Repeated Games with Incomplete Information

Jérôme Renault
Toulouse School of Economics, Université Toulouse 1 Capitole, Toulouse, France

Article Outline

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.

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 $\{p = (p_s)_{s \in S} \in \mathbb{R}^S,\ \forall s \in S\ p_s \geq 0$ and $\sum_{s \in S} p_s = 1\}$. Given s in S, the Dirac measure on s will be denoted by δ_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, $\|p - q\| = \sum_{s \in S} |p_s - q_s|$.

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 do not 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

R. A. Meyers (ed.), *Encyclopedia of Complexity and Systems Science*,
https://doi.org/10.1007/978-3-642-27737-5_452-2

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 (Harsanyi 1967; Mertens and Zamir 1985). Repeated games with incomplete information have been introduced in the sixties by Aumann and Maschler (1995), and we present here the basic and fundamental results of the domain. Let us start with a few well-known elementary examples (Aumann and Maschler 1995; Zamir 1992).

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 nonrevealing, 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: $K \times A \to \mathbb{R}$. There is a signaling function q: $K \times A \to \Delta(U)$, 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 nonempty 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 $\left(k, \left(u_0^i\right)_i\right)$ is selected according to π: 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 $a_t = \left(a_t^i\right)_i$ 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 $\left(u_i^t\right)_i$ is selected according to $q(k, a_t)$, and each player i learns u_i^t (and nothing more than u_i^t) 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 signaling 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."

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 $\sigma^i = \left(\sigma_t^i\right)_{t \geq 1}$, where for each t, σ_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 Σ^i.

A history of length t in Γ is a sequence $(k, u_0, a_1, u_1, \ldots, a_t, u_t)$, 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, and the set of plays is denoted by $\Omega = K \times U \times (A \times U)^\infty$ and is endowed with the product σ-algebra. A strategy profile $\sigma = (\sigma^i)_i$ naturally induces, together with the initial probability π, 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 σ 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 Γ_T where simultaneously each player i chooses a strategy σ^i in Σ^i, then receives the payoff $\gamma_T^i\left((\sigma^j)_{j\in N}\right)$.

Given a discount factor λ in (0, 1], the λ-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 λ-discounted game is the game Γ_λ where simultaneously, each player i chooses a strategy σ^i in Σ^i, then receives the payoff $\gamma_\lambda^i\left((\sigma^j)_{j\in N}\right)$.

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 σ-algebra). Kuhn's theorem (see Aumann (1964), Kuhn (1953) or Sorin (2002) for a modern presentation) states that mixed strategies or behavior strategies are equivalent, in the following sense: for each behavior strategy σ^i, there exists a mixed strategy τ^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 σ^{-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 Γ_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 (1958) to show that the λ-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 Γ_T with T sufficiently large (or similarly uniformly in the discount factor, i.e., simultaneously in all game Γ_λ with λ sufficiently low).

Definitions 1 Player 1 can guarantee the real number υ in the repeated game Γ if: $\forall\varepsilon > 0$, $\exists\sigma^1 \in \Sigma^1$, $\exists T_0$, $\forall T \geq T_0$, $\forall\sigma^2 \in \Sigma^2$, $\gamma_T^1(\sigma^1,\sigma^2) \geq \upsilon - \varepsilon$. Similarly, Player 2 can guarantee υ in Γ if $\forall\varepsilon > 0$, $\exists\sigma^2 \in \Sigma^2$, $\exists T_0$, $\forall T \geq T_0$, $\forall\sigma^1 \in \Sigma^1$, $\gamma_T^1(\sigma^1,\sigma^2) \leq \upsilon + \varepsilon$. If both player 1 and player 2 can guarantee υ, then υ is called the uniform value of the repeated game. A strategy σ^1 of player 1 satisfying $\exists T_0$, $\forall T \geq T_0$, $\forall\sigma^2 \in \Sigma^2$, $\gamma_T^1(\sigma^1,\sigma^2) \geq \upsilon$ 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 υ_T, as T goes to infinity, and υ_λ, as λ goes to zero, converge to the uniform value.

Equilibria of General-Sum Games

In the general case, the T-stage game Γ_T can be seen as the mixed extension of a finite game and consequently possesses a Nash equilibrium. Similarly, the discounted game Γ_λ 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 Γ if: (1) $\forall\varepsilon > 0$, σ is an ε-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 $\left((\gamma_T^i(\sigma))_{i\in N}\right)_T$ converges to a limit payoff $(\gamma^i(\sigma))_{i\in N}$ in $\mathbb{R}^N$.

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

The Standard Model of Aumann and Maschler

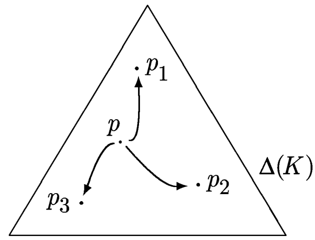
This famous model has been introduced in the sixties by Aumann and Maschler (see the reedition (Aumann and Maschler 1995). It deals with zero-sum games with lack of information on one side and observable actions, as in the basic examples previously presented. There is a finite set of states K, an initial 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 $\max_{k,i,j} |G^k(i, j)|$.

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:

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

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.



Repeated Games with Incomplete Information, Fig. 1 Splitting

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 $p^k > 0$. (Fig. 1)

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 σ 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 $t \geq 0$, $h_t = (i_1, j_1, \ldots, i_t, j_t) \in (I \times J)^t$ 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 σ 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 σ^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 σ^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 cav f, and we have:

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

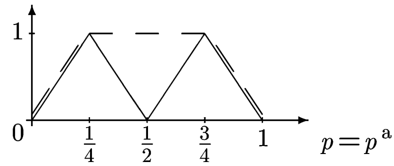
Concavification Lemma 3 If for any initial probability p, the informed player can guarantee $f(p)$ in the game $\Gamma(p)$, then for any p this player can also guarantee $\operatorname{cav} f(p)$ in $\Gamma(p)$.

Nonrevealing 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 cannot 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 Lispchitz 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:



Repeated Games with Incomplete Information, Fig. 2 u and cav$u$

Proposition 1 Player 1 can guarantee cav$u(p)$ in the game $\Gamma(p)$.

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

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 $1/2\, u(3/4) + 1/2\, u(1/4) = \operatorname{cav}u(1/2) = 1$ to player 1.

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 lim $\sup_T$ $v_T(p) \leq \inf_T v_T(p)$. 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 $\Gamma(p)$.

Uniform Value: cav*u* Theorem

We will see here that $\lim_T v_T(p)$ is nothing but cav$u(p)$, 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 nonrevealing way, so will not be able to have a stage payoff greater than $u(q)$, where q if a "limit *a posteriori.*" Since the expectation of the *a posteriori* is the *a priori* p, player 1 cannot guarantee more than $\max\{\sum_{s \in S} \lambda_s u(p_s),$ S 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 cav$u(p)$. Let us now proceed to the formal proof.

Fix a strategy σ^1 of player 1, and define the strategy σ^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 $\|.\|1$. 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

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

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

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

and the result follows. □

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 ht 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 ht: $\bar{\sigma}_{t+1}^1(h_t) = \sum_{k \in K} p_t^k(h_t)\sigma_{t+1}^1(k, h_t) \in \Delta(I)$. $\bar{\sigma}_{t+1}(h_t)$ can be seen as the average action played by player 1 after ht and will be used as a nonrevealing approximation for $\left(\sigma_{t+1}^1(k, h_t)\right)_k$. 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

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

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:

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

Consequently,

$$\mathbb{E}\left(\|p_{t+1} - p_t\| \;|\, h_t\right) = \sum_{i_{t+1} \in I} \bar{\sigma}^1_{t+1}(h_t)(i_{t+1})$$
$$\sum_{k \in K} | \, p^k_{t+1}(h_t, i_{t+1}) - p^k_t(h_t) \, | \, .$$
$$= \sum_{i_{t+1} \in I} \sum_{k \in K} | \, p^k_t(h_t)\sigma^1_{t+1}(k, h_t)(i_{t+1})$$
$$- \bar{\sigma}^1_{t+1}(h_t)(i_{t+1}) p^k_t(h_t) \, |$$
$$= \sum_{k \in K} p^k_t(h_t) \|\sigma^1_{t+1}(k, h_t) - \bar{\sigma}^1_{t+1}(h_t)\|$$
$$= \mathbb{E}\left(\|\sigma^1_{t+1}\left(\tilde{k}, h_t\right) - \bar{\sigma}^1_{t+1}(h_t)\| \;|\, h_t\right).$$

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

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

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

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

Applying Jensen's inequality yields:

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

We now apply Lemma 4 and obtain:

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

This is true for any strategy σ^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 $T \geq 1$,

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

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

Theorem 1 Aumann and Maschler (1995). The game $\Gamma(p)$ has a uniform value which is $\operatorname{cav} u(p)$.

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:

$$v_T(p) = \inf_{\sigma^2 \in \Sigma^2} \sup_{\sigma^1 \in \Sigma^1} \gamma^{1,p}_T(\sigma),$$
$$= \inf_{\sigma^2 \in \Sigma^2} \sup_{\sigma^1 \in \Sigma^1} \sum_{k \in K} p^k \gamma^{1,\delta_k}_T(\sigma),$$
$$= \inf_{\sigma^2 \in \Sigma^2} \sum_{k \in K} p^k \left(\sup_{\sigma^1 \in \Sigma^1} \gamma^{1,\delta_k}_T(\sigma)\right).$$

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: $\forall T \geq 1$, $\operatorname{cav} u(p) \leq v_T(p) \leq \operatorname{cav} u(p) + \frac{M \Sigma_{k \in K} \sqrt{p^k(1 - p^k)}}{\sqrt{T}}.$

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

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

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 $v_{T+1} \le v_T$ by induction on T, using the concavity of v_T).

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

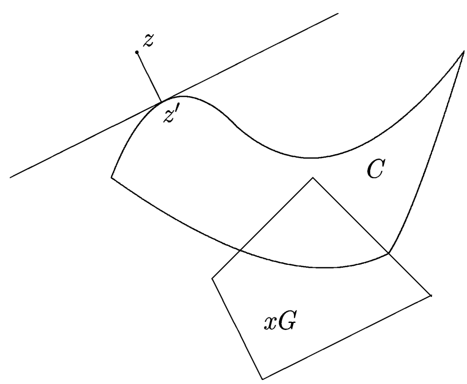
Vector Payoffs and Approachability

The following model has been introduced by D. Blackwell (1956) and is, strictly speaking, not part of the general definition given in section "Definition of the Subject and Its Importance." 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 would not 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 "Vector Payoffs and Approachability," 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 $\|u\| \le M$ for each u in F. A strategy for player 1, resp. player 2, is an element $\sigma = (\sigma_t)_{t\ge 1}$, where σ_t maps F^{t-1} into $\Delta(I)$, resp. $\Delta(J)$. Strategy spaces for player 1 and 2 are, respectively, denoted by Σ and $\mathcal{T}$. A strategy profile (σ, τ) 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: $\forall \varepsilon > 0,\ \exists \sigma \in \Sigma,\ \exists T,\ \forall \tau \in \mathcal{T},\ \forall t \ge T,\ \mathbb{E}_{\sigma,\tau}(d_t) \le \varepsilon$. C is excludable by player 1 if there exist $\delta > 0$ such that $\{z \in \mathbb{R}^K, d(z, C) \ge \delta\}$ 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 $\mathbb{E}_{\sigma,\tau}(d_t) \le \varepsilon$, so the average payoff will be ε-close to C with high



Repeated Games with Incomplete Information, Fig. 3 The Blackwell property 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 $\max_{x\in\Delta(I)} \min_{y\in\Delta(J)} G(x, y) = \min_{y\in\Delta(J)} \max_{x\in\Delta(I)} G(x, y)$.

Necessary and Sufficient Conditions for Approachability

Given a mixed action x in $\Delta(I)$, we write xG for the set of possible vector payoffs 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 (Fig. 3).

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 σ 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 $t = 0$, σ plays arbitrarily at stage $t + 1$. Otherwise, σ 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 σ adapted to C satisfies:

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

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

Proof Assume that player 1 plays σ adapted to C, whereas player 2 plays some strategy τ. 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:

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

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

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

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

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

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) $\forall q \in \mathbb{R}^K$, $\max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k \in K} q^k G^k(x,y) \geq \inf_{c \in C} < q,c >$.

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:

$$< z' - z, G(x,y) > \geq \min_{c \in c} < z' - z, c > = < z' - z, z' > .$$

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

(iv) means that any half-space containing C is approachable by player 1.

(iii) $\Rightarrow$ (iv) is thus clear. (iv) $\Rightarrow$ (i) is similar to (ii) $\Rightarrow$ (i). ▢

Up to minor formulation differences, Theorems 2 and 3 are due to Blackwell (1956). Later on, X. Spinat (2002) 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 $d_t \rightarrow_{t \to \infty} 0 \quad \mathbb{P}_{\sigma,\tau} a.s$ in the definition of approachability does not modify the notion.

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 Sorin 2002).

Weak Approachability

On 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 (1992) 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.

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 cav$u(p)$, and Blackwell approachability will allow for the construction of an *explicit* optimal strategy for the uninformed player. Considering a hyperplane which is tangent to cav$u$ at p, we can find a vector l in $\mathbb{R}^K$ such that

$$< l, p > = \text{cav}u(p) \quad \text{and} \quad \forall q \in \Delta(K), < l, q > \geq \text{cav}u(q) \geq u(q).$$

Define now the orthant $C = \{z \in \mathbb{R}^K, z^k \le l^k \, \forall k \in K\}$. Recall that player 2 does not know the selected state, and an optimal strategy for him cannot 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 $\inf_{c \in C} < q, c > = -\infty \le \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} \sum_{k \in K} q^k G^k(x, y)$. Assume now that $q^k \le 0$ for each k, with $q \ne 0$. Write $s = \sum_k (-q^k)$.

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

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

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

(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,

$$
\begin{aligned}
\gamma_T^{1,p}(\sigma, \tau) &= \sum_{k \in K} p^k \left(\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\sigma,\tau}\left(G^k\left(\tilde{i}_t, \tilde{j}_t\right)\right)\right) \\
&\le < p, l > + \frac{2M}{\sqrt{T}} = \mathrm{cav}u(p) + \frac{2M}{\sqrt{T}}.
\end{aligned}
$$

As shown by Kohlberg (1975), the approachability strategy τ is thus an optimal strategy for player 2 in the repeated game $\Gamma(p)$.

No-Regret Strategies

The theory of approachability can also be used to prove the existence of "no- regret strategies."

Consider a decision-maker, who has to select at each stage n some action i_n in a finite set I. The environment (nature, adversary, other agents following their own goals) will select a stochastic process $(j_n)_{n \ge 1}$ with values in a finite set J, and the decision-maker knows a priori nothing about the way the sequence $(j_n)_n$ is chosen. There is a given payoff function $g: I \times J \to \mathbb{R}$, known by the decision-maker, and at the end of each stage n the decision-maker observes jn and receives the payoff $g(in, jn)$.

Basic Example $I = J = \{0, 1\}$, and $g(i, j) = 1$ if and only if $i = j$: player 1 tries to guess at each stage n the value j_n (a stage could correspond to a day, $j_n = 0$ meaning there is no rain on day n, $j_n = 1$ meaning there is some rain on day n).

What means to play well for the decision-maker? In the basic example, is it good to guess correctly 90% of the stages? Probably not if it happens that $j_n = 1$ for each n.

A strategy for the decision-maker is an element $\sigma = (\sigma_t)_{t \ge 1}$, where for each t σ_t is a mapping from $(I \times J)^{t-1}$ to $\Delta(I)$. A strategy for Nature is an element $\tau = (\tau_t)_{t \ge 1}$, where for each t τ_t is a mapping from $(I \times J)^{t-1}$ to $\Delta(J)$. The sets of strategies of the decision-maker and Nature are, respectively, denoted by Σ and $\mathcal{T}$, and a strategy profile (σ, τ) naturally induces a unique probability on $(I \times J)^\infty$ denoted by $\mathbb{P}_{\sigma,\tau}$. For each stage n, $g_n = g(i_n, j_n)$ is the random variable of the payoff of stage n, and $\bar{g}_n = \frac{1}{n}\sum_{t=1}^{n} g_t$.

Suppose that at the end of some stage n, $(i_1, j_1, \ldots, i_n, j_n)$ has been played. The average payoff for the decision-maker is $\bar{g}_n$, and he can compare this payoff with the payoff he would have got if he had played constantly some action i in I. The difference $\frac{1}{n}\sum_{t=1}^{n} g(i, j_t) - \bar{g}_n$ is called the regret of the decision-maker for not having played constantly action i.

Definition 6 A strategy σ of the decision-maker has no external regret if for all strategy τ of Nature,

$$\limsup_{n\to\infty}\left(\max_{i\in I}\frac{1}{n}\sum_{t=1}^{n}g(i,j_t)-\bar{g}_n\right)\le 0 \quad \mathbb{P}_{\sigma,\tau}a.s.$$

We now define the stronger notion of internal regret: for each pair of actions i and l in I, we do not want the decision-maker to regret to have played action 18 l at each stage where he actually played action i. For $n \geq 1$, i and l in I, let us introduce the random variable:

$$\bar{R}_n(i,l)=\frac{1}{n}\sum_{t\in\{1,..,n\},i_t=i}(g(l,j_t)-g(i_t,j_t)).$$

Definition 7 A strategy σ of the decision-maker has no internal regret if for each strategy τ of Nature,

$$\max_{i\in I,l\in I}\bar{R}_n(i,l)\underset{n\to\infty}{\longrightarrow}0 \quad \mathbb{P}_{\sigma,\tau}a.s.$$

Theorem 6 There exists a strategy σ of the decision-maker with no internal regret.

Proof Define $K = I \times I$. Consider the dynamic game with vector payoffs where at each stage n player 1 chooses i_n in I, player 2 chooses j_n in J, and the vector payoff in $\mathbb{R}^K$ is $(r(i_n, j_n))_{i,l}$ with:

$$\forall (i,l)\in K,\quad r(i_n,j_n)_{i,l}=$$
$$\begin{cases} g(l,j_n)-g(i,j_n) & \text{if} \quad i_n=i\\ 0 & \text{if} \quad i_n\neq i\end{cases}$$

If $i_n = i$, $r(i_n, j_n)i,l$ is the difference between the payoff that the decision-maker could have got at stage n by playing l and what he actually got. We denote by $r_n = r(i_n, j_n)$ the vector of regrets at stage n in $\mathbb{R}^K$, and we write $\bar{r}_n = \frac{1}{n}\sum_{t=1}^{n} r_t$. Notice that $\bar{r}_n$ is nothing but the regret vector $\bar{R}_n$.

Let $C = \mathbb{R}^K_-$ be the negative orthant of $\mathbb{R}^K$. We use the Euclidean norm in this proof. For each r in $\mathbb{R}^K$, the projection of r to C is given by $\pi_C(r) = (\min\{r(k), 0\})_{k\in K}$, and $d(r, C) = \|r^+\|$, where $r^+ = (\max\{r(k), 0\})_{k \in K}$.

We now show that condition (ii) of Theorem 3 is satisfied. Fix $y = (y_j)_{j\in J}$ in $\Delta(J)$ and consider i^* achieving $\max_{l\in I} g(l, y)$. $r(i^*, y)_{i,l} = 0$ if $i \neq i^*$, and $r(i^*, y)_{i*,l} = g(l, y) - g(i^*, y) \leq 0$. So C is a B-set by Theorem 3, and by Theorem 2 we have the existence of a strategy σ of the decision-maker such that for each strategy τ of Nature:

$$\forall n\ge 1,\ \mathbb{E}_{\sigma,\tau}\left(\|\bar{r}_n^+\|\right)\le\frac{2M}{\sqrt{n}}\quad\text{and}$$
$$\|\bar{r}_n^+\|\underset{n\to\infty}{\longrightarrow}0\quad \mathbb{P}_{\sigma,\tau}\ a.s.$$

where M is the constant $\max_{i,j}\|r(i, j)\|$. This concludes the proof of Theorem 6: for each strategy τ of Nature, stage n and all i and l in I,

$$\mathbb{E}_{\sigma,\tau}\bar{R}_n(i,l)\le\frac{2M}{\sqrt{n}}\quad\text{and}$$
$$\max_{i\in I,l\in I}\bar{R}_n(i,l)\underset{n\to\infty}{\longrightarrow}0\quad\mathbb{P}_{\sigma,\tau}.$$

Zero-Sum Games with Lack of Information on Both Sides

The following model has also been introduced by Aumann and Maschler (1995). 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: $\gamma_T^{1,p,q}(\sigma^1,\sigma^2) = \mathbb{E}^{p,q}_{\sigma^1,\sigma^2}\left(\frac{1}{T}\sum_{t=1}^{T} G^{\tilde{k},\tilde{l}}\left(\tilde{i}_t,\tilde{j}_t\right)\right)$, and the T -stage value is written vT (p, q). Similarly, the λ-discounted value of the game will be written $v_\lambda(p, q)$.

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

Maxmin and Minmax of the Repeated Game

Theorem 1 generalizes as follows.

Theorem 7 Aumann and Maschler (1995) In the repeated game $\Gamma(p, q)$, the greatest quantity which can be guaranteed by player 1 is $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p, q)$, and the smallest quantity which can be guaranteed by player 2 is $\mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p, q)$.

Aumann, Maschler, and Stearns also showed that $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p, q)$ can be *defended* by player 2, uniformly in time, i.e., that $\forall \varepsilon > 0, \forall \sigma^1, \exists T_0, \exists \sigma^2, \forall T \ge T_0, \gamma_T^{p,q}(\sigma^1, \sigma^2) \le \mathrm{cav}_{\mathrm{I}} \mathrm{vex}_{\mathrm{II}} u(p, q) v + \varepsilon$. Similarly, $\mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p, q)$ 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 σ^1 of the first player, it is defined as: $I(\sigma^1) = \sup_{\sigma^2} \mathbb{E}^{p,q}_{\sigma^1,\sigma^2}\left(\sum_{k\in K}\sum_{t=0}^{\infty}\left(p_{t+1}^k(\sigma^1) - p_t^k(\sigma^1)\right)^2\right)$, where $p_t(\sigma^1)$ is the *a posteriori* on K of player 2 after stage t given that player 1 uses σ^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 7 implies that $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p,q) = \sup_{\sigma^1\in\Sigma^1} \liminf_T \left(\inf_{\sigma^2\in\Sigma^2} \gamma_T^{1,p,q}(\sigma^1,\sigma^2)\right)$, and $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p, q)$ is called the *maxmin* of the repeated game $\Gamma(p, q)$. Similarly, $\mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p,q) = \inf_{\sigma^2\in\Sigma^2} \limsup_T \left(\sup_{\sigma^1\in\Sigma^1} \gamma_T^1(\sigma^1,\sigma^2)\right)$ 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: $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p, q) = \mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p, q)$. 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.

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

Mertens and Zamir (1971) have shown that here, $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p,q) = -\frac{1}{4} < 0 = \mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p,q)$.

Limit Values

It is easy to see that for each T and λ, the value functions v_T and v_λ 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_λ are viewed as elements of the set $\mathcal{C}$ of continuous mappings from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$.

Theorem 8 Mertens and Zamir (1971) $(v_T)_T$, as T goes to infinity, and $(v_\lambda)_\lambda$, as λ goes to zero, both uniformly converge to the unique solution f of the following system:

$$\begin{cases} f = \mathrm{vex_{II}} \max\{u,f\} \\ f = \mathrm{cav_I} \min\{u,f\} \end{cases}$$

And the convergence of (v_T), resp. (v_λ) is in $O\left(1\sqrt{T}\right)$, resp. $O(\lambda)$.

The above system can be fruitfully studied without reference to repeated games (see Laraki 2001a, b; Mertens and Zamir 1977; Sorin 1984b).

Remark Let U be the set of all nonrevealing value functions, i.e., of functions from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$ satisfying Eq. (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.

Correlated Initial Information

A more general model can be written, where it is no longer assumed that the initial information of the players is 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., $(r \to x^r)$ 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:

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

$\Pi^{\mathrm{I}}(p)$ 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 $\Pi^{\mathrm{I}}(p)$ is concave. And given g: $\Delta(R) \to \mathbb{R}$ which is bounded from above, we define the concavification $\mathrm{cav_I}\, g$ as the smallest function above g which is I- concave. Similarly one can define the set $\Pi^{\mathrm{II}}(p)$ and the notions of II-convexity and II-convexification. With these generalized definitions, the results of Theorem 7 and 8 perfectly extend (Mertens and Zamir 1971).

Nonzero-sum Games with Lack of Information on One Side

We now consider the generalization of the standard model of section "The Standard Model of Aumann and Maschler" to the nonzero-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 (σ^1, σ^2), it is here convenient to denote the expected payoffs up to stage T by:

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

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

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 $B(q)(\gamma) = \sum_{(i,j)\in I\times J} \gamma(i,j)B(q)(i,j)$.

Existence of Equilibria

The question of existence of an equilibrium has remained unsolved for long. Sorin (1983) proved the existence of an equilibrium for two states of nature, and the general case has been solved by Simon et al. (1995).

Exactly as in the zero-sum case, a strategy pair σ 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 8 A joint plan is a triple (S, λ, γ), where:

- S is a finite non empty set (of messages),
- $\lambda = (\lambda^k)_{k\in K}$ (signaling strategy) with for each k, $\lambda^k \in \Delta(S)$ and for each s, $\lambda_s =_{def} \sum_{k\in K} p^k \lambda_s^k > 0$,
- $\gamma = (\gamma_s)_{s\in S}$ (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 λ^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, λ, γ), we define:

- $\forall s \in S$, $p_s = \left(p_s^k\right)_{k\in K} \in \Delta(K)$, with $p_s^k = \frac{p^k\lambda_s^k}{\lambda_s}$ for each k. ps 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 9 A joint plan (S, λ, γ) is an equilibrium joint plan if:

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

Condition (*ii*) can be seen as an incentive condition for player 1 to choose s according to λ^k. Given an equilibrium joint plan (S, λ, γ), one define a strategy pair $(\sigma^{1*}, \sigma^{2*})$ adapted to it. For each message s, first fix a sequence $\left(i_t^s, j_t^s\right)_{t\geq 1}$ of elements in $I \times J$ such that for each (i,j), the empirical frequencies converge to the corresponding probability: $\frac{1}{T}\left|\left\{t, 1\leq t\leq T, \left(i_t^s, j_t^s\right) = (i,j)\right\}\right| \to_{T\to\infty} \gamma_s(i,j)$. 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. σ^{1*} is precisely defined as follows. Player 1 observes the selected state k, then chooses s according to λ^k, and announces s to player 2 by playing $f(s)$ at the first l stages. Finally, σ^{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 ps and payoffs for player 1 given by $(-B^k)_{k\in K}$. We now define σ^{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 σ^{-2} such that: $\forall \varepsilon > 0$, $\exists T_0$, $\forall T \geq T_0$, $\forall \sigma^1 \in \Sigma^1$, $\forall k \in K$, $\alpha_k(\sigma^1, \sigma^{-2}) \leq \varphi^k + \varepsilon$. Such a strategy σ^{-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 "Back to the Standard Model").

Lemma 7 Sorin (1983) 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:

$$\forall k, \alpha_T^k(\sigma^{1*}, \sigma^{2*}) \to_{T\to\infty} \sum_{s\in S} \lambda_s^k A^k(\gamma_s) = \varphi^k$$

because of (*ii*), and $\beta_T^p(\sigma^{1*}, \sigma^{2*}) \to_{T\to\infty} \sum_{k\in K} p^k \sum_{s\in S} \lambda_s^k B^k(\gamma_s) = \psi$.

Assume that player 2 plays σ^{2*}. The existence of σ^{-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 φ^k. The proof can be made uniform in σ^1 and we obtain: $\forall \varepsilon > 0\ \exists T_0\ \forall T \geq T_0, \forall k \in K, \forall \sigma^1 \in \Sigma^1, \alpha_T^k(\sigma^1, \sigma^{2*}) \leq \varphi^k + \in$. Finally assume that player 1 plays σ^{1*}. Condition (i) implies that if player 2 uses σ^{2*}, the payoff of this player will be at least vex $v(p_s)$ if the message is s. Since vex $v(p_s)$ ($= -$ 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 $\forall \varepsilon > 0, \exists T_0, \forall T \geq T_0, \forall \sigma^2 \in \Sigma^2, \beta_T^p(\sigma^{1*}, \sigma^2) \leq \sum_{s \in S} \lambda_s \psi_s + \in = \psi + \varepsilon$. □

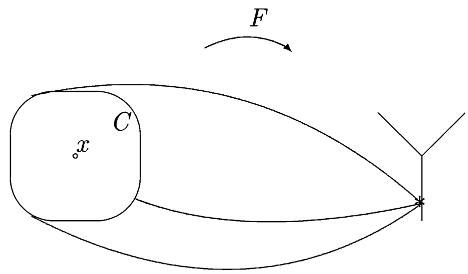
To prove the existence of equilibria, we then look for equilibrium joint plans. The first idea is to consider, for each probability r on K, the set of payoff vectors φ compatible with r being an *a posteriori*. This leads to the consideration of the following correspondence (for each r, $\Phi(r)$ is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

$$\Phi : \Delta(K) \rightrightarrows \mathbb{R}^K$$

$r \mapsto \{(A^k(\gamma))_{k \in K}$, where $\gamma \in \Delta(I \times J)$ satisfies $B(r)(\gamma) \geq \text{vex}\ v(r)\}$. It is easy to see that the graph of Φ, i.e., the set $\{(r, \varphi) \in \Delta(K) \times \mathbb{R}^K, \varphi \in \Phi(r)\}$, is compact that Φ has nonempty convex values and satisfies: $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), < \varphi, q > \geq u(q)$.

Assume now that one can find a finite family $(p_s)_{s \in S}$ of probabilities on K, as well as vectors φ and, for each s, φ_s in $\mathbb{R}^K$ such that: (1) $p \in$ conv $\{p_s, s \in S\}$, (2) $< \varphi, q > \geq u(q)\ \forall q \in \Delta(K)$, (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: $\Delta(K) \to \mathbb{R}$ be a continuous mapping, and Φ: $\Delta(K) \rightrightarrows \mathbb{R}^K$ be a correspondence with compact graph and nonempty convex values such that: $\forall r \in \Delta(K), \forall q \in \Delta(K), \exists \varphi \in \Phi(r), < \varphi, q > \geq u(q)$. Then there exists a finite family $(p_s)_{s \in S}$ of elements of $\Delta(K)$, as well as vectors φ and, for each s, φ_s in $\mathbb{R}^K$ such that:



Repeated Games with Incomplete Information, Fig. 4 A Borsuk-Ulam type theorem by Simon, Spież, and Toruńczyk

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

The proof of Proposition 4 relies, as explained in Renault (2000) or Simon (2002), on a fixed point theorem of Borsuk-Ulam type proved by Simon et al. (1995) via tools from algebraic geometry. A simplified version of this fixed point theorem can be written as follows:

Theorem 9 Simon et al. (1995)**:** 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 nonempty convex values. Then there exists $L \subset \partial C$ and $y \in Y$ such that: $\forall l \in L, y \in F(l)$, and $x \in$ conv (L) (Fig. 4).

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 10 Simon et al. (1995)**:** There exists an equilibrium joint plan. Thus, there exists a uniform equilibrium in the repeated game $\Gamma(p)$.

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 10 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) $\forall \varepsilon > 0\ \exists T_0\ \forall T \geq T_0$, $\forall k \in K$, $\forall \sigma^1 \in \Sigma^1$, $\alpha_T^k(\sigma^1,\sigma^{2*}) \leq \alpha_T^k(\sigma^{1*},\sigma^{2*}) + \varepsilon$, $\forall \varepsilon > 0\ \exists T_0$ $\forall T \geq T_0$, $\forall \sigma^2 \in \Sigma^2$, $\beta_T^{p0}(\sigma^{1*},\sigma^2) \leq \beta_T^p(\sigma^{1*},\sigma^{2*})$ $+\varepsilon$, and (ii) $\left(\alpha_T^k(\sigma^{1*},\sigma^{2*})\right)_{k,T}$ and $\left(\beta_T^{p0}(\sigma^{1*},\sigma^{2*})\right)_T$ 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 "Back to the Standard Model" one can obtain that:

$$< a,\ q > \ \geq u(q)\ \ \forall q \in \Delta(K) \qquad (4)$$

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: $\forall \varepsilon > 0\ \exists T_0\ \forall T \geq T_0$, $\forall \sigma^2 \in \Sigma^2$, $\beta_T^{p0}(\sigma^{1*},\sigma^2) \leq \beta + \varepsilon$, so by Theorem 1:

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

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 σ^{1*} is a nonrevealing 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 11 Let G be the set of triples $(a,\beta,p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ satisfying:

1. $\forall q \in \Delta(\mathrm{K})$, $<\mathrm{a, q}> \geq \mathrm{u(q)}$,
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 nonrevealing equilibria of $\Gamma(p_0)$. The importance of the following definition will appear with Theorem 11 below (which unfortunately has not led to a proof of existence of equilibrium payoffs).

Definition 12 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-σ-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_∞ 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 11 Hart (1985) Let (a, β) be in $\mathbb{R}^K \times \mathbb{R}$.

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

Theorem 11 is too elaborate to be proved here, but let us give a few ideas about the proof. First consider the implication $\Rightarrow$ and fix an equilibrium $\sigma^* = (\sigma^{1*}, \sigma^{2*})$ of $\Gamma(p_0)$ with payoff (a, β). 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 nonrevealing equilibrium at a "limit *a posteriori*," so the convergence will be towards elements of G.

Consider now the converse implication $\Leftarrow$. Let (a, β) be such that $(a, \beta, p_0) \in G^*$ and assume for simplification that the associated bi-martingale (a_n, β_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, β) along the following lines. For each index n, (a_n, β_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 nonrevealing equilibrium of the repeated game $\Gamma(p_N)$ with payoff (a_N, β_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, β_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, β_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 (1995), 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, β^1) or to play the equilibrium $E2$ with payoff (a^2, β^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 similarly with j and j'. $\begin{array}{cc} & \begin{array}{cc} j & j' \end{array} \\ \begin{array}{c} i \\ i' \end{array} & \begin{pmatrix} \times & \\ & \times \end{pmatrix} \end{array}$. 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 signaling and jointly controlled lotteries to construct an equilibrium of $\Gamma_\infty(p_0)$ with payoff (a, β).

Biconvexity and Bimartingales

The previous analysis has led to the introduction and study of biconvexity phenomena. The reference here is Aumann and Hart (1986). Let X and Y be compact convex subsets of Euclidean spaces, and let $(\Omega, \mathcal{F}, \mathcal{P})$ be an atomless probability space.

Definition 13 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: $B \to \mathbb{R}$ is called biconvex if for each $(x, y) \in X \times Y, f(., y)$ and $f(x,.)$ are convex.

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

Definition 14 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:

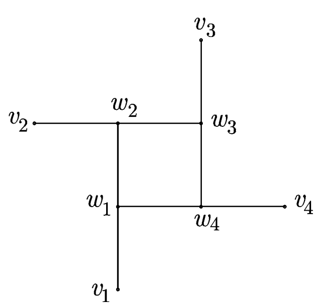
(1) There exists an increasing sequence $(\mathcal{F}_n)_{n \geq 1}$ of finite sub-σ-algebra of $\mathcal{F}$ such that $(Z_n)_n$ is a $(\mathcal{F}_n)_{n \geq 1}$-martingale.
(2) $\forall n \geq 1, X_n = X_{n+1}$ a.s. or $Y_n = Y_{n+1}$ a.s.
(3) Z_1 is a.s. constant.

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

Definition 15 Let A be a measurable subset of $X \times Y$.

$A^* = \{z \in X \times Y$, there exists a bimartingale $(Z_n)_{n \geq 1}$ converging to a limit Z_∞ 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, induces the same set A^*. One can also substitute condition (2) by: $\forall n \geq 1, (X_n = X_{n+1}$ or $Y_n = Y_{n+1})$ a.s. Notice that without condition (2), the set A^* would just be the convex hull of A.



Repeated Games with Incomplete Information, Fig. 5 The "four frogs" example of Aumann and Hart: $A^* \neq$ biconv (A)

We always have $A \subset A^* \subset$ 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 $A^* = \{(x, y) \in [0, 1] \times [0, 1], x = 0$ or $y = 0\}$. A^* always is biconvex and thus contains biconv (A), which is defined as the smallest biconvex set which contains A. The inclusion biconv $(A) \subset A^*$ can also be strict, as shown by the following example:

Example 5 Put $X = Y = [0, 1]$, $v1 = (1/3, 0)$, $v2 = (0, 2/3)$, $v3 = (2/3, 1)$, $v4 = (1, 1/3)$, $w1 = (1/3, 1/3)$, $w2 = (1/3, 2/3)$, $w3 = (2/3, 2/3)$ et $w4 = (2/3, 1/3)$, and $A = \{v1, v2, v3, v4\}$ (Fig. 5).

A is biconvex, so $A =$ 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 $Z_{n+1} = w_{i+1(mod\ 4)}$ 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 $w_1 \in A^* \backslash$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 cannot be separated from A by a continuous bounded biconvex function on A. More precisely, $nsc(B) = \{z \in B, \forall f$: $B \to \mathbb{R}$ bounded biconvex, and continuous on $A, f(z) \leq \sup\{f(z'), z' \in A\}\}$.

Theorem 12 Aumann and Hart (1986): 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 "Characterization of Equilibrium Payoffs." 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 $\mathbb{R}^K_M \times \mathbb{R}_M \times \Delta(K)$. For $B \subset \mathbb{R}^K_M \times \mathbb{R}_M \times \Delta(K)$, B is biconvex if for each a in $\mathbb{R}^K_M$ 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 $\forall a, \forall p, f(a,.,.)$ and $f(.,., p)$ are convex.

Theorem 13 Aumann and Hart (1986): G^* is the largest biconvex set B containing G such that: $\forall z \in B, \forall f: B \to \mathbb{R}$ bounded biconvex, and continuous on A, $f(z) \le \sup\{f(z'), z' \in G\}$.

Nonobservable Actions

We now consider the case where, as in the general definition of section "Definition of the Subject and Its Importance," there is a signaling function q: $K \times A \to \Delta(U)$ 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 signaling*. 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 (Aumann and Maschler 1995) to the general case of signaling function. We keep the notations of section "The Standard Model of Aumann and Maschler." 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 nonrevealing strategies of player 1 is then defined as: $NR(p) = \left\{x = \left(x^k\right)_{k \in K} \in \Delta(I)^K,\ \forall k \in K, \forall k' \in K \text{ s.t. } p^k p^{k'} > 0, \forall j \in J, Q\left(k, x^k, j\right) = Q(k', x^{k'}, j\right\}$. 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 nonrevealing game becomes:

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

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 $\mathrm{cav}\,u(p)$.

The explicit construction of an optimal strategy of player 2 (see section "Back to the Standard Model" here) has also been generalized to the general signaling case (see Kohlberg 1975; Mertens et al. 1994, part B, p.234 for random signals).

Regarding zero-sum games with lack of information on both sides, the results of section "Zero-Sum Games with Lack of Information on Both Sides" have been generalized to the case of state independent signaling (see Mertens 1972; Mertens and Zamir 1971, 1977). 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 signaling: 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 (Zamir 1971, 1973). For state-dependent signaling and lack of information on one side, it was shown by Mertens (1998) 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 signaling has been studied (games with no signals, see (Mertens and Zamir 1976b; Sorin 1989; Waternaux 1983). 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 nonzero-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 signaling (Renault 2000) and more generally to the case where "player 1 can send non revealing signals to player 2" (Simon et al. 2002). The existence of a uniform equilibrium in the general signaling case is still an open question (see Simon et al. 2008).

Advances

1. Zero-sum games with lack of information on one and a half side

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. Sorin and Zamir (1985) provide 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.

2. v_T and v_λ as a function of p

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 Ponssard and Sorin 1980 for more generality). On the contrary, the discounted value v_λ can be quite a complex function of p: in Example 2 of section "Definition of the Subject and Its Importance," Mayberry (1967) has proved that for $2/3 < \lambda < 1$, υ_λ is, at each rational value of p, nondifferentiable.

3. $\lim_T \sqrt{T}(\upsilon_T(p) - \mathrm{cav}u(p))$ and the normal distribution

Convergence of the value functions $(\upsilon_T)_T$ and $(\upsilon_\lambda)_\lambda$ has been widely studied. We have already mentioned the speed of convergence in section "Non-Observable Actions," 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 (Mertens and Zamir 1976a) that for each $p \in [0, 1]$, viewed as the initial probability of state a, the sequence $\sqrt{T}\,\upsilon_T(p)$ 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 $\sqrt{T}\ \upsilon_T(p)$ 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 (de Meyer 1996a, b, 1998, 1999; de Meyer and Moussa Saley 2003, ...).

4. The dual game

B. de Meyer introduced the notion of "dual game" (see the previous references and also de Meyer and Marino 2005; de Meyer and Rosenberg 1999; Laraki 2002; Rosenberg 1998). Let us now illustrate this on the standard model of section "The Standard Model of Aumann and Maschler."

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 $t \le T$, the players choose as usual actions it and jt 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:

$$w_T(z) = \max_{p \in \Delta(K)} (\upsilon_T(p) - < p, z >), \text{ and}$$

$$\upsilon_T(p) = \inf_{z \in \mathbb{R}^K} (w_T(z) + < p, z >).$$

And $(wT)T$ satisfies the dual recursive formula:

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

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 Heuer 1992 on this topic).

5. Approachability

Blackwell's approachability theorem has been extended to infinite dimensional spaces by Lehrer (2003a). As we saw in Theorem 6, approachability theory has strong links with the existence of no-regret strategies (first studied in Hart and Mas-Colell 2000), see also Cesa-Bianchi et al. (2006); Foster and Vohra (1999); Hart (2005); Lehrer (2003b); Rustichini (1999) and the book Cesa-Bianchi and Lugosi (2006), but also with convergence of simple procedures to the set of correlated equilibria (Hart and Mas-Colell 2000) and calibration (Foster 1999; Lehrer 2001). The links between merging, reputation phenomena, and repeated games with incomplete information have been studied in (Sorin 1997), where several existing results are unified. And no-regret and approachability have also been studied when the players have bounded computational capacities (finite automata, bounded recall strategies) (Lehrer and Solan 2003, 2006).

6. Markov chain games with lack of information

In Renault (2006), the standard model of lack of information, 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 Neyman (2008) for non-observable actions). The limit value is however difficult to compute, as shown by the following example from Renault (2006): $K = \{a, b\}$, the payoff matrices are $G^a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $G^b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, the initial probability is $(1/2, 1/2)$, and the state evolves according to the Markov chain $M = \begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}$ with parameter α. If $\alpha = 1$ this is Example 2, and the limit value is 1/4 by Theorem 1.

For $\alpha \in [1/2, 2/3]$, the limit value is $\frac{\alpha}{4\alpha-1}$ (Marino 2005) for $\alpha = 2/3$, (Hörner et al. 2010).

For $\alpha \in [2/3, .719]$, the limit value υ satisfies (Bressaud and Quas 2006): $\frac{1}{\upsilon} = u_0 + u_0 u_1 + u_0 u_1 u_2 + \ldots$, where (u_n) is defined by $u_0 = 1$ and $u_{n+1} = \max\{\psi(u_n), 1 - \psi(u_n)\}$, with $\psi(u) = 3\alpha - 1 - \frac{2\alpha - 1}{u}$. What is the value for $\alpha = 0.9$?

In Markov chain games with lack of information on both sides, each player privately observes his state variable, and both state variables follow exogenous and independent Markov chains. For such games, the existence of the limit value $\lim_T v_T = \lim_\lambda v_\lambda$ has been proved in Gensbittel and Renault (2015). In the case of recurrent and aperiodic chains, the limit value is identified as the unique solution of the Mertens-Zamir system of Theorem 8, with an appropriate nonrevealing function $\hat{u}$ instead of u, corresponding to the limit value of the auxiliary dynamic game where each player is restricted to play strategies that reveal no information on the recurrence class of his own state (such a nonrevealing function was already considered in Renault (2006)).

In the nonzero-sum context, dynamic sender-receiver games are Markov chain games with lack of information on one side, where the payoffs only depend on the actions of the uninformed player. The set of equilibrium payoffs has been characterized under a homothety (random shocks) assumption on the Markov chain (Renault et al. 2013).

7. Extension to zero-sum dynamic games with state process controlled and observed by player 1

It is known since (Sorin 1984a) that the uniform value may not exist in general for stochastic games with lack of information on one side on the payoff matrices (where the payoff matrices of the stochastic game to be played are first randomly selected and announced to player 1 only). Rosenberg et al. (2004) studied stochastic games with a single controller and lack of information on one side on the payoff matrices, showing the existence of the uniform value if the informed player controls the transition, and providing a counter-example if the uninformed player controls the transitions. One can also consider the model of general repeated games with an informed controller (Renault 2012), generalizing the model of Markov chain games with lack of information on one side), i.e., dynamic games with finitely many states, actions and signals, and state processes controlled and observed by player 1.

A general repeated game is given by: 5 non empty finite sets: a set of states or parameters K, a set I of actions for player 1, a set J of actions for player 2, a set C of signals for player 1, and a set D of signals for player 2, an initial distribution $\pi \in \Delta(K \times C \times D)$, a payoff function g: $K \times I \times J$ to $[0, 1]$ for player 1, and a transition function q: $K \times I \times J$ to $\Delta(K \times C \times D)$. The progress of the game is the following: Initially, (k_1, c_1, d_1) is selected according to π, player 1 learns c_1 and player 2 learns d_1. Then simultaneously player 1 chooses i_1 in I and player 2 chooses j_1 in J, and the payoff for player 1 at stage 1 is $g(k_1, i_1, j_1)$, etc. At any stage $t \geq 2$, (k_t, c_t, d_t) is selected according to $q(k_{t-1}, i_{t-1}, j_{t-1})$, player 1 learns ct and player 2 learns dt. Simultaneously, player 1 chooses it in I and player 2 chooses jt in J. The stage payoffs are $g(kt, it, jt)$ for player 1 and the opposite for player 2, and the play proceeds to stage $t + 1$.

In repeated games with an informed controller, it is moreover assumed that:

1) Player 1 is fully informed, in the sense that π and q are such that he can always deduce the state and player 2's signal from his own signal.
2) Player 1 controls the transition, in the sense that the marginal $\bar{q}$ of the transition q on $K \times D$ does not depend on player 2's action.

In this setup, one can prove (Renault 2012) the existence of the uniform value $\upsilon^*(\pi)$, satisfying:

$$\upsilon^*(\pi) = \inf_{n\geq 1} \sup_{m\geq 0} \upsilon_{m,n}(\pi) = \sup_{m\geq 0} \inf_{n\geq 1} \upsilon_{m,n}(\pi).$$

where $v_{m,n}(\pi)$ is the value of the game with payoff $\mathbb{E}_{\pi,\sigma,\tau} \frac{1}{n}\left(\sum_{t=m+1}^{m+n} g_t\right)$, g_t being the payoff of stage t.

Moreover, one can prove for such games the existence of the stronger notion of "general uniform value." Let us first define the values $v_\theta(\pi)$ of the dynamic game with payoff $\gamma_\theta(\pi,\sigma,\tau) = \mathbb{E}_{\pi,\sigma,\tau}\left(\sum_{t\geq 1}\theta_t g_t\right)$, where θ is an evaluation $(\theta_t)_{t\geq 1}$ with nonnegative weights satisfying $\sum_{t\geq 1}\theta_t = 1$, and total variation denoted by $TV(\theta) = \sum_t \mid \theta_{t+1} - \theta_t|$. And $\upsilon^*(\pi)$ is the general uniform value of the game with initial probability π if for each $\varepsilon > 0$ one can find $\alpha > 0$ and a couple of strategies σ^* and τ^* such that for all evaluations θ with $TV(\theta) \leq \alpha$:

$$\forall \tau, \gamma_\theta(\pi,\sigma^*,\tau) \geq v^*(\pi) - \varepsilon \quad \text{and} \quad \forall \sigma, \gamma_\theta(\pi,\sigma,\tau^*) \leq v^*(\pi) + \varepsilon.$$

Considering only Cesaro-evaluations (i.e., of the type $\theta_t = 1/n$ for $t \leq n$, $=0$ for $t > n$ for some n) recovers our Definition 1. Renault and Venel (2017) introduce a new distance (compatible with the weak topology) on the belief space $\Delta(\Delta(K))$ of Borel probabilities over the simplex $X = \Delta(K)$ and prove the existence of the general uniform value in general repeated games with an informed controller. Clearly, the values only depend on player 2's belief p on the initial state, and the limit value υ^* can be characterized as:

$$\forall p \in X, v^*(p) = \inf\{w(p), w: \Delta(X) \to [0,1] \text{ affine } C^0 s.t.$$

1. $\forall p' \in X, w(p') \geq \sup_{a\in\Delta(I)^K} w(\bar{q}(p',a))$
2. $\forall (z, y) \in RR, w(z) \geq y\}$.

where $\bar{q}(p,a) = \sum_{k\in K} p^k \bar{q}\left(k,a^k\right) \in \Delta(K \times D)$ gives the marginal of q on $K \times D$, and $RR = \{(z, y) \in \Delta(X) \times [0, 1]$, there exists a: $X \to \Delta(I)^K$ measurable $s.t.$ $\int_{p\in X} \bar{q}(p, a(p))dz(p) = z$ and $\int_{p\in X} \min_{j\in J}(\sum_{k\in K} p^k g(k, a^k, j))\ dz(p) = y\}$ can be seen as the set of invariant measures and associated payoffs. In the standard model of Aumann and Maschler, (1) is equivalent to w being a concave function on $\Delta(K)$ and (2) is equivalent to w being not lower than the non-revealing function u: so v^* is the smallest concave function above u, and we recover the cav$u$ theorem (Theorem 1).

Finally, the existence of the uniform value has been generalized to the case where Player

1 controls the transitions and is more informed than player 2 (but player 1 does not necessarily observe the current state) in Gensbittel et al. (2014).

8. Symmetric information

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 Neyman and Sorin 1998, the zero-sum case being solved in Forges 1982; Kohlberg and Zamir 1974).

This result does not extend to the case where the stage evolves from stage to stage, i.e., to stochastic games with incomplete information. In the zero-sum symmetric information case where at the end of each stage, the players observe both actions but receive no further information on the current state (*hidden stochastic games*), B. Ziliotto provided in his PhD thesis an example where $\lim_T v_T$ and $\lim_\lambda v_\lambda$ may fail to exist (Ziliotto 2016).

One can also consider zero-sum general repeated games with payoffs defined by the expectation of a Borel function over plays. In the public case where he players have the same information at the end of every stage, the value exists (Gimbert et al. 2016).

9. Continuous-time approach

A continuous time approach can also be used to prove convergence results in general zero-sum repeated games, and in particular Theorem 7, embedding the discrete repeated game into a continuous time game and using viscosity solution tools (Cardaliaguet et al. 2012). A generalization of the cav$u$ theorem (Theorem 1) to infinite action spaces and partial information can be found in Gensbittel (2015), using a probabilistic method based on martingales and a functional method based on approximation schemes for viscosity solutions of Hamilton Jacobi equations.

10. The operator approach for zero-sum games

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 (Laraki 2001b; Rosenberg and Sorin 2001; Sorin 2002).

11. Uncertain duration

One can consider zero-sum repeated games with incomplete information on both sides and *uncertain duration*. In these games, the payoff to the players is the sum of their stage payoffs, up to some stopping time θ which may depend on plays, divided by the expectation of θ. Theorem 8 here generalizes to the case of *public* uncertain duration process (as $\mathbb{E}(\theta) \to \infty$, with a convergence in $O\left(1/\mathbb{E}\left(\sqrt{\theta}\right)\right)$, see Neyman and Sorin (2010). The situation is different if one allows for private uncertain duration processes: any number between the maxmin $\text{cav}_I\text{vex}_{II}u(p, q)$ and the minmax $\text{vex}_{II}\text{cav}_I u(p, q)$ is the value of a long finitely repeated game Γ_T where players' information about the uncertain number of repetitions T is asymmetric (Neyman 2012).

12. Frequent actions

One can consider a repeated game with incomplete information and fixed discount factor, where the time span between two consecutive stages is $1/n$. In the context of zero-sum Markov chain games of lack of information on one side, Cardaliaguet et al. (2016) show the existence of a limit value when n goes to $+\infty$; this value is characterized through an auxiliary stochastic optimization problem and, independently, as the solution of an Hamilton-Jacobi equation.

13. Repeated market games with incomplete information

De Meyer and Moussa Saley studied the modelization via Brownian motions in financial models (de Meyer and Moussa Saley 2003). 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 also de Meyer and Marino 2004 for incomplete information on both sides, and de Meyer 2010).

14. Cheap-talk and communication

In the nonzero-sum setup of section "Non Zero-Sum Games with Lack of Information on One Side," 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 Aumann and Hart 1986; Forges 1984, 1990; Aumann and Maschler 1995). Let us mention also that F. Forges (1988) gave a similar characterization of equilibrium payoffs, for a larger notion of equilibria called communication equilibria (see also Forges 1985 for correlated equilibria). Amitai (1996b) studied the set of equilibrium payoffs in case of lack of information on both sides. Aumann and Hart (2003) characterized the equilibrium payoffs in two player games with lack of information on one side when long, payoff-irrelevant, preplay communication is allowed (see Amitai 1996a for incomplete information on both sides).

15. Known own payoffs

The particular nonzero-sum case where each player knows his own payoffs is particularly worthwhile studying. 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 (1994) showed that any equilibrium payoff can be obtained as the payoff of an equilibrium which is completely revealing. This result generalizes to the nonzero-sum case of lack of information of both sides (see the unpublished manuscript Koren 1992), but uniform equilibria may fail to exist even though both players known their own payoffs.

16. More than 2 players

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 (Renault 2001a), 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 (Renault 2001b; Renault et al. 2014; Renault and Tomala 2004, 2008), with cryptographic considerations. As an application, the existence of completely revealing equilibria, i.e., equilibria where each player eventually learns the state with probability one, is obtained in particular cases (see also Ḧorner et al. 2011 for the related notion of "belief-free" equilibria).

17. Perturbations of repeated games with complete information

Repeated games with incomplete information have been used to study perturbations of repeated games with complete information (see Cripps and Thomas 2003; Fudenberg and Maskin 1986) for Folk theorem-like results (Aumann and Sorin 1989), for enforcing cooperation in games with a Paretodominant outcome, and (Israeli 2010) for a perturbation with known own payoffs). The case where the players have different discount factors has also been investigated (Cripps and Thomas 2003; Lehrer and Yariv 1999).

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 signaling or in n-player repeated games with lack of information on one side? Does the limit value always exist in zero-sum repeated games with incomplete information and signals? 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, learning, and sequential decision problems.

Bibliography

Primary Literature

Amitai M (1996a) Cheap-talk with incomplete information on both sides. PhD thesis, The Hebrew University of Jerusalem. http://ratio.huji.ac.il/dp/dp90.pdf

Amitai M (1996b) Repeated games with incomplete information on both sides. PhD thesis, The Hebrew University of Jerusalem. http://ratio.huji.ac.il/dp/dp105.pdf

Aumann RJ (1964) Mixed and behaviour strategies in infinite extensive games. In: Dresher M, Shapley LS, Tucker AW (eds) Advances in game theory. Annals of Mathematics Study 52. Princeton University Press, pp 627–650

Aumann RJ, Hart S (1986) Bi-convexity and bi-martingales. Israel J Math 54:159–180

Aumann RJ, Hart S (2003) Long cheap talk. Econometrica 71:1619–1660

Aumann RJ, Sorin S (1989) Cooperation and bounded recall. Games Econom Behav 1:5–39

Blackwell D (1956) An analog of the minmax theorem for vector payoffs. Pac J Math 65:1–8

Bressaud X, Quas A (2006) Dynamical analysis of a repeated game with incomplete information. Math Oper Res 31:562–580

Cardaliaguet P, Rainer C, Rosenberg D, Vieille N (2016) Markov games with frequent actions and incomplete information? The limit case. Math Oper Res 41:49–71

Cardaliaguet P, Laraki R, Sorin S (2012) A continuous time approach for the asymptotic value in two-person zero-sum repeated games. SIAM J Control Optim 50:1573–1596

Cesa-Bianchi N, Lugosi G (2006) Prediction, learning and games. Cambridge University Press, Cambridge

Cesa-Bianchi N, Lugosi G, Stoltz G (2006) Regret minimization under partial monitoring. Math Oper Res 31:562–580

Cripps MW, Thomas JP (2003) Some asymptotic results in discounted repeated games of one-sided incomplete information. Math Oper Res 28:433–462

de Meyer B (1996a) Repeated games and partial differential equations. Math Oper Res 21:209–236

de Meyer B (1996b) Repeated games, duality and the central limit theorem. Math Oper Res 21:237–251

de Meyer B (1998) The maximal variation of a bounded martingale and the central limit theorem. Ann Inst Henri Poincaré Probab Stat 34:49–59

de Meyer B (1999) From repeated games to Brownian games. Annales de l'Institut Henri Poincaré, Probabilités et statistiques 35:1–48

de Meyer B (2010) Price dynamics on a stock market with asymmetric information. Games Econom Behav 69:42–71

de Meyer B, Marino A (2004) Repeated market games with lack of information on both sides. DP 2004.66, MSE Université Paris I

de Meyer B, Marino A (2005) Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides. DP 2005.27, MSE Université Paris I

de Meyer B, Moussa Saley H (2003) On the strategic origin of Brownian motion in finance. Int J Game Theory 31:285–319

de Meyer B, Rosenberg D (1999) "Cavu" and the dual game. Math Oper Res 24:619–626

Forges F (1982) Infinitely repeated games of incomplete information: symmetric case with random signals. Int J Game Theory 11:203–213

Forges F (1984) A note on Nash equilibria in repeated games with incomplete information. Int J Game Theory 13:179–187

Forges F (1985) Correlated equilibria in a class of repeated games with incomplete information. Int J Game Theory 14:129–149

Forges F (1988) Communication equilibria in repeated games with incomplete information. Math Oper Res 13:191–231

Forges F (1990) Equilibria with communication in a job market example. Q J Econ 105:375–398

Foster D (1999) A proof of calibration via Blackwell's approachability theorem. Games Econom Behav 29:73–78

Foster D, Vohra R (1999) Regret in the on-line decision problem. Games Econom Behav 29:7–35

Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. Econometrica 54:533–554

Gensbittel F (2015) Extensions of the Cav(u) theorem for repeated games with one-sided information. Math Oper Res 40(1):80–104

Gensbittel F, Renault J (2015) The value of Markov Chain Games with incomplete information on both sides. Math Oper Res 40(4):820–841

Gensbittel F, Oliu-Barton M, Venel X (2014) Existence of the uniform value in repeated games with a more informed controller. J Dynam Games 1(3):411–445

Gimbert H, Renault J, Sorin S, Zielonka W (2016) On values of repeated games with signals. Ann Appl Probab 26:402–424

Harsanyi J (1967-68) Games with incomplete information played by 'Bayesian' players, parts I-III. Manag Sci 8:159–182, 320–334, 486–502
Hart S (1985) Nonzero-sum two-person repeated games with incomplete information. Math Oper Res 10:117–153
Hart S (2005) Adaptative Heuristics. Econometrica 73:1401–1430
Hart S, Mas-Colell A (2000) A simple adaptative procedure leading to correlated equilibrium. Econometrica 68:1127–1150
Heuer M (1992) Optimal strategies for the uninformed player. Int J Game Theory 20:33–51
Ḧorner J, Lovo S, Tomala T (2011) Belief-free equilibria in games with incomplete information: characterization and existence. J Econ Theory (5):1770–1795
Ḧorner J, Rosenberg D, Solan E, Vieille N (2010) On a Markov game with one-sided incomplete information. Oper Res 58:1107–1115
Israeli E (2010) Sowing doubt optimally in two-person repeated games. Games Econom Behav 28:203–216. 1999
Kohlberg E (1975) Optimal strategies in repeated games with incomplete information. Int J Game Theory 4:7–24
Kohlberg E, Zamir S (1974) Repeated games of incomplete information: the symmetric case. Ann Stat 2:40–41
Koren G (1992) Two-person repeated games where players know their own payoffs, master thesis, Tel-Aviv University. http://www.ma.huji.ac.il/hart/papers/koren.pdf
Kuhn HW (1953) Extensive games and the problem of information. In: Kuhn and Tucker (eds) Contributions to the theory of games, vol II. Annals of Mathematical Studies 28. Princeton University Press, pp 193–216
Laraki R (2001a) Variational inequalities, system of functional equations and incomplete information repeated games. SIAM J Control Optim 40:516–524
Laraki R (2001b) The splitting game and applications. Int J Game Theory 30:359–376
Laraki R (2002) Repeated games with lack of information on one side: the dual differential approach. Math Oper Res 27:419–440
Lehrer E (2001) Any inspection is manipulable. Econometrica 69:1333–1347
Lehrer E (2003a) Approachability in infinite dimensional spaces. Int J Game Theory 31:253–268
Lehrer E (2003b) A wide range no-regret theorem. Games Econom Behav 42:101–115
Lehrer E, Solan E (2003) No regret with bounded computational capacity, *DP 1373*, Center for Mathematical Studies in Economics and Management Science, Northwestern University
Lehrer E, Solan E (2006) Excludability and bounded computational capacity. Math Oper Res 31:637–648
Lehrer E, Yariv L (1999) Repeated games with lack of information on one side: the case of different discount factors. Math Oper Res 24:204–218
Marino A (2005) The value of a particular Markov chain game. Chapters 5 and 6, PhD thesis, Universite Paris I, 2005. http://alexandre.marino.free.fr/theseMarino.pdf
Mayberry J-P (1967) Discounted repeated games with incomplete information, Report of the U.S. Arms control and disarmament agency, ST116, chapter V, Mathematica, Princeton, pp 435–461
Mertens J-F (1972) The value of two-person zero-sum repeated games: the extensive case. Int J Game Theory 1:217–227
Mertens J-F (1998) The speed of convergence in repeated games with incomplete information on one side. Int J Game Theory 27:343–357
Mertens J-F, Zamir S (1971) The value of two-person zero-sum repeated games with lack of information on both sides. Int J Game Theory 1:39–64
Mertens J-F, Zamir S (1976a) The normal distribution and repeated games. Int J Game Theory 5:187–197
Mertens J-F, Zamir S (1976b) On a repeated game without a recursive structure. Int J Game Theory 5:173–182
Mertens J-F, Zamir S (1977) A duality theorem on a pair of simultaneous functional equations. J Math Anal Appl 60:550–558
Mertens J-F, Zamir S (1985) Formulation of Bayesian analysis for games with incomplete information. Int J Game Theory 14:1–29
Neyman A (2008) Existence of optimal strategies in Markov games with incomplete information. Int J Game Theory 37:581–596
Neyman A (2012) The value of two-person zero-sum repeated games with incomplete information and uncertain duration. Int J Game Theory 41:95–207
Neyman A, Sorin S (1998) Equilibria in repeated games with incomplete information: the general symmetric case. Int J Game Theory 27:201–210
Neyman A, Sorin S (2010) Repeated games with public uncertain duration processes. Int J Game Theory 39:29–52
Ponssard JP, Sorin S (1980) The LP formulation of finite zero-sum games with incomplete information. Int J Game Theory 9:99–105
Renault J (2000) 2-player repeated games with lack of information on one side and state independent signalling. Math Oper Res 4:552–572
Renault J (2001a) 3-player repeated games with lack of information on one side. Int J Game Theory 30:221–246
Renault J (2001b) Learning sets in state dependent signalling game forms: a characterization. Math Oper Res 26:832–850
Renault J (2006) The value of Markov chain games with lack of information on one side. Math Oper Res 31:490–512
Renault J (2012) The value of repeated games with an informed controller. Math Oper Res 37:154–179
Renault J, Tomala T (2004) Learning the state of nature in repeated games with incomplete information and signals. Games Econom Behav 47:124–156

Renault J, Tomala T (2008) Probabilistic reliability and privacy of communication using multicast in general neighbor networks. J Cryptol 21(2):250–279

Renault J, Venel X (2017) A distance for probability spaces, and long-term values in Markov decision processes and repeated games. Math Oper Res 42(2):349–376

Renault J, Solan E, Vieille N (2013) Dynamic sender-receiver games. J Econ Theory 148:502–534

Renault J, Renou L, Tomala T (2014) Secure message transmission on directed networks. Games Econom Behav 85:1–18

Rosenberg D (1998) Duality and Markovian strategies. Int J Game Theory 27:577–597

Rosenberg D, Sorin S (2001) An operator approach to zero- sum repeated games. Israel J Math 121:221–246

Rosenberg D, Solan E, Vieille N (2004) Stochastic games with a single controller and incomplete information. SIAM J Control Optim 43:86–110

Rustichini A (1999) Minimizing regret: the general case. Games Econom Behav 29:224–243

Shalev J (1994) Nonzero-sum two-person repeated games with incomplete information and known-own payoffs. Games Econom Behav 7:246–259

Simon RS (2002) Separation of joint plan equilibrium payoffs from the min-max functions. Games Econom Behav 1:79–102

Simon RS, Spież S, Toruńczyk H (1995) The existence of equilibria in certain games, separation for families of convex functions and a theorem of Borsuk- Ulam type. Israel J Math 92:1–21

Simon RS, Spież S, Toruńczyk H (2002) Equilibrium existence and topology in some repeated games with incomplete information. Trans AMS 354:5005–5026

Simon RS, Spież S, Toruńczyk H (2008) Equilibria in a class of games and topological results implying their existence. Rev R Acad Cien Serie A Mat 102:161–179

Sion M (1958) On general minimax theorems. Pac J Math 8:171–176

Sorin S (1983) Some results on the existence of Nash equilibria for non- zero sum games with incomplete information. Int J Game Theory 12:193–205

Sorin S (1984a) Big match with lack of information on one side (Part I). Int J Game Theory 13:201–255

Sorin S (1984b) On a pair of simultaneous functional equations. J Math Anal Appl 98:296–303

Sorin S (1989) On recursive games without a recursive structure: existence of limv_n. Int J Game Theory 18:45–55

Sorin S (1997) Merging, reputation, and repeated games with incomplete information. Games Econom Behav 29:274–308

Sorin S, Zamir S (1985) A 2-person game with lack of information on 1 and 1/2 sides. Math Oper Res 10:17–23

Spinat X (2002) A necessary and sufficient condition for approachability. Math Oper Res 27:31–44

Vieille N (1992) Weak approachability. Math Oper Res 17:781–791

Waternaux C (1983) Solution for a class of repeated games without recursive structure. Int J Game Theory 12:129–160

Zamir S (1971) On the relation between finitely and infinitely repeated games with incomplete information. Int J Game Theory 1:179–198

Zamir S (1973) On repeated games with general information function. Int J Game Theory 21:215–229

Ziliotto B (2016) Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture maxmin=limvn. Ann Probab 44:1107–1133

Books and Reviews

Aumann RJ, Maschler M (1995) 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)

Forges F (1992) Repeated games of incomplete information: non-zero sum. In: Aumann RJ, Hart S (eds) Handbook of game theory, vol I. Elsevier, North-Holland, pp 155–177

Laraki R, Sorin S (2014) Chapter 2: Advances in zero-sum dynamic games. In: Zamir S, Young P (eds) Handbook of game theory, vol IV. Elsevier, North-Holland, pp 27–93

Laraki R, Renault J, Tomala T (2006) 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)

Mertens J-F (1987) Repeated games. In: Proceedings of the international congress of mathematicians, Berkeley 1986. American Mathematical Society, Dordrecht, pp 1528–1577

Mertens J-F, Sorin S, Zamir S (1994) Repeated games. CORE discussion paper 9420-9422

Sorin S (2002) A first course on zero-sum repeated games. Mathématiques et Applications. Springer-Verlag Berlin Heidelberg

Zamir S (1992) Repeated games of incomplete information: zero-sum. In: Aumann RJ, Hart S (eds) Handbook of game theory, vol I. Elsevier, North-Holland, pp 109–154