Stubborn learning

Jean-François Laslier
CNRS and Ecole Polytechnique

Bernard Walliser
Paris School of Economics (ENPC, ENS)

Abstract

The paper studies 2×2 games when both players use a specific reinforcement learning rule applied to mixed strategies. According to this rule, a player keeps on incrementing the probability of a pure strategy if and only if her utility increases. We study the possible trajectories of such a system and the asymptotic states for symmetric, zero-sum, and twin games.

1 Introduction

The paper examines a specific learning rule in the class of reinforcement models (Sutton and Barto 1998) . In reinforcement models, at each point in time, the decision-maker observes only her past utility and chooses which action to play according to her past performances. The model we study, we call *stubborn learning*, is based on the following principles:

-at each period, the decision maker is able to shift his action of an incremental quantity, in one direction or the other,

-he observes the utility obtained in the two past moves

-if the preceding shift induced an increase (decrease) in utility, he keeps on going in the same direction (he reverses direction).

These principles can only be applied in situations where the agent's strategy space is one-dimensional. It was already applied to the Cournot duopoly by Huck, Norman and Oechssler (2004) and independently by Tregouet (2004). It was applied to a continuous version of the Prisoner's Dilemma by Huck, Norman and Oechssler (2005).

A natural application is the case when the action space is the interval $[0, 1]$, interpreted as the probability of chosing one of two possible actions. In the present paper we consider 2×2 games (two players, two actions per player) in mixed strategies. Note already that the above principles need then to be completed in order to define the behavior at the boundary of the strategy space; this will be done in section 2.

The rule has two notable features, which derive directly from the stated principles. First, the rule calls for weak cognitive capacities of the player, for computing as well as for memorizing. Second, in a one-player settings, the rule induces the player to follow the familiar gradient-descent method. In a game setting the rule describes the behavior of an adaptive agent which acts as if she was alone.

In the interior of the strategy space, the system is essentially driven by collective optimality considerations. When both players see they utilities increase, they both continue, hence they generate locally a welfare-increasing path. On the border of the strategy space, two logics interfer: the optimality logic and the equilibrium logic. The precise resultant effect depends on the details of the game.

The rule is applied in this paper to three classes of generic 2×2 games, namely. symmetric, zero-sum and twin games. In each of these classes, we describe all the possible trajectories and their asymptotic behaviors.

Two examples of the obtained results are the following.

In the Prisoner's Dilemma, the system first move in the direction of the Pareto optimum. When it reaches a border of the action space, it is stuck. The system thus escapes the curse of sub-optimal Nash equilibrium.

In Matching Pennies, the system circles around the mixed Nash equilibrium following a slowly expanding square. When it reaches a border of the action space, it cycles around the action space. The system thus tends to avoid the mixed Nash equilibrium.

Our study provides all elements which can be re-used for stuying other classes of 2×2 games. Extension to more than two players, keeping only two strategies per player, is also straightforward. But the extension to the case of more than two possible pure strategies raises different problems since the strategy space of the player is no longer one-dimensional. The rule should thus be generalized in its definition and this is out of the scope of the present paper.

The next section provides the precise definition of the Stubborn learning rule. Section 3 studies the behavior of the system at interior points, on the borders and at the corners of the strategy space, according to the parameters of the 2×2 game. Section 4 is devoted to symmetric games, Section 5 to zero-sum games and Section 6 to twin games.

2 The learning rule

2.1 Framework: 2x2 games

In the general 2x2 game played by players 1 and 2,the payoffs are described in the following matrix:

$1 \setminus 2$	A_2	B_2
A_1	(a_1, a_2)	(b_1, b_2)
B_1	(c_1, c_2)	(d_1, d_2)

Player i plays action A_i with probability α_i. The strategy space is thus the square $[0, 1] \times [0, 1]$. The expected utility of player i is:

$$u_i(\alpha_1, \alpha_2) = a_i\alpha_1\alpha_2 + b_i\alpha_1(1-\alpha_2) + c_i(1-\alpha_1)\alpha_2 + d_i(1-\alpha_1)(1-\alpha_2).$$

For each period t, denote by $\alpha_1(t)$ and $\alpha_2(t)$ the current strategies and by

$$\widetilde{u}_i(t) = u_i(\alpha_1(t), \alpha_2(t))$$

the average utility of player i. In this paper, we assume that each player is able to observe the expected utility at each period. A more realistic variant would assume that the player only observes the actual utility at each period.

2.2 Informal definition

We will now describe the learning rule followed by each player i at each period t. At period t, the player holds in his memory the levels of utility he got and the strategies he chose in the last two periods. From one period to the next, the player increments his strategy by a small amount $\pm\varepsilon$.

The basic principle states that the agent keeps on changing his strategy in the same direction as long as his utility is increasing, but changes in the reversed direction if his utility is decreasing. Such a principle can be applied as long as he is strictly inside the strategy space. The principle fails and must be completed in two cases.

First if the agent is at a border of the strategy space and the previous principle recommends that he goes outside, the agent cannot do it. In such a case, we stipulate that when a player wants to chose, as his strategy, a probability smaller than 0, he chooses the probability 0 (and likewise for 1).Hence when the player wishes to, but cannot, cross the border, he stays on it (and he keeps in his memory the fact that he wanted to cross the border).

Second, if the agent's utility does not change, the previous principle is mute. Such a case generically does not happen at interior points, but it cannot be neglected for border points (in fact this generically can happen only at corner points). In such a case, we stipulate that the player explores in the sense that he chooses at random to increase or decrease his strategy. Hence the player cannot be stuck forever at the same place.

Note that this rule is purely individual and can be used by the decision-maker without knowledge of his natural or strategic environment. Moreover, the rule is deterministic except in the last case.

2.3 Formal definition

The rule is defined recursively for player i at period t. The state variables are the strategy $\alpha_i(t)$ and the observed utility level $\widetilde{u}_i(t)$. We introduce an auxiliary variable $v_i(t)$, which takes value $+1$ and -1 and which indicates in which direction the player intends to increment his strategy; $v_i(t) = +1$ (resp. -1) means that the player wants to increase (resp. decrease) the probability $\alpha_i(t)$.

- The player has in his memory four numbers:

- $\widetilde{u}_i(t-2)$ is the level of utility he obtained at the penultimate period
- $\widetilde{u}_i(t-1)$ is the level of utility he obtained at the last period
- $\alpha_i(t-1)$ is the strategy he played at the last period
- $v_i(t-1)$ is direction he was intending to go in the last period.

- The player computes his intended direction by:

$$v_i(t) = S[\widetilde{u}_i(t-1) - \widetilde{u}_i(t-2)] \cdot v_i(t-1)$$

where

$$S[\delta] = \begin{cases} +1 \text{ if } \delta > 0 \\ -1 \text{ if } \delta < 0 \\ \pm 1 \text{ at random if } \delta = 0 \end{cases}$$

("at random" here means with equiprobability.) This means that the player keeps his direction unchanged when his utility has increased and reverse direction when his utility has decreased. In the case when his utility has not changed, the new intended direction is chosen at random.

- The player computes his strategy

$$\alpha_i(t) = H[\alpha_i(t-1) + \varepsilon \cdot v_i(t)]$$

where

$$H[x] = \begin{cases} 0 \text{ if } x < 0 \\ x \text{ if } 0 \leq x \leq 1 \\ 1 \text{ if } x > 1 \end{cases}$$

This computation simply states that the player implements his intended strategy $\alpha_i(t-1)+\varepsilon \cdot v_i(t)$ whenever that is physically possible (probability is between 0 and 1), and sticks to the border if not.

The initial conditions to be specified are $\widetilde{u}_i(0), \widetilde{u}_i(1), \alpha_i(1)$, and $v_i(1)$.

For convenience, we make the technical assumption that $\varepsilon = 1/N$ for some integer N and that $N\alpha_i(0)$ is an integer.It follows that $N\alpha_i(t)$ is an integer for all t. Consequently–, when a player reaches a border of the strategy space, he reaches it exactly.

2.4 Payoff variations

In this section, we study the instantaneous differential variations of the payoffs. The derivatives of the utility function for player i are the following:

$$\begin{aligned} \partial u_i/\partial \alpha_1 &= b_i - d_i + E_i\alpha_2 \\ \partial u_i/\partial \alpha_2 &= c_i - d_i + E_i\alpha_1 \end{aligned}$$

with

$$E_i = a_i - b_i - c_i + d_i.$$

The expression of the payoff variation depends on the current state being interior or on the border of the strategy space.

Payoff variation at interior points. Denote by $k(t)$ the indicator of covariant ($k(t) = +1$) or contravariant ($k(t) = -1$) evolution of the players' strategies defined by.

$$k(t) = v_1(t)v_2(t).$$

The first order approximation for the utility difference (omitting the period index t) of player 1 can be written in the following way since $d\alpha_2 = kd\alpha_1$:

$$du_1 = \frac{\partial u_1}{\partial \alpha_1} d\alpha_1 + \frac{\partial u_1}{\partial \alpha_2} d\alpha_2 = U_1 d\alpha_1.$$

Then the intended strategy variation can be expressed in a compact way:

$$v_1(t) = v_1(t-1) \cdot \text{sign} U_1(t-1) \cdot v_1(t-1) = \text{sign} U_1(t-1).$$

Here, we denote:

$$U_1 = \begin{cases} U_1^+ \text{ if } k = +1 \\ U_1^- \text{ if } k = -1 \end{cases}$$

hence

$$\begin{aligned} U_1^+ &= \frac{\partial u_1}{\partial \alpha_1} + \frac{\partial u_1}{\partial \alpha_2} = b_1 + c_1 - 2d_1 + E_1(\alpha_1 + \alpha_2) \\ U_1^- &= \frac{\partial u_1}{\partial \alpha_1} - \frac{\partial u_1}{\partial \alpha_2} = b_1 - c_1 + E_1(-\alpha_1 + \alpha_2). \end{aligned}$$

These numbers are interpreted as follow: U_1^+ characterizes the utility variation of player 1 when the system moves parallel to the first diagonal ($d\alpha_2 = d\alpha_1$) while U_1^- characterizes the utility variation of player 1 when the system moves parallel to the second diagonal ($d\alpha_2 = -d\alpha_1$)

The same can be computed for the second player:

$$du_2 = \frac{\partial u_2}{\partial \alpha_1} d\alpha_1 + \frac{\partial u_2}{\partial \alpha_2} d\alpha_2 = U_2 d\alpha_2 = \begin{cases} U_2^+ d\alpha_2 \text{ if } k = +1 \\ U_2^- d\alpha_2 \text{ if } k = -1 \end{cases},$$

where

$$\begin{aligned} U_2^+ &= \frac{\partial u_2}{\partial \alpha_1} + \frac{\partial u_2}{\partial \alpha_2} = b_2 + c_2 - 2d_2 + E_2(\alpha_1 + \alpha_2) \\ U_2^- &= -\frac{\partial u_2}{\partial \alpha_1} + \frac{\partial u_2}{\partial \alpha_2} = -b_2 + c_2 + E_2(\alpha_1 - \alpha_2). \end{aligned}$$

Payoff variation on the border of the strategy space. On the border $\alpha_1 = 0$, when the first player does not move, the first order approximation for the utility difference is :

$$\begin{aligned} du_1 &= \frac{\partial u_1}{\partial \alpha_2} d\alpha_2 = \left(U_1^+ - U_1^-\right) d\alpha_2, \qquad (1) \\ du_2 &= \frac{\partial u_2}{\partial \alpha_2} d\alpha_2 = \left(U_2^+ + U_2^-\right) d\alpha_2. \end{aligned}$$

Similar expressions hold for $\alpha_1 = 1$ and $\alpha_2 = 0$ or 1.

Synthesis. Notice that, from an interior point, the system can move in the four directions parallel to the two diagonals. From a point on a border, the system can also move in four directions, either horizontally or vertically. The possible utility variations can therefore be depicted with a rosace. In each of the eight possible directions, the rosace indicates the sign of the players' payoff variations. For example:

$$\begin{bmatrix} +- & & +- & & -+ \\ & \nwarrow & \uparrow & \nearrow & \\ +- & \leftarrow & & \rightarrow & -+ \\ & \swarrow & \downarrow & \searrow & \\ +- & & -+ & & -+ \end{bmatrix}$$

reads as follows: The upper left corner corresponds to a North-West direction. The utility variations are positive for player 1 and negative for player 2.

Notice that the rosace is constrained by the following rule: for any player, by cycling around the table the signs must be in turn four times positive and four times negative (this leaves 64 possible schemes). In particular, signs in opposite directions are opposite, hence it is sufficient to know the signs in four successive directions. For the sake of simplicity the above scheme will be depicted as:

$$\begin{bmatrix} +- & +- & -+ \\ +- & & -+ \\ +- & -+ & -+ \end{bmatrix}$$

A game is covariant (resp. contravariant) in some direction at a given point if the signs in the rosace are the same (resp. opposite). By convention, an asterix $*$ instead of a sign will mean that the sign can be $+$ or $-$.

Remark 1 *This learning rule differs profoundly from the 'gradient learning rule' which is sometimes considered (Sutton and Barto 1998). In the gradient case, the increment of the probability of a player only depends on the impact of this player's move, the other player's move being implicitely fixed. The utility variation for the players are then:*

$$\begin{aligned} du_1 &= \frac{\partial u_1}{\partial \alpha_1} d\alpha_1, \\ du_2 &= \frac{\partial u_2}{\partial \alpha_2} d\alpha_2. \end{aligned}$$

For instance, in an all-or-nothing version, where each player has a constant increment, this increment is such that: $\delta_i(t) = \text{sign } \hat{U}_i(t-1)$ *where* $\hat{U}_i = U_i^+ + U_i^-$*. A stubborn learner follows this gradient rule only when the other agent stays on a border. The rule differs at interior points.*

3 System evolution

3.1 Partition of the strategy space

We now study the trajectories of the system which result from the two players applying the previous rule. From the above analysis, the individual behavior may change qualitatively when the system reaches two types of lines: the borders of the strategy space and the "separating lines" $U_1^+ = 0$, $U_1^- = 0$, $U_2^+ = 0$, $U_2^- = 0$. Note that separating lines are parallel to the first or second diagonal.

When α_1 and α_2 are not constrained, the four separating lines define at most 9 areas in the plane (α_1, α_2). In some special classes of games, the number of areas is reduced since some separting lines may coincide (in that case we call them double separating lines). The strategy space is a square in the plane (α_1, α_2), which may intersect one or several of these areas.

In the sequel, we will describe the evolution of the system at interior points and on separating lines, borders and corners.

3.2 Evolution at an interior point

Proposition 1. At interior points, after the first move, only two types of trajectories are possible

(i) trajectories going in concert (each player is moving in a fixed direction) South-West, South-East, North-West or North East.

(ii) trajectories going crabwise (one player is moving in a fixed direction and the other player alternates) North, East, South or West.

The last case happens iff $U_1^- U_2^- \succ 0$ and $U_1^+ U_2^+ \prec 0$.

Proof:

According to the expression of sign $v_i(t)$ given by Equation 2.4, the first player moves in some direction independently of the preceding increment $v_i(t-1)$, but according to $k(t-1)$. Since both players act in the same way, the following Table gives, for each value of $k(t-1)$, the direction of evolution of the system in each region of constant signs for $U_1(t)$ and $U_2(t)$ (Table 2.) In this Table, the arrows depict in the usual way the direction of evolution, for instance the South-East arrow $\searrow$ means that $v_1(t) = +1$ and $v_2(t) = -1$.

$k = +1$	$U_2^+ \prec 0$	$U_2^+ \succ 0$	$k = -1$	$U_2^- \prec 0$	$U_2^- \succ 0$
$U_1^+ \prec 0$	$\swarrow$	$\nwarrow$	$U_1^- \prec 0$	$\swarrow$	$\nwarrow$
$U_1^+ \succ 0$	$\searrow$	$\nearrow$	$U_1^- \succ 0$	$\searrow$	$\nearrow$

(2)

Consider an initial value for k. The sign of k indicates whether U_i^+or U_i^- is the relevant expression for U_i. Then Table 2 provides the direction of evolution of (α_1, α_2). This leads to a new value of k. If this new value is the same as the preceding one, the system keeps the same direction. If the sign of k has changed, then the relevant expression for U_i changes and the Table precisely records these changes. Except for the initial period, the system evolution can be described qualitatively as long as the system stays inside an area of the strategy space where U_1 and U_2 have constant signs. In most cases, the direction of evolution is well defined independently of the initial value $k(0)$. In some cases, two directions are possible according to this initial value. The general Table 3 defines the one or two possible *regimes* for each configuration of parameters.

	$U_1^- \prec 0$ $U_2^- \prec 0$	$U_1^- \succ 0$ $U_2^- \succ 0$	$U_1^- \prec 0$ $U_2^- \succ 0$	$U_1^- \succ 0$ $U_2^- \prec 0$
$U_1^+ \prec 0$ $U_2^+ \prec 0$	$\swarrow$	$\swarrow$	$\swarrow$ if $k(0) > 0$ $\nwarrow$ if $k(0) < 0$	$\swarrow$ if $k(0) > 0$ $\searrow$ if $k(0) < 0$
$U_1^+ \succ 0$ $U_2^+ \succ 0$	$\nearrow$	$\nearrow$	$\nearrow$ if $k(0) > 0$ $\nwarrow$ if $k(0) < 0$	$\nearrow$ if $k(0) > 0$ $\searrow$ if $k(0) < 0$
$U_1^+ \prec 0$ $U_2^+ \succ 0$	$\longleftarrow$	$\uparrow$	$\nwarrow$	$\searrow$
$U_1^+ \succ 0$ $U_2^+ \prec 0$	$\downarrow$	$\longrightarrow$	$\nwarrow$	$\searrow$

(3)

This Table has to be read as follows:

- in the four North-West regions and in the four South-East regions, the unique arrow indicates the direction in which the system steadily evolves. For instance the arrow $\swarrow$ indicates that α_1 and α_2 are both decreasing.

-in the four North-East regions, the movement is always in the same direction, but this direction depends on the initial value of k.

-in the four South-West regions, the unique arrow depicts the average evolution, since the system evolves in fact crabwise along a trend. For instance, the North arrow $\uparrow$ means that one move out of two goes North-East while every other move goes North-West. **QED**

From this proposition follow two restrictions about the possible evolutions. Firstly, the direction (v_1, v_2) is either constant or cyclic of order two. The reason is that the direction is completely determined by the sign of k which can only be constant ot alternate. Secondly, a cycle cannot be made of two opposite directions (going back and forth from one state to another). The reason is that the utility increment would change sign at each move but it only changes when it is negative.

To summarize, only two types of trajectories are possible as long as they do not reach a border. Either both directions are constant, hence the two utilities increase. We will then say that the players are moving *in concert*. Or one direction is constant and the other alternates, hence one player sees her utility increase and the other player sees her utility decrease. We then say that the players are moving *crabwise*. This is typically the case if the first player's strategy has a strong negative impact on the second player's payoff while the the second player's strategy has little effect on both players' payoffs.

3.3 Evolution on a separating line

Proposition 2. When crossing a separating line (simple or double) between two areas, only three kinds of trajectories are possible, depending on the relevant regimes on both sides of the line:

(i) The system continues in the interior of the new area. This happens in all cases where the regime in the reached area drives the system away from the separating line.

(ii) The system is scotched in the ε-neighborhood of its impact point. This happens only when the directions on both sides are strictly opposed.

(iii) The system slides along the separating line, in the ε-neighborhood of the line. This happens in the remaining cases.

Proof. Without restriction we consider a separating $U_i^- = 0$. This line is parallel to the main diagonal. It may be the line corresponding to one player only, $U_1^- = 0$ or $U_2^- = 0$, however for some classes of games these two lines may be identical and we shall refer to the case $U_1^- = U_2^- = 0$.

Without restriction, we consider the case where the system was previously at the North-West of the separating line. Then, the separating line can be reached by three types of trajectories, namely $\searrow$, $\longrightarrow$ and $\downarrow$. There is no prior restriction on the possible types of trajectories in the area South-East of the separating line.

Reading the Table 3, crossing the separating line $U_i^- = 0$ means shifting from one cell to another in the same line. Each one of the three other cells can be reached, changing the sign of U_1^- only (label 1) of U_2^- only (label 2), or of both (label 3).

The following Table records the possible shifts between types of trajectories. The numbers in the cells of this table indicates the player(s) involved in the shift.

		after							
		$\searrow$	$\longrightarrow$	$\nearrow$	$\uparrow$	$\nwarrow$	$\longleftarrow$	$\swarrow$	$\downarrow$
	$\downarrow$	1A	3C			1F			
before	$\searrow$		2B	123E	2D	3G	1D	123E	1B
	$\longrightarrow$	2A				2F			3C

This Table indicates that some shifts are impossible (blank cells). Moreover some can be considered as similar for symmetry reasons. Seven different cases remain, denoted A to G. For convenience, we assume that at time t the system is exactly on the separating line. Let $(\alpha_1(t), \alpha_2(t))$ be the point where the trajectory first reaches the line $U_1^- = 0$ and/or $U_2^- = 0$. The preceding point, at $t-1$, thus was $(\alpha_1(t) - \varepsilon, \alpha_2(t) + \varepsilon)$ The next point is denoted:

$$(\alpha_1(t+1), \alpha_2(t+1)) = (\alpha_1(t) + \varepsilon\delta_1(t+1), \alpha_2(t) + \varepsilon\delta_2(t+1)).$$

case A . Transition from $\downarrow$ to $\searrow$. (thus through $U_1^- = 0$) .Coming from a $\downarrow$ move and reaching the line, the next point, at $t+1$, is obtained for $\delta_1(t+1) = -1$, and $\delta_2(t+1) = -1$. Note that the point at $t+1$ is still on the line $U_1^- = 0$. For player 2, the utility variation is positive like before. For player 1, the utility variation is still negative, given by U_1^+. Hence the next move is: $\delta_1(t+2) = +1$, and $\delta_2(t+2) = -1$. After that the system continues in the same direction. To sum up, the system crosses the separating line and keeps on going in concert South East.

case B. Transition from $\searrow$ to $\downarrow$ (thus through $U_1^- = 0$) Coming from a $\searrow$ move, the system crosses the border and stays in the same direction: $\delta_1(t+1) = +1$ and $\delta_2(t+1) = -1$. Since it is now completely on the other side of the line, it continues in the new direction $\downarrow$.

case C. Transition from $\longrightarrow$ to $\downarrow$ (thus through $U_1^- = U_2^- = 0$). Coming from a $\longrightarrow$ move, and reaching the line, the next point is given by $\delta_1(t+1) = +1$ and $\delta_2(t+1) = +1$,which is still on the line. The utility variations (given by U_1^+ and U_2^+) are still positive for player 1 and negative for player 2. Hence the next move is: $\delta_1(t+2) = +1$, and $\delta_2(t+2) = -1$. Ths system is now completely on the other side of the line and continues in the new direction$\downarrow$.

case D. Transition from $\searrow$ to $\longleftarrow$ (thus through $U_1^- = 0$) In that case, the system goes globally South-West, staying ε-close to the separating line. The precise path is the repetition of a pattern of six consecutive moves, starting from the separating line:

$\delta_1(t+1) = +1$ and $\delta_2(t+1) = -1$,
$\delta_1(t+2) = -1$ and $\delta_2(t+1) = -1$,
$\delta_1(t+3) = -1$ and $\delta_2(t+3) = +1$,
$\delta_1(t+4) = -1$ and $\delta_2(t+4) = -1$,
$\delta_1(t+5) = -1$ and $\delta_2(t+5) = +1$,
$\delta_1(t+6) = +1$ and $\delta_2(t+6) = -1$.

case E Transition from $\searrow$ to $\nearrow$ (thus through $U_1^- = 0$, with or without $U_2^- = 0$). In that case, the system crosses the separating line, turns left and keeps on going North-East, ε-close to the separating line (on the same side of the line).

case F Transition from $\longrightarrow$ to $\nwarrow$ (thus through $U_1^- = 0$). In that case, the system goes globally North-East, staying ε-close to the separating line. The precise path is the repetition of a pattern of six consecutive moves, starting from the separating line:

$\delta_1(t+1) = +1$ and $\delta_2(t+1) = +1$,
$\delta_1(t+2) = +1$ and $\delta_2(t+1) = -1$,
$\delta_1(t+3) = -1$ and $\delta_2(t+3) = +1$,
$\delta_1(t+4) = -1$ and $\delta_2(t+4) = +1$,
$\delta_1(t+5) = +1$ and $\delta_2(t+5) = +1$,
$\delta_1(t+6) = +1$ and $\delta_2(t+6) = -1$.

case G. Transition from $\searrow$ to $\nwarrow$ (thus through $U_1^- = U_2^- = 0$). In that case, the system is locked around the point where it crosses the line The precise cycle is made of four consecutive moves:

$\delta_1(t+1) = +1$, $\delta_2(t+1) = -1$
$\delta_1(t+2) = -1$, $\delta_2(t+1) = +1$
$\delta_1(t+3) = -1$, $\delta_2(t+3) = +1$
$\delta_1(t+4) = +1$, $\delta_2(t+4) = -1$.

3.4 Evolution on a border line

Proposition 3. When reaching a border form the interior of the state space,

(i) If the system was moving in concert, the trajectory is of one of two kinds (depending on the payoffs):

i-a). The system is stuck in an 2ε-neighborhood of its impact point. This happens when the game is contravariant on the border and the relevant player (player 2 on a vertical border) is attracted by the border (this payer's utility increases in the three directions pointing to the border).

i-b) The systems slides in an ε-neighborhood of the border, with an angle of $\pi/4$ with respect to its inital direction. This happens in all other cases

(ii) If the system was moving crabwise, the system slides in an ε-neighborhood of the border in the direction corresponding to the best response of the relevant player on this border.

Proof. Note that the trajectory on the border only depends on the signs of payoff variations in three directions, because the direction orthogonal to the border does not matter. Without loss of generality we study the vertical East border $\alpha_1 = 0$. Without loss of generality we suppose that the trajectory first meets the border going South-West, at t.

Case A: The system was moving in concert before reaching the border. In that case the sign of the payoff variation in the South-West direction is $++$. Five subcases are possible depending on the signs of the payoff variations.

Case A1.

$$\begin{bmatrix} ** & -- & -- \\ ** & & ** \\ ++ & ++ & ** \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = -1$$

which means that the system slides downward on the border.

Case A2.

$$\begin{bmatrix} +- & +- & -- \\ ** & & ** \\ ++ & -+ & -+ \end{bmatrix}$$

The successives moves are:

$$\begin{array}{c} \delta_1(t+1) = 0 \ , \ \delta_2(t+1) = -1 \\ \delta_1(t+2) = +1 \ , \ \delta_2(t+2) = -1 \\ \delta_1(t+3) = -1 \ , \ \delta_2(t+3) = -1 \end{array}$$

which means that the system slides downward in an ε-neighborhood of the border.

Case A3

$$\begin{bmatrix} ++ & ++ & -- \\ ** & & ** \\ ++ & -- & -- \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = -1$$
$$\delta_1(t+2) = +1 \ , \ \delta_2(t+2) = +1$$
$$\delta_1(t+3) = -1 \ , \ \delta_2(t+3) = -1$$

which means that the system slides downward in an ε-neighborhood of the border. (Note that the pattern is different from the one in case A2. Here the "speed" downward is 1/3.)

Case A4.

$$\begin{bmatrix} *+ & -+ & -- \\ ** & & ** \\ ++ & +- & *- \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = -1$$
$$\delta_1(t+2) = 0 \ , \ \delta_2(t+2) = +1$$
$$\delta_1(t+3) = +1 \ , \ \delta_2(t+3) = +1$$
$$\delta_1(t+4) = -1 \ , \ \delta_2(t+4) = -1$$

which means that the system is stuck in an ε-neighborhood of the impact point on the border in a 4-cycle.

Case A5

$$\begin{bmatrix} ++ & +- & -- \\ ** & & ** \\ ++ & -+ & -- \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = -1$$
$$\delta_1(t+2) = +1 \ , \ \delta_2(t+2) = -1$$
$$\delta_1(t+3) = -1 \ , \ \delta_2(t+3) = +1$$
$$\delta_1(t+4) = 0 \ , \ \delta_2(t+4) = +1$$
$$\delta_1(t+5) = 0 \ , \ \delta_2(t+5) = -1$$

which means that the system is stuck in an 2ε-neighborhood of the impact point on the border in a 4-cycle.

Synthesis for case A: The following Table indicates the trajectory for each possible case of the rosace. The horizontal lines correspond to the signs in the South direction of the rosace. The columns corresponds to the signs in the South East direction of the rosace. The sympol ⊠ means that the system is scotched around some point at the border. The sign $\downarrow$ indicates that the system slides down along the vertical border. A blank cell indicates that the case is impossible (remember that the South-West direction of the rosace is $++$).

S \ SE	$++$	$-+$	$+-$	$--$
$++$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
$-+$		$\downarrow$		⊠
$+-$			⊠	⊠
$--$				$\downarrow$

Case B: The system was moving crabwise West before reaching the border. In that case the sign of the payoff variation in the South-West direction is $+-$. and the sign in the North-West direction is $+-$. Four subcases are possible depending on the signs of the payoff variations.

Case B1:

$$\begin{bmatrix} +- & -- & -+ \\ ** & & ** \\ +- & ++ & -+ \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = +1$$
$$\delta_1(t+2) = +1 \ , \ \delta_2(t+2) = -1$$
$$\delta_1(t+3) = -1 \ , \ \delta_2(t+3) = -1$$

which means that the system slides downwads in a ε-neighborhood of the border, at "speed" 1:3.

Case B2

$$\begin{bmatrix} +- & +- & -+ \\ ** & & ** \\ +- & -+ & -+ \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = +1$$
$$\delta_1(t+2) = 0 \ , \ \delta_2(t+2) = -1$$
$$\delta_1(t+3) = +1 \ , \ \delta_2(t+3) = -1$$
$$\delta_1(t+4) = -1 \ , \ \delta_2(t+4) = -1$$

which means that the system slides downwards in a ε-neighborhood of the border, at "speed" 1/2.

Case B3:

$$\begin{bmatrix} +- & -+ & -+ \\ ** & & ** \\ +- & +- & -+ \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = +1$$
$$\delta_1(t+2) = +1 \ , \ \delta_2(t+2) = +1$$
$$\delta_1(t+3) = -1, \ \delta_2(t+3) = +1$$
$$\delta_1(t+4) = 0 \ , \ \delta_2(t+4) = -1$$
$$\delta_1(t+5) = 0 \ , \ \delta_2(t+5) = +1$$

which means that the system slides upwards in a ε-neighborhood of the border, at "speed" 1/2.

Case B4

$$\begin{bmatrix} +- & ++ & -+ \\ ** & & ** \\ +- & -- & -+ \end{bmatrix}$$

The successives moves are:

$$\delta_1(t+1) = 0 \ , \ \delta_2(t+1) = +1$$
$$\delta_1(t+2) = 0 \ , \ \delta_2(t+2) = +1$$

which means that the system slides upwards in a ε-neighborhood of the border, at "speed" 1.

Synthesis for case B: The following Table indicates the trajectory for each possible case of the rosace. Again, the horizontal lines correspond to the signs in the South direction of the rosace. But the signs in the South East direction of the rosace is imposed to $-+$ (remember that the signs in the South West direction of the rosace are $-+$). Moreover the result does not depend on the way the crabwise trajectory reaches the border.

S \ SE	$-+$
$++$	$\downarrow$
$-+$	$\downarrow$
$+-$	$\uparrow$
$--$	$\uparrow$

From this sudy of case B, we can conclude that the trajectory on the border always follows the direction of the best response for player 2.

QED

3.5 Evolution at a corner

Proposition: When reaching a corner, the trajectory follows one of two patterns:

(i) It escapes from the corner following a neiborhood of a border. This happens for games covariant in all directions at this corner. If the system arrives at the corner following a border it can either make a U-turn (cases *A1a* or *A3* in the proof); or turn at right angle (case *A1d* in the proof.)

(ii) It stays in a 2ε-neighborhood of the corner. This happens for all other games.

Proof. Consider, without restriction, the corner $(0, 0)$. Consider also provisionaly that the system arrives at the corner along the border $\alpha_1 = 0$. From the preeceding section, there are five cases to be distinguished: A1, A2, A3, B1, and B2, which we study successively.

Case A1a:

$$\begin{bmatrix} ++ & -- & -- \\ ++ & & -- \\ ++ & ++ & -- \end{bmatrix}$$

The system reaches the corner where it is blocked (it does not move for one period, then uses a random device). Then, whatever its departing move from the corner, the system goes back North along the border $\alpha_1 = 0$

Case A1b:

$$\begin{bmatrix} +- & -- & -- \\ +* & & *- \\ ++ & ++ & -+ \end{bmatrix}$$

The system,first blocked at the corner, engages in a 3-cycle around it.
Case A1c:

$$\begin{bmatrix} -+ & -- & -- \\ -+ & & +- \\ ++ & ++ & +- \end{bmatrix}$$

The system engages finally in a 4-cycle around the corner
Case A1c':

$$\begin{bmatrix} -+ & -- & -- \\ ++ & & -- \\ ++ & ++ & +- \end{bmatrix}$$

The system is blocked in successive ways.
Case A1d:

$$\begin{bmatrix} -- & -- & -- \\ ++(--) & & --(++) \\ ++ & ++ & ++ \end{bmatrix}$$

The system turns and goes Right along the border $\alpha_2 = 0$
Case A1d':

$$\begin{bmatrix} -- & -- & -- \\ +-(-+) & & -+(+-) \\ ++ & ++ & ++ \end{bmatrix}$$

The system is scotched in a 4-cycle around the corner.
Case A2a:

$$\begin{bmatrix} +- & +- & -- \\ ++ & & -- \\ ++ & -+ & -+ \end{bmatrix}$$

For all three ways of arriving at the corner, the system is scotched around the corner in a 5-cycle.

Case A2b:

$$\begin{bmatrix} +- & +- & -- \\ +- & & -+ \\ ++ & -+ & -+ \end{bmatrix}$$

There are three ways of arriving at the corner, but each of them leads to the same conclusion: the system is scotched around the corner in a 6-cycle.
Case A3:

$$\begin{bmatrix} ++ & ++ & -- \\ ++ & & -- \\ ++ & -- & -- \end{bmatrix}$$

The system comes to the corner and goes back along the border towards the North.

Case B1:

$$\begin{bmatrix} +- & -- & -+ \\ +- & & -+ \\ +- & ++ & -+ \end{bmatrix}$$

The system comes to the corner and is scotched around in a 3-cycle.

Case B2:

$$\begin{bmatrix} +- & +- & -+ \\ +- & & -+ \\ +- & -+ & -+ \end{bmatrix}$$

The system comes to the corner and is scotched around in a 3-cycle.

Another possibility is to arrive at the corner along the diagonal. Two subcases are possible: either coming crabwise or in concert.

If coming crabwise, the system in fact was sliding in an ε-neighborhood of a border. The last positions of the system were: for instance:

$$\alpha_1(t-2) = 2\varepsilon,\ \alpha_2(t-2) = 0$$
$$\alpha_1(t-1) = \varepsilon,\ \alpha_2(t-1) = \varepsilon$$
$$\alpha_1(t) = 0,\ \alpha_2(t) = 0$$
$$\alpha_1(t+1) = 0,\ \alpha_2(t+1) = \varepsilon$$

The next moves have been described in the paragraph on behavior on a border line. We have already studied what happens to the system when reachning a corner anywhen during its sliding on a border.

If coming in concert on the diagonal, The South-West direction is an improvement for both players therefore the system is blocked at the corner and the next move is random. There are 16 rosaces compatible with the South-West direction being an improvement for both players. An exhaustive study of these cases shows that in the four covariant rosaces the system leaves the corner North or South, and in all other cases the system is stuck in a neighborhood of the corner.

4 Symmetric games

4.1 Potential attractors

In a symmetric game: $a_1 = a_2 = a$, $b_1 = c_2 = b$, $c_1 = b_2 = c$, $d_1 = d_2 = d$. Hence

$$E_1 = E_2 = E = a - b - c + d.$$

The parameter $E = \frac{\partial^2 u_1}{\partial\alpha_1\partial\alpha_2} = \frac{\partial^2 u_2}{\partial\alpha_1\partial\alpha_2}$ will be called the *coupling* parameter. We restrict attention to the cases $E \neq 0$[1]. A game with $E > 0$ will be called a *coupling* game and a game with $E < 0$ will be called a *decoupling* game. If $E > 0$, when one player goes in one direction (say $d\alpha_1 > 0$), the other player is all the more induced to go in the same direction ($\frac{\partial u_2}{\partial \alpha_2}$ increases).

For convenience and without restriction, it can be assumed that $b \geq c$ (if $b < c$, an equivalent game is obtained by exchanging rows and exchanging columns). Since the utility levels are defined up to increasing affine transformation, we can fix two of the four parameters a, b, c, d. We restrict attenbtion to the case $b \neq c$[2]. It appears that the most convenient normalization is to set the values of b and c:

$$b = +1,\ c = -1$$

so that

$$E = a + d.$$

Then the various games to distinguish will be described in the plane (a, d).

Natural candidates for attractors of the dynamic process are Nash equilibria and Bentham optima.

Nash equilibria.

As concern the *pure equilibria* (defined by the values of α_1 and α_2), three types of games can be considered:

- If $a > -1$ and $d < 1$ or if $a < -1$ and $d > 1$, there is only one equilibrium, respectively $(1, 1)$ or $(0, 0)$. Notice that the equilibrium $(1, 1)$ is Pareto optimal if and only if $a > d$ and that the equilibrium $(0, 0)$ is Pareto optimal if and only if $a < d$. For instance, the *Appointment game* is obtained with $a = 2, d = 0$ (hence $E = 2$). Likewise, the *Coupling Prisoner's Dilemma* is obtained with $a = 0, d = 0.5$ (hence $E = 0.5$) and the *Decoupling Prisoner's Dilemma* is obtained with $a = -0.5, d = 0$ (hence $E = -0.5$).

- If $a > -1$ and $d > 1$, there are two symmetric equilibria. In this case, $E \succ 0$. For instance, the *Stag-Hunt game* corresponds to: $a = 1, d = 3$ (hence $E = 4$).

- If $a < -1$ and $d < 1$, there are two assymetric equilibria $(1, 0)$ and $(0, 1)$. In that case $E \prec 0$. For instance, the *Battle of Sexes* corresponds to: $a = -2, d = -3$ (hence $E = -5$)

In the last two cases, there is moreover a *mixed equilibrium* obtained for $\widehat{\alpha_1} = \widehat{\alpha_2} = (d - 1)/(a + d)$.

Bentham optima

[1] Within the class of symmetric games, generically, $E \neq 0$. Note that this rules out symmetric and zero-sum games. Zero-sum games will be treated in the next section..

[2] Within the class of symmetric games, generically, $b \neq c$. Note that this rules out symmetric and twin games. Twin games will be treated in another section.

Consider the maximization of the sum W of players utilities over the whole strategy space:

$$\begin{aligned} W &= u_1 + u_2 \\ &= 2a\alpha_1\alpha_2 + (b+c)\left[\alpha_1(1-\alpha_2) + \alpha_2(1-\alpha_1)\right] + 2d(1-\alpha_1)(1-\alpha_2) \\ &= 2a\alpha_1\alpha_2 + 2d(1-\alpha_1)(1-\alpha_2) \\ dW &= (b+c-2d)(d\alpha_1 + d\alpha_2) + 2E(\alpha_2 d\alpha_1 + \alpha_1 d\alpha_2) \\ &= -2d(d\alpha_1 + d\alpha_2) + 2(a+d)(\alpha_2 d\alpha_1 + \alpha_1 d\alpha_2) \\ d^2W &= 2E\ d\alpha_1\ d\alpha_2 \\ &= 2(a+d)\ d\alpha_1\ d\alpha_2 \end{aligned}$$

From the expression of d^2W, one can see that a maximum of W is never interior. On the borders of the square, W is an affine function, hence a maximum of W can only be at a corner of the square.

We are firstly interested in the *global Bentham optima*; where W is maximized. They are given by the the largest of three values:

$$\begin{aligned} 2d, \text{ obtained for } \alpha_1 &= \alpha_2 = 0, \\ 0, \text{ obtained for } \alpha_1 &= 0,\ \alpha_2 = 1, \text{ or } \alpha_1 = 1,\ \alpha_2 = 0, \\ 2a, \text{ obtained for } \alpha_1 &= \alpha_2 = 1, \end{aligned}$$

We secondly introduce the notion of a *local Bentham optimum.* This is a local maximum of the function W.; again they can onlybe found at corners. They are given by the following conditions:

- $(\alpha_1, \alpha_2) = (0, 0)$ is a local Bentham optimum iff $d > 0$,
- $(\alpha_1, \alpha_2) = (0, 1)$ and $(0, 1)$ are local Bentham optima iff $a < 0$ and $d < 0$,
- $(\alpha_1, \alpha_2) = (1, 1)$ is a local Bentham optimum iff $a > 0$.

According to the payoffs, five cases can be distinguished:

If $0 < a < d$, $(0, 0)$ is a global Bentham optimum and $(1, 1)$ is a local one.

If. $0 < d < a$, $(1, 1)$ is a global Bentham optimum and $(0, 0)$ is a local one.

If $a < 0 < d$. $(0, 0)$ is a global Bentham optimum.

If $d < 0 < a$, $(1, 1)$ is a global Bentham optimum.

If $a, d < 0$, $(0, 1)$ and $(1, 0)$ are both global Bentham optima.

In the same spirit, define a *diagonal Bentham optimum* to be a (global or local) maximum of W on a line parallel to the main diagonal. Such a line L has equation

$$\alpha_1 - \alpha_2 = r.$$

Denote $\widetilde{M}$ the point on the main diagonal

$$\widetilde{M} = (\widetilde{\alpha_1}, \widetilde{\alpha_2}) = (\frac{d}{a+d}, \frac{d}{a+d}).$$

Consider the line L_0 of equation $\alpha_1 + \alpha_2 = \alpha_{10} + \alpha_{20}$. Graphically, L_0 is the line parallel to the second diagonal and passing through M_0. Let M be the intersection of L_0 and L. The coordinates (α_1, α_2) of M are such that

$$\begin{aligned} \alpha_1 + \alpha_2 &= \alpha_{10} + \alpha_{20} \\ \alpha_1 - \alpha_2 &= r. \end{aligned}$$

The bilinear function W of α_1 and α_2 is easy to maximize on L, and one obtains the following conclusions:

- For decoupling games ($E < 0$), on the line L, W has its maximum at M. Note that M may be outside the strategy space, in which case the diagonal maximum is on a border of the strategy space.
- For coupling games ($E > 0$), on the line L, W has its minimum at M. Hence a diagonal local Bentham optimum is always on a border of the strategy space.

The following Table summarizes the findings about Nash, Global and Local Bentham optima. To read this Table: NE stands for Nash Equilibrium, GO stands for Global Bentham optimum, and LO stands for a local Bentham optimum which is not global. The Table provides the coordinates of the point in the action space. For Bentham optima, the indication "$(0,0)$ or $(1,1)$" means that the optimum is $(0,0)$ if $d > a$ and $(1,1)$ if $d < a$.

	$a < -1$	$-1 < a < 0$	$0 < a$
$1 < d$	**A** NE: (0,0) GO: (0,0) LO: none	**B** NE: (0,0) and (1,1) GO: (0,0) LO: none	**C** NE: (0,0) and (1,1) GO: (0,0) or (1,1) LO: (1,1) or (0,0)
$0 < d < 1$	**D** NE: (0,1) and (1,0) GO: (0,0) LO: none	**E** NE: (1,1) GO: (0,0) LO: none	**F** NE: (1,1) GO: (0,0) or (1,1) LO: (1,1) or (0,0)
$d < 0$	**G** NE: (0,1) and (1,0) GO: (0,1) and (1,0) LO: none	**H** NE: (1,1) GO: (0,1) and (1,0) LO: none	**I** NE: (1,1) GO: (1,1) LO: none

4.2 State transition diagram

For a symmetric game, the strategy space is a square centered on the main diagonal, and any such square is the strategy space of some symmetric game. Hence, in this section, we study the state transition diagram without taking into account the border lines

There are only three separating lines, two of them are parallell, and they define six regions in the (α_1, α_2) plane separated by the lines of equations $U_1^+ =$

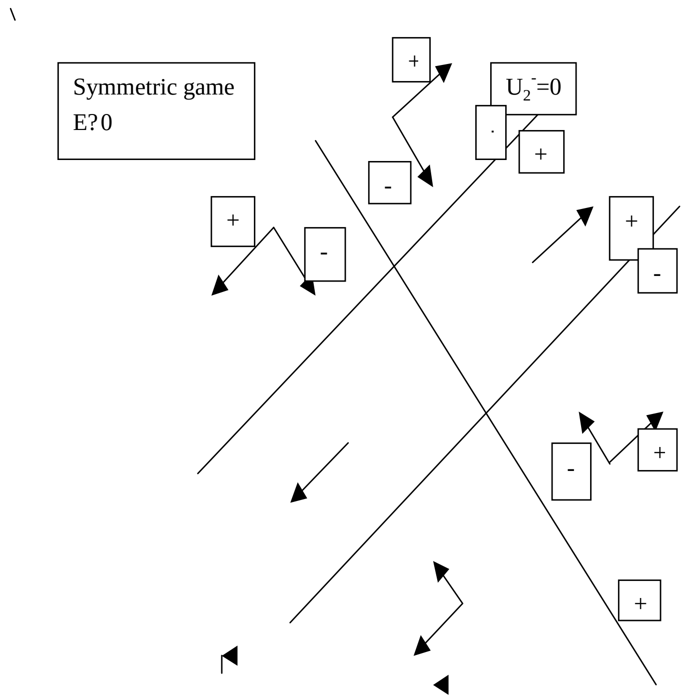


Figure 1: Symmetric coupling game ($E > 0$).

$U_2^+ = 0$, $U_1^- = 0$ and $U_2^- = 0$, with:

$$\begin{aligned}
U_1^+ &= U_2^+ = b + c - 2d + E(\alpha_1 + \alpha_2) \\
&= -2d + (a+d)(\alpha_1 + \alpha_2) \\
U_1^- &= b - c + E(\alpha_2 - \alpha_1) \\
&= 2 + (a+d)(\alpha_2 - \alpha_1) \\
U_2^- &= b - c - E(\alpha_2 - \alpha_1) \\
&= 2 - (a+d)(\alpha_2 - \alpha_1)
\end{aligned}$$

Hence, the relative positions of these lines depend only on the sign of E. The two corresponding diagrams are depicted in Figures 1 and 2.

By looking at the Figures 1 and 2, one can see the possible trajectories of the system (as long as it does not reach a border of the action space). Inside the zones defined by the separating lines, the system follows the arrows; (as we discussed above, there are two possible direction of movement in the zones where two arrows are depicted). We now discuss the trajectory of the system when reaching a separating line.

If $E > 0$, this happens (up to some symmetry) in the upper part of Figure

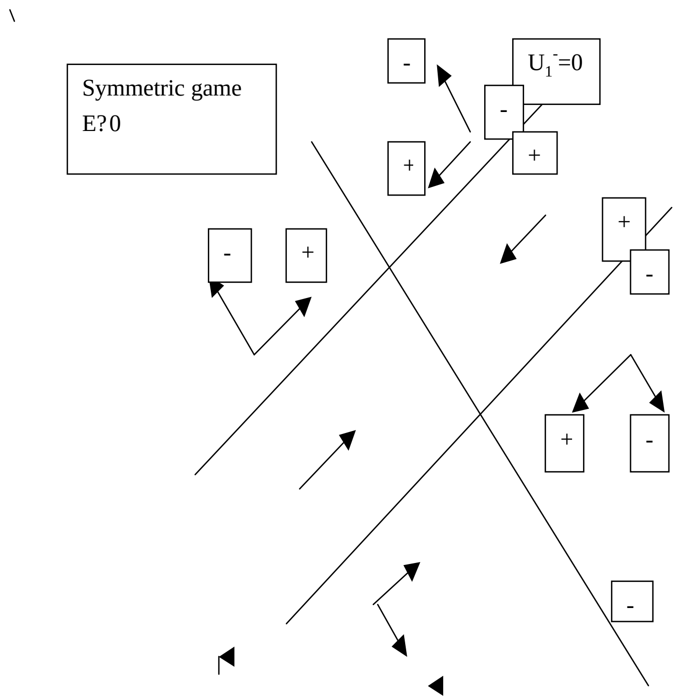


Figure 2: Symmetric decoupling game ($E < 0$).

1 when the system is moving South-East above the separating line $U_2^- = 0$, and would be moving North East on the other side of this line. Then, according to the case (1) of the paragraph Evolution on separating lines, the system initially moving South East then slides North East along the separating line.

If $E < 0$, this happens in two cases. Firstly, in the upper part of Figure 2 when the system is moving South West above the separating line $U_1^+ = U_2^+ = 0$ and is reaching a zone where it may be going either North-East or North-West. But it can be easily shown that, in a symmetric game, when the system is on a trajectory parallel to the first diagonal, the utility variations of the two players are the same, and therefore the system remains on that diagonal. It follows that in this subcase, the system is stuck (in a cycle) on the separating line.

Secondly, in the central part of Figure 2 when the system is moving South West above the separating line $U_1^+ = U_2^+ = 0$ and would be moving North East on the other side of this line. Then, according to the case (5) of the paragraph Evolution on separating lines, the system initially moving South West then is stuck (in a cycle) on the separating line.

4.3 Properties of the trajectories

The location of the state space in the plane (α_1, α_2) was neglected in the previous paragraph, we now take it into account in order to exhibit the convergence properties of the process.

It can be observed that the point M_0 is at the intersection of the two lines of equation $U_1^+ = U_2^+ = 0$ (that is L_0) and $\alpha_1 = \alpha_2$ (that is the main diagonal).

A careful inspection shows that M_0 belongs to the strategy space iff $a > 0$ and $d > 0$ (in the case $E \geq 0$) or $a < 0$ and $d < 0$ (case $E \leq 0$). The strategy space lies entirely below M_0 if and only iff $a < 0$ (when $E > 0$) and $a > 0$ (when $E < 0$). The strategy space lies entirely above M_0 if and only if $d < 0$ (when $E > 0$) and $d > 0$ (when $E < 0$).

The strategy space lies entirely between the lines L_1^- and L_2^- iff $0 \leq a+d \leq 2$ (case $E \geq 0$) or $-2 \leq a+d \leq 0$ (case $E \leq 0$), which simply boils to $-2 \leq a+d \leq 2$

Finally, for the first player, notice that the line $U_1^+ - U_1^- = 0$ is vertical, the line $U_1^+ + U_1^- = 0$ is horizontal, and they cross at the intersection of the lines $U_1^+ = 0$ and $U_1^- = 0$, that is the point N_1:

$$N_1 = \left(\frac{d+1}{a+d}, \frac{d-1}{a+d} \right).$$

Likewise, for the second player, the line $U_2^+ - U_2^- = 0$ is horizontal, the line $U_2^+ + U_2^- = 0$ is vertical, and they intersect at the point N_2:

$$N_2 = \left(\frac{d-1}{a+d}, \frac{d+1}{a+d} \right).$$

When reaching a border line, the system can either stop or slide along the border line (see section “Evolution on border lines”). The system is stuck in cases A4 and A5.

In case A4, the stated conditions for the system to stop on the border $\alpha_1 = 0$ become, for a symmetric game:

$$-1 < d < 1.$$

By symmetry, the condition is the same for the border line $\alpha_2 = 0$. Note that the system is never stuck on borders $\alpha_1 = 1$ or $\alpha_2 = 1$ since the corresponding conditions would be $b = 1 < a < c = -1$, which was excluded when we set the condition $b > c$.

In case A5, the conditions are never met for a symmetric game.

In the next two sections, depending on the sign of E, we use the previous results in order to describe the trajectory of the system for any initial state. The behavior of the system when it slides along a border and meets a separating line was not examined, but is easily considered when happening.

4.4 Convergence results

When considering the whole behavior of the system, the following result holds:

Theorem. *If there exists a unique local Bentham optimum, then the system points towards it. If this point is also a Nash equilibrium then the system converges to it. If not, the system is stuck at the point where it first reached the boundary of the action space.*

If there exist two local Bentham optima, then the system points towards one of them, depending on the initial point. If the local optimum is also a Nash equilibrium, it converges to it. If not, the system is stuck at the point where it first reached the boundary of the action space.

If there exist diagonal Bentham optima, the system points towards one of them, depending on the initial state. If it reaches a border in the direction of a Nash equilibrium, then it converges towards the diagonal Bentham optimum on that border. If it reaches a border in a direction which does not point towards a Nash equilibrium, then it is stuck on the border.

Note that, crudely expressed, this results mean that the system behavior is driven by the notion of Bentham optimality inside the action space and by the notion of Nash equilibrium on the border of the action space.

Proof:

We need to distinguish the nine cases, A to I. Within each case, the subcases to be distinguished correspond to the sign of E (that is $a + d$ being positive or negative), the position of M_0 with regard to the action space (a and d being positive or negative) and the position of the action space with respect to the separating lines (here two conditions are involved: $a + d$ being larger or smaller than -2, 0 and 2). Note that in some of these cases, the sign of E is given, and in some others, it is not.

Case A: $a < -1$ and $d > 1$.

Subcase A1: $-2 < a + d < 0$ (hence $E < 0$). The action space is entirely above M_0 and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points South-West untill it reaches a border, it then slides along the border until it reaches the corner $(0, 0)$.

Subcase A2: $0 < a+d < 2$ (hence $E > 0$). The action space is entirely below M_0 and between the separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory points South-West untill it reaches a border, it then slides along the border until it reaches the corner $(0,0)$.

Subcase A3: $2 < a+d$ (hence $E > 0$). The action space is entirely below M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory depends on the initial point. If the initial point is between the separating lines L_1^- and L_2^-, the situartion is similar to subcase A2. If the initial point is above L_1^-, the trajectory depends on the inital direction. If the initial direction is South-East, then the system goes South-East until it reaches the separating line L_1^-, then slides along L_1^- until it reaches the border $\alpha_1 = 0$, then slides along this border until reaching the corner $(0,0)$. If the initial direction is South-West, the system goes South-West until it reaches the border $\alpha_1 = 0$, then slides along this border until reaching the corner $(0,0)$.

Subcase A4: $a+d < -2$ (hence $E < 0$). The action space is entirely above M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory depends on the initial point. If the initial point is between the separating lines L_1^- and L_2^-, the situation is similar to subcase A1. If the initial point is above L_1^-, the trajectory depends on the inital direction. If the initial direction is South-West, the system goes South-West until it reaches the border $\alpha_1 = 0$, then slides down along this border, crosses the separating line and reaches the corner $(0,0)$ where it stops. If the initial direction is North-West, then the system goes North-West until it reaches a border, either $a_1 = 0$, or $\alpha_2 = 1$. In both cases, it slides along the border in untill it reaches the corner $(0,1)$. From this corner, it slides down along the border $a_1 = 0$, crosses the separating line and reaches the corner $(0,0)$ where it is stops. (Remark that this is one case in which the system has two opposite trajectories: sliding up or down on a border).

To sum up, in case A, the system ultimately reaches the corner $(0,0)$ which is here the unique Nash Equilibrium and the unique global Bentham optimum

Case B: $-1 < a < 0$ and $d > 1$.

Subcase B1 $a+d < 2$: Identical to A2.

Subcase B2 $a+d > 2$: Identical to A3.

To sum up, in case B, the system ultimately reaches the corner $(0,0)$ which is here one of the two Nash Equilibria and the unique global Bentham optimum

Case C: $0 < a$ and $d > 1$. (Hence $E > 0$.)

Subcase C1: $a+d < 2$. The action space, includes M_0 and lies between the separating lines L_1^- and L_2^-. According to Figure 1, if the initial point is below L^+, the trajectory goes South-West until it reaches a border then slides along the boder until it rreaches the point $(0,0)$; if the intial point is above L_2, the trajectory goes North-East until it reaches a border then slides along the border until it rreaches the point $(1,1)$.

Subcase C2: $a+d > 2$. The action space, includes M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory depends on the initial point. If the initial point is between the separating lines L_1^- and L_2^-, the situartion is similar to the previous subcase C1. If the initial point is above L_1^-, the trajectory also depends on the inital direction. (i) If the initial point is below L^+ and the initial direction is South-East, then the system goes South-East until it reaches the separating line L_1^-, then slides along L_1^- until it reaches the border $\alpha_1 = 0$, then slides along this border until reaching the corner $(0, 0)$. (ii) If the initial point is below L^+ and the initial direction is South-West, the system goes South-West until it reaches the border $\alpha_1 = 0$, then slides along this border until reaching the corner $(0, 0)$. (iii) If the initial point is above L^+ and the initial direction is South-West, then the system goes South-West until it reaches the separating line L_1^-, then slides along L_1^- until it reaches the border $\alpha_2 = 1$, then slides along this border until reaching the corner $(1, 1)$. (iv) If the initial point is above L^+ and the initial direction is North-West, the system goes North-West until it reaches the border $\alpha_2 = 1$, then slides along this border until reaching the corner $(1, 1)$.

To sum up, in case C, if the initial point is below the line L^+, the system goes towards the corner $(0, 0)$ which is a (local or global) Bentham optimum and a Nash equilibrium; if he initial point is above the line L^+, the system goes towards the corner $(1, 1)$ which is a (local or global) Bentham optimum and a Nash equilibrium.

Case D: $a < -1$ and $0 < d < 1$ (hence $E < 0$).

Subcase D1: $-2 < a + d < 0$ The action space is entirely above M_0 and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points South-West untill it reaches a border, where it is scotched.

Subcase D2: $a+d < -2$ The action space is entirely above M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory depends on the initial point. If the initial point is between the separating lines L_1^- and L_2^-, the situation is similar to subcase D1: the system is scotched on the border. If the initial point is above L_1^-, the trajectory depends on the inital situation and direction. If the initial direction is South-West, the system goes South-West until it reaches the border $\alpha_1 = 0$, where it is scotched. If the initial direction is North-West, then the system goes North-West until it reaches a border $\alpha_1 = 0$ or $\alpha_2 = 1$. If it reaches the border $\alpha_1 = 0$, it is scotched. If it reaches the border $\alpha_2 = 1$ then it slides along this border until it reaches the corner $(0, 1)$, where it is scotched.

To sum up, in case D, the system ends up being scotched on one of the borders adjacent to the global optimum $(0, 0)$, either directly or indirectly after sliding along another border.

Case E: $-1 < a < 0$ and $0 < d < 1$.

Subcase E1: $a + d > 0$ (hence $E > 0$) The action space is entirely below M_0 and between the separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory points South-West untill it reaches a border, where it is scotched.

Subcase E2: $a + d < 0$ (hence $E < 0$) The action space is entirely above M_0 and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points South-West untill it reaches a border, where it is scotched.

To sum up, in case E, the system goes South-West and is ultimately scotched on a border.

Case F: $0 < a$ and $0 < d < 1$ (hence $E > 0$)

Subcase F1: $a + d < 2$ The action space, includes M_0 and lies between the separating lines L_1^- and L_2^-. According to Figure 1, if the initial point is below L^+, the trajectory goes South-West until it reaches a border where it is scotched; if the intial point is above L_2, the trajectory goes North-East until it reaches a border then slides along the border until it rreaches the point $(1, 1)$.

Subcase F2: $a + d > 2$ The action space includes M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory depends on the initial point. If the initial point is between the separating lines L_1^- and L_2^-, the situartion is similar to the previous subcase F1. If the initial point is above L_1^-, the trajectory also depends on the inital direction. (i) If the initial point is below L^+ and the initial direction is South-East, then the system goes South-East until it reaches the separating line L_1^-, then slides along L_1^- until it reaches the border $\alpha_1 = 0$, where it is stuck. (ii) If the initial point is below L^+ and the initial direction is South-West, the system goes South-West until it reaches the border $\alpha_1 = 0$, where it is stuck. (iii) If the initial point is above L^+ and the initial direction is South-West, then the system goes South-West until it reaches the separating line L_1^-, then slides along L_1^- until it reaches the border $\alpha_2 = 1$, then slides along this border until reaching the corner $(1, 1)$. (iv) If the initial point is above L^+ and the initial direction is North-West, the system goes North-West until it reaches the border $\alpha_2 = 1$, then slides along this border until reaching the corner $(1, 1)$.

To sum up, in case F, if the initial point is below the line L^+, the system goes in the direction of the corner $(0, 0)$ which is a (local or global) Bentham optimum but not a Nash equilibrium, and the system is stuck before reaching this corner; if the initial point is above the line L^+, the system goes towards the corner $(1, 1)$ which is a (local or global) Bentham optimum and a Nash equilibrium.

Case G: $a < -1$ and $d < 0$ (hence $E < 0$)

Subcase G1: $a + d > -2$, The action space, includes M_0 and lies between the separating lines L_1^- and L_2^-. According to Figure 2, if the initial point is below L^+, the trajectory goes North-East until it reaches the separating line L^+ where it stops. If the initial point is above L^+, the trajectory goes South-West and it either reaches first the border $\alpha_1 = 0$ where it is stuck, or it reaches first the separating line L^+, where it stops.

Subcase G2: $a + d < -2$ The action space, includes M_0 and intersects both separating lines L_1^- and L_2^-. According to Figure 2, if the initial point is between L_1^- and L_2^-, then the trajectory is as in subcase G1. If the initial point is above L_1^-, the trajectory also depends on the inital direction. (i) If the initial point is below L^+ and the initial direction is North-East, then the system goes North-East until it reaches the border $\alpha_2 = 1$ then slides East along this border until it reaches the separating line L^+ where it is stuck. (ii) If the initial point is below L^+ and the initial direction is North-West, the system goes North-West until it reaches the border $\alpha_1 = 0$, where it slides North**Il+ va prendre des virages** (iii) , (iv) .

Case H: $-1 < a < 0$ and $d < 0$. (hence $E < 0$)

Subcase H1: $a + d > -2$. This subcase is identical to G1

Subcase H2: $a + d < -2$. This subcase is identical to G2

Case I: $0 < a$ and $d < 0$.

Subcase I1: $0 < a + d < 2$ (hence $E > 0$) The action space is entirely above M_0 and between the separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory points North-East untill it reaches a border, slides along that border and reaches the corner $(1, 1)$.

Subcase I2: $-2 < a+d < 0$ (hence $E < 0$) The action space is entirely below M_0 and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points North-East untill it reaches a border, slides along that border and reaches the corner $(1, 1)$.

Subcase I3: $a + d > 2$ (hence $E > 0$) The action space is entirely above M_0 and intersects the separating lines L_1^- and L_2^-. According to Figure 1, if the initial point is between the separating lines L_1^- and L_2^-, the situartion is similar to the previous subcase I1. If the initial direction is South-East, the trajectory first reaches the separating line L_1^- then slides along that seaparating line until it reaches the border $\alpha_2 = 1$ then slides along this border until it reaches the corner $(1, 1)$. If the intitial direction is North-East, the system reaches the border $\alpha_2 = 1$ then slides along that border and reaches the corner $(1, 1)$.

Subcase I4: $a+d < -2$ (hence $E < 0$) The action space is entirely below M_0 and intersects the separating lines L_1^- and L_2^-. According to Figure 2, if the initial point is between the separating lines L_1^- and L_2^-, the situation is similar to the previous subcase I2. If the initial direction is North-West, the trajectory first reaches the border $\alpha_1 = 0$ and then slides North, reaches the corner $(0, 1)$ and then ?. If the intitial direction is North-East, the system reaches the border $\alpha_2 = 1$ then slides along that border and reaches the corner $(1, 1)$.

To sum up, the system always goes to the corner $(1, 1)$ which is a global optimum and the unique Nash equilibrium.

Quod Erat Demonstrandum

5 Zero-sum games

5.1 Potential attractors

In a zero-sum game: $a_1 = -a_2 = a$, $b_1 = -b_2 = b$, $c_1 = -c_2 = c$, $d_1 = -d_2 = d$. Hence $E = E_1 = -E_2$. Without restriction (exchanging players), we can suppose that $E > 0$

As concerns the pure Nash equilibria, two cases are possible:

- one pure equilibrium, which can be at any corner.

- no pure equilibrium when $a \succ c, a \succ b, d \succ b, d \succ c$ (case $E > 0$) or when $a \prec c, a \prec b, d \prec b, d \prec c$ (case $E < 0$, that we rule out).

Consider the point

$$M_0 = (\alpha_{10}, \alpha_{20}) = \left(\frac{d-c}{E}, \frac{d-b}{E}\right).$$

The point M_0 can be located anywhere in the plane (α_1, α_2) even with the constraint $E > 0$. Two relevant cases are possible. If M_0 belongs to the action space, then M_0 corresponds to the unique mixed equilibrium. If M_0 does not belong to the action space, then the unique equilibrium is at a corner (and is pure).

Proposition. The equilibrium is then given by the following rule:
- (0,0) if $\alpha_{10} \prec 0, \alpha_{20} \succ 0$
- (0,1) if $\alpha_{10} \succ 0, \alpha_{20} \succ 1$
- (1,1) if $\alpha_{10} \succ 1, \alpha_{20} \prec 1$
- (1,0) if $\alpha_{10} \prec 1, \alpha_{20} \prec 0$

The Bentham optima are obviously degenerated.

5.2 State transition diagram

There are only two separating lines defining 4 areas:
$U_1^+ = -U_2^+ = b + c - 2d + E(\alpha_1 + \alpha_2) = 0$
$U_1^- = U_2^- = b - c + E(\alpha_2 - \alpha_1) = 0$

They intersect at the point M_0.

The phase diagram is the following (note that at interior points the system is always moving crabwise):

Consider for instance matching pennies, obtained for $a = d = 1, b = c = -1$.The separating lines are respectively: $\alpha_1 + \alpha_2 - 1 = 0$ and $\alpha_2 - \alpha_1 = 0$ crossing at $\alpha_1 = \alpha_2 = 1/2$.The action space is centered around the same point.

5.3 Properties of the trajectories

We first examine the behavior of the system when reaching a separating line. Without loss of generality, assume that the system is coming crabwise from the East and intersects the separating line parallel to the first diagonal. According to Figure **??**, this intersection is South-West of M_0. According to the paragraph Evolution on Separating Lines, case 2, the system turns right and continues crabwise to the North . The important point is that the trajectory after its turn on the separating line has made a small step away from M_0, will be stated in the next proosition. For that we need a definition. For a crabwise trajectoray the associated mean line is the line joining the middle of its constituting segments.

Proposition. When crossing a separating line, the trajectory of the system is such that its mean line is further away from point M_0 after the crossing than before.

Proof. Consider first the case in which the system arrives axactly on the separating line. (Such is the case if the payoffs are integers and N is a multiple of E.) According the Figure "Virage à droite", coming from A then B, the system reaches the separating line in C. Since the segment BC is entirely in the

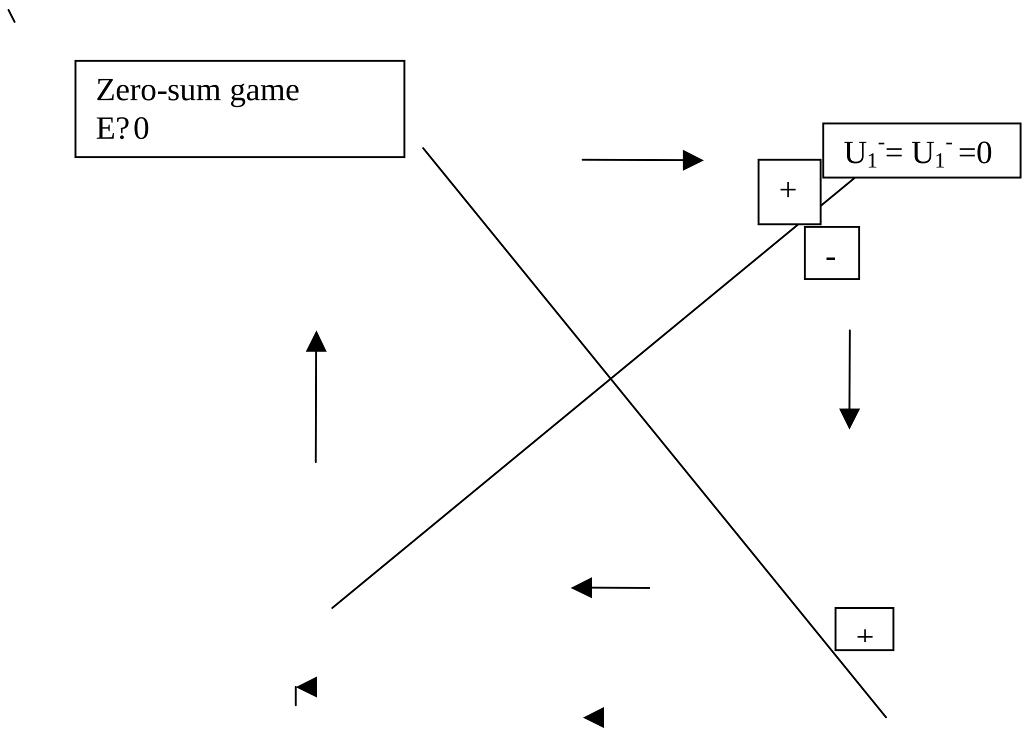


Figure 3: Zero-sum game, $E > 0$

initial zone, the system goes from C to D, exactly on the separating line. after D, it goes to E, for the following reason:

The utility variations from A to B is $(+,-)$ (positive for the first player and negative for the second). The utility variation from E to F is $(-,+)$ because E and F are on the other side of the separating line, hence the utility variation from F to E is $(+,-)$. By continuity, the utility variation from C to D is $(+,-)$ too. Hence the system goes from D to EIt then continues from E to F. Notice that the mean line is further away from M_0, by $1/N$.

Consider now the the case in which the systems does not arrive exactly on the separating line. According to the figure "Virage à droire bis", coming from A and B, the system goes to C and crosses the separating line between B and C. The utility variation from B to C is not straightforward and has to be computed. Let $B = (\alpha_1 - 1/N, \alpha_2)$. Then $C = (\alpha_1, \alpha_2 - 1/N)$. The utility variation (for player 1) $U(C) - U(B)$ is:

$$(1/N)(\alpha_1 - \alpha_2)(-a + b + c - d) + (1/N)(-b + c)$$

This variation is thus zero on the line of equation

$$\alpha_1 - \alpha_2 = \frac{b - c}{E}$$

This is precisely the separating line we consider. Hence, if the middle of BC is under the separating line, the system turns left after C and goes to D,then turns right to E, being completely above the separating line. If the middle of BC is above the separating line, the system turns right after C and goes to D', then turns left to E'.But in both cases, the mean line is further away from the point M_0 after having crossed the separating line than before.

QED

We now examine the behavior of the system when reaching a border. Without loss of generality, assume again that the system is coming crabwise from the East and is reaching the border $\alpha_1 = 0$. According to the general analysis, the system goes North along the border (up to ε) if $d \succ c$ and goes South along the border if $d \prec c$. Similar conditions hold for $\alpha_2 = 0$ (going West if $d \succ b$ and going East if $d \prec b$), $\alpha_1 = 1$ (going South if $b \succ a$ and going North if $b \prec a$), and $\alpha_2 = 1$ (going East if $c < a$ and going West if $c > a$)

5.4 Convergence result

The following result can be stated:

Theorem: (i) *When the game has a pure Nash equilibrium, the system converges towards it. (ii) When the system has no pure Nash equilibrium, the system asymptotically cycles around the greatest square situated in the strategy space and centered on the mixed Nash equilibrium.*

Proof. Recall that $M_0 = (\alpha_{10}, \alpha_{20})$ is the intersection of the two separating lines.

(i) Here M_0 is outside the action space. For this part of the proof we suppose w.l.o.g. that $\alpha_{10} > 0$ and $\alpha_{20} > 1$. Hence $d > c$ and $a < c$; the equilibrium is at the point $(0, 1)$.

Four cases will be distinguished.

Case A: the action space does not intersect any separating line. For instance, the action space lies entirely in the quarter of plane South of M_0.

From any interior initial point, the system goes crabwise West until it reaches the border $\alpha_1 = 0$. According to the theorem "Comportement au bord", the system goes in the direction of the best response for player 2. This best response is given by the sign of $d - c$. Because $d > c$ the system goes North. Thus the system goes towards the pure equilibrium.

Then, when the system is sliding along the border towards the equilibrium, it reaches the equilibrium To prove that the system is stuck around the equilibrium, first notice that the rosace is fully determined:

$$\begin{bmatrix} +- & +- & +- \\ +- & & -+ \\ -+ & -+ & -+ \end{bmatrix}$$

and the same reasoning as usual shows that any trajectory ends in a cycle around the corner.

Case B: The action space intersects only the separating line parallel to the first diagonal. If the initial point is below the separating line, the system goes West until it reaches either the border $\alpha_1 = 0$, or the separating line. In the first case, it then slides North along the border, crosses the separating line, continues North untill it reaches the equilibrium $(0, 1)$. In the second case, the system reaches the separating line, goes North until it reaches the border $\alpha_2 = 1$, then goes West until it reaches the equilibrium $(0, 1)$. If the initial point is above the separating line, the system first goes North untill it reaches the border $\alpha_2 = 1$, then goes West until it reaches the equilibrium $(0, 1)$.

Case C: The action space intersects only the separating line parallel to the second diagonal. If the initial point is below the separating line, the system goes West until it reaches the border $\alpha_1 = 0$, then slides North along the border untill it reaches the equilibrium $(0, 1)$. If the initial point is above the separating line, the system first goes South untill it reaches the separating line, then goes West until it reaches the border line, then goes North until it reaches the equilibrium $(0, 1)$.

Case D: the action space intersects both separating lines.It is just a superposition of the two former cases

(ii) Here M_0 is inside the action space. It is the mixed-straegy equilibrium of the game. By symmetry we can consider only the case where:

$$0 < \alpha_{10} < \alpha_{20} < 1/2.$$

These conditions imply that $b < c < d$. Then the largest square centered on M_0 and included in the action space will be denoted by S. It has summits:

$$(0, \alpha_{20} - \alpha_{10}), (2\alpha_{10}, \alpha_{20} - \alpha_{10}), (2\alpha_{10}, \alpha_{20} + \alpha_{10}), (0, \alpha_{20} + \alpha_{10}).$$

Two cases have to be considered according to the initial point.

Case A: If the initial point is outside S but not on a border then, when the system reaches a separating line, it turns Right (if $E > 0$, which we now suppose). Moreover, according to proposition ??, the system is one step further away from M_0 after this turn. After zero, one or two such right-turn, the system reaches a border. If the border is the border $\alpha_{20} = 0$, the system goes West, according to section describing the behavior at a border. If the border is the border $\alpha_{10} = 0$, the system goes North. In all cases, it turns right. Moreover, if it reaches again a separating line while moving on a border (this is possible if and only if the system slides along the border $\alpha_{20} = 0$), the system continues straight on the border.

Therefore the system asyptotically cycles in a 3ε-neighborhood of S.

Case B: If the initial point is inside S, then the system turns always right each time it reaches a separating line. Since the system is one step further away from M_0 after each turn, this holds until the system reaches the square S. In fact, it goes even outside the square till reaching a border. But, this was already considered in case A.

6 Twin games

6.1 Potential attractors

In a twin game: $a_1 = a_2 = a$, $b_1 = b_2 = b$, $c_1 = c_2 = c$, $d_1 = d_2 = d$. Hence $E_1 = E_2 = E$.Without loss of generality, it can be stated that $E \succ 0$.

The game has:

-two pure Nash equilibria if $a, d \succ b, c$.

-one pure Nash equilibrium otherwise

Consider the point

$$M_0 = (\alpha_{10}, \alpha_{20}) = \left(\frac{d-c}{E}, \frac{d-b}{E} \right).$$

The point M_0 can be located anywhere in the plane (α_1, α_2) even with the constraint $E > 0$. It is situated inside the strategy space when there are two pure Nash equilibria and represents a mixed Nash equilibrium. It is situated out of the strategy space when there is only one pure Nash equilibrium.Precisely, the unique pure equilibrium is:

$(0, 0)$ if $\alpha_{10} > 0$ and $\alpha_{20} > 0$, and one of them is > 1,

$(0, 1)$ if $\alpha_{10} < 0$ and $\alpha_{20} > 1$,

$(1, 1)$ if $\alpha_{10} < 1$ and $\alpha_{20} < 1$, and one of them is < 0,

$(1, 0)$ if $\alpha_{10} > 1$ and $\alpha_{20} < 0$.

As concern the Bentham values, they coincide with each'player's payoff. Hence, there is a global optimum at one corner (the unique Nash one or the the Pareto-dominating in case of two) and local Bentham optimum at the other ones.

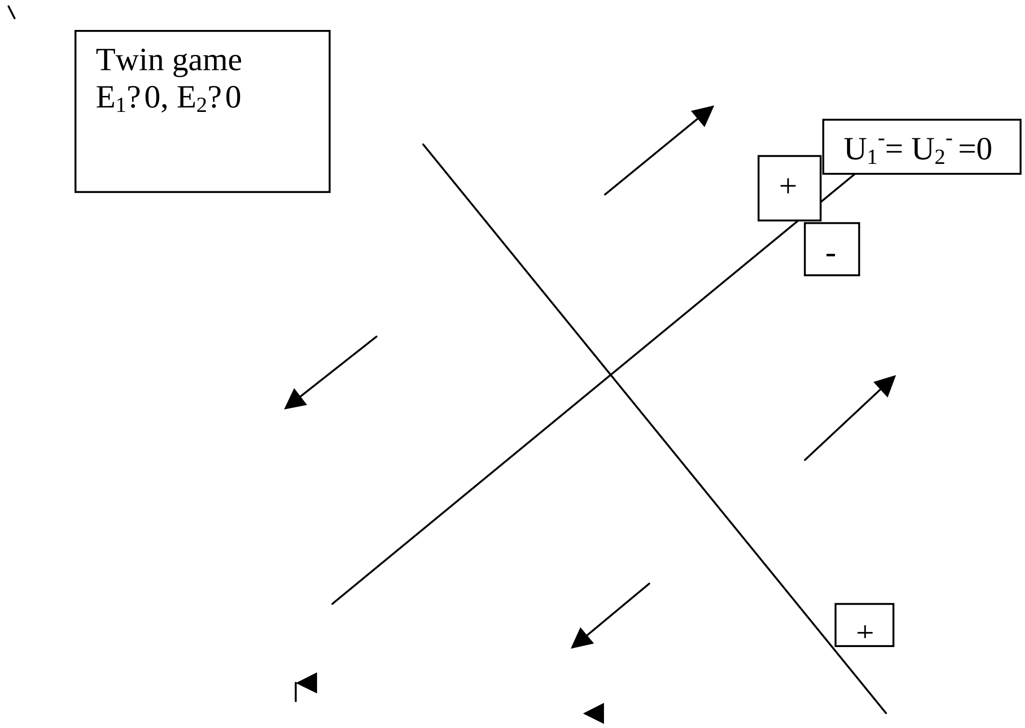


Figure 4: Twin game $E > 0$

6.2 State transition diagram

There are only two separating lines defining 4 areas:

$U_1^+ = U_2^+ = b + c - 2d + E(\alpha_1 + \alpha_2) = 0$
$U_1^- = U_2^- = b - c + E(\alpha_2 - \alpha_1) = 0$
which croos at M_0.

The state transition diagram is the following:

6.3 Trajectory properties

As can be seen, the system never reaches a separating line, except maybe on borders.

When coming on a border line, the system is never stuck. When crossing a separating line on a border, the system

6.4 Convergence results

We can state the following theorem:

Theorem: (i) *If there is a unique pure Nash equilibrium the system converges towards it. (ii) If there are two pure Nash equilibria, the system converges towards one of them, depending on the initial state.*

Proof.

(i) The most general case is to have the two separating lines intersecting the state space and deviding this space into 3 areas. Without loss of generality we suppose that M_0 is inside the triangle at the left of the state space, defined by the three conditions:

$a_{10} < 0$

$a_{10} + a_{20} > 0$

$a_{20} - a_{10} < 1$

See Figure [TwinPureTroisZones]

In that case the unique Nash equilibrium is at the corner $(1, 1)$.

If the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 < \alpha_{10} + \alpha_{20}$ and $\alpha_1 > \alpha_2$. Then the system goes South-West until it reaches the border $\alpha_2 = 0$. According to Proposition [Reaching a Border], it follows the border towards the West (according to player 1's best response) until reaching the corner $(0, 0)$. Then, according tp proposition [Reaching a corner] the system follows the border $\alpha_1 = 0$, crosses the two separating lines and reaches the corner $(0, 1)$. Here it turns right ,follows the border $\alpha_2 = 1$, crosses a separating line and reaches the pure Nash equilibrium $(1, 1)$, where it is blocked.

If the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 < \alpha_{10} + \alpha_{20}$ but $\alpha_1 < \alpha_2$. Then the system goes South West until it reaches the border $\alpha_1 = 0$. Then proposition [Reaching a Border] still applies: the system makes a $3\pi/4$ right turn and continues North like in the previous case.

If the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 > \alpha_{10} + \alpha_{20}$ Then system moves North-East until reaching one the two borders $\alpha_1 = 1$ or $\alpha_2 = 1$, follows that border until reachning the Nash equilibrium $(1, 1)$.

The other cases for M_0 outside the state space are in fact sub-cases of the previous ones and symmetric cases in which the pure Nash equilibrium is another corner..

(ii) When M_0 is inside the state space, there are two pure Nash equilibrium, located (since E is supposed to be > 0) at $(0, 0)$ and $(1, 1)$.

If the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 < \alpha_{10} + \alpha_{20}$ the system goes South-West, reaches a border $\alpha_1 = 0$ or $\alpha_2 = 0$, then follows the boredr towards the pure Nash equilibrium $(0, 0)$, where it is blocked. If the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 > \alpha_{10} + \alpha_{20}$ the system likewise goes to a border, then to the pure Nash equlibrium $(1, 1)$.

7 References

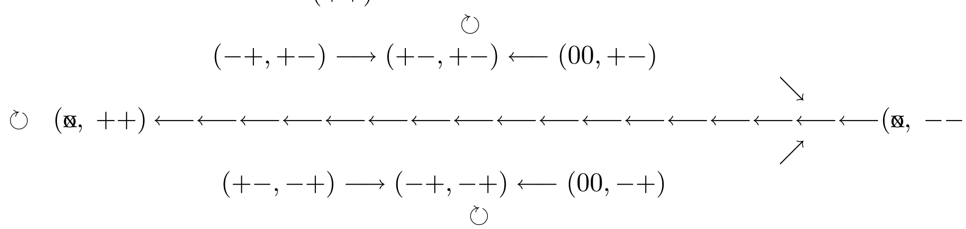
Huck, S.- Norman, H.T.- Oechssler,J. (2004): GLAD: a simple adaptive strategy that yields cooperation in dilemma games, *Physica D*.

Huck, S.- Norman, H.T.- Oechssler, J. (2005): Through trial and error to collusion, *International Economic Review*, 45(1)

Sutton, R.S. - A. G. Barto (1998) *Reinforcement learning* MIT Press.

Tregouet, T. (2004): Le duopole, l'apprentissage de la concurrence, document ENSAE.

These principles can be illustrated in the following 'utility transition diagram' which is read as follows. The symbol $(+-)$ means that utility of player 1 is increasing and utility of player 2 is decreasing. The symbol $(+-,\ -+)$ stands for the state made of the two successive utility variations: $(+-)$ followed by $(-+)$. The symbol $(\boxtimes,\ ++)$ stands for anyone of the four states in which the final variations are $(++)$. The symbol $(00,\ ++)$ stands for the initial state in which the first variations are $(++)$.



On this diagram, one can see the two types of trajectories which appear as absorbing states of the transition diagram: Type 1 trajectories are the repetition of $(++)$ whereas Type 2 trajectories are the repetition of $(-+)$ or of $(+-)$.

The preceding principles allow moreover to contruct a 'state transition diagram' for any game. The state transition diagram gives, for each interior point of the strategy space (α_1, α_2), the local direction of the movement (after the two initial periods).

$$\begin{bmatrix} +- & & ++ & & -+ \\ & \nwarrow & \uparrow & \nearrow & \\ -- & \leftarrow & & \rightarrow & ++ \\ & \swarrow & \downarrow & \searrow & \\ +- & & -- & & -+ \end{bmatrix}$$

$$\begin{bmatrix} +- & ++ & -+ \\ -- & & ++ \\ +- & -- & -+ \end{bmatrix}$$