents is non-coalitionist, both use one of the two fixed pure strategies. In each model, the imputation is a fixed parameter. The asymptotic behaviour of these models is studied and a numerical illustration is also given.

2. BIMATRIX GAMES AND TWO-PERSON COALITIONS

Let us consider a bimatrix game with pay-off matrices

$$\mathbf{A} := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{B} := \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

The state space, i.e. the Cartesian product of mixed strategy sets, is then $S_2 \times S_2$ where $S_2 = \{p = (p_1, p_2) : p_1 + p_2 = 1, p_1 \geq 0, p_2 \geq 0\}$. First we look for an interior (totally mixed) Nash equilibrium, i.e., a strategy pair $(\mathbf{p}^*, \mathbf{q}^*) \in \mathrm{int}\, S_2 \times S_2$ with $\mathbf{p}^*\mathbf{A}\mathbf{q}^* \geq \mathbf{p}\mathbf{A}\mathbf{q}^*$ and $\mathbf{q}^*\mathbf{B}\mathbf{p}^* \geq \mathbf{q}\mathbf{B}\mathbf{p}^*$ for all $(\mathbf{p}, \mathbf{q}) \in S_2 \times S_2$. Throughout the paper, we will assume that $sign(a_{22} - a_{12}) = sign(a_{11} - a_{21}) \neq 0$ and $sign(b_{22} - b_{12}) = sign(b_{11} - b_{21}) \neq 0$. Then the unique interior Nash equilibrium is $(\mathbf{p}^*, \mathbf{q}^*)$ with

$$\mathbf{p}^* := \left(\frac{b_{22} - b_{12}}{b_{11} - b_{12} - b_{21} + b_{22}}, \frac{b_{11} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}}\right), \quad (2.1)$$

and

$$\mathbf{q}^* := \left(\frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}, \frac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}\right). \quad (2.2)$$

Let us consider now the two-person co-operative (coalition) game corresponding to the above bimatrix game. In addition we shall assume that no entries of the matrix $\mathbf{A} + \mathbf{B}^{\mathrm{T}}$ are equal. Then it is easy to see that the coalition game is *NE-essential*[1] in the sense that

$$\max_{\mathbf{p},\mathbf{q} \in S_2} \mathbf{p}(\mathbf{A} + \mathbf{B}^{\mathrm{T}})\mathbf{q} = \max_{i,j \in \{1,2\}} (a_{ij} + b_{ji}) > \mathbf{p}^*(\mathbf{A} + \mathbf{B}^{\mathrm{T}})\mathbf{q}^*.$$

If the strictly maximal entry of matrix $\mathbf{A} + \mathbf{B}^{\mathrm{T}}$ is $a_{kl} + b_{lk}$, then the strategy pair providing the maximum pay-off to the two-person coalition is the corresponding pure strategy pair denoted by $(\mathbf{e}_k, \mathbf{e}_l)$. For a given NE-essential game a pair $(u, v) \in \mathbf{R}^2$ will be called an *NE-imputation*, if

$$u + v = \max_{\mathbf{p},\mathbf{q} \in S_2} \mathbf{p}(\mathbf{A} + \mathbf{B}^{\mathrm{T}})\mathbf{q} \quad \text{and} \quad \mathbf{p}^*\mathbf{A}\mathbf{q}^* \leq \mathrm{u}, \mathbf{q}^*\mathbf{B}\mathbf{p}^* \leq \mathrm{v},$$

where at least one of the inequalities is strict, and (u, v) is said to be a *strict NE-imputation* if both inequalities are strict.

3. MONOMORPHIC REPLICATOR DYNAMICS

Let us consider populations I and II consisting only of non-cooperative individuals playing their corresponding interior Nash equilibrium strategies. In evolutionary game theory the populations in such a situation are called monomorphic, since all players in the same population have the same behavioural phenotype. In this section we shall consider whether the coalitionist behaviour would propagate in such populations.

Let us assume that there are two possible behaviour types for the agents: coalitionist and non-coalitionist. In a pair-wise conflict of two coalitionists a two-person coalition is formed with the fixed NE-imputation (u, v). If at least one of the agents is non-coalitionist, then both agents play their respective NE strategy. It is assumed that the relative rate of change in frequency of a behaviour in each population is equal to the difference between the expected payoff to the given behaviour type and the average payoff over the corresponding population (social or economic unit). Denote by x_1 and x_2 the frequencies of coalitionist and non-coalitionist behaviours in population I, respectively, and put $\mathbf{x} = (x_1, x_2)$. The meaning of $\mathbf{y} = (y_1, y_2)$ is similar for population II. Then the corresponding replicator dynamics is

$$\begin{aligned}
\dot{x}_1 &= x_1[uy_1 + \mathbf{p}^*\mathbf{A}\mathbf{q}^*y_2 - V(\mathbf{x},\mathbf{y})],\\
\dot{x}_2 &= x_2[\mathbf{p}^*\mathbf{A}\mathbf{q}^* - V(\mathbf{x},\mathbf{y})],\\
\dot{y}_1 &= y_1[vx_1 + \mathbf{q}^*\mathbf{B}\mathbf{p}^*x_2 - W(\mathbf{x},\mathbf{y})],\\
\dot{y}_2 &= y_2[\mathbf{q}^*\mathbf{B}\mathbf{p}^* - W(\mathbf{x},\mathbf{y})]
\end{aligned} \tag{3.1}$$

where $V(\mathbf{x},\mathbf{y}) = ux_1y_1 + \mathbf{p}^*\mathbf{A}\mathbf{q}^*(x_1y_2 + x_2)$ and $W(\mathbf{x},\mathbf{y}) = vx_1y_1 + \mathbf{q}^*\mathbf{B}\mathbf{p}^*(y_1x_2 + y_2)$ are the average payoffs of the players in population I and II, respectively. The set $S_2 \times S_2$ and its "faces" are positively invariant. A rest point must satisfy the following equalities

$$\begin{aligned}
(\mathbf{p}^*\mathbf{A}\mathbf{q}^* - u)y_1x_1x_2 &= 0,\\
(\mathbf{q}^*\mathbf{B}\mathbf{p}^* - v)x_1y_1y_2 &= 0.
\end{aligned}$$

Thus, since at least one of the inequalities $\mathbf{p}^*\mathbf{A}\mathbf{q}^* \leq u$, $\mathbf{q}^*\mathbf{B}\mathbf{p}^* \leq v$ is strict, there is no interior rest point.

THEOREM 1. *For any strict NE-imputation* (u, v) , *the coalitionist state* ((1,0), (1,0)) *is globally asymptotically stable for dynamics* (3.1).

Proof. Observe that

$$\begin{aligned}\mathbf{p}^*\mathbf{A}\mathbf{q}^* - V(\mathbf{x},\mathbf{y}) &= \mathbf{p}^*\mathbf{A}\mathbf{q}^* - ux_1y_1 - \mathbf{p}^*\mathbf{A}\mathbf{q}^*(x_1y_2 + x_2)\\ &= \mathbf{p}^*\mathbf{A}\mathbf{q}^*(1 - (x_1y_2 + x_2)) - ux_1y_1\\ &= (\mathbf{p}^*\mathbf{A}\mathbf{q}^* - u)x_1y_1,\end{aligned}$$

and similarly

$$\mathbf{q}^*\mathbf{B}\mathbf{p}^* - W(\mathbf{x},\mathbf{y}) = (\mathbf{q}^*\mathbf{B}\mathbf{p}^* - v)x_1y_1.$$

Thus $\mathbf{p}^*\mathbf{A}\mathbf{q}^* < u$ and $\mathbf{q}^*\mathbf{B}\mathbf{p}^* < v$ imply $\dot{x}_2 < 0$ and $\dot{y}_2 < 0$, respectively. Therefore, the coalitionist agents with imputation greater than the corresponding NE payoff will win in the sense that, in the long term, non-coalitionists are eliminated from the corresponding population. □

4. POLYMORPHIC REPLICATOR DYNAMICS

Let us consider populations I and II where pairs of individuals who are not involved in a cooperative game must play one of their pure strategies. In evolutionary game theory, the populations in such a situation are called polymorphic, since within each population several behavioural phenotypes may be present at a given time. In this section we shall study whether the coalitionist behaviour would propagate in such populations.

Let us start out from an NE-essential bimatrix game as given above, and fix a NE-imputation (u, v) .

We again assume that, in a pair-wise conflict of two coalitionists, a two-person coalition is formed with imputation (u, v). However, unlike the monomorphic model, if at least one of the agents is a non-coalitionist, then fixed pure strategies ($\mathbf{e}_1$ or $\mathbf{e}_2$) are played. Denote by x_1 and x_2 the relative frequencies of coalitionists of population I playing $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively against non-coalitionists of population II. Furthermore, let x_3 and x_4 be the relative frequencies of non-coalitionists of population I playing $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively. We use similar notation y_1, y_2, y_3, y_4 for population II. This situation can be described by a 4-player bimatrix game with the following partitioned payoff matrices

$$\bar{\mathbf{A}} := \begin{pmatrix} \mathbf{U} & \mathbf{A} \\ \mathbf{A} & \mathbf{A} \end{pmatrix} \in \mathbf{R}^{4\times 4}, \qquad \overline{\mathbf{B}} := \begin{pmatrix} \mathbf{V} & \mathbf{B} \\ \mathbf{B} & \mathbf{B} \end{pmatrix} \in \mathbf{R}^{4\times 4},$$

where

$$\mathbf{U} := \begin{pmatrix} u & u \\ u & u \end{pmatrix} \quad \text{and} \quad \mathbf{V} := \begin{pmatrix} v & v \\ v & v \end{pmatrix}.$$

For this 4-person game the corresponding polymorphic replicator dynamics is

$$\begin{aligned}
\dot{x}_1 &= x_1[u(y_1 + y_2) + e_1\mathbf{A}(y_3, y_4) - V(\mathbf{x}, \mathbf{y})],\\
\dot{x}_2 &= x_2[u(y_1 + y_2) + e_2\mathbf{A}(y_3, y_4) - V(\mathbf{x}, \mathbf{y})],\\
\dot{x}_3 &= x_3[e_1\mathbf{A}(y_1 + y_3, y_2 + y_4) - V(\mathbf{x}, \mathbf{y})],\\
\dot{x}_4 &= x_4[e_2\mathbf{A}(y_1 + y_3, y_2 + y_4) - V(\mathbf{x}, \mathbf{y})],\\
\dot{y}_1 &= y_1[v(x_1 + x_2) + e_1\mathbf{B}(x_3, x_4) - W(\mathbf{x}, \mathbf{y})],\\
\dot{y}_2 &= y_2[v(x_1 + x_2) + e_2\mathbf{B}(x_3, x_4) - W(\mathbf{x}, \mathbf{y})],\\
\dot{y}_3 &= y_3[e_1\mathbf{B}(x_1 + x_3, x_2 + x_4) - W(\mathbf{x}, \mathbf{y})],\\
\dot{y}_4 &= y_4[e_2\mathbf{B}(x_1 + x_3, x_2 + x_4) - W(\mathbf{x}, \mathbf{y})],
\end{aligned} \tag{4.1}$$

where

$$\begin{aligned} V(\mathbf{x}, \mathbf{z}) = u(x_1 + x_2)(y_1 + y_2) + (x_1, x_2)\mathbf{A}(y_3, y_4) + \\ + (x_3, x_4)\mathbf{A}(y_1 + y_3, y_2 + y_4) \end{aligned}$$

and

$$\begin{aligned} W(\mathbf{x}, \mathbf{z}) = v(x_1 + x_2)(y_1 + y_2) + (y_1, y_2)\mathbf{B}(x_3, x_4) + \\ + (y_3, y_4)\mathbf{B}(x_1 + x_3, x_2 + x_4) \end{aligned}$$

denote the average payoffs in populations I and II, respectively.

The set $S_4 \times S_4$ and its "faces" are positively invariant under the above dynamics. It is easy to see that the "coalitionist face", $C := \{(\mathbf{x}, \mathbf{y}) \in (S_4 \times S_4) | x_3 = x_4 = 0 \text{ and } y_3 = y_4 = 0\}$ consists only of rest points of dynamics (4.1). Since under our conditions the initial bimatrix game has a unique interior NE, it is easily seen that the unique non-trivial rest point of dynamics (4.1) in the "non-coalitionist face" $D := \{(\mathbf{x}, \mathbf{y}) \in (S_4 \times S_4)\ | x_1 = x_2 = 0 \text{ and } y_1 = y_2 = 0\}$ is $(\mathbf{x}^*, \mathbf{y}^*)$ with

$$\mathbf{x}^* = \left(0, 0, \frac{b_{22} - b_{12}}{b_{11} - b_{12} - b_{21} + b_{22}}, \frac{b_{11} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}}\right),$$

$$\mathbf{y}^* = \left(0, 0, \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}, \frac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}\right).$$

Observe that the non-zero coordinates provide the Nash equilibrium of the original bimatrix game of the non-coalitionist agents of the different populations. We notice that, within the non-coalitionist face, either the interior Nash equilibrium is neutrally stable with periodic orbits or two of the vertices are locally asymptotically stable (Hofbauer and Sigmund, 1998).

Now we shall find conditions implying that the interior of the "non-coalitionist face" is a repellor.

THEOREM 2. *Suppose that for each payoff matrix the arithmetic mean of each column is not greater than the corresponding NE-imputation component:*

$$\begin{aligned} &u \geq \frac{1}{2}(a_{11} + a_{21}), \quad u \geq \frac{1}{2}(a_{12} + a_{22}), \\ &v \geq \frac{1}{2}(b_{11} + b_{21}), \quad v \geq \frac{1}{2}(b_{12} + b_{22}) \end{aligned} \tag{4.2}$$

with at least one strict inequality.

Then the interior of D is a repellor (*i.e. no trajectory in the interior of $S_4 \times S_4$ has an ω-limit point in the interior of D*).

Proof. Define a function

$$L : \mathrm{int}(S_4 \times S_4) \to \mathbf{R},$$

$$\begin{aligned} L(\mathbf{x}, \mathbf{y}) := &\ln x_1 + \ln x_2 - \ln x_3 - \ln x_4 + \ln y_1 + \\ &+ \ln y_2 - \ln y_3 - \ln y_4. \end{aligned}$$

Then along any trajectory of dynamics (4.1) starting from the interior of the strategy set, the function L is strictly increasing. Indeed, the derivative of L with respect to the considered dynamics is

$$DL(\mathbf{x},\mathbf{y}) = (1,1)[\mathbf{U}-\mathbf{A}](y_1,y_2) + (1,1)[\mathbf{V}-\mathbf{B}](x_1,x_2)$$
$$(\mathbf{x},\mathbf{y}) \in \mathrm{int}(S_4 \times S_4).$$

Condition (4.2) on the imputation implies that for all $y_1 > 0$, $y_2 > 0$, $x_1 > 0$ and $x_2 > 0$ we have

$$(2u - a_{11} - a_{21})y_1 + (2u - a_{12} - a_{22})y_2 + (2v - b_{11} - b_{21}) x_1 + (2v - b_{12} - b_{22})x_2 > 0$$

Hence $DL(\mathbf{x},\mathbf{y}) > 0$, implying that function L is strictly increasing along each trajectory starting from the interior. Since the function L approaches $-\infty$ near any point of the interior of D, the latter is a repellor. □

Remark 1. The condition (4.2) of Theorem 2 can be related to the concept of dominance used by Akin (1980) for symmetric two-person matrix games.

Following Akin (1980) and Akin and Hofbauer (1982) we can prove

THEOREM 3. *Under the conditions of Theorem 2 we have* $x_3(t)x_4(t)y_3(t)y_4(t) \to 0$ as $t \to +\infty$ *for any trajectory in the interior of* $S_4 \times S_4$.

Proof. Let (x,y) be a solution of dynamics (4.1) with $(x(0),y(0)) \in \mathrm{int}(S_4 \times S_4)$. Since the "faces" of $S_4 \times S_4$ are invariant, so is $\mathrm{int}(S_4{\times}S_4)$. By the proof of Theorem 2, $L(x(t),y(t))$ is strictly increasing. First we show that

$$\lim_{t\to+\infty} L(x(t),y(t)) = +\infty. \tag{4.4}$$

Suppose the contrary: for some $c \in \mathbf{R}$ we have $\lim_{t\to+\infty} L(x(t), y(t)) = c$. This implies $\lim_{t\to+\infty} d/dt\, L(x(t),y(t)) = 0$. By the compactness of $S_4 \times S_4$ the solution (x,y) has an ω-limit point $(x^0,y^0) \in S_4 \times S_4$. Then, for an appropriate $t_n \to +\infty$, we have $\lim_{n\to+\infty} d/dt\, L(x(t_n),y(t_n)) = 0$. By inequality (4.2), the latter implies $x_1^0 = x_2^0 = y_1^0 = y_2^0 = 0$. Hence $\lim_{n\to+\infty} L(x(t_n),y(t_n)) = +\infty$ which is a contradiction to the supposed finite limit.

Now from (4.4) we conclude

$$\lim_{t\to+\infty} \frac{x_1(t)x_2(t)}{x_3(t)x_4(t)}\frac{y_1(t)y_2(t)}{y_3(t)y_4(t)} = +\infty.$$

Since the numerator in the quotient is bounded, we have $x_3(t)x_4(t)y_3(t)y_4(t) \to 0$ as $t \to \infty$.

The interpretation of Theorem 3 is that, in at least one of the populations, the frequency of at least one of the non-coalitionist types arbitrarily approaches zero, risking dying out. On the other hand, as we will see from the following example, Theorems 2 and 3 combined do not imply the coalitionist face is asymptotically stable (or even attracting) for the polymorphic model. This contrasts with Theorem 1 which showed the coalitionist vertex is asymptotically stable for the monomorphic model.

EXAMPLE 1. Let

$$\mathbf{A} := \begin{pmatrix} 10 & 1 \\ 2 & 4 \end{pmatrix}, \qquad \mathbf{B} := \begin{pmatrix} 2 & 4 \\ 1 & 6 \end{pmatrix}.$$

This is a *Coordination Game* (as a non-cooperative bimatrix game) and we will show that the coalitionist face is not globally attractive. This game has one mixed Nash equilibrium, $(\mathbf{p}^*, \mathbf{q}^*) = \left(\left(\frac{2}{3}, \frac{1}{3}\right), \left(\frac{3}{11}, \frac{8}{11}\right)\right)$ as well as two pure Nash equilibria, ((1,0),(1,0)) and ((0,1),(0,1)) . Since the matrix $\mathbf{A} + \mathbf{B}^{\mathbf{T}}$ has its strictly maximal entry

$$a_{11} + b_{11} = 12 \quad \text{and} \quad (\mathbf{p}^*\mathbf{A}\mathbf{q}^*, \mathbf{q}^*\mathbf{B}\mathbf{p}^*) = \left(\frac{38}{11}, \frac{8}{3}\right),$$

$(u, v) = (6.5, 5.5)$ is a strict NE-imputation. Furthermore, since maximal column sums of $\mathbf{A}$ and $\mathbf{B}$ are 12 and 10, respectively, this imputation satisfies the inequalities in (4.2) of Theorem 2. Thus the interior of the non-coalitionist face is a repellor.

For any $(\mathbf{x}, \mathbf{y})$ in the interior of $S_4 \times S_4$ near ((0,0,1,0), (1,0,0,0)), we have $(\bar{\mathbf{A}}y)_3 > (\bar{\mathbf{A}}\mathbf{y})_i$ for $i = 1, 2, 4$ and $(\overline{\mathbf{B}}\mathbf{x})_1 > (\overline{\mathbf{B}}\mathbf{x})_j$ for $j = 2, 3, 4$. Thus x_3 and y_1 are monotonically increasing in a neighbourhood of ((0,0,1,0),(1,0,0,0)) and so this rest point of (4.1) is neutrally stable. In fact, every interior trajectory of (4.1) initially sufficiently close to ((0,0,1,0),(1,0,0,0)) converges to a point on the edge E joining ((0,0,1,0),(1,0,0,0)) to

((0,0,1,0),(0,0,1,0)). Moreover, any initial interior point near the coalitionist pure strategy ((1,0,0,0),(1,0,0,0)) also evolves to a point on E near ((0,0,1,0),(1,0,0,0)). Thus the coalitionist face is not even locally attractive for this example.

The reason the coalitionist face is not asymptotically stable in Example 1 is that coalitionist populations, whose mean strategies are not near the interior "NE" in the coalitionist face when playing in a non-cooperative game (e.g. when they are near the pure coalitionist strategy ((1,0,0,0),(1,0,0,0))), evolve away from the coalitionist face. The following result shows this cannot happen when the coalitionist mean strategies are near the interior "NE" in the coalitionist face. Theorem 4 is then the counterpart of Theorem 1 for the monomorphic model in that in Section 3 the coalitionists were forced to play the interior NE strategy in their non-cooperative encounters.

Finally we notice that if we take $u = 5$ and $v = 7$ in the above example, then conditions of Theorem 4 hold but those of Theorems 2 and 3 do not.

THEOREM 4. *If* (u, v) *is a strict NE-imputation then each point of C near enough the point* $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) := ((p_1^*, \; p_2^*, \; 0, \; 0), (q_1^*, \; q_2^* \; 0, \; 0))$ *is locally attractive for* $\mathrm{int}(S_4 \times S_4)$.

Proof. Clearly we have

$$\hat{\mathbf{x}} = \left(\frac{b_{22} - b_{12}}{b_{11} - b_{12} - b_{21} + b_{22}}, \frac{b_{11} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}}, 0, 0\right),$$

$$\hat{\mathbf{y}} = \left(\frac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}, \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}, 0, 0\right).$$

Consider the function $L^* : \mathrm{int}(S_4 \times S_4) \to \mathbf{R}$

$$L^*(\mathbf{x}, \mathbf{y}) := \hat{x}_1 \ln x_1 + \hat{x}_2 \ln x_2 - \hat{x}_1 \ln x_3 - \hat{x}_2 \ln x_4 + \hat{y}_1 \ln y_1 + \\ + \hat{y}_2 \ln y_2 - \hat{y}_1 \ln y_3 - \hat{y}_2 \ln y_4.$$

Its derivative with respect to the polymorphic replicator dynamics (4.1) is

$$DL^*(\mathbf{x}, \mathbf{y}) = u(y_1 + y_2) - \mathbf{p}^* \mathbf{A}(y_1, y_2) + v(x_1 + x_2) - \mathbf{q}^* \mathbf{B}(x_1, x_2).$$

Now let us consider function L^* near the Nash-equilibrium lying in C. We claim that the derivative with respect to the dynamics (4.1) is positive. To see this, take $\boldsymbol{\varepsilon}, \boldsymbol{\delta} \in \mathbf{R}^4$ and define

$$\mathbf{x} = \hat{\mathbf{x}} + \boldsymbol{\delta} \text{ and } \mathbf{y} = \hat{\mathbf{y}} + \boldsymbol{\varepsilon}.$$

Using this notation we have

$$\begin{aligned} DL^*(\mathbf{x},\mathbf{y}) =& u - \mathbf{p}^*\mathbf{A}\mathbf{q}^* + v - \mathbf{q}^*\mathbf{B}\mathbf{p}^* + u(\varepsilon_1 + \varepsilon_2) - \\ & - \mathbf{p}^*\mathbf{A}(\delta_1, \delta_2) + v(\delta_1 + \delta_2) - \mathbf{q}^*\mathbf{B}(\varepsilon_1, \varepsilon_2). \end{aligned}$$

Since (u, v) is a strict NE-imputation, for $|\boldsymbol{\varepsilon}|$ and $|\boldsymbol{\delta}|$ sufficiently small we have $DL^*(\mathbf{x},\mathbf{y}) > 0$. Thus the coalitionist face is attractive in the neighbourhood of the Nash equilibrium.

Observe the conclusion of Theorem 3 remains valid under the weaker assumption that the coalition game is NE-essential for a given imputation (i.e. $u + v > \mathbf{p}^*\mathbf{A}\mathbf{q}^* + \mathbf{q}^*\mathbf{B}\mathbf{p}^*$). □

Remark 2. The proof of the above theorem could also be carried out by linearizing the dynamics (4.1) about the interior NE within C.

5. DISCUSSION

The aim of this paper was to introduce the replicator dynamics approach into the theory of cooperative games. Cooperative game theory traditionally works with maximin solutions for given coalitions; whereas, the Nash equilibrium plays a central role for the replicator dynamics (e.g. every interior equilibrium of the replicator dynamics is a Nash equilibrium in the bimatrix game). For this reason, we developed our model and theorems based on the concept of Nash equilibrium rather than maximin solutions.

We investigated the propagation of coalitionist behaviour in terms of both monomorphic and polymorphic replicator dynamics for NE-essential games. We have shown that in the monomorphic model the coalitionist players eliminate the non-coalitionists. In the polymorphic model the coalitionists are able to propagate in the population but in general they cannot eliminate the non-coalitionist behaviour.

Finally, we note that a dynamic approach to coalition games may also be conceptually important in other biological contexts, such as those to describe the evolution of symbiotic and mutualistic interactions.

NOTE

1. In the theory of cooperative games, the standard definition of an essential game uses the maximin value instead of the present Nash equilibrium value.

REFERENCES

Akin, E. (1980), Domination or equilibrium, *Mathematical Biosciences* 50, 239–250.

Akin, E. and Hofbauer, J. (1982), Recurrence of the unfit, *Mathematical Biosciences* 60, 51–62.

Cressman, R. (1992), *The Stability Concept of Evolutionary Game Theory*. Berlin: Springer.

Hofbauer, J. and Sigmund, K. (1998), *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press.

Owen, G. (1968), *Game Theory*, Philadelphia: Saunders.

Owen, G. (1995), *Game Theory*, 3rd edn. San Diego: Academic Press.

Weibull, J.W. (1995), *Evolutionary Game Theory*. Cambridge, Mass.: The MIT Press.

Address for correspondence: Ross Cressman, Department of Mathematics, Wilfrid Laurier University, Waterloo, Ont., Canada N2L 3C5, (E-mail: rcressma@wlu.ca)

József Garay, Theoretical Biology and Ecological Modelling Research Group of the Hungarian Academy of Science and Department of Plant Taxonomy and Ecology, L. Eötvös University, Pázmány Péter sétány 1/C H-1117 Budapest, Hungary (E-mail: garayj@ludens.elte.hu)

Antonino Scarelli, Department of ecology and Sustainable Economic Development, University of Tuscia, Via S. Giovanni Decollato, 01100 Viterbo, Italy (E-mail: scarelli@unitus.it)

Zoltán Varga, Institute of Mathematics and Informatics, Szent István University, Páter K. u. 1. Godollo H-2103 Hungary
(E-mail: zvarga@ mszi.gau.hu)