

# Repeated Games with Incomplete Information<sup>1</sup>

Jérôme Renault<sup>2</sup> (Ceremade, Université Paris Dauphine, France)

## Article Outline

Glossary and Notation

I. Definition of the Subject and its Importance

II. Strategies, Payoffs, Value and Equilibria

III. The standard model of Aumann and Maschler

IV. Vector Payoffs and Approachability

V. Zero-sum games with lack of information on both sides

VI. Non zero-sum games with lack of information on one side

VII. Non-observable actions

VIII. Miscellaneous

IX. Future directions

X. Bibliography

## Glossary and Notation

**Repeated game with incomplete information:** a situation where several players repeat the same stage game, the players having different knowledge of the stage game which is repeated.

**Strategy of a player:** a rule, or program, describing the action taken by the player in any possible case which may happen.

**Strategy profile:** a vector containing a strategy for each player.

**Lack of information on one side:** particular case where all the players but one perfectly know the stage game which is repeated.

**Zero-sum games:** 2-player games where the players have opposite payoffs.

**Value:** Solution (or price) of a zero-sum game, in the sense of the fair amount that player 1 should give to player 2 to be entitled to play the game.

**Equilibrium:** Strategy profile where each player's strategy is in best reply against the strategy of the other players.

**Completely revealing strategy:** strategy of a player which eventually reveals to the other players everything known by this player on the selected state.

**Non revealing strategy:** strategy of a player which reveals nothing on the selected state.

---

<sup>1</sup>March 3, 2008

<sup>2</sup>I thank Françoise Forges, Sergiu Hart, Dinah Rosenberg, Robert Simon and Eilon Solan for their comments on a preliminary version of this chapter.

**The simplex of probabilities over a finite set:** for a finite set  $S$ , we denote by  $\Delta(S)$  the set of probabilities over  $S$ , and we identify  $\Delta(S)$  to  $\{p = (p_s)_{s \in S} \in \mathbb{R}^S, \forall s \in S p_s \geq 0 \text{ and } \sum_{s \in S} p_s = 1\}$ . Given  $s$  in  $S$ , the Dirac measure on  $s$  will be denoted by  $\delta_s$ . For  $p = (p_s)_{s \in S}$  and  $q = (q_s)_{s \in S}$  in  $\mathbb{R}^S$ , we will use, unless otherwise specified,  $\|p - q\| = \sum_{s \in S} |p_s - q_s|$ .

## I. Definition of the Subject and its Importance

### Introduction

In a repeated game with incomplete information, there is a basic interaction called stage game which is repeated over and over by several participants called players. The point is that the players do not perfectly know the stage game which is repeated, but rather have different knowledge about it. As illustrative examples, one may think of the following situations: an oligopolistic competition where firms don't know the production costs of their opponents, a financial market where traders bargain over units of an asset which terminal value is imperfectly known, a cryptographic model where some participants want to transmit some information (e.g., a credit card number) without being understood by other participants, a conflict when a particular side may be able to understand the communications inside the opponent side (or might have a particular type of weapons),...

Natural questions arising in this context are as follows. What is the optimal behavior of a player with a perfect knowledge of the stage game ? Can we determine which part of the information such a player should use ? Can we price the value of possessing a particular information ? How should one player behave while having only a partial information ?

Foundations of games with incomplete information have been studied in [27] and [55]. Repeated games with incomplete information have been introduced in the sixties by Aumann and Maschler [87], and we present here the basic and fundamental results of the domain. Let us start with a few well known elementary examples ([87], [93]).

**Basic Examples** In each example, there are two players, and the game is zero-sum, i.e. player 2's payoff always is the opposite of player 1's payoff. There are two states  $a$  and  $b$ , and the possible stage games are given by two real matrices  $G^a$  and  $G^b$  with identical size. Initially a true state of nature  $k \in \{a, b\}$  is selected with even probability between  $a$  and  $b$ , and  $k$  is announced *to player 1 only*. Then the matrix game  $G^k$  is repeated over and over: at every stage, simultaneously player 1 chooses a row  $i$ , whereas player 2 chooses a column  $j$ , the stage payoff for player 1 is then  $G^k(i, j)$  but only  $i$  and  $j$  are publicly announced before proceeding to the next stage. Players are patient and want to maximize their long-run average expected payoffs.

**Example 1:**  $G^a = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$  and  $G^b = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$ .

This example is trivial. In order to maximize his payoff, player 1 just has to play, at any stage, the *Top* row if the state is  $a$  and the *Bottom* row if the state is  $b$ .

**Example 2:**  $G^a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  and  $G^b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ .

A naive strategy for player 1 would be to play at stage 1: *Top* if the state is  $a$ , and *Bottom* if the state is  $b$ . Such a strategy is called completely revealing, or CR, because it allows player 2 to deduce the selected state from the observation of the actions played by player 1. This strategy of player 1 would be optimal here if a single stage was to be played, but it is a very weak strategy on the long run and does not guarantee more than zero at each stage  $t \geq 2$  (because player 2 can play *Left* or *Right* depending on player 1's first action).

On the opposite, player 1 may not use his information and play a non revealing, or NR, strategy, i.e. a strategy which is independent of the selected state. He can consider the average matrix  $\frac{1}{2}G^a + \frac{1}{2}G^b = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ , and play independently at each stage an optimal mixed action in this matrix, i.e. here the unique mixed action  $\frac{1}{2}Top + \frac{1}{2}Bottom$ . It will turn out that this is here the optimal behavior for player 1, and the value of the repeated game is the value of the average matrix, i.e.  $1/4$ .

**Example 3:**  $G^a = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}$  and  $G^b = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}$ .

Playing a CR strategy for player 1 does not guarantee more than zero in the long-run, because player 2 will eventually be able to play *Middle* if the state is  $a$ , and *Left* if the state is  $b$ . But a NR strategy will not do better, because the average matrix  $\frac{1}{2}G^a + \frac{1}{2}G^b$  is  $\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix}$ , hence has value 0.

We will see later that an optimal strategy for player 1 in this game is to play as follows. Initially, player 1 chooses an element  $s$  in  $\{T, B\}$  as follows: if  $k = a$ , then  $s = T$  with probability  $3/4$ , and thus  $s = B$  with probability  $1/4$ ; and if  $k = b$ , then  $s = T$  with probability  $1/4$ , and  $s = B$  with probability  $3/4$ . Then at each stage player 1 plays row  $s$ , independently of the actions taken by player 2. The conditional probabilities satisfy:  $P(k = a | s = T) = 3/4$ , and  $P(k = a | s = B) = 1/4$ . At the end of stage 1, player 2 will have learnt, from the action played by his opponent, something about the selected state: his belief on the state will move from  $\frac{1}{2}a + \frac{1}{2}b$  to  $\frac{3}{4}a + \frac{1}{4}b$  or to  $\frac{1}{4}a + \frac{3}{4}b$ . But player 2 still does not know perfectly the selected state. Such a strategy of player 1 is called *partially revealing*.

### General Definition

Formally, a repeated game with incomplete information is given by the following data. There is a set of players  $N$ , and a set of states  $K$ . Each player  $i$  in  $N$  has a set of actions  $A^i$  and a set of signals  $U^i$ , and we denote by  $A = \prod_{i \in N} A^i$  the set of action profiles and by  $U = \prod_{i \in N} U^i$  the set of signal profiles. Every

player  $i$  has a payoff function  $g^i : K \times A \rightarrow \mathbb{R}$ . There is a signalling function  $q : K \times A \rightarrow \Delta(U)$ , and an initial probability  $\pi \in \Delta(K \times U)$ . In what follows, we will always assume the sets of players, states, actions and signals to be non empty and finite.

A repeated game with incomplete information can thus be denoted by  $\Gamma = (N, K, (A^i)_{i \in N}, (U^i)_{i \in N}, (g^i)_{i \in N}, q, \pi)$ . The progress of the game is the following.

- Initially, an element  $(k, (u_0^i)_i)$  is selected according to  $\pi$ :  $k$  is the realized state of nature and will remain fixed, and each player  $i$  learns  $u_0^i$  (and nothing more than  $u_0^i$ ).

- At each integer stage  $t \geq 1$ , simultaneously every player  $i$  chooses an action  $a_t^i$  in  $A^i$ , and we denote by  $a_t = (a_t^i)_i$  the action profile played at stage  $t$ . The stage payoff of a player  $i$  is then given by  $g^i(k, a_t)$ . A signal profile  $(u_t^i)_i$  is selected according to  $q(k, a_t)$ , and each player  $i$  learns  $u_t^i$  (and nothing more than  $u_t^i$ ) before proceeding to the next stage.

Remarks:

1. The players do not necessarily know their stage payoff after each stage (as an illustration, imagine the players bargaining over units of an asset which terminal value will only be known “at the end” of the game). This is without loss of generality, because it is possible to add hypotheses on  $q$  so that each player will be able to deduce his stage payoff from his realized stage signal.

2. Repeated games with complete information are a particular case, corresponding to the situation where each initial signal  $u_0^i$  reveals the selected state. Such games are studied in the chapter “Repeated games with complete information”.

3. Games where the state variable  $k$  evolve from stage to stage, according to the actions played, are called stochastic games. These games are not covered here, but in a specific chapter entitled “Stochastic games”.

4. The most standard case of signalling function is when each player exactly learns, at the end of each stage  $t$ , the whole action profile  $a_t$ . Such games are usually called games with “perfect monitoring”, “full monitoring”, “perfect observation” or with “observable actions”.

## II. Strategies, Payoffs, Value and Equilibria

### Strategies

A (behavior) strategy for player  $i$  is a rule, or program, describing the action taken by this player in any possible case which may happen. These actions may be chosen at random, so a strategy for player  $i$  is an element  $\sigma^i = (\sigma_t^i)_{t \geq 1}$ , where for each  $t$ ,  $\sigma_t^i$  is a mapping from  $U^i \times (U^i \times A^i)^{t-1}$  to  $\Delta(A^i)$  giving the lottery played by player  $i$  at stage  $t$  as a function of the past signals and actions of player  $i$ . The set of strategies for player  $i$  is denoted by  $\Sigma^i$ .

A history of length  $t$  in  $\Gamma$  is a sequence  $(k, u_0, a_1, u_1, \dots, a_t, u_t)$ , and the set of

such histories is the finite set  $K \times U \times (A \times U)^t$ . An infinite history is called a play, the set of plays is denoted by  $\Omega = K \times U \times (A \times U)^\infty$  and is endowed with the product  $\sigma$ -algebra. A strategy profile  $\sigma = (\sigma^i)_i$  naturally induces, together with the initial probability  $\pi$ , a probability distribution over the set of histories of length  $t$ . This probability uniquely extends to a probability over plays, and is denoted by  $\mathbb{P}_{\pi, \sigma}$ .

### Payoffs

Given a time horizon  $T$ , the average expected payoff of player  $i$ , up to stage  $T$ , if the strategy profile  $\sigma$  is played is denoted by:

$$\gamma_T^i(\sigma) = \mathbb{E}_{\mathbb{P}_{\pi, \sigma}} \left( \frac{1}{T} \sum_{t=1}^T g^i(k, a_t) \right).$$

The  $T$ -stage game is the game  $\Gamma_T$  where simultaneously, each player  $i$  chooses a strategy  $\sigma^i$  in  $\Sigma^i$ , then receives the payoff  $\gamma_T^i((\sigma^j)_{j \in N})$ .

Given a discount factor  $\lambda$  in  $(0, 1]$ , the  $\lambda$ -discounted payoff of player  $i$  is denoted by:

$$\gamma_\lambda^i(\sigma) = \mathbb{E}_{\mathbb{P}_{\pi, \sigma}} \left( \lambda \sum_{t=1}^{\infty} (1 - \lambda)^{t-1} g^i(k, a_t) \right).$$

The  $\lambda$ -discounted game is the game  $\Gamma_\lambda$  where simultaneously, each player  $i$  chooses a strategy  $\sigma^i$  in  $\Sigma^i$ , then receives the payoff  $\gamma_\lambda^i((\sigma^j)_{j \in N})$ .

**Remark:** A strategy for player  $i$  is called *pure* if it always plays in a deterministic way. A *mixed strategy* for player  $i$  is defined as a probability distribution over the set of pure strategies (endowed with the product  $\sigma$ -algebra). Kuhn's theorem (see [37], [3] or [92] for a modern presentation) states that mixed strategies or behavior strategies are equivalent, in the following sense: for each behavior strategy  $\sigma^i$ , there exists a mixed strategy  $\tau^i$  of the same player such that  $\mathbb{P}_{\pi, \sigma^i, \sigma^{-i}} = \mathbb{P}_{\pi, \tau^i, \sigma^{-i}}$  for any strategy profile  $\sigma^{-i}$  of the other players, and *vice-versa* if we exchange the words "behavior" and "mixed". Unless otherwise specified, the word strategy will refer here to a behavior strategy, but we will also sometimes equivalently use mixed strategies, or even mixtures of behavior strategies.

### Value of zero-sum games

By definition the game is zero-sum if there are two players, say player 1 and player 2, with opposite payoffs. The  $T$ -stage game  $\Gamma_T$  can then be seen as a matrix game, hence by the minmax theorem it has a value  $v_T = \sup_{\sigma^1} \inf_{\sigma^2} \gamma_T^1(\sigma^1, \sigma^2) = \inf_{\sigma^2} \sup_{\sigma^1} \gamma_T^1(\sigma^1, \sigma^2)$ . Similarly, one can use Sion's theorem ([71]) to show that the  $\lambda$ -discounted game has a value  $v_\lambda = \sup_{\sigma^1} \inf_{\sigma^2} \gamma_\lambda^1(\sigma^1, \sigma^2) = \inf_{\sigma^2} \sup_{\sigma^1} \gamma_\lambda^1(\sigma^1, \sigma^2)$ .

To study long term strategic aspects, it is also important to consider the following notion of uniform value. Players are asked to play well uniformly in the

time horizon, i.e. simultaneously in all game  $\Gamma_T$  with  $T$  sufficiently large (or similarly uniformly in the discount factor, i.e. simultaneously in all game  $\Gamma_\lambda$  with  $\lambda$  sufficiently low).

**Definitions 1:** Player 1 can guarantee the real number  $v$  in the repeated game  $\Gamma$  if:  $\forall \varepsilon > 0, \exists \sigma^1 \in \Sigma^1, \exists T_0, \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \gamma_T^1(\sigma^1, \sigma^2) \geq v - \varepsilon$ . Similarly, Player 2 can guarantee  $v$  in  $\Gamma$  if  $\forall \varepsilon > 0, \exists \sigma^2 \in \Sigma^2, \exists T_0, \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \gamma_T^1(\sigma^1, \sigma^2) \leq v + \varepsilon$ . If both player 1 and player 2 can guarantee  $v$ , then  $v$  is called the uniform value of the repeated game. A strategy  $\sigma^1$  of player 1 satisfying  $\exists T_0, \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \gamma_T^1(\sigma^1, \sigma^2) \geq v$  is then called an optimal strategy of player 1 (optimal strategies of player 2 are defined similarly).

The uniform value, whenever it exists, is necessarily unique. Its existence is a strong property, which implies that both  $v_T$ , as  $T$  goes to infinity, and  $v_\lambda$ , as  $\lambda$  goes to zero, converge to the uniform value.

### Equilibria of general-sum games

In the general case, the  $T$ -stage game  $\Gamma_T$  can be seen as the mixed extension of a finite game, and consequently possesses a Nash equilibrium. Similarly, the discounted game  $\Gamma_\lambda$  always has, by the Nash Glicksberg theorem, a Nash equilibrium. Concerning uniform notions, couples of optimal strategies are generalized as follows.

**Definitions 2:** A strategy profile  $\sigma = (\sigma^i)_{i \in N}$  is a uniform Nash equilibrium of  $\Gamma$  if: 1)  $\forall \varepsilon > 0, \sigma$  is an  $\varepsilon$ -Nash equilibrium in every finitely repeated game sufficiently long, that is:  $\exists T_0, \forall T \geq T_0, \forall i \in N, \forall \tau^i \in \Sigma^i, \gamma_T^i(\tau^i, \sigma^{-i}) \leq \gamma_T^i(\sigma) + \varepsilon$ , and 2) the sequence of payoffs  $((\gamma_T^i(\sigma))_{i \in N})_T$  converges to a limit payoff  $(\gamma^i(\sigma))_{i \in N}$  in  $\mathbb{R}^N$ .

Remark: The initial probability  $\pi$  will play a great role in the following analyses, so we will often write  $\gamma_T^{i,\pi}(\sigma)$  for  $\gamma_T^i(\sigma)$ ,  $v_T(\pi)$  for the value  $v_T$ , etc...

## III. The standard model of Aumann and Maschler

This famous model has been introduced in the sixties by Aumann and Maschler (see the reedition [87]). It deals with zero-sum games with lack of information on one side and observable actions, as in the basic examples previously presented. There is a finite set of states  $K$ , an initial probability  $p = (p^k)_{k \in K}$  on  $K$ , and a family of matrix games  $G^k$  with identical size  $I \times J$ . Initially, a state  $k$  in  $K$  is selected according to  $p$ , and announced to player 1 (called the informed player) only. Then the matrix game  $G^k$  is repeated over and over: at every stage, simultaneously player 1 chooses a row  $i$  in  $I$ , whereas player 2 chooses a column  $j$  in  $J$ , the stage payoff for player 1 is then  $G^k(i, j)$  but only  $i$  and  $j$  are publicly announced

before proceeding to the next stage. Denote by  $M$  the constant  $\max_{k,i,j} |G^k(i,j)|$ .

### III.1 Basic tools: Splitting, Martingale, Concavification, and the Recursive Formula

The following aspects are simple but fundamental. The initial probability  $p = (p^k)_{k \in K}$  represents the initial belief, or *a priori*, of player 2 on the selected state of nature. Assume that player 1 chooses his first action (or more generally a message or signal  $s$  from a finite set  $S$ ) according to a probability distribution depending on the state, i.e. according to a transition probability  $x = (x^k)_{k \in K} \in \Delta(S)^K$ . For each signal  $s$ , the probability that  $s$  is chosen is denoted  $\lambda(x, s) = \sum_k p^k x^k(s)$ , and given  $s$  such that  $\lambda(x, s) > 0$  the conditional probability on  $K$ , or *a posteriori* of player 2, is  $\hat{p}(x, s) = \left( \frac{p^k x^k(s)}{\lambda(x, s)} \right)_{k \in K}$ . We clearly have:

$$p = \sum_{s \in S} \lambda(x, s) \hat{p}(x, s). \quad (1)$$

So the *a priori*  $p$  lies in the convex hull of the *a posteriori*. The following lemma expresses a reciprocal: player 1 is able to induce any family of *a posteriori* containing  $p$  in its convex hull.

**Splitting lemma 1.** *Assume that  $p$  is written as a convex combination  $p = \sum_{s \in S} \lambda_s p_s$  with positive coefficients. There exists a transition probability  $x \in \Delta(S)^K$  such that  $\forall s \in S$ ,  $\lambda_s = \lambda(x, s)$  and  $p_s = \hat{p}(x, s)$ .*

**Proof:** just put  $x^k(s) = \frac{\lambda_s p_s^k}{p^k}$  if  $p^k > 0$ . □

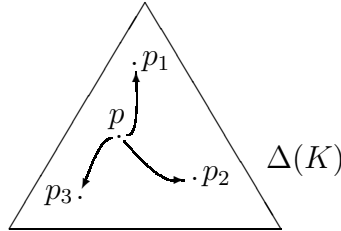


Figure 1: Splitting.

Equation 1 not only tells that the *a posteriori* contains  $p$  in their convex hull, but also that the expectation of the *a posteriori* is the *a priori*. We are here in a repeated context, and for every strategy profile  $\sigma$  one can define the process  $(p_t(\sigma))_{t \geq 0}$  of the *a posteriori* of player 2. We have  $p_0 = p$ , and  $p_t(\sigma)$  is the random variable of player 2's belief on the state after the first  $t$  stages. More precisely, we define for any  $t \geq 0$ ,  $h_t = (i_1, j_1, \dots, i_t, j_t) \in (I \times J)^t$  and  $k$  in  $K$ :

$$p_t^k(\sigma, h_t) = \mathbb{P}_{p,\sigma}(k|h_t) = \frac{p^k \mathbb{P}_{\delta^k,\sigma}(h_t)}{\mathbb{P}_{p,\sigma}(h_t)}.$$

$p_t(\sigma, h_t) = (p_t^k(\sigma, h_t))_{k \in K} \in \Delta(K)$  (arbitrarily defined if  $\mathbb{P}_{p,\sigma}(h_t) = 0$ ) is the conditional probability on the state of nature given that  $\sigma$  is played and  $h_t$  has occurred in the first  $t$  stages. It is easy to see that as soon as  $\mathbb{P}_{p,\sigma}(h_t) > 0$ ,  $p_t(\sigma, h_t)$  does not depend on player 2's strategy  $\sigma^2$ , nor on player 2's last action  $j_t$ . It is fundamental to notice that:

**Martingale of a posteriori lemma 2.**  $(p_t(\sigma))_{t \geq 0}$  is a  $\mathbb{P}_{p,\sigma}$ -martingale with values in  $\Delta(K)$ .

This is indeed a general property of Bayesian learning of a fixed unknown parameter: *the expectation of what I will know tomorrow is what I know today*. This martingale is controlled by the informed player, and the splitting lemma shows that this player can essentially induce any martingale issued from the *a priori*  $p$ . Notice that, to be able to compute the realizations of the martingale, player 2 needs to know the strategy  $\sigma^1$  used by player 1.

The splitting lemma also easily gives the following concavification result. Let  $f$  be a continuous mapping from  $\Delta(K)$  to  $\mathbb{R}$ . The smallest concave function above  $f$  is denoted by  $\text{cav} f$ , and we have:

$$\text{cav} f(p) = \max \left\{ \sum_{s \in S} \lambda_s f(p_s), S \text{ finite}, \forall s \lambda_s \geq 0, p_s \in \Delta(K), \sum_{s \in S} \lambda_s = 1, \sum_{s \in S} \lambda_s p_s = p \right\}.$$

**Concavification lemma 3.** If for any initial probability  $p$ , the informed player can guarantee  $f(p)$  in the game  $\Gamma(p)$ , then for any  $p$  this player can also guarantee  $\text{cav} f(p)$  in  $\Gamma(p)$ .

### III.2 Non revealing games

As soon as player 1 uses a strategy which depends on the selected state, the martingale of *a posteriori* will move and player 2 will have learnt something on the state. This is the dilemma of the informed player: he can not use the information on the state without revealing information. Imagine now that player 1 decides to reveal no information on the selected state, and plays independently of it. Since payoffs are defined via expectations, it is as if the players were repeating the average matrix game  $G(p) = \sum_{k \in K} p^k G^k$ . Its value is:

$$u(p) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{i,j} x(i)y(j)G(p)(i,j) = \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} \sum_{i,j} x(i)y(j)G(p)(i,j).$$

$u$  is a Lipschitz function, with constant  $M$ , from  $\Delta(K)$  to  $\mathbb{R}$ . Clearly, player 1 can guarantee  $u(p)$  in the game  $\Gamma(p)$  by playing i.i.d. at each stage an optimal strategy in  $G(p)$ . By the concavification lemma, we obtain:



**Proposition 1.** Player 1 can guarantee  $\text{cavu}(p)$  in the game  $\Gamma(p)$ .

Let us come back to the examples. In example 1, we have  $u(p) = \text{Val} \begin{pmatrix} -(1-p) & 0 \\ 0 & -p \end{pmatrix} = -p(1-p)$ , where  $p \in [0, 1]$  stands here for the probability of state  $a$ . This is a convex function of  $p$ , and  $\text{cavu}(p) = 0$  for all  $p$ . In example 2,  $u(p) = p(1-p)$  for all  $p$ , hence  $u$  is already concave and  $\text{cavu} = u$ . Regarding example 3, the following picture show the functions  $u$  (regular line), and  $\text{cavu}$  (dashed line).

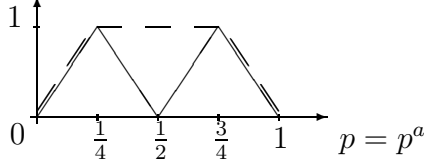


Figure 2:  $u$  and  $\text{cavu}$ .

Let us consider again the partially revealing strategy previously described. With probability  $1/2$ , the  $a$  *posteriori* will be  $\frac{3}{4}a + \frac{1}{4}b$ , and player 1 will play *Top* which is optimal in  $\frac{3}{4}G^a + \frac{1}{4}G^b = \begin{pmatrix} 3 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix}$ . Similarly with probability  $1/2$ , the  $a$  *posteriori* will be  $\frac{1}{4}a + \frac{3}{4}b$  and player 1 will play an optimal strategy in  $\frac{1}{4}G^a + \frac{3}{4}G^b$ . Consequently, this strategy guarantees  $1/2 u(3/4) + 1/2 u(1/4) = \text{cavu}(1/2) = 1$  to player 1.

### III.3 Player 2 can guarantee the limit value

In the infinitely repeated game with initial probability  $p$ , player 2 can play as follows:  $T$  being fixed, he can play an optimal strategy in the  $T$ -stage game  $\Gamma_T(p)$ , then forget everything and play again an optimal strategy in the  $T$ -stage game  $\Gamma_T(p)$ , etc... By doing so, he guarantees  $v_T(p)$  in the game  $\Gamma(p)$ . So he can guarantee  $\inf_T v_T(p)$  in this game, and this implies that  $\limsup_T v_T(p) \leq \inf_T v_T(p)$ . As a consequence, we obtain:

**Proposition 2.** The sequence  $(v_T(p))_T$  converges to  $\inf_T v_T(p)$ , and this limit can be guaranteed by player 2 in the game  $\Gamma(p)$ .

### III.4 Uniform value: $\text{cavu}$ theorem

We will see here that  $\lim_T v_T(p)$  is nothing but  $\text{cavu}(p)$ , and since this quantity can be guaranteed by both players this is the uniform value of the game  $\Gamma(p)$ . The idea of the proof is the following. The martingale  $(p_t(\sigma))_{t \geq 0}$  is bounded, hence will converge almost surely, and we have a bound on its  $L^1$  variation (see lemma 4 below). This means that after a certain stage the martingale will essentially remain constant, so approximately player 1 will play in a non revealing way, so

will not be able to have a stage payoff greater than  $u(q)$ , where  $q$  is a “limit *a posteriori*”. Since the expectation of the *a posteriori* is the *a priori*  $p$ , player 1 can not guarantee more than  $\max\{\sum_{s \in S} \lambda_s u(p_s), S \text{ finite}, \forall s \in S \lambda_s \geq 0, p_s \in \Delta(K), \sum_{s \in S} \lambda_s = 1, \sum_{s \in S} \lambda_s p_s = p\}$ , that is more than  $\text{cavu}(p)$ . Let us now proceed to the formal proof.

Fix a strategy  $\sigma^1$  of player 1, and define the strategy  $\sigma^2$  of player 2 as follows: play at each stage an optimal strategy in the matrix game  $G(p_t)$ , where  $p_t$  is the current *a posteriori* in  $\Delta(K)$ . Assume that  $\sigma = (\sigma^1, \sigma^2)$  is played in the repeated game  $\Gamma(p)$ . To simplify notations, we write  $\mathbb{P}$  for  $\mathbb{P}_{p,\sigma}$ ,  $p_t(h_t)$  for  $p_t(\sigma, h_t)$ , etc.. We use everywhere norms  $\|\cdot\|_1$ . To avoid confusion between variables and random variables in the following computations, we will use tildes to denote random variables, e.g.  $\tilde{k}$  will denote the random variable of the selected state.

**Lemma 4.**

$$\forall T \geq 1, \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(\|p_{t+1} - p_t\|) \leq \frac{\sum_{k \in K} \sqrt{p^k(1-p^k)}}{\sqrt{T}}.$$

**Proof:** This is a property of martingales with values in  $\Delta(K)$  and expectation  $p$ . We have for each state  $k$  and  $t \geq 0$ :  $\mathbb{E}((p_{t+1}^k - p_t^k)^2) = \mathbb{E}(\mathbb{E}((p_{t+1}^k - p_t^k)^2 | \mathcal{H}_t))$ , where  $\mathcal{H}_t$  is the  $\sigma$ -algebra on plays generated by the first  $t$  action profiles. So  $\mathbb{E}((p_{t+1}^k - p_t^k)^2) = \mathbb{E}(\mathbb{E}((p_{t+1}^k)^2 + (p_t^k)^2 - 2p_{t+1}^k p_t^k | \mathcal{H}_t)) = \mathbb{E}((p_{t+1}^k)^2) - \mathbb{E}((p_t^k)^2)$ . So  $\mathbb{E}(\sum_{t=0}^{T-1} (p_{t+1}^k - p_t^k)^2) = \mathbb{E}((p_T^k)^2) - (p^k)^2 \leq p^k(1-p^k)$ . By Cauchy-Schwartz inequality, we also have for each  $k$ ,

$$\mathbb{E}\left(\frac{1}{T} \sum_{t=0}^{T-1} |p_{t+1}^k - p_t^k|\right) \leq \sqrt{\frac{1}{T} \mathbb{E}\left(\sum_{t=0}^{T-1} (p_{t+1}^k - p_t^k)^2\right)}$$
 and the result follows.  $\square$

For  $h_t$  in  $(I \times J)^t$ ,  $\sigma_{t+1}^1(k, h_t)$  is the mixed action in  $\Delta(I)$  played by player 1 at stage  $t+1$  if the state is  $k$  and  $h_t$  has previously occurred, and we write  $\bar{\sigma}_{t+1}^1(h_t)$  for the law of the action of player 1 of stage  $t+1$  after  $h_t$ :  $\bar{\sigma}_{t+1}^1(h_t) = \sum_{k \in K} p_t^k(h_t) \sigma_{t+1}^1(k, h_t) \in \Delta(I)$ .  $\bar{\sigma}_{t+1}^1(h_t)$  can be seen as the average action played by player 1 after  $h_t$ , and will be used as a non revealing approximation for  $(\sigma_{t+1}^1(k, h_t))_k$ . The next lemma precisely links the variation of the martingale  $(p_t(\sigma))_{t \geq 0}$ , i.e. the information revealed by player 1, and the dependence of player 1's action on the selected state, i.e. the information used by player 1.

**Lemma 5.**

$$\forall t \geq 0, \forall h_t \in (I \times J)^t, \mathbb{E}(\|p_{t+1} - p_t\| | h_t) = \mathbb{E}\left(\left\| \sigma_{t+1}^{\tilde{k}}(h_t) - \bar{\sigma}_{t+1}^1(h_t) \right\| | h_t\right).$$

**Proof:** Fix  $t \geq 0$  and  $h_t$  in  $(I \times J)^t$  s.t.  $\mathbb{P}_{p,\sigma}(h_t) > 0$ . For  $(i_{t+1}, j_{t+1})$  in  $I \times J$ ,

one has:

$$\begin{aligned}
p_{t+1}^k(h_t, i_{t+1}, j_{t+1}) &= \mathbb{P}(\tilde{k} = k | h_t, i_{t+1}) \\
&= \frac{\mathbb{P}(\tilde{k} = k | h_t) \mathbb{P}(i_{t+1} | k, h_t)}{\mathbb{P}(i_{t+1} | h_t)} \\
&= \frac{p_t^k(h_t) \sigma_{t+1}^1(k, h_t)(i_{t+1})}{\bar{\sigma}_{t+1}^1(h_t)(i_{t+1})}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathbb{E}(\|p_{t+1} - p_t\| | h_t) &= \sum_{i_{t+1} \in I} \bar{\sigma}_{t+1}^1(h_t)(i_{t+1}) \sum_{k \in K} |p_{t+1}^k(h_t, i_{t+1}) - p_t^k(h_t)| \\
&= \sum_{i_{t+1} \in I} \sum_{k \in K} |p_t^k(h_t) \sigma_{t+1}^1(k, h_t)(i_{t+1}) - \bar{\sigma}_{t+1}^1(h_t)(i_{t+1}) p_t^k(h_t)| \\
&= \sum_{k \in K} p_t^k(h_t) \|\sigma_{t+1}^1(k, h_t) - \bar{\sigma}_{t+1}^1(h_t)\| \\
&= \mathbb{E}(\|\sigma_{t+1}^1(\tilde{k}, h_t) - \bar{\sigma}_{t+1}^1(h_t)\| | h_t).
\end{aligned}$$

□

We can now control payoffs. For  $t \geq 0$  and  $h_t$  in  $(I \times J)^t$ :

$$\begin{aligned}
\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1}, \tilde{j}_{t+1}) | h_t\right) &= \sum_{k \in K} p_t^k(h_t) G^k(\sigma_{t+1}^1(k, h_t), \sigma_{t+1}^2(h_t)) \\
&\leq \sum_{k \in K} p_t^k(h_t) G^k(\bar{\sigma}_{t+1}^1(h_t), \sigma_{t+1}^2(h_t)) \\
&\quad + M \sum_{k \in K} p_t^k(h_t) \|\sigma_{t+1}^1(k, h_t) - \bar{\sigma}_{t+1}^1(h_t)\| \\
&\leq u(p_t(h_t)) + M \sum_{k \in K} p_t^k(h_t) \|\sigma_{t+1}^1(k, h_t) - \bar{\sigma}_{t+1}^1(h_t)\|,
\end{aligned}$$

where  $u(p_t(h_t))$  comes from the definition of  $\sigma^2$ . By lemma 5, we get:

$$\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1}, \tilde{j}_{t+1}) | h_t\right) \leq u(p_t(h_t)) + M \mathbb{E}(\|p_{t+1} - p_t\| | h_t).$$

Applying Jensen's inequality yields:

$$\mathbb{E}\left(G^{\tilde{k}}(\tilde{i}_{t+1}, \tilde{j}_{t+1})\right) \leq \text{cav}u(p) + M \mathbb{E}(\|p_{t+1} - p_t\|).$$

We now apply lemma 4 and obtain:

$$\begin{aligned}
\gamma_T^{1,p}(\sigma^1, \sigma^2) &= \mathbb{E}\left(\frac{1}{T} \sum_{t=0}^{T-1} G^{\tilde{k}}(\tilde{i}_{t+1}, \tilde{j}_{t+1})\right) \\
&\leq \text{cav}u(p) + \frac{M}{\sqrt{T}} \sum_{k \in K} \sqrt{p^k(1-p^k)}.
\end{aligned}$$

This is true for any strategy  $\sigma^1$  of player 1. Considering the case of an optimal strategy for player 1 in the  $T$ -stage game  $\Gamma_T(p)$ , we have shown:

**Proposition 3.** For  $p$  in  $\Delta(K)$  and  $T \geq 1$ ,

$$v_T(p) \leq \text{cavu}(p) + \frac{M \sum_{k \in K} \sqrt{p^k(1-p^k)}}{\sqrt{T}}.$$

It remains to conclude about the existence of the uniform value. We have seen that player 1 can guarantee  $\text{cavu}(p)$ , that player 2 can guarantee  $\lim_T v_T(p)$ , and we obtain from proposition 3 that  $\lim_T v_T(p) \leq \text{cavu}(p)$ . This is enough to deduce Aumann and Maschler's celebrated "cavu" theorem.

**Theorem 1 (Aumann and Maschler [87]).** The game  $\Gamma(p)$  has a uniform value which is  $\text{cavu}(p)$ .

### III.5 $T$ -stage values and the recursive formula

As the  $T$ -stage game is a zero-sum game with incomplete information where player 1 is informed, we can write:

$$\begin{aligned} v_T(p) &= \inf_{\sigma^2 \in \Sigma^2} \sup_{\sigma^1 \in \Sigma^1} \gamma_T^{1,p}(\sigma), \\ &= \inf_{\sigma^2 \in \Sigma^2} \sup_{\sigma^1 \in \Sigma^1} \sum_{k \in K} p^k \gamma_T^{1,\delta_k}(\sigma), \\ &= \inf_{\sigma^2 \in \Sigma^2} \sum_{k \in K} p^k \left( \sup_{\sigma^1 \in \Sigma^1} \gamma_T^{1,\delta_k}(\sigma) \right). \end{aligned}$$

This shows that  $v_T$  is the infimum of a family of affine functions of  $p$ , hence is a concave function of  $p$ . This concavity represents the advantage of player 1 to possess the information on the selected state. Clearly, we have  $v_T(p) \geq u(p)$ , hence we get the inequalities:  $\forall T \geq 1, \text{cavu}(p) \leq v_T(p) \leq \text{cavu}(p) + \frac{M \sum_{k \in K} \sqrt{p^k(1-p^k)}}{\sqrt{T}}$ .

It is also easy to prove that the  $T$ -stage value functions satisfy the following recursive formula:

$$\begin{aligned} v_{T+1}(p) &= \frac{1}{T+1} \max_{x \in \Delta(I)^K} \min_{y \in \Delta(J)} \left( G(p, x, y) + T \sum_{i \in I} x(p)(i) v_T(\hat{p}(x, i)) \right), \\ &= \frac{1}{T+1} \min_{y \in \Delta(J)} \max_{x \in \Delta(I)^K} \left( G(p, x, y) + T \sum_{i \in I} x(p)(i) v_T(\hat{p}(x, i)) \right), \end{aligned}$$

where  $x = (x^k(i))_{i \in I, k \in K}$ , with  $x^k$  the mixed action used at stage 1 by player 1 if the state is  $k$ ,  $G(p, x, y) = \sum_{k,i,j} p^k G^k(x^k(i), y(j))$  is the expected payoff of stage 1,  $x(p)(i) = \sum_{k \in K} p^k x^k(i)$  is the probability that action  $i$  is played at stage 1, and  $\hat{p}(x, i)$  is the conditional probability on  $K$  given  $i$ .

The next property interprets easily: the advantage of the informed player can only decrease as the number of stages increases (for a proof, one can show that  $v_{T+1} \leq v_T$  by induction on  $T$ , using the concavity of  $v_T$ ).

**Lemma 6.** The  $T$ -stage value  $v_T(p)$  is non increasing in  $T$ .

## IV. Vector Payoffs and Approachability

The following model has been introduced by D. Blackwell ([7]) and is, strictly speaking, not part of the general definition given in section I. We still have a family of  $I \times J$  matrices  $(G^k)_{k \in K}$ , where  $K$  is a finite set of parameters. At each stage  $t$ , simultaneously player 1 chooses  $i_t \in I$  and player 2 chooses  $j_t \in J$ , and the stage “payoff” is the full vector  $G(i_t, j_t) = (G^k(i_t, j_t))_{k \in K}$  in  $\mathbb{R}^K$ . Notice that there is no initial probability or true state of nature here, and both players have a symmetric role. We assume here that after each stage both players observe exactly the stage vector payoff (but one can check that assuming that the action profiles are observed wouldn’t change the results). A natural question is then to determine the sets  $C$  in  $\mathbb{R}^K$  such that player 1 (for example) can force the average long term payoff to belong to  $C$ ? Such sets will be called *approachable* by player 1.

In section IV, we use Euclidean distances and norms. Denote by  $F = \{(G^k(i, j))_{k \in K}, i \in I, j \in J\}$  the finite set of possible stage payoffs, and by  $M$  a constant such that  $\|u\| \leq M$  for each  $u$  in  $F$ . A strategy for player 1, resp. player 2, is an element  $\sigma = (\sigma_t)_{t \geq 1}$ , where  $\sigma_t$  maps  $F^{t-1}$  into  $\Delta(I)$ , resp.  $\Delta(J)$ . Strategy spaces for player 1 and 2 are respectively denoted by  $\Sigma$  and  $\mathcal{T}$ . A strategy profile  $(\sigma, \tau)$  naturally induces a unique probability on  $(I \times J \times F)^\infty$  denoted by  $\mathbb{P}_{\sigma, \tau}$ . Let  $C$  be a “target” set, that will always be assumed, without loss of generality, a closed subset of  $\mathbb{R}^K$ . We denote by  $g_t$  the random variable, with value in  $F$ , of the payoff of stage  $t$ , and we use  $\bar{g}_t = \frac{1}{t} \sum_{t'=1}^t g_{t'} \in \text{conv}(F)$ , and finally  $d_t = d(\bar{g}_t, C)$  for the distance from  $\bar{g}_t$  to  $C$ .

**Definition 3:**  $C$  is approachable by player 1 if :  $\forall \varepsilon > 0, \exists \sigma \in \Sigma, \exists T, \forall \tau \in \mathcal{T}, \forall t \geq T, \mathbb{E}_{\sigma, \tau}(d_t) \leq \varepsilon$ .  $C$  is excludable by player 1 if there exist  $\delta > 0$  such that  $\{z \in \mathbb{R}^K, d(z, C) \geq \delta\}$  is approachable by player 1.

Approachability and excludability for player 2 are defined similarly.  $C$  is approachable by player 1 if for each  $\varepsilon > 0$ , this player can force that for  $t$  large we have  $\mathbb{E}_{\sigma, \tau}(d_t) \leq \varepsilon$ , so the average payoff will be  $\varepsilon$ -close to  $C$  with high probability. A set cannot be approachable by a player as well as excludable by the other player. In the usual case where  $K$  is a singleton, we are in dimension 1 and the Minmax theorem implies that for each  $t$ , the interval  $[t, +\infty[$  is either approachable by player 1, or excludable by player 2, depending on the comparison between

$t$  and the value  $\max_{x \in \Delta(I)} \min_{y \in \Delta(J)} G(x, y) = \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} G(x, y)$ .

#### IV.1. Necessary and sufficient conditions for Approachability

Given a mixed action  $x$  in  $\Delta(I)$ , we write  $xG$  for the set of possible vector pay-offs when player 1 uses  $x$ , i.e.  $xG = \{G(x, y), y \in \Delta(J)\} = \text{conv} \{\sum_{i \in I} x_i G(i, j), j \in J\}$ . Similarly, we write  $Gy = \{G(x, y), x \in \Delta(I)\}$  for  $y$  in  $\Delta(J)$ .

**Definition 4:** The set  $C$  is a  $B$ (lackwell)-set for player 1 if for every  $z \notin C$ , there exists  $z' \in C$  and  $x \in \Delta(I)$  such that: (i)  $\|z' - z\| = d(z, C)$ , and (ii) the hyperplane containing  $z'$  and orthogonal to  $[z, z']$  separates  $z$  from  $xG$ .

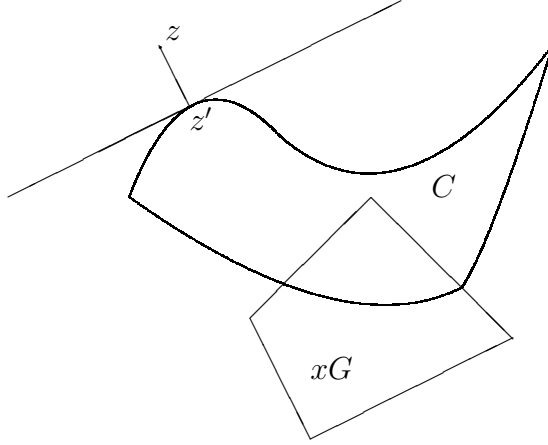


Figure 3: The Blackwell property.

For example, any set  $xG$ , with  $x$  in  $\Delta(I)$ , is a  $B$ -set for player 1. Given a  $B$ -set for player 1, we now construct a strategy  $\sigma$  adapted to  $C$  as follows. At each positive stage  $t + 1$ , player 1 considers the current average payoff  $\bar{g}_t$ . If  $\bar{g}_t \in C$ , or if  $t = 0$ ,  $\sigma$  plays arbitrarily at stage  $t + 1$ . Otherwise,  $\sigma$  plays at stage  $t + 1$  a mixed action  $x$  satisfying the previous definition for  $z = \bar{g}_t$ .

**Theorem 2** If  $C$  is a  $B$ -set for player 1, a strategy  $\sigma$  adapted to  $C$  satisfies:

$$\forall \tau \in \mathcal{T}, \forall t \geq 1 \quad \mathbb{E}_{\sigma, \tau}(d_t) \leq \frac{2M}{\sqrt{t}} \quad \text{and} \quad d_t \xrightarrow{t \rightarrow \infty} 0 \quad \mathbb{P}_{\sigma, \tau} \text{ a.s.}$$

As an illustration, in dimension 1 and for  $C = \{0\}$ , this theorem implies that a bounded sequence  $(x_t)_t$  of reals, such that the product  $x_{T+1} \left( \frac{1}{T} \sum_{t=1}^T x_t \right)$  is non-positive for each  $T$ , Cesaro converges to zero.

**Proof:** Assume that player 1 plays  $\sigma$  adapted to  $C$ , whereas player 2 plays some strategy  $\tau$ . Fix  $t \geq 1$ , and assume that  $\bar{g}_t \notin C$ . Consider  $z' \in C$  and  $x \in \Delta(I)$

satisfying (i) and (ii) of definition 4 for  $z = \bar{g}_t$ . We have:

$$\begin{aligned}
d_{t+1}^2 = d(\bar{g}_{t+1}, C)^2 &\leq \|\bar{g}_{t+1} - z'\|^2 \\
&= \left\| \frac{1}{t+1} \sum_{l=1}^{t+1} g_l - z' \right\|^2 \\
&= \left\| \frac{1}{t+1} (g_{t+1} - z') + \frac{t}{t+1} (\bar{g}_t - z') \right\|^2 \\
&= \left( \frac{1}{t+1} \right)^2 \|g_{t+1} - z'\|^2 + \left( \frac{t}{t+1} \right)^2 d_t^2 \\
&\quad + \frac{2t}{(t+1)^2} \langle g_{t+1} - z', \bar{g}_t - z' \rangle.
\end{aligned}$$

By hypothesis, the expectation, given the first  $t$  action profiles  $h_t \in (I \times J)^t$ , of the above scalar product is non-positive, so  $\mathbb{E}(d_{t+1}^2 | h_t) \leq \left(\frac{t}{t+1}\right)^2 d_t^2 + \left(\frac{1}{t+1}\right)^2 \mathbb{E}(\|g_{t+1} - z'\|^2 | h_t)$ . Since  $\mathbb{E}(\|g_{t+1} - z'\|^2 | h_t) \leq \mathbb{E}(\|g_{t+1} - \bar{g}_t\|^2 | h_t) \leq (2M)^2$ , we have:

$$\mathbb{E}(d_{t+1}^2 | h_t) \leq \left(\frac{t}{t+1}\right)^2 d_t^2 + \left(\frac{1}{t+1}\right)^2 4M^2. \quad (2)$$

Taking the expectation, we get, whether  $\bar{g}_t \notin C$  or not:  $\forall t \geq 1$ ,  $\mathbb{E}(d_{t+1}^2) \leq \left(\frac{t}{t+1}\right)^2 \mathbb{E}(d_t^2) + \left(\frac{1}{t+1}\right)^2 4M^2$ . By induction, we obtain that for each  $t \geq 1$ ,  $\mathbb{E}(d_t^2) \leq \frac{4M^2}{t}$ , and  $\mathbb{E}(d_t) \leq \frac{2M}{\sqrt{t}}$ .

Put now, as in Sorin 2002 ([92]),  $e_t = d_t^2 + \sum_{t' > t} \frac{4M^2}{t'^2}$ . Inequality (2) gives  $\mathbb{E}(e_{t+1} | h_t) \leq e_t$ , so  $(e_t)$  is a non-negative supermartingale which expectation goes to zero. By a standard probability result, we obtain  $e_t \xrightarrow{t \rightarrow \infty} 0$   $\mathbb{P}_{\sigma, \tau}$  a.s., and finally  $d_t \xrightarrow{t \rightarrow \infty} 0$   $\mathbb{P}_{\sigma, \tau}$  a.s.  $\square$

This theorem implies that any  $B$ -set for player 1 is approachable by this player. The converse is true for convex sets.

**Theorem 3** Let  $C$  be a closed convex subset of  $\mathbb{R}^K$ .

- (i)  $C$  is a  $B$ -set for player 1,
- $\Leftrightarrow$  (ii)  $\forall y \in \Delta(J), Gy \cap C \neq \emptyset$ ,
- $\Leftrightarrow$  (iii)  $C$  is approachable by player 1,
- $\Leftrightarrow$  (iv)  $\forall q \in \mathbb{R}^K, \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k \in K} q^k G^k(x, y) \geq \inf_{c \in C} \langle q, c \rangle$ .

**Proof:** The implication (i)  $\Rightarrow$  (iii) comes from theorem 2. Proof of (iii)  $\Rightarrow$  (ii): assume there exists  $y \in \Delta(J)$  such that  $Gy \cap C = \emptyset$ . Since  $Gy$  is approachable by player 2, then  $C$  is excludable by player 2 and thus  $C$  is not approachable by player 1. Proof of (ii)  $\Rightarrow$  (i): Assume that  $Gy \cap C \neq \emptyset \forall y \in \Delta(J)$ . Consider

$z \notin C$  and define  $z'$  as its projection onto  $C$ . Define the matrix game where payoffs are projected towards the direction  $z' - z$ , i.e. the matrix game  $\sum_{k \in K} (z'^k - z^k) G^k$ . By assumption, one has:  $\forall y \in \Delta(J), \exists x \in \Delta(I)$  such that  $G(x, y) \in C$ , hence such that:

$$\langle z' - z, G(x, y) \rangle \geq \min_{c \in C} \langle z' - z, c \rangle = \langle z' - z, z' \rangle.$$

So  $\min_{y \in \Delta(J)} \max_{x \in \Delta(I)} \langle z' - z, G(x, y) \rangle \geq \langle z' - z, z' \rangle$ . By the minmax theorem, there exists  $x$  in  $\Delta(I)$  such that  $\forall y \in \Delta(J), \langle z' - z, G(x, y) \rangle \geq \langle z' - z, z' \rangle$ , that is  $\langle z' - z, z' - G(x, y) \rangle \leq 0$ .

(iv) means that any half-space containing  $C$  is approachable by player 1. (iii)  $\implies$  (iv) is thus clear. (iv)  $\implies$  (i) is similar to (ii)  $\implies$  (i).  $\square$

Up to minor formulation differences, theorems 2 and 3 are due to Blackwell ([7]). More recently, X. Spinat ([82]) has proved the following characterization.

**Theorem 4** A closed set is approachable for player 1 if and only if it contains a  $B$ -set for player 1.

As a consequence, it shows that adding the condition  $d_t \xrightarrow{t \rightarrow \infty} 0$   $\mathbb{P}_{\sigma, \tau}$ -a.s. in the definition of approachability does not modify the notion.

## IV.2. Approachability for player 1 versus Excludability for player 2

As a corollary of theorem 3, we obtain that: *A closed convex set in  $\mathbb{R}^K$  is either approachable by player 1, or excludable by player 2.*

One can show that when  $K$  is a singleton, then any set is either approachable by player 1, or excludable by player 2. A simple example of a set which is neither approachable for player 1 nor excludable by player 2 is given in dimension 2 by:  $G = \begin{pmatrix} (0, 0) & (0, 0) \\ (1, 0) & (1, 1) \end{pmatrix}$ , and  $C = \{(1/2, v), 0 \leq v \leq 1/4\} \cup \{(1, v), 1/4 \leq v \leq 1\}$  (see [92]).

## IV.3. Weak Approachability

One can weaken the definition of approachability by giving up time uniformity.

**Definition 5:**  $C$  is weakly approachable by player 1 if:  $\forall \varepsilon > 0, \exists T, \forall t \geq T, \exists \sigma \in \Sigma, \forall \tau \in \mathcal{T}, \mathbb{E}_{\sigma, \tau}(d_t) \leq \varepsilon$ .  $C$  is weakly excludable by player 1 if there exists  $\delta > 0$  such that  $\{z \in \mathbb{R}^K, d(z, C) \geq \delta\}$  is weakly approachable by player 1.

N. Vieille ([83]) has proved, via the consideration of certain differential games:

**Theorem 5** A subset of  $\mathbb{R}^K$  is either weakly approachable by player 1 or weakly excludable by player 2.



#### IV.4. Back to the standard model

Let us come back to Aumann and Maschler's model with a finite family of matrices  $(G^k)_{k \in K}$ , and an initial probability  $p$  on  $\Delta(K)$ . By theorem 1, the repeated game  $\Gamma(p)$  has a uniform value which is  $\text{cavu}(p)$ , and Blackwell approachability will allow for the construction of an *explicit* optimal strategy for the uninformed player. Considering a hyperplane which is tangent to  $\text{cavu}$  at  $p$ , we can find a vector  $l$  in  $\mathbb{R}^K$  such that

$$\langle l, p \rangle = \text{cavu}(p) \text{ and } \forall q \in \Delta(K), \langle l, q \rangle \geq \text{cavu}(q) \geq u(q).$$

Define now the orthant  $C = \{z \in \mathbb{R}^K, z^k \leq l^k \forall k \in K\}$ . Recall that player 2 does not know the selected state, and an optimal strategy for him can not depend on player 1' strategy, and consequently on a martingale of a posteriori. He will play in a way such that player 1's long term payoff is, *simultaneously for each  $k$  in  $K$* , not greater than  $l^k$  if the state is  $k$ .

Fix  $q = (q^k)_k$  in  $\mathbb{R}^K$ . If there exists  $k$  with  $q^k > 0$ , we clearly have  $\inf_{c \in C} \langle q, c \rangle = -\infty \leq \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} \sum_{k \in K} q^k G^k(x, y)$ . Assume now that  $q^k \leq 0$  for each  $k$ , with  $q \neq 0$ . Write  $s = \sum_k (-q^k)$ .

$$\begin{aligned} \inf_{c \in C} \langle q, c \rangle &= \sum_{k \in K} q^k l^k \\ &= -s \langle l, \frac{-q}{s} \rangle \\ &\leq -s u\left(\frac{-q}{s}\right) \\ &\leq -s \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k \in K} \frac{-q^k}{s} G^k(x, y) \\ &= \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} \sum_{k \in K} q^k G^k(x, y) \end{aligned}$$

This is condition (iv) of theorem 3, adapted to player 2. So  $C$  is a  $B$ -set for player 2, and a strategy  $\tau$  adapted to  $C$  satisfies by theorem 2:  $\forall \sigma \in \Sigma, \forall k \in K$ ,

$$\mathbb{E}_{\sigma, \tau} \left( \frac{1}{T} \sum_{t=1}^T G^k(\tilde{i}_t, \tilde{j}_t) - l^k \right) \leq \mathbb{E}_{\sigma, \tau} \left( d \left( \frac{1}{T} \sum_{t=1}^T G^k(\tilde{i}_t, \tilde{j}_t), C \right) \right) \leq \frac{2M}{\sqrt{T}},$$

(where  $M$  is here an upper bound for the Euclidean norms of the vectors  $(G^k(i, j))_{k \in K}$ , with  $i \in I$  and  $j \in J$ .) So,

$$\begin{aligned} \gamma_T^{1,p}(\sigma, \tau) &= \sum_{k \in K} p^k \left( \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{\sigma, \tau}(G^k(\tilde{i}_t, \tilde{j}_t)) \right) \\ &\leq \langle p, l \rangle + \frac{2M}{\sqrt{T}} = \text{cavu}(p) + \frac{2M}{\sqrt{T}}. \end{aligned}$$

As shown by Kohlberg ([34]), the approachability strategy  $\tau$  is thus an optimal strategy for player 2 in the repeated game  $\Gamma(p)$ .

## V. Zero-sum games with lack of information on both sides

The following model has also been introduced by Aumann and Maschler ([87]). We are still in the context of zero-sum repeated games with observable actions, but it is no longer assumed that one of the players is fully informed. The set of states is here a product  $K \times L$  of finite sets, and we have a family of matrices  $(G^{k,l})_{(k,l) \in K \times L}$  with size  $I \times J$ , as well as initial probabilities  $p$  on  $K$ , and  $q$  on  $L$ . In the game  $\Gamma(p, q)$ , a state of nature  $(k, l)$  is first selected according to the product probability  $p \otimes q$ , then  $k$ , resp.  $l$ , is announced to player 1, resp. player 2 only. Then the matrix game  $G^{k,l}$  is repeated over and over: at every stage, simultaneously player 1 chooses a row  $i$  in  $I$ , whereas player 2 chooses a column  $j$  in  $J$ , the stage payoff for player 1 is  $G^{k,l}(i, j)$  but only  $i$  and  $j$  are publicly announced before proceeding to the next stage.

The average payoff for player 1 in the  $T$ -stage game is written:  $\gamma_T^{1,p,q}(\sigma^1, \sigma^2) = E_{\sigma^1, \sigma^2}^{p,q} \left( \frac{1}{T} \sum_{t=1}^T G^{\tilde{k}_t, \tilde{l}_t}(\tilde{i}_t, \tilde{j}_t) \right)$ , and the  $T$ -stage value is written  $v_T(p, q)$ . Similarly, the  $\lambda$ -discounted value of the game will be written  $v_\lambda(p, q)$

The non revealing game now corresponds to the case where player 1 plays independently of  $k$  and player 2 plays independently of  $l$ . Its value is denoted by:

$$u(p, q) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k,l} p^k q^l G^{k,l}(x, y). \quad (3)$$

Given a continuous function  $f : \Delta(K) \times \Delta(L) \rightarrow \mathbb{R}$ , we denote by  $\text{cav}_I f$  the concavification of  $f$  with respect to the first variable: for each  $(p, q)$  in  $\Delta(K) \times \Delta(L)$ ,  $\text{cav}_I f(p, q)$  is the value at  $p$  of the smallest concave function from  $\Delta(K)$  to  $\mathbb{R}$  which is above  $f(\cdot, q)$ . Similarly, we denote by  $\text{vex}_{II} f$  the convexification of  $f$  with respect to the second variable. It can be shown that  $\text{cav}_I f$  and  $\text{vex}_{II} f$  are continuous, and we can compose  $\text{cav}_I \text{vex}_{II} f$  and  $\text{vex}_{II} \text{cav}_I f$ . These functions are both concave in the first variable and convex in the second variable, and they satisfy  $\text{cav}_I \text{vex}_{II} f(p, q) \leq \text{vex}_{II} \text{cav}_I f(p, q)$ .

### V.1. Maxmin and Minmax of the repeated game

Theorem 1 generalizes as follows.

**Theorem 6** ([87]) In the repeated game  $\Gamma(p, q)$ , the greatest quantity which can be guaranteed by player 1 is  $\text{cav}_I \text{vex}_{II} u(p, q)$ , and the smallest quantity which can be guaranteed by player 2 is  $\text{vex}_{II} \text{cav}_I u(p, q)$ .

Aumann, Maschler and Stearns also showed that  $\text{cav}_I \text{vex}_{II} u(p, q)$  can be *defended* by player 2, uniformly in time, i.e. that  $\forall \varepsilon > 0, \forall \sigma^1, \exists T_0, \exists \sigma^2, \forall T \geq$

$T_0$ ,  $\gamma_T^{p,q}(\sigma^1, \sigma^2) \leq \text{cav}_I \text{vex}_{II} u(p, q)v + \varepsilon$ . Similarly,  $\text{vex}_{II} \text{cav}_I u(p, q)$  can be *defended* by player 1.

The proof uses the martingales of *a posteriori* of each player, and a useful notion is that of the informational content of a strategy: for a strategy  $\sigma^1$  of the first player, it is defined as:  $I(\sigma^1) = \sup_{\sigma^2} \mathbb{E}_{\sigma^1, \sigma^2}^{p,q} \left( \sum_{k \in K} \sum_{t=0}^{\infty} (p_{t+1}^k(\sigma^1) - p_t^k(\sigma^1))^2 \right)$ , where  $p_t(\sigma^1)$  is the *a posteriori* on  $K$  of player 2 after stage  $t$  given that player 1 uses  $\sigma^1$ . By linearity of the expectation, the supremum can be restricted to strategies of player 2 which are both pure and independent of  $l$ .

Theorem 6 implies that  $\text{cav}_I \text{vex}_{II} u(p, q) = \sup_{\sigma^1 \in \Sigma^1} \liminf_T \left( \inf_{\sigma^2 \in \Sigma^2} \gamma_T^{1,p,q}(\sigma^1, \sigma^2) \right)$ , and  $\text{cav}_I \text{vex}_{II} u(p, q)$  is called the *maxmin* of the repeated game  $\Gamma(p, q)$ . Similarly,  $\text{vex}_{II} \text{cav}_I u(p, q) = \inf_{\sigma^2 \in \Sigma^2} \limsup_T \left( \sup_{\sigma^1 \in \Sigma^1} \gamma_T^1(\sigma^1, \sigma^2) \right)$  is called the *minmax* of  $\Gamma(p, q)$ . As a corollary, we obtain that the repeated game  $\Gamma(p, q)$  has a uniform value if and only if:  $\text{cav}_I \text{vex}_{II} u(p, q) = \text{vex}_{II} \text{cav}_I u(p, q)$ . This is not always the case, and there exist counter-examples to the existence of the uniform value.

**Example 4:**  $K = \{a, a'\}$ , and  $L = \{b, b'\}$ , with  $p$  and  $q$  uniform.

$$\begin{aligned} G^{a,b} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix} & G^{a,b'} &= \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ G^{a',b} &= \begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} & G^{a',b'} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \end{aligned}$$

Mertens and Zamir ([51]) have shown that here,  $\text{cav}_I \text{vex}_{II} u(p, q) = -\frac{1}{4} < 0 = \text{vex}_{II} \text{cav}_I u(p, q)$ .

## V.2. Limit values

It is easy to see that for each  $T$  and  $\lambda$ , the value functions  $v_T$  and  $v_\lambda$  are concave in the first variable, and convex in the second variable. They are all Lipschitz functions, with the same constant  $M = \max_{i,j,k,l} |G^{k,l}(i, j)|$ , and here also, recursive formulae can be given. In the following result,  $v_T$  and  $v_\lambda$  are viewed as elements of the set  $\mathcal{C}$  of continuous mappings from  $\Delta(K) \times \Delta(L)$  to  $\mathbb{R}$ .

**Theorem 7 (Mertens and Zamir ([51]))**  $(v_T)_T$ , as  $T$  goes to infinity, and  $(v_\lambda)_\lambda$ , as  $\lambda$  goes to zero, both uniformly converge to the unique solution  $f$  of the following system:

$$\begin{cases} f &= \text{vex}_{II} \max\{u, f\} \\ f &= \text{cav}_I \min\{u, f\} \end{cases}$$

The above system can be fruitfully studied without reference to repeated games (see [54], [78], [38], [39]).

Remark: Let  $U$  be the set of all non revealing value functions, i.e. of functions from  $\Delta(K) \times \Delta(L)$  to  $\mathbb{R}$  satisfying equation (3) for some family of matrices  $(G^{k,l})_{k,l}$ . One can easily show that any mapping in  $\mathcal{C}$  is a uniform limit of elements in  $U$ .

### V.3. Correlated initial information

A more general model can be written, where it is no longer assumed that the initial information of the players are independent. The set of states is now denoted by  $R$  (instead of  $K \times L$ ), initially a state  $r$  in  $R$  is chosen according to a known probability  $p = (p^r)_{r \in R}$ , and each player receives a deterministic signal depending on  $r$ . Equivalently, each player  $i$  has a partition  $R^i$  of  $R$  and observes the element of his partition which contains the selected state.

After the first stage, player 1 will play an action  $x = (x^r)_{r \in R}$  which is measurable with respect to  $R^1$ , i.e.  $(r \rightarrow x^r)$  is constant on each atom of  $R^1$ . After having observed player 1's action at the first stage, the conditional probability on  $R$  necessarily belongs to the set:

$$\Pi^I(p) = \left\{ (\alpha^r p^r)_{r \in R}, \forall r \alpha^r \geq 0, \sum_r \alpha^r p^r = 1 \text{ and } (\alpha^r)_r \text{ is } R^1\text{-measurable} \right\}.$$

$\Pi^I(p)$  contains  $p$ , and is a convex compact subset of  $\Delta(R)$ . A mapping  $f$  from  $\Delta(R)$  to  $\mathbb{R}$  is now said to be I-concave if for each  $p$  in  $\Delta(R)$ , the restriction of  $f$  to  $\Pi^I(p)$  is concave. And given  $g : \Delta(R) \rightarrow \mathbb{R}$  which is bounded from above, we define the concavification  $\text{cav}_I g$  as the smallest function above  $g$  which is I-concave. Similarly one can define the set  $\Pi^{II}(p)$  and the notions of II-convexity and II-convexification. With these generalized definitions, the results of theorem 6 and 7 perfectly extend ([51]).

## VI. Non zero-sum games with lack of information on one side

We now consider the generalization of the standard model of section III to the non-zero sum case. Hence two players infinitely repeat the same bimatrix game, with player 1 only knowing the bimatrix. Formally, we have a finite set of states  $K$ , an initial probability  $p$  on  $K$ , and families of  $I \times J$ -payoff matrices  $(A^k)_{k \in K}$  and  $(B^k)_{k \in K}$ . Initially, a state  $k$  in  $K$  is selected according to  $p$ , and announced to player 1 only. Then the bimatrix game  $(A^k, B^k)$  is repeated over and over: at every stage, simultaneously player 1 chooses a row  $i$  in  $I$ , whereas player 2 chooses a column  $j$  in  $J$ , the stage payoff for player 1 is then  $A^k(i, j)$ , the stage payoff for player 2 is  $B^k(i, j)$ , but only  $i$  and  $j$  are publicly announced before proceeding to the next stage. Without loss of generality, we assume that  $p^k > 0$  for each  $k$ , and that each player has at least 2 actions.

Given a strategy pair  $(\sigma^1, \sigma^2)$ , it is here convenient to denote the expected payoffs up to stage  $T$  by:

$$\alpha_T^p(\sigma^1, \sigma^2) = \mathbb{E}_{p, \sigma^1, \sigma^2} \left( \frac{1}{T} \sum_{t=1}^T A^{\tilde{k}}(\tilde{i}_t, \tilde{j}_t) \right) = \sum_{k \in K} p^k \alpha_T^k(\sigma^1, \sigma^2).$$

$$\beta_T^p(\sigma^1, \sigma^2) = \mathbb{E}_{p, \sigma^1, \sigma^2} \left( \frac{1}{T} \sum_{t=1}^T B^{\tilde{k}}(\tilde{i}_t, \tilde{j}_t) \right) = \sum_{k \in K} p^k \beta_T^k(\sigma^1, \sigma^2).$$

Given a probability  $q$  on  $K$ , we write  $A(q) = \sum_k q^k A^k$ ,  $B(q) = \sum_k q^k B^k$ ,  $u(q) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} A(q)(x, y)$  and  $v(q) = \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} B(q)(x, y)$ . If  $\gamma = (\gamma(i, j))_{(i, j) \in I \times J} \in \Delta(I \times J)$ , we put  $A(q)(\gamma) = \sum_{(i, j) \in I \times J} \gamma(i, j) A(q)(i, j)$  and similarly  $B(q)(\gamma) = \sum_{(i, j) \in I \times J} \gamma(i, j) B(q)(i, j)$ .

## VI.1 Existence of Equilibria

The question of existence of an equilibrium has remained unsolved for long. Sorin ([76]) proved the existence of an equilibrium for two states of nature, and the general case has been solved by Simon *et al.* ([73]).

Exactly as in the zero-sum case, a strategy pair  $\sigma$  induces a sequence of *a posteriori*  $(p_t(\sigma))_{t \geq 0}$  which is a  $\mathbb{P}_{p, \sigma}$ -martingale with values in  $\Delta(K)$ . We will concentrate on the cases where this martingale moves only once.

**Definition 6:** A joint plan is a triple  $(S, \lambda, \gamma)$ , where:

- $S$  is a finite non empty set (of messages),
- $\lambda = (\lambda^k)_{k \in K}$  (signalling strategy) with for each  $k$ ,  $\lambda^k \in \Delta(S)$  and for each  $s$ ,  $\lambda_s =_{\text{def}} \sum_{k \in K} p^k \lambda_s^k > 0$ ,
- $\gamma = (\gamma_s)_{s \in S}$  (contract) with for each  $s$ ,  $\gamma_s \in \Delta(I \times J)$ .

The idea is due to Aumann, Maschler and Stearns. Player 1 observes  $k$ , then chooses  $s \in S$  according to  $\lambda^k$  and announces  $s$  to player 2. Then the players play pure actions corresponding to the frequencies  $\gamma_s(i, j)$ , for  $i$  in  $I$  and  $j$  in  $J$ . Given a joint plan  $(S, \lambda, \gamma)$ , we define:

- $\forall s \in S$ ,  $p_s = (p_s^k)_{k \in K} \in \Delta(K)$ , with  $p_s^k = \frac{p^k \lambda_s^k}{\lambda_s}$  for each  $k$ .  $p_s$  is the *a posteriori* on  $K$  given  $s$ .
- $\varphi = (\varphi^k)_{k \in K} \in \mathbb{R}^K$ , with for each  $k$ ,  $\varphi^k = \max_{s \in S} A^k(\gamma_s)$ .
- $\forall s \in S$ ,  $\psi_s = B(p_s)(\gamma_s)$  and  $\psi = \sum_{k \in K} p^k \sum_{s \in S} \lambda_s^k B^k(\gamma_s) = \sum_{s \in S} \lambda_s \psi_s$ .

**Definition 7:** A joint plan  $(S, \lambda, \gamma)$  is an equilibrium joint plan if:

- (i)  $\forall s \in S$ ,  $\psi_s \geq \text{vex} v(p_s)$ ,
- (ii)  $\forall k \in K$ ,  $\forall s \in S$  s.t.  $p_s^k > 0$ ,  $A^k(\gamma_s) = \varphi^k$ , and
- (iii)  $\forall q \in \Delta(K)$ ,  $\langle \varphi, q \rangle \geq u(q)$ .

Condition (ii) can be seen as an incentive condition for player 1 to choose  $s$  according to  $\lambda^k$ . Given an equilibrium joint plan  $(S, \lambda, \gamma)$ , one define a strategy pair  $(\sigma^{1*}, \sigma^{2*})$  adapted to it. For each message  $s$ , first fix a sequence  $(i_t^s, j_t^s)_{t \geq 1}$  of elements in  $I \times J$  such that for each  $(i, j)$ , the empirical frequencies converge to the corresponding probability:  $\frac{1}{T} |\{t, 1 \leq t \leq T, (i_t^s, j_t^s) = (i, j)\}| \xrightarrow{T \rightarrow \infty} \gamma_s(i, j)$ . We also fix an injective mapping  $f$  from  $S$  to  $I^l$ , where  $l$  is large enough, corresponding to a code between the players to announce an element in  $S$ .  $\sigma^{1*}$  is precisely defined as follows. Player 1 observes the selected state  $k$ , then chooses  $s$  according to  $\lambda^k$ , and announces  $s$  to player 2 by playing  $f(s)$  at the first  $l$  stages. Finally,  $\sigma^{1*}$  plays  $i_t^s$  at each stage  $t > l$  as long as player 2 plays  $j_t^s$ . If at some stage  $t > l$  player 2 does not play  $j_t^s$  then player 1 punishes his opponent by playing an optimal strategy in the zero-sum game with initial probability  $p_s$  and payoffs for player 1 given by  $(-B^k)_{k \in K}$ . We now define  $\sigma^{2*}$ . Player 2 arbitrarily plays at the beginning of the game, then compute at the end of stage  $l$  the message  $s$  sent by player 1. Next he plays at each stage  $t > l$  the action  $j_t^s$  as long as player 1 plays  $i_t^s$ . If at some stage  $t > l$ , player 1 does not play  $i_t^s$ , or if the first  $l$  actions of player 1 correspond to no message, then player 2 plays a punishing strategy  $\bar{\sigma}^2$  such that:  $\forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \forall k \in K, \alpha_T^k(\sigma^1, \bar{\sigma}^2) \leq \varphi^k + \varepsilon$ . Such a strategy  $\bar{\sigma}^2$  exists because of condition (iii): it is an approachability strategy for player 2 of the orthant  $\{x \in \mathbb{R}^K, \forall k \in K x^k \leq \varphi^k\}$  (see section IV.4).

**Lemma 7 ([76]):** A strategy pair adapted to an equilibrium joint plan is a uniform equilibrium of the repeated game.

**Proof:** The payoffs induced by  $(\sigma^{1*}, \sigma^{2*})$  can be easily computed:  $\forall k, \alpha_T^k(\sigma^{1*}, \sigma^{2*}) \xrightarrow{T \rightarrow \infty} \sum_{s \in S} \lambda_s^k A^k(\gamma_s) = \varphi^k$  because of (ii), and  $\beta_T^p(\sigma^{1*}, \sigma^{2*}) \xrightarrow{T \rightarrow \infty} \sum_{k \in K} p^k \sum_{s \in S} \lambda_s^k B^k(\gamma_s) = \psi$ . Assume that player 2 plays  $\sigma^{2*}$ . The existence of  $\bar{\sigma}^2$  implies that no detectable deviation of player 1 is profitable, so if the state is  $k$ , player 1 will gain no more than  $\max_{s' \in S} A^k(\gamma_{s'})$ . But this is just  $\varphi^k$ . The proof can be made uniform in  $\sigma^1$  and we obtain:  $\forall \varepsilon > 0 \exists T_0 \forall T \geq T_0, \forall k \in K, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq \varphi^k + \varepsilon$ . Finally assume that player 1 plays  $\sigma^{1*}$ . Condition (i) implies that if player 2 uses  $\sigma^{2*}$ , the payoff of this player will be at least  $v(p_s)$  if the message is  $s$ . Since  $\text{vex } v(p_s) (= -\text{cav}(-v(p_s)))$  is the value, from the point of view of player 2 with payoffs  $(B^k)_k$ , of the zero-sum game with initial probability  $p_s$ , player 2 fears the punition by player 1, and  $\forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^p(\sigma^{1*}, \sigma^2) \leq \sum_{s \in S} \lambda_s \psi_s + \varepsilon = \psi + \varepsilon$ .  $\square$

To prove the existence of equilibria, we then look for equilibrium joint plans. The first idea is to consider, for each probability  $r$  on  $K$ , the set of payoff vectors  $\varphi$  compatible with  $r$  being an *a posteriori*. This leads to the consideration of the following correspondence (for each  $r$ ,  $\Phi(r)$  is a subset of  $\mathbb{R}^K$ ):

$$\begin{aligned} \Phi : \Delta(K) &\rightrightarrows \mathbb{R}^K \\ r &\mapsto \{(A^k(\gamma))_{k \in K}, \text{ where } \gamma \in \Delta(I \times J) \text{ satisfies } B(r)(\gamma) \geq \text{vex } v(r)\} \end{aligned}$$

It is easy to see that the graph of  $\Phi$ , i.e. the set  $\{(r, \varphi) \in \Delta(K) \times \mathbb{R}^K, \varphi \in \Phi(r)\}$ , is compact, that  $\Phi$  has non empty convex values, and satisfies:  $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), \langle \varphi, q \rangle \geq u(q)$ .

Assume now that one can find a finite family  $(p_s)_{s \in S}$  of probabilities on  $K$ , as well as vectors  $\varphi$  and, for each  $s$ ,  $\varphi_s$  in  $\mathbb{R}^K$  such that: 1)  $p \in \text{conv} \{p_s, s \in S\}$ , 2)  $\langle \varphi, q \rangle \geq u(q) \forall q \in \Delta(K)$ , 3)  $\forall s \in S, \varphi_s \in \Phi(p_s)$ , and 4)  $\forall s \in S, \forall k \in K, \varphi_s^k \leq \varphi^k$  with equality if  $p_s^k > 0$ . It is then easy to construct an equilibrium joint plan. Thus we get interested in proving the following result.

**Proposition 4:** Let  $p$  be in  $\Delta(K)$ ,  $u : \Delta(K) \rightarrow \mathbb{R}$  be a continuous mapping, and  $\Phi : \Delta(K) \rightrightarrows \mathbb{R}^K$  be a correspondence with compact graph and non empty convex values such that:  $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), \langle \varphi, q \rangle \geq u(q)$ . Then there exists a finite family  $(p_s)_{s \in S}$  of elements of  $\Delta(K)$ , as well as vectors  $\varphi$  and, for each  $s$ ,  $\varphi_s$  in  $\mathbb{R}^K$  such that:

- $p \in \text{conv} \{p_s, s \in S\}$ ,
- $\langle \varphi, q \rangle \geq u(q) \forall q \in \Delta(K)$ ,
- $\forall s \in S, \varphi_s \in \Phi(p_s)$ ,
- $\forall s \in S, \forall k \in K, \varphi_s^k \leq \varphi^k$  with equality if  $p_s^k > 0$ .

The proof of proposition 4 relies, as explained in [59] or [72], on a fixed point theorem of Borsuk-Ulam type proved by Simon, Spieź and Toruńczyk ([73]) via tools from algebraic geometry. A simplified version of this fixed point theorem can be written as follows:

**Theorem 8 ([73]):** Let  $C$  be a compact subset of an  $n$ -dimensional Euclidean space,  $x \in C$  and  $Y$  be a finite union of affine subspaces of dimension  $n - 1$  of an Euclidean space. Let  $F$  be a correspondence from  $C$  to  $Y$  with compact graph and non empty convex values. Then there exists  $L \subset \partial C$  and  $y \in Y$  such that:  $\forall l \in L, y \in F(l)$  and  $x \in \text{conv}(L)$ .

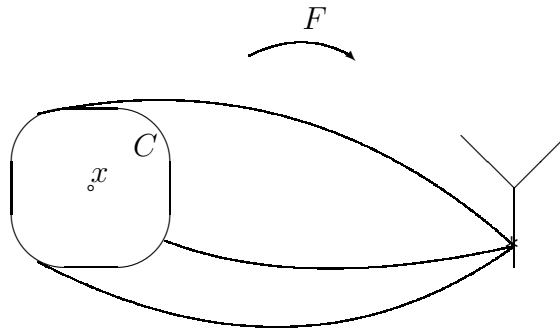


Figure 4: A Borsuk-Ulam type theorem by Simon, Spieź and Toruńczyk.

Notice that for  $n = 1$  (corresponding to 2 states of nature), the image by  $F$  of the connected component of  $C$  containing  $x$  necessarily is a singleton, hence

the result is clear. In the general case, one finally obtains:

**Theorem 9 ([73]):** There exists an equilibrium joint plan. Thus there exists a uniform equilibrium in the repeated game  $\Gamma(p)$ .

## VI.2 Characterization of equilibrium payoffs

Characterizing equilibrium payoffs, as the Folk theorem does for repeated games with complete information, has been a challenging problem. We denote here by  $p_0$  the initial probability in the interior of  $\Delta(K)$ . We are interested in the set of equilibrium payoffs, in the convenient following sense:

**Definition 8:** A vector  $(a, b)$  in  $\mathbb{R}^K \times \mathbb{R}$  is called an equilibrium payoff of the repeated game  $\Gamma(p_0)$  if there exists a strategy pair  $(\sigma^{1*}, \sigma^{2*})$  satisfying:

- (i)  $\forall \varepsilon > 0 \exists T_0 \forall T \geq T_0, \forall k \in K, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq \alpha_T^k(\sigma^{1*}, \sigma^{2*}) + \varepsilon,$   
 $\forall \varepsilon > 0 \exists T_0 \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^{p_0}(\sigma^{1*}, \sigma^2) \leq \beta_T^p(\sigma^{1*}, \sigma^{2*}) + \varepsilon,$  and
- (ii)  $(\alpha_T^k(\sigma^{1*}, \sigma^{2*}))_{k,T}$  and  $(\beta_T^{p_0}(\sigma^{1*}, \sigma^{2*}))_T$  respectively converge to  $a$  and  $b$ .

Since  $p$  lies in the interior of  $\Delta(K)$ , the first line of (i) is equivalent to:  $\forall \varepsilon > 0 \exists T_0 \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \alpha_T^p(\sigma^1, \sigma^{2*}) \leq \alpha_T^p(\sigma^{1*}, \sigma^{2*}) + \varepsilon$ . The strategy pair  $(\sigma^{1*}, \sigma^{2*})$  is thus a uniform equilibrium of the repeated game, with the additional requirement that expected average payoffs of player 1 converge in each state  $k$ . In some sense, player 1 is viewed here as  $|K|$  different types or players, and we require the existence of the limit payoff of each type. We will only consider such uniform equilibria in the sequel.

Notice that the above definition implies:  $\forall k \in K, \forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq a^k + \varepsilon$ . So the orthant  $\{x \in \mathbb{R}^K, x^k \leq a^k \forall k \in K\}$  is approachable by player 2, and by theorem 3 and subsection IV.4 one can obtain that:

$$\langle a, q \rangle \geq u(q) \forall q \in \Delta(K) \quad (4)$$

Condition (4) is called the individual rationality condition for player 1, and does not depend on the initial probability in the interior of  $\Delta(K)$ . Regarding player 2, we have:  $\forall \varepsilon > 0 \exists T_0 \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^{p_0}(\sigma^{1*}, \sigma^2) \leq \beta + \varepsilon$ , so by theorem 1:

$$\beta \geq \text{vex } v(p_0). \quad (5)$$

Condition (5) is the individual rationality condition for player 2: at equilibrium, this player should have at least the value of the game where player 1's plays in order to minimize player 2's payoffs.

Imagine now that  $\sigma^{1*}$  is a non revealing strategy for player 1, and that the players play actions with empirical frequencies corresponding to a given probability distribution  $\pi = (\pi_{i,j})_{(i,j) \in I \times J} \in \Delta(I \times J)$ . We will have:  $\forall k \in K,$



$a^k = \sum_{i,j} \pi_{i,j} A^k(i,j)$  and  $\beta = \sum_k p_0^k \sum_{i,j} \pi_{i,j} B^k(i,j)$ , and if the individual rationality conditions are satisfied, no detectable deviation of a player can be profitable. This leads to the definition of the following set, where  $M$  is the constant  $\max\{|A^k(i,j)|, |B^k(i,j)|, (i,j) \in I \times J\}$ , and  $\mathbb{R}_M = [-M, M]$ .

**Definition 9:** Let  $G$  be the set of triples  $(a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$  satisfying:

1.  $\forall q \in \Delta(K), \langle a, q \rangle \geq u(q)$ ,
2.  $\beta \geq \text{vex}v(p)$ ,
3.  $\exists \pi \in \Delta(I \times J)$  s.t.  $\beta = \sum_k p^k \sum_{i,j} \pi_{i,j} B^k(i,j)$  and  $\forall k \in K, a^k \geq \sum_{i,j} \pi_{i,j} A^k(i,j)$  with equality if  $p^k > 0$ .

We need to considerate every possible initial probability because the main state variable of the model is, here also, the belief, or *a posteriori*, of player 2 on the state of nature.  $\{(a, \beta), (a, \beta, p_0) \in G\}$  is the set of payoffs of non revealing equilibria of  $\Gamma(p_0)$ . The importance of the following definition will appear with theorem 10 below (which unfortunately has not led to a proof of existence of equilibrium payoffs).

**Definition 10:**  $G^*$  is defined as the set of elements  $g = (a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$  such that there exist a probability space  $(\Omega, \mathcal{A}, Q)$ , an increasing sequence  $(\mathcal{F}_n)_{n \geq 1}$  of finite sub- $\sigma$ -algebras of  $\mathcal{A}$ , and a sequence of random variables  $(g_n)_{n \geq 1} = (a_n, \beta_n, p_n)_{n \geq 1}$  defined on  $(\Omega, \mathcal{A})$  with values in  $\mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$  satisfying: (i)  $g_1 = g$  a.s., (ii)  $(g_n)_{n \geq 1}$  is a martingale adapted to  $(\mathcal{F}_n)_{n \geq 1}$ , (iii)  $\forall n \geq 1, a_{n+1} = a_n$  a.s. or  $p_{n+1} = p_n$  a.s., and (iv)  $(g_n)_n$  converges a.s. to a random variable  $g_\infty$  with values in  $G$ .

Let us forget for a while the component of player 2's payoff. A process  $(g_n)_n$  satisfying (ii) and (iii) may be called a bi-martingale, it is a martingale such that at every stage, one of the two components remains a.s. constant. So the set  $G^*$  can be seen as the set of starting points of converging bi-martingales with limit points in  $G$ .

**Theorem 10** (Hart, [28]) Let  $(a, \beta)$  be in  $\mathbb{R}^K \times \mathbb{R}$ .

$$(a, \beta) \text{ is an equilibrium payoff of } \Gamma(p_0) \iff (a, \beta, p_0) \in G^*.$$

Theorem 10 is too elaborate to be proved here, but let us give a few ideas about the proof. First consider the implication  $\implies$ , and fix an equilibrium  $\sigma^* = (\sigma^{1*}, \sigma^{2*})$  of  $\Gamma(p_0)$  with payoff  $(a, \beta)$ . The sequence of *a posteriori*  $(p_t(\sigma^*))_{t \geq 0}$  is a  $\mathbb{P}_{p_0, \sigma^*}$ -martingale. Modify now slightly the time structure so that at each stage, player 1 plays first, and then player 2 plays without knowing the action chosen by player 1. At each half-stage where player 2 plays, his *a posteriori* remains constant. At each half-stage where player 1 plays, the “expectation of player

1's future payoff" (which can be properly defined) remains constant. Hence, the heuristic apparition of the bimartingale. And since bounded martingale converge, for large stages everything will be fixed and the players will approximately play a non revealing equilibrium at a "limit *a posteriori*", so the convergence will be towards elements of  $G$ .

Consider now the converse implication  $\Leftarrow$ . Let  $(a, \beta)$  be such that  $(a, \beta, p_0) \in G^*$ , and assume for simplification that the associated bi-martingale  $(a_n, \beta_n, p_n)$  converges in a fixed number  $N$  of stages:  $\forall n \geq N, (a_n, \beta_n, p_n) = (a_N, \beta_N, p_N) \in G$ . One can construct an equilibrium  $(\sigma^{1*}, \sigma^{2*})$  of  $\Gamma(p_0)$  with payoff  $(a, \beta)$  along the following lines. For each index  $n$ ,  $(a_n, \beta_n)$  will be an equilibrium payoff of the repeated game with initial probability  $p_n$ . Eventually, player 1 will play independently of the state, the *a posteriori* of player 2 will be  $p_N$ , and the players will end up playing a non revealing equilibrium of the repeated game  $\Gamma(p_N)$  with payoff  $(a_N, \beta_N)$ . What should be played before ? Since we are in an undiscounted setup, any finite number of stages can be used for communication without influencing payoffs. Let  $n < N$  be such that  $a_{n+1} = a_n$ . To move from  $(a_n, \beta_n, p_n)$  to  $(a_n, \beta_{n+1}, p_{n+1})$ , player 1 can simply use the splitting lemma (lemma 1) in order to signal part of the state to player 2. Let now  $n < N$  be such that  $p_{n+1} = p_n$ , so that we want to move from  $(a_n, \beta_n, p_n)$  to  $(a_{n+1}, \beta_{n+1}, p_n)$ . Player 1 will play independently of the state, and both players will act so as to convexify their future payoffs. This convexification is done through procedures called "jointly controlled lotteries" and introduced in the sixties by Aumann and Maschler ([87]), with the following simple and brilliant idea. Imagine that the players have to decide with even probability whether to play the equilibrium  $E1$  with payoff  $(a^1, \beta^1)$  or to play the equilibrium  $E2$  with payoff  $(a^2, \beta^2)$ . The players may not be indifferent between  $E1$  and  $E2$ , e.g. player 1 may prefer  $E1$  whereas player 2 prefers  $E2$ . They will proceed as follows, with  $i$  and  $i'$ , respectively  $j$  and  $j'$ , denoting two distinct actions of player 1, resp. player 2. Simultaneously and independently, player 1 will select  $i$  or  $i'$  with probability  $1/2$ , whereas player 2 will behave sim-

ilarly with  $j$  and  $j'$ . 
$$\begin{matrix} & j & j' \\ i & \times & \\ i' & & \times \end{matrix}$$

. Then the equilibrium  $E1$  will be played if the diagonal has been reached, i.e. if  $(i, j)$  or  $(i', j')$  has been played, and otherwise the equilibrium  $E2$  will be played. This procedure is robust to unilateral deviations: none of the players can deviate and prevent  $E1$  and  $E2$  to be chosen with probability  $1/2$ . In general, jointly controlled lotteries are procedures allowing to select an alternative among a finite set according to a given probability (think of binary expansions if necessary), in a way which is robust to deviations by a single player. S. Hart has precisely shown how to combine steps of signalling and jointly controlled lotteries to construct an equilibrium of  $\Gamma_\infty(p_0)$  with payoff  $(a, \beta)$ .

### VI.3 Biconvexity and bimartingales

The previous analysis has led to the introduction and study of biconvexity phenomena. The reference here is [4]. Let  $X$  and  $Y$  be compact convex subsets of Euclidean spaces, and let  $(\Omega, \mathcal{F}, \mathcal{P})$  be an atomless probability space.

**Definition 11:** A subset  $B$  of  $X \times Y$  is biconvex if for every  $x$  in  $X$  and  $y$  in  $Y$ , the sections  $B_{x.} = \{y' \in Y, (x, y') \in B\}$  and  $B_{.y} = \{x' \in X, (x', y) \in B\}$  are convex. If  $B$  is biconvex, a mapping  $f : B \rightarrow \mathbb{R}$  is called biconvex if for each  $(x, y) \in X \times Y$ ,  $f(\cdot, y)$  and  $f(x, \cdot)$  are convex.

As in the usual convexity case, we have that if  $f$  is biconvex, then for each  $\alpha$  in  $\mathbb{R}$ , the set  $\{(x, y) \in B, f(x, y) \leq \alpha\}$  is biconvex.

**Definition 12:** A sequence of random variables  $Z_n = (X_n, Y_n)_{n \geq 1}$  with values in  $X \times Y$  is called a bimartingale if:

(1) there exists an increasing sequence  $(\mathcal{F}_n)_{n \geq 1}$  of finite sub- $\sigma$ -algebra of  $\mathcal{F}$  such that  $(Z_n)_n$  is a  $(\mathcal{F}_n)_{n \geq 1}$ -martingale.

(2)  $\forall n \geq 1, X_n = X_{n+1}$  a.s. or  $Y_n = Y_{n+1}$  a.s.

(3)  $Z_1$  is a.s. constant.

Notice that  $(Z_n)_{n \geq 1}$  being a bounded martingale, it converges almost surely to a limit  $Z_\infty$ .

**Definition 13:** Let  $A$  be a measurable subset of  $X \times Y$ .

$A^* = \{z \in X \times Y, \text{ there exists a bimartingale } (Z_n)_{n \geq 1} \text{ converging to a limit } Z_\infty$   
such that  $Z_\infty \in A$  a.s. and  $Z_1 = z$  a.s. }.

One can show that any atomless probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ , or any product of convex compact spaces  $X \times Y$  containing  $A$ , induce the same set  $A^*$ . One can also substitute condition (2) by:  $\forall n \geq 1, (X_n = X_{n+1} \text{ or } Y_n = Y_{n+1})$  a.s. Notice that without condition (2), the set  $A^*$  would just be the convex hull of  $A$ .

We always have  $A \subset A^* \subset \text{conv}(A)$ , and these inclusions can be strict. For example, if  $X = Y = [0, 1]$  and  $A = \{(0, 0), (1, 0), (0, 1)\}$ , it is possible to show that  $A^* = \{(x, y) \in [0, 1] \times [0, 1], x = 0 \text{ or } y = 0\}$ .  $A^*$  always is biconvex and thus contains  $\text{biconv}(A)$ , which is defined as the smallest biconvex set which contains  $A$ . The inclusion  $\text{biconv}(A) \subset A^*$  can also be strict, as shown by the following example:

**Example 5:** Put  $X = Y = [0, 1]$ ,  $v_1 = (1/3, 0)$ ,  $v_2 = (0, 2/3)$ ,  $v_3 = (2/3, 1)$ ,  $v_4 = (1, 1/3)$ ,  $w_1 = (1/3, 1/3)$ ,  $w_2 = (1/3, 2/3)$ ,  $w_3 = (2/3, 2/3)$  et  $w_4 = (2/3, 1/3)$ , and  $A = \{v_1, v_2, v_3, v_4\}$ .

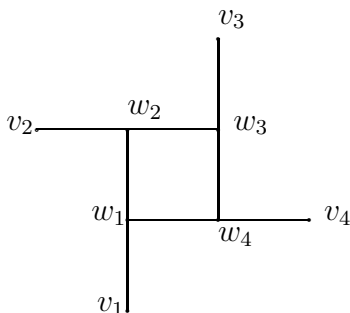


Figure 5: The “four frogs” example of Aumann and Hart:  $A^* \neq \text{biconv}(A)$ .

$A$  is biconvex, so  $A = \text{biconv}(A)$ . Consider now the following Markov process  $(Z_n)_{n \geq 1}$ , with  $Z_1 = w_1$ . If  $Z_n \in A$ , then  $Z_{n+1} = Z_n$ . If  $Z_n = w_i$  for some  $i$ , then  $Z_{n+1} = w_{i+1(\text{mod } 4)}$  with probability  $1/2$ , and  $Z_{n+1} = v_i$  with probability  $1/2$ .  $(Z_n)_n$  is a bimartingale converging a.s. to a point in  $A$ , hence  $w_1 \in A^* \setminus \text{biconv}(A)$ .

We now present a geometric characterization of the set  $A^*$ , and assume here that  $A$  is closed. For each biconvex subset  $B$  of  $X \times Y$  containing  $A$ , we denote by  $nsc(B)$  the set of elements of  $B$  which can not be separated from  $A$  by a continuous bounded biconvex function on  $A$ . More precisely,  $nsc(B) = \{z \in B, \forall f : B \rightarrow \mathbb{R} \text{ bounded biconvex, and continuous on } A, f(z) \leq \sup\{f(z'), z' \in A\}\}$ .

**Theorem 11** ([4]):  $A^*$  is the largest biconvex set  $B$  containing  $A$  such that  $nsc(B) = B$ .

Let us now come back to repeated games and to the notations of subsection VI.2. To be precise, we need to add the component of player 2’s payoff, and consequently to slightly modify the definitions.  $G$  is closed in  $\mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ . For  $B \subset \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ ,  $B$  is biconvex if for each  $a$  in  $\mathbb{R}_M^K$  and for each  $p$  in  $\Delta(K)$ , the sections  $\{(\beta, p'), (a, \beta, p') \in B\}$  and  $\{(a', \beta), (a', \beta, p) \in B\}$  are convex. A real function  $f$  defined on a biconvex set  $B$  is said to be biconvex if  $\forall a, \forall p$ ,  $f(a, \cdot, \cdot)$  and  $f(\cdot, \cdot, p)$  are convex.

**Theorem 12** ([4]):  $G^*$  is the largest biconvex set  $B$  containing  $G$  such that:  $\forall z \in B, \forall f : B \rightarrow \mathbb{R}$  bounded biconvex, and continuous on  $A$ ,  $f(z) \leq \sup\{f(z'), z' \in G\}$ .

## VII. Non-observable actions

We now consider the case where, as in the general definition of section I, there is a signalling function  $q : K \times A \rightarrow \Delta(U)$  giving the distributions of the signals received by the players as a function of the state of nature and the action profile just played. The particular case where  $q(k, a)$  does not depend on  $k$  is called *state*

*independent signalling.* The previous models correspond to the particular case of perfect observation, where the signals received by the players exactly reveal the action profile played.

Theorem 1 has been generalized ([87]) to the general case of signalling function. We keep the notations of section III. Given a mixed action  $x \in \Delta(I)$ , an action  $j$  in  $J$  and a state  $k$ , we denote by  $Q(k, x, j)$  the marginal distribution on  $U^2$  of the law  $\sum_{i \in I} x(i) q(k, i, j)$ , i.e.  $Q(k, x, j)$  is the law of the signal received by player 2 if the state is  $k$ , player 1 uses  $x$  and player 2 plays  $j$ . The set of non revealing strategies of player 1 is then defined as:  $NR(p) =$

$$\{x = (x^k)_{k \in K} \in \Delta(I)^K, \forall k \in K, \forall k' \in K \text{ s.t. } p^k p^{k'} > 0, \forall j \in J, Q(k, x^k, j) = Q(k', x^{k'}, j)\}.$$

If the initial probability is  $p$  and player 1 plays a strategy  $x$  in  $NR(p)$  (i.e. plays  $x^k$  if the state is  $k$ ), the *a posteriori* of player 2 will remain a.s. constant: player 2 can deduce no information on the selected state  $k$ . The value of the non revealing game becomes:

$$u(p) = \max_{x \in NR(p)} \min_{y \in \Delta(J)} \sum_{k \in K} p^k G^k(x^k, y) = \min_{y \in \Delta(J)} \max_{x \in NR(p)} \sum_{k \in K} p^k G^k(x^k, y),$$

where  $G^k(x^k, y) = \sum_{i,j} x^k(i) y(j) G^k(i, j)$ , and the convention  $u(p) = -\infty$  if  $NR(p) = \emptyset$ . Theorem 1 perfectly extends here: The repeated game with initial probability  $p$  has a uniform value given by  $\text{cav}u(p)$ .

The explicit construction of an optimal strategy of player 2 (see IV.4 here) has also been generalized to the general signalling case (see [34], and [91], part B, p.234 for random signals).

Regarding zero-sum games with lack of information on both sides, the results of section V have been generalized to the case of state independent signalling (see [49], [51] and [54]). Attention has been paid to the speed of convergence of the value function  $(v_T)_T$ , and bounds are identical for both models of lack of information on one side and on both sides, if we assume state independent signalling: this speed is of order  $1/T^{1/2}$  for games with perfect observation, and of order  $1/T^{1/3}$  for games with signals (these orders are optimal, both for lack of information on one side and lack of information on both sides, see [85], [86]). For state dependent signalling and lack of information on one side, it was shown by Mertens [50] that the convergence occurs with worst case error  $\sim (\ln n/n)^{1/3}$ .

A particular class of zero-sum repeated games with state dependent signalling has been studied (games with no signals, see [53], [84] and [79]). In these games, the state  $k$  is first selected according to a known probability and is not announced to the players; then after each stage both players receive the same signal which is either “nothing” or “the state is  $k$ ”. It was shown that the maxmin and the minmax may differ, although  $\lim_T v_T$  always exists.

In non zero-sum repeated games with lack of information on one side, the existence of “joint plan” equilibria have been generalized to the case of state independent signalling ([59]), and more generally to the case where “player 1 can send non revealing signals to player 2” ([74]). The existence of a uniform equilibrium in the general signalling case is still an open question (see [75]).

## VIII. Miscellaneous

### VIII.1 Zero-sum games

In games with lack of information on one side, it is important that player 1 knows not only the selected state  $k$ , but also the *a priori*  $p$ . [81] provides an example of a game with lack of information on “one and a half” side with no uniform value. More precisely, in this example nature first chooses  $p$  in  $\{p_1, p_2\}$  according to a known probability, and announces  $p$  to player 2 only; then  $k$  is selected according to  $p$ , and announced to player 1 only; finally the matrix game  $G^k$  is played.

For games with lack of information on one side, the value function  $v_T$  is a concave piecewise linear function of the initial probability  $p$  (see [58] for more generality). On the contrary, the discounted value  $v_\lambda$  can be quite a complex function of  $p$ : in example 2 of section I, Mayberry ([48]) has proved that for  $2/3 < \lambda < 1$ ,  $v_\lambda$  is, at each rational value of  $p$ , non differentiable.

Convergence of the value functions  $(v_T)_T$  and  $(v_\lambda)_\lambda$  have been widely studied. We have already discussed the speed of convergence in section VII, but much more can be said.

**Example 6:** Standard model of lack of information on one side and observable actions.  $K = \{a, b\}$ ,  $G^a = \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix}$  and  $G^b = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$ . One can show ([52]) that for each  $p \in [0, 1]$ , viewed as the initial probability of state  $a$ , the sequence  $\sqrt{T} v_T(p)$  converges to  $\varphi(p)$ , where  $\varphi(p) = \frac{1}{\sqrt{2\pi}} e^{-x_p^2/2}$ , and  $x_p$  satisfies  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_p} e^{-x^2/2} dx = p$ . So the limit of  $\sqrt{T} v_T(p)$  is the standard normal density function evaluated at its  $p$ -quantile.

The apparition of the normal distribution is by no way an isolated phenomenon, but rather an important property of some repeated games ([11], [12], [13], [17], [14], ...).

B. de Meyer introduced the notion of “dual game” (see the previous references and also [66], [18], [40], [16]). Let us now illustrate this on the standard model of section III.

Let  $z$  be a parameter in  $\mathbb{R}^K$ . In the dual game  $\Gamma_T^*(z)$ , player 1 first secretly chooses the state  $k$ . Then at each stage  $t \leq T$ , the players choose as usual actions

$i_t$  and  $j_t$  which are announced before proceeding to the next stage. With time horizon  $T$ , player 1's payoff finally is  $\frac{1}{T} \sum_{t=1}^T G^k(i_t, j_t) - z^k$ . This player is thus now able to fix the state equal to  $k$ , but has to pay  $z^k$  for it. It can be shown that the  $T$ -stage dual game  $\Gamma_T^*(z)$  has a value  $w_T(z)$ .  $w_T$  is convex, and is linked to the value of the primal game by the conjugate formula:

$$w_T(z) = \max_{p \in \Delta(K)} (v_T(p) - \langle p, z \rangle), \text{ and } v_T(p) = \inf_{z \in \mathbb{R}^K} (w_T(z) + \langle p, z \rangle).$$

And  $(w_T)_T$  satisfies the dual recursive formula:

$$w_{T+1}(z) = \min_{y \in \Delta(J)} \max_{i \in I} \frac{T}{T+1} w_T \left( \frac{T+1}{T} z - \frac{1}{T} \sum_{j \in J} y_j (G^k(i, j))_k \right).$$

There are also strong relations between the optimal strategies of the players in the primal and dual games, and this gives a way to compute recursively optimal strategies of the uninformed player in the finite game (see also [31] on this topic).

Repeated games with incomplete information, as well as stochastic games, can also be studied in a functional analysis setup called the operator approach. This general approach is based on the study of the recursive formula ([67], [39], [92]).

In [62], the standard model, as well as the proof of theorem 1, is generalized to the case where the state is not fixed at the beginning of the game, but evolves according to a Markov chain uniquely observed by player 1 (see also [56] for non observable actions, [47] and [32] for the difficulty of computing the value, [64] for the generalization to a state process controlled and observed by player 1, and [68] for several kinds of stochastic games with lack of information on one side). It is known since [77] that the uniform value may not exist in general for stochastic games with lack of information on one side (where the stochastic game to be played is first randomly selected and announced to player 1 only).

Blackwell's approachability theorem has been extended to infinite dimensional spaces by Lehrer ([42]). Approachability theory has strong links with the existence of no-regret strategies (first studied in [29], see also [25], [69], [43], [30], [8] and the recent book [9]), convergence of simple procedures to the set of correlated equilibria ([29]), and calibration ([24], [41]). The links between merging, reputation phenomena and repeated games with incomplete information have been studied in [80], where several existing results are unified. Finally, no-regret and approachability have also been studied when the players have bounded computational capacities (finite automata, bounded recall strategies) ([45], [44]).

Let us mention also that de Meyer and Moussa Saley studied the modelization via Brownian motions in financial models ([17]). They introduced a marked game based on a repeated game with lack of information on one side, and showed the

endogenous apparition of a Brownian motion (see [15] for incomplete information on both sides).

## VIII.2 Non zero-sum games

In the setup of section VI, it is interesting to study the number of communication stages which is needed to construct the different equilibria. This number is linked with the convergence of the associated bimartingales (see [87], [4], [20], [23]). Let us mention also that F. Forges ([22]) gave a similar characterization of equilibrium payoffs, for a larger notion of equilibria called communication equilibria (see also [21] for correlated equilibria). Amitai ([2]) studied the set of equilibrium payoffs in case of lack of information on both sides. Aumann and Hart ([5]) characterized the equilibrium payoffs in two player games with lack of information on one side when long, payoff-irrelevant, preplay communication is allowed (see [1] for incomplete information on both sides).

The particular case where each player knows his own payoffs is particularly worthwhile studying (*known own payoffs*). In the two-player case with lack of information on one side, this amounts to say that player 2's payoffs do not depend on the selected state. In this case, Shalev ([70]) showed that any equilibrium payoff can be obtained as the payoff of an equilibrium which is completely revealing. This result generalizes to the non zero-sum case of lack of information of both sides (see the unpublished manuscript [36]), but unfortunately uniform equilibria may fail to exist even though both players know their own payoffs.

Another model deals with the symmetric case, where the players have an incomplete, but identical, knowledge of the selected state. After each stage they receive the same signal, which may depend on the state. A. Neyman and S. Sorin have proved the existence of equilibrium payoffs in the case of two players (see [57], the zero-sum case being solved in [35] and [19]).

Few papers study the case of more than 2 players. The existence of uniform equilibrium has been studied for 3 players and lack of information on one side ([60]), and in the case of two states of nature it appears that a completely revealing equilibria, or a joint plan equilibria by one of the informed players, always exists. Concerning  $n$ -player repeated games with incomplete information and signals, several papers study how the initial information can be strategically transmitted, independently of the payoffs ([63], [61], and [65] for an application to a cryptographic model). As an application, the existence of completely revealing equilibria is obtained in particular cases.

Repeated games with incomplete information have been used to study perturbations of repeated games with complete information (see [26] and [10] for Folk theorem-like results, [6] for enforcing cooperation in games with a Pareto-



dominant outcome, and [33] for a perturbation with known own payoffs). The case where the players have different discount factors has also been investigated ([46], [10]).

## IX. Future directions

Several open problems are well formulated and deserve attention. Does a uniform equilibrium always exist in two-player repeated games with lack of information on one side and general signalling, or in  $n$ -player repeated games with lack of information on one side ? More conceptually, one should look for classes of  $n$ -player repeated games with incomplete information which allow for the existence of equilibria, and/or for a tractable description of equilibrium payoffs (or at least of some of these payoffs). Regarding applications, there is certainly a lot of room in the vast fields of financial markets, cryptology, and sequential decision problems.

## X. Bibliography

### References

- [1] M. Amitai. Cheap-Talk with Incomplete Information on Both Sides. PhD Thesis, The Hebrew University of Jerusalem, 1996. <http://ratio.huji.ac.il/dp/dp90.pdf>
- [2] M. Amitai. Repeated games with incomplete information on Both Sides. PhD Thesis, The Hebrew University of Jerusalem, 1996. <http://ratio.huji.ac.il/dp/dp105.pdf>
- [3] R. J. Aumann. Mixed and behaviour strategies in infinite extensive games. in *Advances in game theory*, Dresher, Shapley and Tucker (eds.), Annals of Mathematics Study 52, Princeton University Press, 627–650, 1964.
- [4] R. J. Aumann and S. Hart. Bi-convexity and bi-martingales, *Israel Journal of Mathematics*, 54:159–180, 1986.
- [5] R. J. Aumann and S. Hart. Long cheap talk, *Econometrica*, 71:1619–1660, 2003.
- [6] R. J. Aumann and S. Sorin. Cooperation and bounded recall, *Games and Economic Behavior*, 1, 5-39, 1989.
- [7] D. Blackwell. An analog of the minmax theorem for vector payoffs. *Pacific Journal of Mathematics*, 65:1–8, 1956.

- [8] N. Cesa-Bianchi, G. Lugosi and G. Stoltz. Regret minimization under partial monitoring. *Mathematics of Operations Research*, 31:562–580, 2006.
- [9] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning and Games*. Cambridge University Press, 2006.
- [10] M.W. Cripps and J.P. Thomas. Some asymptotic results in discounted repeated games of one-sided incomplete information. *Mathematics of Operations Research*, 28:433–462, 2003.
- [11] B. de Meyer. Repeated games and partial differential equations, *Mathematics of Operations Research*, 21, 209-236, 1996 a.
- [12] B. de Meyer. Repeated games, duality and the central limit theorem, *Mathematics of Operations Research*, 21, 237-251, 1996 b.
- [13] B. de Meyer. The maximal variation of a bounded martingale and the central limit theorem, *Annales de l'Institut Henri Poincaré, Probabilités et statistiques*, 34, 49–59, 1998.
- [14] B. de Meyer. From repeated games to Brownian games, *Annales de l'Institut Henri Poincaré, Probabilités et statistiques*, 35, 1–48, 1999.
- [15] B. de Meyer and A. Marino, Repeated market games with lack of information on both sides. DP 2004.66, MSE Université Paris I, 2004.
- [16] B. de Meyer and A. Marino, Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides. DP 2005.27, MSE Université Paris I, 2005.
- [17] B. de Meyer and H. Moussa Saley On the strategic origin of Brownian motion in finance, *International Journal of Game Theory*, 31, 285-319, 2003.
- [18] B. de Meyer and D. Rosenberg. “Cavu” and the dual game, *Mathematics of Operations Research*, 24, 619-626, 1999.
- [19] F. Forges. Infinitely repeated games of incomplete information: symmetric case with random signals, *International Journal of Game Theory*, 11, 203–213, 1982.
- [20] F. Forges. A note on Nash equilibria in repeated games with incomplete information, *International Journal of Game Theory*, 13, 179–187, 1984.
- [21] F. Forges. Correlated equilibria in a class of repeated games with incomplete information, *International Journal of Game Theory*, 14, 129–149, 1985.
- [22] F. Forges. Communication equilibria in repeated games with incomplete information, *Mathematics of Operations Research*, 13, 191-231, 1988.

- [23] F. Forges. Equilibria with communication in a job market example, *Quarterly Journal of Economics*, 105, 375–398, 1990.
- [24] D. Foster, A Proof of Calibration via Blackwell’s Approachability Theorem, *Games and Economic Behavior*, 29:73–78, 1999.
- [25] D. Foster and R. Vohra, Regret in the On-Line Decision Problem, *Games and Economic Behavior*, 29:7–35, 1999.
- [26] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54:533–554, 1986.
- [27] J. Harsanyi. Games with incomplete information played by ‘Bayesian’ players, parts I-III, *Management Science*, 8, 159–182, 320–334, 486–502, 1967–68.
- [28] S. Hart. Nonzero-sum two-person repeated games with incomplete information. *Mathematics of Operations Research*, 10, 117–153, 1985.
- [29] S. Hart and A. Mas-Colell A simple adaptative procedure leading to Correlated Equilibrium. *Econometrica*, 68, 1127–1150, 2000.
- [30] S. Hart Adaptative Heuristics, *Econometrica*, 73, 1401–1430, 2005.
- [31] M. Heuer. Optimal strategies for the uninformed player. *International Journal of Game Theory*, 20:33–51, 1992.
- [32] J. Hörner, D. Rosenberg, E. Solan and N. Vieille. *On Markov Games with Incomplete Information on One Side*, Discussion paper 1412, Center for Mathematical Studies in Economics and Management Science, Northwestern University, 2006.
- [33] E. Israeli. Sowing Doubt Optimally in Two-Person Repeated Games. *Games and Economic Behavior*, 28:203–216, 1999.
- [34] E. Kohlberg. Optimal strategies in repeated games with incomplete information, *International Journal of Game Theory*, 4:7–24, 1975.
- [35] E. Kohlberg and S. Zamir. Repeated games of Incomplete information: The Symmetric Case. *Annals of Statistics*, 2:40–41, 1974.
- [36] G. Koren. Two-person repeated games where players know their own payoffs, master thesis, Tel-Aviv University, 1992. <http://www.ma.huji.ac.il/hart/papers/koren.pdf>
- [37] H.W. Kuhn. Extensive games and the problem of information. in *Contributions to the Theory of Games* (Kuhn and Tucker, eds.), vol. II, 193–216. Annals of Mathematical Studies 28, Princeton University Press, 1953.

- [38] R. Laraki. Variational inequalities, system of functional equations and incomplete information repeated games, *SIAM Journal on control and optimization*, 40:516-524, 2001 a.
- [39] R. Laraki. The splitting game and applications, *International Journal of Game Theory*, 30:359-376, 2001 b.
- [40] R. Laraki. Repeated games with lack of information on one side : the dual differential approach, *Mathematics of Operations Research*, 27:419-440, 2002.
- [41] E. Lehrer, Any inspection is manipulable, *Econometrica*, 69:1333-1347, 2001.
- [42] E. Lehrer, Approachability in infinite dimensional spaces, *International Journal of Game Theory*, 31:253-268, 2003.
- [43] E. Lehrer, A wide range no-regret theorem, *Games and Economic Behavior*, 42:101-115, 2003.
- [44] E. Lehrer and E. Solan, No regret with bounded computational capacity, *DP 1373*, Center for Mathematical Studies in Economics and Management Science, Northwestern University, 2003.
- [45] E. Lehrer and E. Solan, Excludability and bounded computational capacity, *Mathematics of Operations Research*, 31:637-648, 2006.
- [46] E. Lehrer and L. Yariv, Repeated games with lack of information on one side : the case of different discount factors, *Mathematics of Operations Research*, 24:204-218, 1999.
- [47] A. Marino. The value of a particular Markov chain game. Chapters 5 and 6, PhD thesis, Université Paris I, 2005. <http://alexandre.marino.free.fr/theseMarino.pdf>
- [48] J.-P. Mayberry. Discounted repeated games with incomplete information, *Report of the U.S. Arms control and disarmament agency*, ST116, chapter V, Mathematica, Princeton, 435-461.
- [49] J.-F. Mertens. The value of two-person zero-sum repeated games: the extensive case, *International Journal of Game Theory*, 1: 217-227, 1972.
- [50] J.-F. Mertens. The speed of convergence in repeated games with incomplete information on one side, *International Journal of Game Theory*, 27: 343-357, 1998.
- [51] J.-F. Mertens and S. Zamir. The value of two-person zero-sum repeated games with lack of information on both sides, *International Journal of Game Theory*, 1:39-64, 1971.

- [52] J.-F. Mertens and S. Zamir. The normal distribution and repeated games, *International Journal of Game Theory*, 5:187–197, 1976.
- [53] J.-F. Mertens and S. Zamir. On a repeated game without a recursive structure, *International Journal of Game Theory*, 5:173–182, 1976.
- [54] J.-F. Mertens and S. Zamir. A duality theorem on a pair of simultaneous functional equations, *Journal of Mathematical Analysis and Applications*, 60:550–558, 1977.
- [55] J.-F. Mertens and S. Zamir. Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory*, 14:1–29, 1985.
- [56] A. Neyman. Existence of Optimal Strategies in Markov Games with Incomplete Information. preprint, Institute of Mathematics and Center for the Study of Rationality, Hebrew University, Jerusalem, december 2005.
- [57] A. Neyman and S. Sorin. Equilibria in Repeated Games with Incomplete Information: The General Symmetric Case, *International Journal of Game Theory*, 27:201–210, 1998.
- [58] J.P. Ponsard and S. Sorin. The LP formulation of finite zero-sum games with incomplete information, *International Journal of Game Theory*, 9:99–105, 1980.
- [59] J. Renault. 2-player repeated games with lack of information on one side and state independent signalling. *Mathematics of Operations Research*, 4, 552–572, 2000.
- [60] J. Renault. 3-player repeated games with lack of information on one side, *International Journal of Game Theory*, 30:221–246, 2001 a.
- [61] J. Renault. Learning sets in state dependent signalling game forms: a characterization, *Mathematics of Operations Research*, 26, 832–850, 2001 b.
- [62] J. Renault. The value of Markov chain games with lack of information on one side, *Mathematics of Operations Research*, 31, 490–512, 2006.
- [63] J. Renault and T. Tomala. Learning the state of nature in repeated games with incomplete information and signals, *Games and Economic Behavior*, 47:124–156, 2004 b.
- [64] J. Renault. The value of Repeated Games with an informed controller. Technical report, Université Paris-Dauphine, Ceremade 2007-02.
- [65] J. Renault and T. Tomala. Probabilistic reliability and privacy of communication using multicast in general neighbor networks. To appear in *Journal of Cryptology*.

- [66] D. Rosenberg. Duality and Markovian strategies, *International Journal of Game Theory*, 27:577–597, 1998.
- [67] D. Rosenberg and S. Sorin. An operator approach to zero- sum repeated games, *Israel Journal of Mathematics*, 121:221–246, 2001.
- [68] D. Rosenberg, E. Solan and N. Vieille. Stochastic games with a single controller and incomplete information. *SIAM Journal on Control and Optimization*, 43, 86–110, 2004.
- [69] A. Rustichini. Minimizing Regret: The General Case, *Games and Economic Behavior*, 29:224–243, 1999.
- [70] J. Shalev. Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs, *Games and Economic Behavior*, 7:246–259, 1994.
- [71] M. Sion. On General Minimax Theorems. *Pacific Journal of Mathematics*, 8, 171–176, 1958.
- [72] R.S. Simon. Separation of joint plan equilibrium payoffs from the min-max functions, *Games and Economic Behavior*, 1:79–102, 2002.
- [73] R.S. Simon, S. Spież and H. Toruńczyk. The existence of equilibria in certain games, separation for families of convex functions and a theorem of Borsuk-Ulam type, *Israel Journal of Mathematics*, 92:1–21, 1995.
- [74] R.S. Simon, S. Spież and H. Toruńczyk. Equilibrium existence and topology in some repeated games with incomplete information, *Transactions of the AMS*, 354:5005–5026, 2002.
- [75] R.S. Simon, S. Spież and H. Toruńczyk. Equilibria in a class of games and topological results implying their existence, *RACSAM, Rev. R. Acad. Cien. Serie A. Mat.*, 102, 161–179, 2008.
- [76] S. Sorin. Some results on the existence of Nash equilibria for non- zero sum games with incomplete information *International Journal of Game Theory*, 12:193–205, 1983.
- [77] S. Sorin. Big match with lack of information on one side (Part I). *International Journal of Game Theory*, 13, 201–255, 1984.
- [78] S. Sorin. On a pair of simultaneous functional equations, *Journal of Mathematical Analysis and Applications*, 98:296–303, 1984.
- [79] S. Sorin. On recursive games without a recursive structure: existence of  $\lim v_n$ . *International Journal of Game Theory*, 18, 45–55, 1989.

- [80] S. Sorin. Merging, reputation, and repeated games with incomplete information *Games and Economic Behavior*, 29:274-308, 1997.
- [81] S. Sorin and S. Zamir. A 2-person game with lack of information on 1 and 1/2 sides, *Mathematics of operations research*, 10:17-23, 1985.
- [82] X. Spinat. A necessary and sufficient condition for approachability, *Mathematics of operations research*, 27:31-44, 2002.
- [83] N. Vieille. Weak approachability, *Mathematics of operations research*, 17:781-791, 1992.
- [84] C. Waternaux. Solution for a class of repeated games without recursive structure. *International Journal of Game Theory*, 12, 129–160, 1983.
- [85] S. Zamir. On the relation between finitely and infinitely repeated games with incomplete information, *International Journal of Game Theory*, 1:179–198, 1971.
- [86] S. Zamir. On repeated games with general information function, *International Journal of Game Theory*, 21:215–229, 1973.

## Books and Reviews

- [87] R. J. Aumann and M. Maschler. *Repeated games with incomplete information*, with the collaboration of R.E. Stearns. M.I.T. Press, 1995 (contains a reedition of chapters of Reports to the U.S. Arms Control and Disarmament Agency ST-80, 116 and 143, *Mathematica*, 1966-1967-1968).
- [88] F. Forges. Repeated Games of Incomplete Information: Non-zero sum. in *Handbook of Game Theory, I* (R.J. Aumann and S. Hart, Eds.), 155-177, Elsevier Science Publishers, 1992.
- [89] R. Laraki, J. Renault and T. Tomala. Théorie des Jeux, Introduction à la théorie des jeux répétés. *Editions de l'Ecole Polytechnique, journées X-UPS 2006*, ISBN: 978-2-7302-1366-0, in French (Chapter 3 deals with repeated games with incomplete information).
- [90] J.-F. Mertens. Repeated games. *Proceedings of the International Congress of Mathematicians, Berkeley 1986*, 1528–1577. *American Mathematical Society*, 1987.
- [91] J.-F. Mertens, S. Sorin, and S. Zamir. Repeated games. CORE discussion paper 9420-9422, 1994.

- [92] S. Sorin. A first course on zero-sum repeated games. *Mathématiques et Applications*, Springer, 2002.
- [93] S. Zamir. Repeated Games of Incomplete Information: zero-sum. in *Handbook of Game Theory, I* (R.J. Aumann and S. Hart, Eds.), 109-154, Elsevier Science Publishers, 1992.