ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

STUBBORN LEARNING

Jean-François LASLIER
Bernard WALLISER

Cahier n° 2011-12

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

Stubborn learning

Jean-François Laslier
CNRS and Ecole Polytechnique

Bernard Walliser
Paris School of Economics (ENPC, ENS)

June 2011

Abstract

The paper studies a specific reinforcement learning rule in two-player games when each player faces a unidimensional strategy set. The essential feature of the rule is that a player keeps on incrementing her strategy in the same direction if and only if her utility increases. The paper concentrates on games on the square $[0,1] \times [0,1]$ with bilinear payoff functions such as the mixed extensions of 2×2 games. It studies the behavior of the system in the interior as well as on the borders of the strategy space. It precisely exhibits the trajectories of the system and the asymptotic states for symmetric, zero-sum, and twin games.

1 Introduction

The paper examines a specific learning rule belonging to the family of reinforcement models (Sutton and Barto 1998). In such 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*, applies only to situations where the agent's strategy space is one-dimensional. It is based on the following principles:
- at each period, the decision maker is able to shift her action of an incremental quantity, in one direction or the other,
- she observes the utility obtained in the two past moves
- if the preceding shift induced an increase (decrease) in utility, she keeps on going in the same direction (she reverses direction).

This rule has three notable features, which derive directly from the stated principles. First, it requires relatively weak cognitive capacities of the player, for computing as well as for memorizing. Second, it is purely individual and can be used by the decision-maker without knowledge of her natural or strategic environment. Third, in a one-player setting, it induces the player to follow the familiar gradient-descent method. In a game setting, it describes the behavior of an adaptive agent which acts as if she was alone.

The rule was applied to the Cournot duopoly independently by Huck, Norman and Oechssler (2004) and by Trégouet (2004). It was applied to a continuous version of the Prisoner's Dilemma by Huck, Norman and Oechssler (2005).

In the present paper, we propose a new application for two-player generic bilinear games on the square: each player has as strategy set a closed interval, say $[0, 1]$ and his payoff is linear with respect to both her own strategy and her opponent's strategy. Some of our results extend easily to more general payoff functions, in particular some local results. We mention them in the text but keep the main focus on the bilinear case. The natural interpretation of such a game is a standard 2×2 game "played in mixed strategies". But it should be clear that our players do not randomize: they chose pure strategies.

We now describe more precisely the learning rule followed by each player. At each period, he 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 for the opposite direction if his utility is decreasing. However, this rule fails and must be completed in two cases.

First, if the agent's action is at a border of the strategy space (probability 0 or 1), the previous rule may prescribe an action outside this space. In such a case, we stipulate that the chosen action is not changed at all. Hence when the player wishes to, but cannot, cross the border, she stays on it. However, she keeps in his memory the fact that she wants 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 and even at usual border points. But it happens at corner points. In such a case, we stipulate that the player explores in the sense that she chooses at random whether to increase or decrease his strategy. Hence the player cannot be stuck forever at the same place. Notice that the rule is deterministic except in this last case.

Applied to general 2×2 games, two main properties of the system trajectories appear. 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 in the same direction, hence they generate locally a welfare-increasing path. On the border of the strategy space, two logics interfere: the optimality logic and the equilibrium logic. The precise resultant effect depends on the details of the game.

Convergence properties are studied specifically for three classes of games, namely symmetric, zero-sum and twin games. Two examples of the obtained results are the following. In the Prisoner's Dilemma, the system first moves in the direction of the Pareto optimum. When it reaches a border of the strategy space, it is stuck in a neighborhood of the impact point. The system thus escapes the curse of a sub-optimal Nash equilibrium. In Matching Pennies, the system circles around the mixed Nash equilibrium following a slowly expanding square. After it reaches a border of the strategy space, it cycles around the strategy space. The system thus tends to avoid the mixed Nash equilibrium.

Extension to more than two players is straightforward. But the extension to

a multi-dimensional action space is much harder. The rule has to be generalized in its definition and this can be done in different ways.

The next section provides a formal 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 game

In the general 2x2 game played by players 1 and 2, player i plays strategy α_i belonging to $[0,1]$.The strategy space is thus the square $[0,1] \times [0,1]$. The utility (or payoff) of player i is bilinear:

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

By analogy with a two-action game allowing mixed strategies, the payoffs can be summarized in the following matrix:

$1 \setminus 2$	$\alpha_2 = 1$	$\alpha_2 = 0$
$\alpha_1 = 1$	(a_1, a_2)	(b_1, b_2)
$\alpha_1 = 0$	(c_1, c_2)	(d_1, d_2)

This matrix is introduced in the paper for two different uses. First, it characterizes the restricted type of game we want to study. A symmetric, zero sum or twin game corresponds to specific bilinear utility functions. Second, it is associated with well-known equilibrium points, either in pure or in mixed strategies. These correspond respectively to corner and interior equilibria for the original game.

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 utility of player i.

2.2 Definition of the rule

The rule is followed by player i recursively at each 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 keeps in his memory four pieces of data:

- $\widetilde{u}_i(t-2)$ 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 follow in the last period.

- The player computes his intended direction :

$$v_i(t) = \begin{cases} v_i(t-1) & \text{if} \quad \widetilde{u}_i(t-1) > \widetilde{u}_i(t-2) \\ -v_i(t-1) & \text{if} \quad \widetilde{u}_i(t-1) < \widetilde{u}_i(t-2) \\ \pm v_i(t-1) \text{ at random} & \text{if} \quad \widetilde{u}_i(t-1) = \widetilde{u}_i(t-2) \end{cases}$$

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

- The player computes his actual strategy:

$$\alpha_i(t) = \begin{cases} 0 & \text{if } \alpha_i(t-1) + \varepsilon \cdot v_i(t) < 0 \\ 1 & \text{if } \alpha_i(t-1) + \varepsilon \cdot v_i(t) > 1 \\ \alpha_i(t-1) + \varepsilon \cdot v_i(t) & \text{in the other cases.} \end{cases}$$

  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. Denoting $\delta_i(t) \in \{-1, 0, +1\}$ the player's actual increment, his actual strategy can be written as:

$$\alpha_i(t) = \alpha_i(t-1) + \varepsilon \cdot \delta_i(t)$$

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.

Remark that the definition of the rule only requires that the strategy space of the decision-maker is uni-dimensionnal. The rule is well-defined for any n-player game, even if the payoff function is not linear with respect to individual strategies.

2.3 Payoff variations

We now study the instantaneous differential variations of the payoffs. The derivatives of the utility function for player i are the following:

$$\partial u_i / \partial \alpha_1 = b_i - d_i + E_i \alpha_2$$
$$\partial u_i / \partial \alpha_2 = c_i - d_i + E_i \alpha_1$$

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 similar ($k(t) = +1$) or opposite ($k(t) = -1$) evolution of the players' intended 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 of player 1 can be expressed in a compact way:

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

Since U_1 depends on k, denote:

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

with:

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

These numbers are interpreted as follows: 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 variation can be computed for the second player:

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

with

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

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 :

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

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

Remark: The 'stubborn learning rule' differs profoundly from the 'gradient learning rule' which is sometimes considered (Sutton and Barto 1998). In the last 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 implicitly fixed. The utility variation for the players are then:

$$du_1 = \frac{\partial u_1}{\partial \alpha_1} d\alpha_1 = \left(U_1^+ + U_1^-\right) d\alpha_1 = \widehat{U_1} d\alpha_1,$$
$$du_2 = \frac{\partial u_2}{\partial \alpha_2} d\alpha_2 = \left(U_2^+ + U_2^-\right) d\alpha_2 = \widehat{U_2} d\alpha_2.$$

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

Remark: In the case we study – bilinear games – the above expressions such as $\frac{\partial u_1}{\partial \alpha_1}$ or U_1^+ are linear. In the general case they would be defined at any point where the payoff function is differentiable, but they would not be linear.

2.4 Synthesis

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 in each point of the strategy space 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 variation is positive for player 1 and negative for player 2.

Notice that the rosace is constrained by the following *continuity rule*: for any player, by cycling around the table the signs must be in turn four times

positive and four times negative. In particular, signs in opposite directions are opposite, hence it is sufficient to know the signs in four successive directions. This leaves open 64 possible schemes.

For the sake of simplicity, the above scheme will be depicted as:

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

By convention, an asterisk $*$ instead of a sign means that the sign can be either $+$ or $-$.

Some definitions grounded on the rosace at a given point will be used in further results:

- a game is "covariant" (resp. "contravariant") if the players' respective signs in the rosace are the same (resp. opposite) in all directions.

- a border is "attractive" (resp. "repulsive") for a player if his utility increases (resp. decreases) in the three directions pointing to it.

- a corner is "attractive" (rep. "repulsive") for a player if his utility increases (resp. decreases) in the three directions pointing to it

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 which define a square in the plane $(\alpha_1, \alpha_2) : \alpha_1 = 0, \alpha_1 = 1, \alpha_2 = 0, \alpha_2 = 1$.

By definition, a player is "relevant" on a border when he is not constrained by that border (for instance player 1 on a horizontal border).

- the "separating lines" $L_1^+(U_1^+ = 0)$, $L_1^-(U_1^- = 0)$, $L_2^+(U_2^+ = 0)$, $L_1^+(U_2^- = 0)$ which 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 separating lines may coincide. Moreover, the strategy space may intersect one or several of these areas.

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

Remark: The separating lines are straight lines because the payoff functions are bilinear. In the general case, they would be separating curves. The results of this section would hold true as well for general payoff functions when separating curves "behave nicely". In this paper we do not discuss this point further.

3.2 Evolution at interior points

At interior points, the system is constrained by two features. First, 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 or alternate. Second, 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.

It follows that only two types of trajectories are possible as long as they do not reach a border or a separating line. Either no player changes his direction, hence both utilities increase. We will then say that the players are moving *in concert* in four possible directions: South-West, South-East, North-West or North East. Or one player keeps a constant direction while the other alternates, hence the first player sees her utility increase while the other sees her utility decrease. We then say that the players are moving *crab-wise* in four possible average directions: North, East, South or West.

The following proposition, proved in Appendix A, makes precise the conditions for moving in concert or crab-wise.

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

(i) The system moves crab-wise (one player is moving in a fixed direction and the other player alternates). This happens iff $U_1^- U_2^- \succ 0$ *and* $U_1^+ U_2^+ \prec 0$.

(ii) The system moves in concert (each player is moving in a fixed direction). This happens in all other cases.

Moreover, in general, the direction followed by the system does not depend on the initial move. The only exception is when two (orthogonal) in concert trajectories are viable from the same initial point. This happens iff $U_1^- U_2^- \prec 0$ *and* $U_1^+ U_2^+ \succ 0$.

3.3 Evolution on a separating line

When reaching a separating line, the system may change direction. This change can be inferred from the phase diagram by looking at the patterns of feasible directions on each side of the separating line.

The following proposition, proved in Appendix B, makes precise the conditions for these changes.

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 stuck 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, in the direction closest to the resultant of the directions on both sides of

the line. This happens in the remaining cases.

3.4 Evolution on a border line

When reaching a border line, the system has to change direction. The system is either stuck on the border, close to the impact point, or slides in a neighborhood of the border. The following proposition, proved in Appendix C, makes precise the conditions for these behaviors.

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

(i) If the system was previously moving in concert, the trajectory is of one of two kinds

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 this border is attractive for the relevant player.

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

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

3.5 Evolution at a corner

When reaching a corner, the system is either stuck in the neighborhood of the corner, or slides away in a neighborhood of a border. This border may be the one the system comes from (it makes a U-turn) or the other one (it makes a L-turn). The following proposition, proved in Appendix D, makes precise the conditions for these behaviors.

Proposition 4. *When reaching a corner, the trajectory follows one of two patterns:*

(i) It escapes from the corner following the neighborhood of a border which is attractive for the relevant player. This happens either when the game is covariant at the corner or when the corner is repulsive for one of the players.

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

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

[1]Within the class of symmetric games, generically, $E \neq 0$. But this rules out games which are both symmetric and zero-sum. Generic zero-sum games will be treated in section 4.

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 $b \geq c$ (since if $b \leq c$, 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 attention 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). The strategy space is a square centered on the main diagonal, and any such square is the strategy space of some symmetric game.

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 which is symmetric, namely $(1, 1)$ 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 asymmetric 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 $E = -5$)

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

Bentham optima

Consider the maximization of the sum W of players utilities over the whole

[2]Within the class of symmetric games, generically, $b \neq c$. But this rules out games which are both symmetric and twin. Generic twin games will be treated in section 5.

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 first interested in the *global Bentham optimum*. These are global maxima of the function W and are determined 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 further introduce the notion of a *local Bentham optimum*. These are local maxima of the W and 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 $a > 0$.

We finally define a *diagonal Bentham optimum*. This is 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 by P the point on the main diagonal

$$P = (\alpha_1^P, \alpha_2^P) = (\frac{d}{a+d}, \frac{d}{a+d}).$$

Consider the line L^P which is parallel to the second diagonal and passes through P. Its equation is:

$$\alpha_1 + \alpha_2 = \alpha_1^P + \alpha_2^P$$

Let M be the intersection of L and L^P.

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.

	$a < -1$	$-1 < a < 0$	$0 < a$
$1 < d$	**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

Table 1: Symmetric games: Nash equilibria and Bentham optima

- 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.

Table 1 summarizes the Nash equilibria, as well as Global and Local Bentham optima (by their coordinates). To read this Table: NE stands for Nash Equilibrium, GO stands for Global Bentham optimum, and LO stands for Local Bentham optimum when it is not global. For Bentham optima, the indication "$(0,0)$ or $(1,1)$" means that the optimum is $(0,0)$ if $d > a$ and $(1,1)$ if $d < a$.

4.2 State transition diagram

For symmetric games, there are only three separating lines, two of them parallel. They define six regions in the (α_1, α_2) plane separated by the lines $L^+(U_1^+ = U_2^+ = 0)$, $L_1^-(U_1^- = 0)$ and $L_2^-(U_2^- = 0)$, with:

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

Note that point P lies in the middle of the segment N_1N_2 defined on L^+ by L_1^- and L_2^-.

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.

In figures 1 and 2, the arrows describe the possible trajectories of the system as long as they do not reach a border of the strategy space or a separating line.

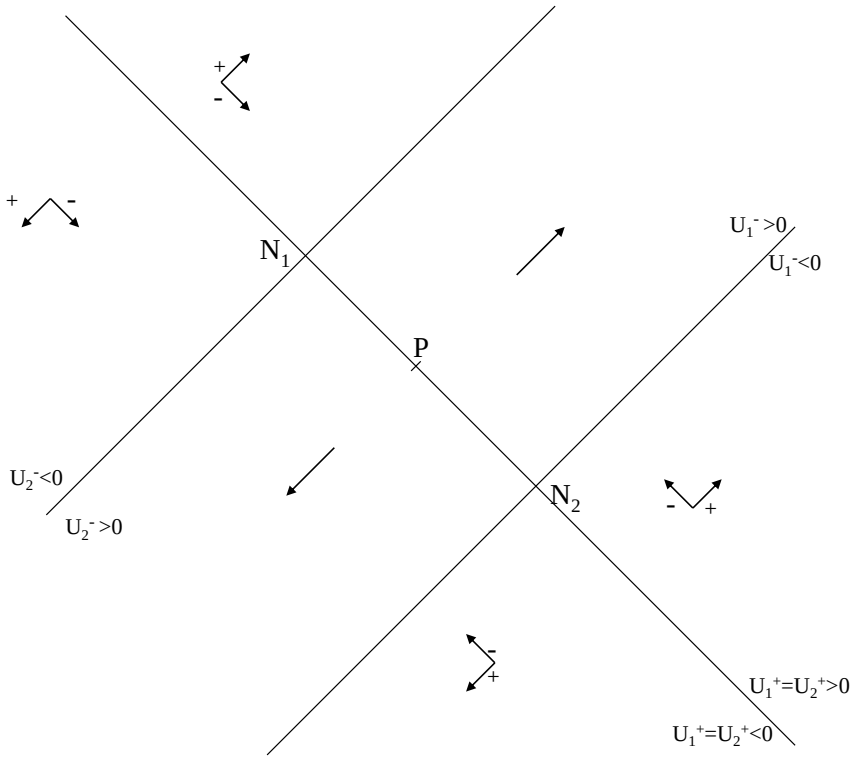

Figure 1: Symmetric coupling game ($E > 0$)

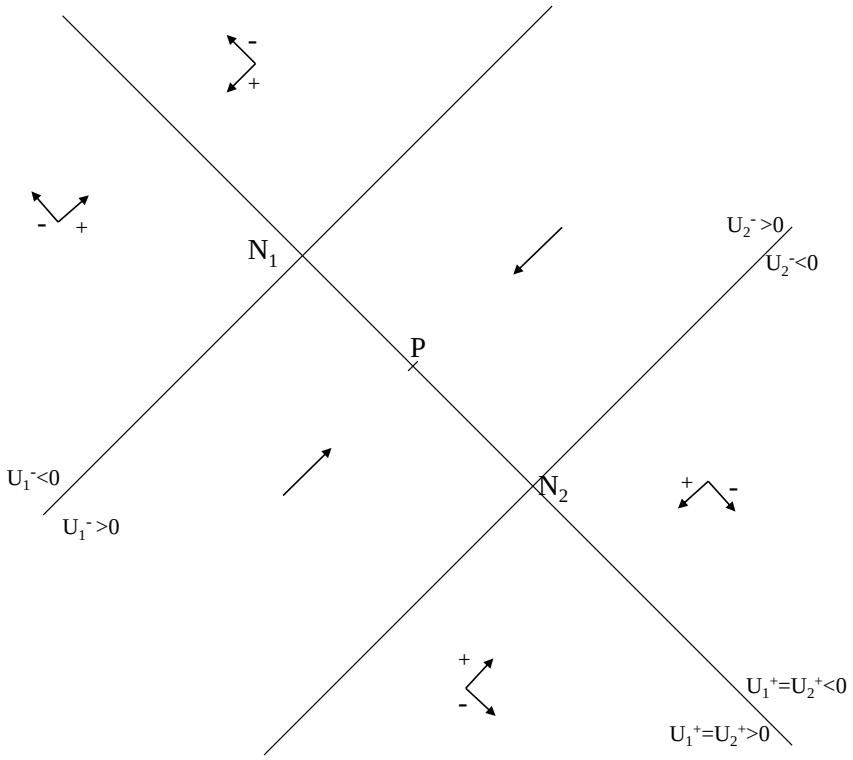

Figure 2: Symmetric de-coupling game ($E < 0$)

Note again that that they are two possible directions of evolution in the zones where two arrows are depicted, according to the initial state.

When reaching a separating line, two cases are possible:

If $E > 0$, this happens (up to some symmetry) in the upper part of Figure 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 rules of evolution on separating lines (case (i)), the system initially moving South East then slides North East along the separating line.

If $E < 0$, this happens in two subcases. Firstly, in the upper part of Figure 2, the system is moving South West above the separating line L^+ and is reaching a zone where it may go either North-East or North-West. But it can easily be shown that, when the system is on a trajectory parallel to the first diagonal, the utility variations of the two players are the same, hence the system remains on that diagonal. It follows that the system is stuck (in a cycle) on the separating line. Secondly, in the central part of Figure 2, the system is moving South West above the separating line L^+ and would be moving North East on the other side of this line. Then, according to the rules of evolution on separating lines (case (ii)), the system is again stuck (in a cycle) on the separating line.

4.3 Properties of the trajectories

We now take into account the borders of the strategy space in order to exhibit the convergence properties of the process.

It can be observed that the point P is at the intersection of L^+ with the main diagonal. A careful inspection shows that P belongs to the strategy space iff $a > 0$ and $d > 0$ (case $E \geq 0$) or $a < 0$ and $d < 0$ (case $E \leq 0$). The strategy space lies entirely below P if and only if $a < 0$ (when $E > 0$) and $a > 0$ (when $E < 0$). The strategy space lies entirely above P if and only if $d < 0$ (when $E > 0$) and $d > 0$ (when $E < 0$).

The strategy space lies entirely between the lines $U_1^- = 0$ and $U_2^- = 0$ iff $0 < a + d < 2$ (case $E < 0$) or $-2 < a + d < 0$ (case $E < 0$), which simply boils to $-2 < a + d < 2$

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 L_1^+ and L_1^-, 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 behaves according to the rules of the section "evolution on border lines". It can either stop or slide along the

border. The stated condition for the system to stop on the border $\alpha_1 = 0$ becomes, 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 $b > c$.

The behavior of the system when it slides along a border and meets a separating line is not examined, but is easily considered when happening.

4.4 Convergence results

When considering the whole behavior of the system, it appears that it is driven by the notion of Bentham optimality inside the strategy space and by the notion of Nash equilibrium on the border of the strategy space. The precise statement is provided in the following Theorem 1, proved in Apprendix E.

Theorem 1. *For symmetric games.*

(i) If there exists a unique symmetric global Bentham optimum, then the system coming from the interior of the action space 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 reaches the border of the strategy space.

(ii) If there exist a local and a global Bentham optima, both symmetric, then the system coming from the interior of the action space 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 reaches the border of the strategy space.

(iii) If there exist two asymmetric global Bentham optima, the system coming from the interior of the action space points towards a diagonal Bentham optimum. If it meets no border before, it converges toward it. If it meets first a border, it follows the border and converges towards the diagonal Bentham optimum on the border.

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 (by just eventually exchanging players), we can assume that $E > 0$

With respect to the pure Nash equilibria, two cases are possible:

- 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 $E < 0$ which is ruled out).

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

Looking for mixed equilibria, consider the point

$$Q = (\alpha_1^Q, \alpha_2^Q) = \left(\frac{d-c}{E}, \frac{d-b}{E} \right).$$

The point Q can be located anywhere in the plane (α_1, α_2) even with the constraint $E > 0$. Two relevant cases are possible. If Q belongs to the strategy space, then it corresponds to the unique mixed equilibrium. If Q does not belong to the strategy space, then the unique equilibrium is at a corner (and is pure). It is even more precisely located according to the following lemma (whose proof is omitted).

Lemma 1. The pure equilibrium is given by the following rule:
- (0,0) if $\alpha_1^Q \prec 0, \alpha_2^Q \succ 0$
- (0,1) if $\alpha_1^Q \succ 0, \alpha_2^Q \succ 1$
- (1,1) if $\alpha_1^Q \succ 1, \alpha_2^Q \prec 1$
- (1,0) if $\alpha_1^Q \prec 1, \alpha_2^Q \prec 0$

In other respects, as concern the Bentham optima, they are obviously degenerated.

5.2 State transition diagram

For zero sum games, there are only two separating lines $L^+(U_1^+ = U_2^+ = 0)$ and $L^-(U_1^- = U_2^- = 0)$ 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 precisely intersect at point Q. The phase diagram is depicted in Figure 3. A vertical or horizontal arrow describes the mean trajectory followed by the system when moving crab-wise.

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 strategy 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 crab-wise from the East and intersects the separating line parallel to the first diagonal. According to Figure 4, this intersection is South-West of Q. According to the rules of evolution on separating lines, case (ii) indicates that the system turns right and continues crab-wise to the North. The important point is that the trajectory after its turn on the separating line has made a small step away from Q, as will be stated in the next lemma.

Lemma 2. When crossing a separating line, the trajectory of the system is such that its mean line (joining the middle of its constituent players) is further away from point Q after the crossing than before.

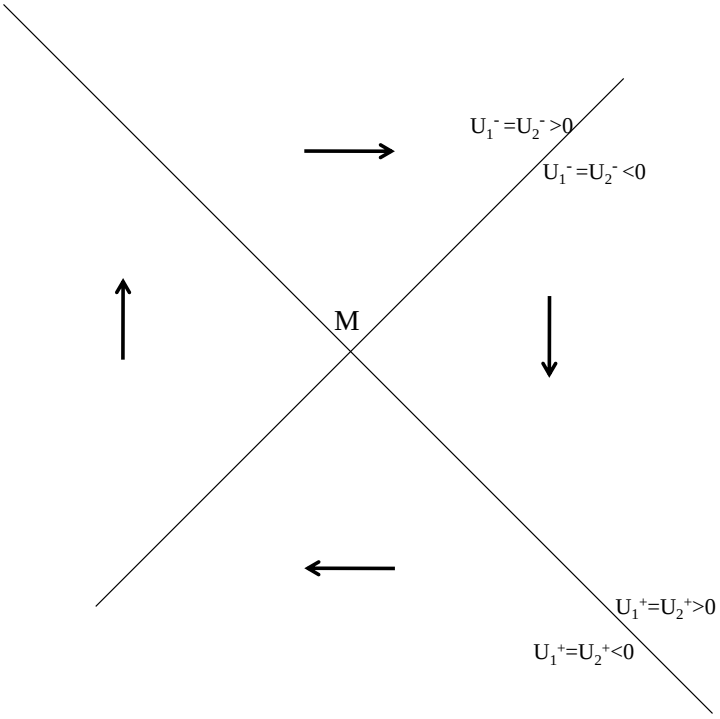

Figure 3: Zero-sum game with $E > 0$

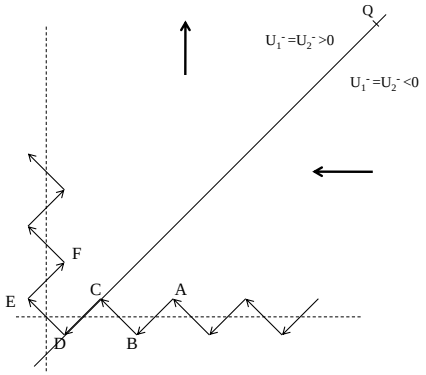

Figure 4: Right turn in a zero-sum game with $E > 0, case1$

Proof. Consider first the case in which the system arrives exactly on the separating line. (This is the case if the payoffs are integers and N is a multiple of E.) According to Figure 4, coming from A then B, the system reaches the separating line in C. Since the segment BC is entirely in the 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 variation obtained by the two players from A to B is positive for the first player and negative for the second, in short: $(+,-)$. 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 E. It then continues from E to F. Notice that the mean line is further away from Q, by magnitude $1/N$.

Consider now the case in which the system does not arrive exactly on the separating line. According to Figures 5 and 6, 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 is:

$$U(C) - U(B) = (1/N)(\alpha_1 - \alpha_2)(-a + b + c - d) + (1/N)(-b + c)$$

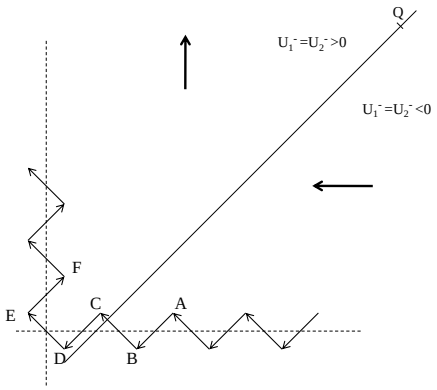

Figure 5: Right turn in a zero-sum game with $E > 0$, case 2

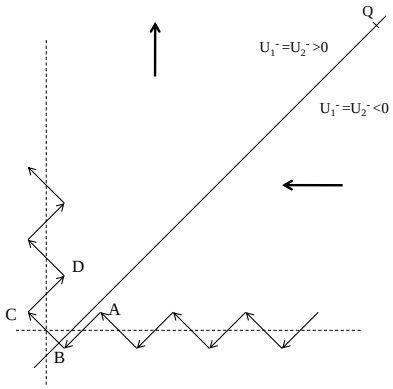

Figure 6: Right turn in a zero-sum game with $E > 0$, case 3

This variation is 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 (Figure 4) 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 (Figure 5) the system turns right after C and goes to D, then turns left to E. In both cases, the succession of moves has the same structure. Moreover, the mean line is further away from the point Q 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 crab-wise 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

For zero-sum games, pure strategy Nash equilibra are point attractors of the system, but the system cycles far away from mixed strategy Nash equilibria. The following result, proved in Appendix F, makes these statements precise:

Theorem 2: *For zero-sum games.*

(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 (interior) Nash equilibrium.

6 Twin games

6.1 Potential attractors

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

$$Q = (\alpha_1^Q, \alpha_2^Q) = \left(\frac{d-c}{E}, \frac{d-b}{E} \right).$$

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

$(0,0)$ if $\alpha_1^Q > 0$ and $\alpha_2^Q > 0$, and one of them is > 1,
$(0,1)$ if $\alpha_1^Q < 0$ and $\alpha_2^Q > 1$,
$(1,1)$ if $\alpha_1^Q < 1$ and $\alpha_2^Q < 1$, and one of them is < 0,
$(1,0)$ if $\widehat{\alpha_1} > 1$ and $\widehat{\alpha_2} < 0$.

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

6.2 State transition diagram

There are only two separating lines $L^+(U_1^+ = U_2^+ = 0)$ and $L^-(U_1^- = U_2^- = 0)$ 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 cross at Q. The state transition diagram is described in Figure 7.

6.3 Trajectory properties

As can be seen, the system never reaches a separating line, except maybe on borders. When coming on a border, according to the rules of evolution at a border, the system is never stuck.

6.4 Convergence results

For twin games, pure Nash equilibria are point attractors. We can state the following theorem, proved in Appendix G:

Theorem 3. *For twin games.*

(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.

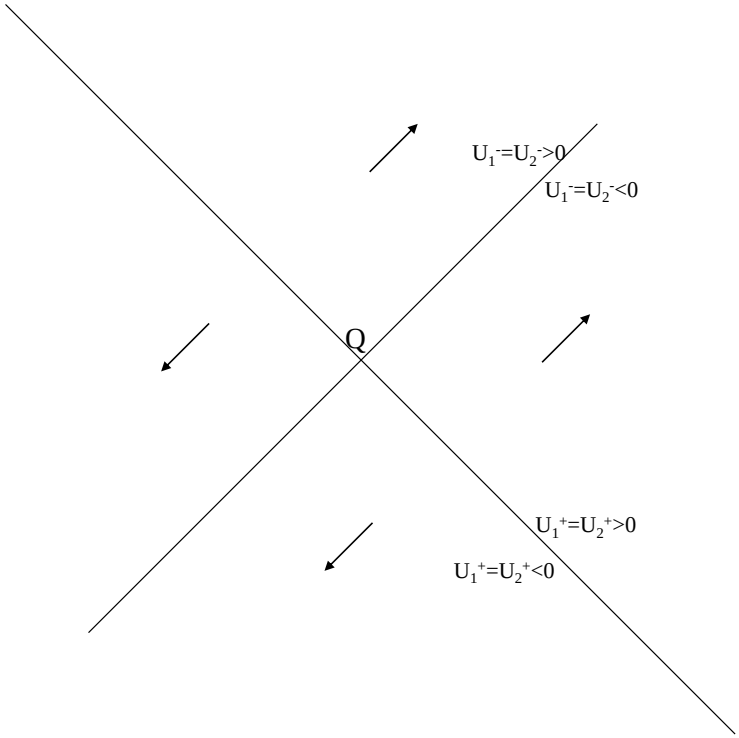

Figure 7: Twin game with $E > 0$

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

Table 2: interior points

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

Table 3: Possible regimes

Appendix: Proofs

Appendix A: Proof of Proposition 1. (Evolution at interior points)

According to the expression of *sign* $v_i(t)$, the first player moves in some direction independently of the preceding increment $v_i(t-1)$, but according to $k(t-1) = v_1(t-1).v_2(t-1)$. Since both players act in the same way, Table 2 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 $U_2(t)$. In Table 2 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 $v_2(t) = -1$.

Consider an initial value for k. The sign of k indicates whether U_i^+or U_i^- is the relevant expression for U_i. Table 1 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 Table 1 indicates the new direction.

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. 3 defines the one or two possible regimes for each configuration of parameters.

Table 3 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 crab-wise 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

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

Table 4: Possible shifts on a separating line

Appendix B: Proof of Proposition 2. (Evolution at a separating line)

Without restriction, we consider a separating line $U_i^- = 0$ which 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 refer then 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 South-East of the separating line.

According to 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). Table 4 records the possible shifts between types of trajectories. The numbers in the cells of this table indicates the player(s) involved in the shift and the letter the case involved.

Table 4 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 $U_2^- = 0$. Since the last move leading to the separating line is the same in all cases, the preceding point, at $t-1$, is such that: $\delta_1(t) = +1$ and $\delta_2(t) = -1$.

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 $\delta_2(t+2) = -1$. The 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 6-step pattern, 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 6-step pattern, 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 4-cycle is made of four consecutive moves:

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

QED

S \ SE	$++$	$-+$	$+-$	$--$
$++$	$A1$	$A1$	$A1$	$A1$
$-+$	$\varnothing$	$A2$	$\varnothing$	$A5$
$+-$	$\varnothing$	$\varnothing$	$A4$	$A4$
$--$	$\varnothing$	$\varnothing$	$\varnothing$	$A3$

Table 5: Possible shifts on a border

Appendix C: Proof of proposition 3. (Evolution on a border line)

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 the rosace, the sign of the payoff variation in the South-West direction is necessarily $++$. Since the signs in the West direction are irrelevant, the only relevant signs are those in the South and South-East directions. That would yield 16 possibilities, but some are impossible thanks to the continuity rule in the rosace. As will be seen, some cases are moreover identical. Table 5 summarizes the different cases to be considered:

Subcase A1.

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

The successive moves are:

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

which means that the system slides downward on the border.

Subcase A2.

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

The successive moves are:

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

which means that the system slides downward in an ε-neighborhood of the border according to a 4-step pattern.

Subcase A3.

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

The successive 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 according to a 2-step pattern.

Subcase A4.

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

The successive moves are:

$$\begin{array}{c}\delta_1(t+1)=0\ ,\ \delta_2(t+1)=-1\\ \delta_1(t+2)=0\ ,\ \delta_2(t+2)=+1\\ \delta_1(t+3)=+1\ ,\ \delta_2(t+3)=+1\\ \delta_1(t+4)=-1\ ,\ \delta_2(t+4)=-1\end{array}$$

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

Subcase A5.

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

The successive moves are:

$$\begin{array}{c}\delta_1(t+1)=0\ ,\ \delta_2(t+1)=-1\\ \delta_1(t+2)=+1\ ,\ \delta_2(t+2)=-1\\ \delta_1(t+3)=-1\ ,\ \delta_2(t+3)=+1\\ \delta_1(t+4)=0\ ,\ \delta_2(t+4)=+1\\ \delta_1(t+5)=0\ ,\ \delta_2(t+5)=-1\end{array}$$

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: Table 6 indicates the trajectory for each possible case of the rosace. The symbol $\varnothing$ means that the case is impossible. The symbol $\bigstar$ means that the system is stuck around some point at the border. The symbol $\downarrow$ indicates that the system slides down near the vertical border. Note that the direction followed on the border is in continuity with the direction before the impact on the border.

Case B: the system was moving crab-wise West before reaching the border. The sign of the payoff variation in the South-West direction is $+-$. and the sign in the South-East direction is $-+$. Hence, only the signs in the South direction are free. When the system, moving crab-wise, first reaches the border, its last movement is assumed, without restriction to be South-West. Table 7 summarizes the four subcases to be considered for the remaining directions:

S - SE	$++$	$-+$	$+-$	$--$
$++$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
$-+$	$\varnothing$	$\downarrow$	$\varnothing$	$\bigstar$
$+-$	$\varnothing$	$\varnothing$	$\bigstar$	$\bigstar$
$--$	$\varnothing$	$\varnothing$	$\varnothing$	$\downarrow$

Table 6: case A (border line)

S - SE	$-+$
$++$	$B1$
$-+$	$B2$
$+-$	$B3$
$-$	$B4$

Table 7: case B (border line)

Subcase B1:

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

The successive 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 downwards in a ε-neighborhood of the border, according to a 3-step pattern.

Subcase B2:

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

The successive moves are:

$$\begin{array}{c} \delta_1(t+1) = 0 \ , \ \delta_2(t+1) = +1 \\ \delta_1(t+2) = 0 \ , \ \delta_2(t+2) = -1 \\ \delta_1(t+3) = +1 \ , \ \delta_2(t+3) = -1 \\ \delta_1(t+4) = -1 \ , \ \delta_2(t+4) = -1 \end{array}$$

which means that the system slides downwards in a ε-neighborhood of the border, with a 4-step pattern.

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

Table 8: results for case B (border line)

Subcase B3:

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

The successive moves are:

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

which means that the system slides upwards in a ε-neighborhood of the border, with a 4-step pattern.

Subcase B4:

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

The successive (similar) moves are

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

which means that the system slides upwards on the border.

Synthesis for case B : Table 8 indicates the trajectory for each possible case of the rosace. It may be observed that the trajectory on the border always follows the direction of the best response for the relevant player, namely player 2. Hence, the result does not depend on the way the crab-wise trajectory reaches the border (which can be computed directly).

QED

1 \ 2	$-$ $+$ $+\ +\ +$	$-$ $-$ $+\ +\ +$	$+$ $+$ $+\ -\ -$	$-$ $+$ $+\ +\ -$
$-$ $+$ $+\ +\ +$	A1	A2	A2'	A3
$-$ $-$ $+\ +\ +$	A2	A1	A4	A2
$+$ $+$ $+\ -\ -$	A3	A4	A1'	A2'
$+$ $+$ $+\ +\ -$	A3	A3	A2'	A1'

Table 9: Possible configurations (corners)

Appendix D: Proof of Proposition 4. (Evolution at a corner)

Consider, without restriction, the corner $(0, 0)$. We study simultaneously what happens when the system reaches a corner coming from a border, or directly reaches a corner coming from the diagonal.

Case A: the system reaches one border or the other coming in concert. Notice that the South -West signs of the rosace have to be $++$. According moreover to the continuity principle of the rosace, sixteen configurations are depicted in Table 9. However, the system behavior reduces to four subcases indicated in the Table. The subcases A2 and A3 are asymmetric with respect to changing players 1 and 2, so we distinguish in the Table cases A2 from A2' and A3 from A3'.

Subcase A1. Example of such a rosace:

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

When the system reaches the West border, it slides South then turns at the corner and moves East along the South border. When the system reaches the South border, it slides West then makes a U-turn at the corner and moves East along the South border.

Subcase A2:

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

When the system reaches the West border, it slides South then turns at the corner and is stuck in a 4-cycle in the neighborhood of the corner. When the system reaches the South border, it is stuck in a cycle around the intersection point with the border.

Subcase A3:

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

Independently of how the corner is reached from a border, the system is stuck a 3-cycle at the corner.

Subcase A4.

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

The system is stuck at the point where it reaches the border (whatever this border is). If it reaches directly the corner, it is stuck in a 4-cycle at the corner.

Synthesis for case A: Table 10 indicates the system behavior for each subcase. The symbol ■ indicates that the system is stuck at the corner and the symbol □ indicates that the system does not stay at the corner. The symbol ★ indicates that the system is stuck on a border when reaching it. The arrows indicate that the system slides along the border. For instance, in the North-West cell, the system attracted by the South border will not stay at the corner and will slide East along the South border. One can notice, from the second diagonal, that the system escapes from the corner if and only if the game is covariant. If the system arrives at the corner following a border, it can either make a U-turn or turn at right angle.

Case B: the system reaches one border or the other coming crab-wise. Without loss of generality, we suppose that the system moves crab-wise horizontally towards the West border. The signs of two directions are imposed and the continuity principle applies. Hence, Table 11 depicts the four subcases to be distinguished:

Subcase B1.

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

The system, after it reaches the West border (be it at the corner or not), goes North and slides in the ε-neighborhood of the West border.

1 \ 2	− / + / + + +	− / − / + + +	+ / + / + − −	+ / + / + + −
− / + / + + +	↓ □ ⇄	↓ ■ ★	★ ■ ⟵	↓ ■ ⟵
− / − / + + +	↓ ■ ★	↓ □ ⇄	★ ■ ★	↓ ■ ★
+ / + / + − −	↓ ■ ⟵	★ ■ ★	↓↑ □ ⟵	★ ■ ⟵
+ / + / + + −	↓ ■ ⟵	↓ ■ ⟵	★ ■ ⟵	↓↑ □ ⟵

Table 10: results for case A(corners)

1 \ 2	− / − / − − +	− / − / − + +
+ / + / + + −	B1	B3
+ / + / + − −	B2	B4

Table 11: subcases B (corners)

1 \ 2	$\begin{matrix}-\\-\\- & - & +\end{matrix}$	$\begin{matrix}-\\-\\- & + & +\end{matrix}$
$\begin{matrix}+\\+\\+ & + & -\end{matrix}$	$\uparrow$ $\square$	$\downarrow$ $\blacksquare$
$\begin{matrix}+\\+\\+ & - & -\end{matrix}$	$\uparrow$ $\square$	$\downarrow$ $\blacksquare$

Table 12: results for case B (corners)

Subcase B2.

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

As in B1, the system, after it reaches the West border (be it at the corner or not), goes North and slides in the ε-neighborhood of the West border.

Subcase B3.

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

The system, after it reaches the West border (be it at the corner or not), goes South, reaches the corner and follows a 3-cycle around the corner.

Subcase B4.

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

The system, after it reaches the West border (be it at the corner or not), goes South, reaches the corner and follows a 4-cycle around the corner.

Synthesis for case B: Table 12 indicates the system behavior for each subcase. One can notice, from the first column, that the system escapes from the corner if and only if the corner is repulsive for a player (player 2).

QED

Appendix E: Proof of Theorem 1. (Symmetric games)

We examine the nine cases distinguished in Table 1 in the main text. 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 P with regard to the strategy space (a and d being positive or negative) and the position of the strategy 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 strategy space is entirely above P and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points South-West until 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 strategy space is entirely below P and between the separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory points South-West until 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 strategy space is entirely below P 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 situation is similar to subcase A2. If the initial point is above L_1^-, the trajectory depends on the initial 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 it 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 strategy space is entirely above P 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 initial 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 $\alpha_1 = 0$, or $\alpha_2 = 1$. In both cases, it slides along the border in until it reaches the corner $(0,1)$. From this corner, it slides down along the border $\alpha_1 = 0$, crosses the separating line and reaches the corner $(0,0)$ where it is stops.

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 strategy space includes P 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 border until it reaches the point $(0, 0)$; if the initial point is above L^+, the trajectory goes North-East until it reaches a border then slides along the border until it reaches the point $(1, 1)$.

Subcase C2. $a + d > 2$. The strategy space includes P 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 situation is similar to subcase C1. If the initial point is above L_1^-, the trajectory also depends on the initial 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 strategy space is entirely above P and between the separating lines L_1^- and L_2^-. According to Figure 2, the system trajectory points South-West until it reaches a border, where it is scotched.

Subcase D2: $a+d < -2$. The action space is entirely above P 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 stuck on the border. If the initial point is above L_1^-, the trajectory depends on the initial 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 stuck. 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 stuck. If it reaches the border $\alpha_2 = 1$ then it slides along this border until it reaches the corner $(0, 1)$, where it is stuck.

To sum up, in case D, the system ends up being stuck 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 strategy space is entirely below P and between the separating lines L_1^- and L_2^-. According to Figure 1, the system trajectory points South-West until it reaches a border, where it is stuck.

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

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

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

Subcase F1: $a + d < 2$. The strategy space includes P 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 initial point is above L^+, the trajectory goes North-East until it reaches a border then slides along the border until it reaches the point $(1, 1)$.

Subcase F2: $a + d > 2$. The strategy space includes P 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 situation is similar to the previous subcase F1. If the initial point is above L_1^-, the trajectory also depends on the initial 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 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 and finally reaches it.

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

Subcase G1: $a + d > -2$. The strategy space includes P 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 separating line L^+ where it stops or reaches first the border $\alpha_1 = 0$, follows it until reaching the separating line L^+ where it is stuck.

Subcase G2: $a + d < -2$. The strategy space includes P , intersects both separating lines L_1^- and L_2^-. The point N_1 may be inside or outside the strategy space, but the subcases are the same. Assume that N_1 is outside, say West of the border $\alpha_1 = 0$. (i) If the initial point is under L_1^- and South enough, the trajectory is as in subcase G1; it goes South-West and reaches L^+ where it is stuck. (ii) If the initial point is directly under L_1^-, the system meets first the border $\alpha_1 = 0$,slides along it and reaches L^+ where it is stuck (iii) If the initial point is above L_1^- and the initial direction is South-West, the system meets the border $\alpha_1 = 0$, slides along it, crosses L_1^- and reaches finally L^+ where it stops.(iv) If the initial point is above L_1^- and the initial direction is North-West, the system meets first the border $\alpha_2 = 1$,slides on it in West direction, then turns along border $\alpha_1 = 0$, slides South around it and reaches finally L^+ where it is stuck.

To sum up, in case G, the system goes on a diagonal Bentham optima, either an interior one when it reaches it directly or on a border when it reaches first a border.

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.

To sum up, case H is analogous to case G.

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

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

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

Subcase I3: $a + d > 2$ (hence $E > 0$). The strategy space is entirely above P 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 situation 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 separating line until it reaches the border $\alpha_2 = 1$, then slides along this border until it reaches the

corner $(1,1)$. If the initial 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 strategy space is entirely below P 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 turns to reach the corner $(1,1)$. If the initial 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, in case I, the system always goes to the corner $(1,1)$ which is a global optimum and the unique Nash equilibrium.

QED

Appendix F: Proof of Theorem 2. (Zero-sum games)

We distinguish two cases related to the position of the intersection Q of the separating lines with regard to the strategy space.

Case A: Q is outside the action space. For this part of the proof, we suppose that $\alpha_1^Q > 0$ and $\alpha_2^Q > 1$. Hence $d > c$ and $a < c$; the equilibrium is at the point $(0, 1)$.

Subcase A1: the action space does not intersect any separating line. For instance, the action space lies entirely in the quarter of plane South of Q. From any interior initial point, the system goes crab-wise West until it reaches the border $\alpha_1 = 0$. According to the theorem about border behavior, 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 and finally reaches it. To prove that the system is stuck around the equilibrium, first notice that the rosace is fully determined:

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

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

Subcase A2: 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 until 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 until it reaches the border $\alpha_2 = 1$, then goes West until it reaches the equilibrium $(0, 1)$.

Subcase A3: 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 until it reaches the equilibrium $(0, 1)$. If the initial point is above the separating line, the system first goes South until it reaches the separating line, then goes West until it reaches the border line, then goes North until it reaches the equilibrium $(0, 1)$.

Subcase A4: the action space intersects both separating lines. It is just a superposition of the two former cases.

To sum up, in case A, the system always converges towards the pure equilibrium.

Case B: Q is inside the action space. It is the mixed-strategy equilibrium of the game. By symmetry, we can consider only the case where:

$$0 < \alpha_1^Q < \alpha_2^Q < 1/2.$$

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

$$(0, \alpha_2^Q - \alpha_1^Q), (2\alpha_1^Q, \alpha_2^Q - \alpha_1^Q), (2\alpha_1^Q, \alpha_2^Q + \alpha_1^Q), (0, \alpha_2^Q + \alpha_1^Q).$$

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

Subcase B1. 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 lemma 2, the system is one step further away from $\widehat{M}$ after this turn. After zero, one or two such right turns, the system reaches a border. If the border is the border $\alpha_2 = 0$, the system goes West, according to the description of the behavior at a border. If the border is the border $\alpha_1 = 0$, 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 asymptotically cycles in a 3ε-neighborhood of S.

Subcase B2. 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 $\widehat{M}$ after each turn, this holds until the system reaches the square S. In fact, it goes even outside the square until reaching a border. But, this was already considered in case A.

QED

Appendix G: Proof of Theorem 3. (Twin games)

Case A: the two separating lines intersect the strategy space and divide this space into 4 areas. Without loss of generality we suppose that Q is inside the triangle at the left of the state space, defined by the three conditions:
$\alpha_1^Q < 0$
$\alpha_1^Q + \alpha_2^Q > 0$
$\alpha_2^Q - \alpha_1^Q < 1$
In that case the unique Nash equilibrium is at the corner $(1,1)$.

Subcase A1. The initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 < \alpha_1^Q + \alpha_2^Q$ and $\alpha_1 > \alpha_2$. Then the system goes South-West until it reaches the border $\alpha_2 = 0$. According to the behaviour at a border, it follows the border towards the West (according to player 1's best response) until reaching the corner $(0,0)$. Then, according the behaviour around 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.

Subcase A2. the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 < \alpha_1^Q + \alpha_2^Q$, but $\alpha_1 < \alpha_2$. Then the system goes South West until it reaches the border $\alpha_1 = 0$. Then behaviour on a border still applies: the system makes a $3\pi/4$ right turn and continues North like in the previous case.

Subcase A3. the initial point (α_1, α_2) is such that $\alpha_1 + \alpha_2 > \alpha_1^Q + \alpha_2^Q$ Then the system moves North-East until reaching one the two borders $\alpha_1 = 1$ or $\alpha_2 = 1$, follows that border until reaching the Nash equilibrium $(1,1)$.

The other cases for Q 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.

When Q is inside the state space, there are two pure Nash equilibrium, located (since we assume $E > 0$) at $(0,0)$ and $(1,1)$.

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

QED

References

[1] Huck, S. - Norman, H.T. - Oechssler, J. (2004) "GLAD: a simple adaptive strategy that yields cooperation in dilemma games" Physica D 200: 133—138

[2] Huck, S. - Norman, H.T. - Oechssler, J. (2005) "Through trial and error to collusion" International Economic Review, 45: 205—224.

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

[4] Trégouet, T. (2004): "Le duopole, l'apprentissage de la concurrence" document ENSAE.