International Journal of Game Theory, Vol. 14 Issue 4, page 249–272

On a Repeated Game with State Dependent Signalling Matrices.

By *S. Sorin*, Strasbourg[1]

Abstract: We prove the existence of the minmax and the maxmin for a repeated game with lack of information on both sides and signalling matrices which depend on the state.

1 Introduction

1. A. When considering two person zero-sum infinitely repeated games with incomplete information, a major distinction has to be made between games with and without a recursive structure.

In the first case the analysis can be pursued by conditionning with respect to some σ-algebra generated by a common knowledge information gathered along the play, and then by using some "state variables".

For the second case where such tools do not work (the state space should increase strictly at each stage of the play) new methods relying on the construction of a one shot auxiliary game have been introduced [see *Mertens/Zamir*, 1976; *Waternaux*].

Coming back to games with recursive structure, three main classes are solved up to now:

a) games with lack of information on one side. [*Aumann/Maschler*].

b) games with lack of information on both sides in which the signalling matrices may depend on the player but not on the state of nature [*Mertens/Zamir*, 1971/72; *Mertens; Mertens/Zamir*, 1980, see condition (ii)* p. 203].

c) games where the players have the same initial information and the same signalling matrices which moreover completely reveal the moves (see (d) below), [*Kohlberg/Zamir*, 1974; *Forges*].

1. B Here an example is treated of a game not fulfilling these hypotheses. It can be described as case (b) without condition (ii)* and keeping the recursive structure: there are four possible states of nature, corresponding to the independent case [see *Mertens/*

[1] *Sylvain Sorin*, Université Louis Pasteur, Départment de Mathématique, 7 Rue René Descartes, 67084 Strasbourg, France.

Zamir, 1980, p. 202] with two types for each player: given each state (μ, ν) with μ and ν in $\{0, 1\}$, Player I knows μ, Player II knows ν and the distribution of (μ, ν) is the product of the marginals with $p = \text{Prob}(\mu = 1)$, $q = \text{Prob}(\nu = 1)$. Accordingly there are four 2×2 pay off matrices $A(\mu, \nu)$ and in addition four signalling matrices (common to the players):

$$H(1,1) = \begin{pmatrix} T & L \\ C & D \end{pmatrix} \qquad H(1,0) = \begin{pmatrix} T & R \\ C & D \end{pmatrix}$$

$$H(0,1) = \begin{pmatrix} B & L \\ C & D \end{pmatrix} \qquad H(0,0) = \begin{pmatrix} B & R \\ C & D \end{pmatrix}$$

After each stage, if (μ, ν) is the state and players I and II play their moves i and j, the letter $h_{ij}(\mu, \nu)$ is told to both players. This ends the description of the game $G(p, q)$. Note that no pay off is announced but that:

(d) any letter reveals the pair of moves used, namely $i \neq i'$ or $j \neq j'$ imply $h_{ij}(\mu, \nu) \neq h_{i'j'}(\mu', \nu')$ for all μ, μ', ν, ν', as in the "symmetric case" (c).

Hence if none of the player has initial information the game belongs to the class (c).

Remark now that the extreme games (*ie* where $p(1-p)q(1-q) = 0$) fall in (a) since in this case only one player is informed. It follows then from *Aumann/Maschler* that the game $G(p, q)$ has a value at these points, denoted by $\mathrm{v}(p, q)$.

Notice finally that as soon as Player I plays top, some type is revealed: if Player II plays left, $\mu = 1$ if T is announced, $\mu = 0$ if B, similary if Player II plays right, $\nu = 1$ if L and $\nu = 0$ if R.

On the contrary if Player I plays bottom the letter announced is independent of the state.

We shall thus call the previous letters, exceptional letters and $\{C, D\}$ regular letters.

1. C. In order to state the results let us introduce some notations. $\sigma = (\sigma^1, \sigma^0)$ (resp. τ^1, τ^0)) is the strategy of Player I (type 1, type 0), (resp. of Player II). (i_n, j_n) and h_n denote the moves and the signal at stage n.

Given the state (μ, ν) the payoff at stage n is thus $g_n = a_{i_n j_n}(\mu, \nu)$. We shall use the expectation (induced by p, q, σ, τ) of the Cesaro mean of the payoffs namely:

$$\bar{\gamma}_n^{pq}(\sigma, \tau) = E_{p,q,\sigma,\tau}\left(\frac{1}{n}\sum_{m=1}^{n} g_m\right)$$

to define a maxmin and a minmax in $G(p, q)$ [see *Mertens/Zamir*, 1980]:

$\underline{v}$ is the max-min of G if, for every (p, q) in $[0, 1]^2$ we have:

i) $\forall\, \epsilon > 0$, $\exists\, \sigma$ and $\exists\, N$ such that $\forall\, \tau$, $\forall\, n \geqslant N$:

$$\bar{\gamma}_n^{p,q}\,(\sigma, \tau) \geqslant \underline{v}\,(p, q) - \epsilon$$

ii) $\forall\, \sigma$, $\forall\, \epsilon > 0$, $\exists\, \tau$ and $\exists\, N$ such that for $n \geqslant N$:

$$\bar{\gamma}_n^{p,q}\,(\sigma, \tau) \leqslant \underline{v}\,(p, q) + \epsilon$$

The first condition says that Player I can obtain $\underline{v}$ (up to some ϵ) in any sufficiently long game, uniformly upon the strategies of his opponent.

We shall refer to it by saying that Player I can *guarantee* $\underline{v}$.

Condition ii) corresponds to the existence of a best reply of Player II to each σ, with a payoff (thought as $\underline{lim}$) less than $\underline{v}$. We shall write that Player II can *defened* $\underline{v}$.

The minmax $\bar{v}$ is obviously defined in a dual way.

The main result of this paper is now: $\underline{v}$ *and* $\bar{v}$ *exist.*

1. D. Since the tools and results of this paper are quite different from the previous ones in this field, let us recall briefly the main idea of the proof in case b).

Let us consider the maxmin (the minmax is similar since the two players are symmetric). Let NR be the set of non-revealing strategies i.e. strategies that induce on the signals a distribution independent of the type.

By playing NR, Player I reduces the situation to a game with lack of information on one side (where Player II is informed) belonging to class a). On the other hand, knowing the strategy of Player I, Player II can first exhaust a maximal amount of information, without revealing anything and then play optimal as if he was the only informed player.

Here both players may have to use revealing strategies and we shall need another subset of strategies. We define NS to be the set of non-separating strategies, i.e. such that $\sigma^1 = \sigma^0$ (resp. $\tau^1 = \tau^0$): the strategy is independent of the type. Note that for the game under consideration in this paper NS is not included in NR (this is the case when (ii)* holds).

Typically the strategies used will be defined in two parts:

1) up to the stage when Player I plays Top for the first time

2) and after conditionally to the letter announced.

In order to prove that one player can garantee a payoff we shall proceed as follows: for step 2) we can introduce conditional absorbing payoffs that this player can guarantee after this stage (since the revealed game belong to class a)). By letting then this player use NS strategies in part 1), the game reduces to a stochastic game with lack of information on one side of a special type studied in previous papers [*Sorin*, 1984, 1985].

To show that one player can defend a payoff we let him first play NS and exhaust the maximal amount of information from his opponent, and then in part 2) let him use an optimal strategy in the reduced game at the given posterior. Note that the previous behaviour may be revealing and thus may induce new aspects in the opponent's strategy. In order to control this phenomena a fixed point argument is used.

1. E. Let us finally introduce some notations which will be used along the proofs.

Recall that given f real function on [0, 1], the concavification of f (Cav f) is the smallest concave function greater than f on [0, 1] and Vex f is defined in a dual way. If f is a function of several variables $\underset{x}{\mathrm{Cav}}\, f$ denotes the concavification of the function restricted to the variable x.

As usual $\tilde{p}_n$ (and similarly $\tilde{q}_n$) will denote the posterior at stage n *ie* the conditional expectation of p (given σ and τ), with respect to the algebra H_n generated by the histories $(h_1, \ldots, h_n)$ up to stage n.

We shall also use:

$$u' = 1 - u, a_{ij}(p, q) = E_{p,q}(a_{ij}(\mu, \nu)), \| A \| = \max_{\substack{i,j \\ \mu,\nu}} | a_{ij}(\mu, \nu) |$$

$$m \wedge n = \min(m, n), m \vee n = \max(m, n).$$

2 Minmax

We prove in this section the existence of minmax and we shall give an explicit formula for it.

We first define an auxiliary game as follows:

Given $\alpha = (\alpha_1, \alpha_0)$ and $\beta = (\beta_1, \beta_0)$ in $\mathbf{R}^2$, $\Gamma(p, q, \alpha, \beta)$ is the infinitely repeated stochastic game with lack of information on one side described by:

$$\begin{array}{ll} p & \begin{pmatrix} \mathrm{v}(1,q)^* & (q\,\alpha_1 + q'\,\beta_1)^* \\ a_{21}(1,q) & a_{22}(1,q) \end{pmatrix} \\ p' & \begin{pmatrix} \mathrm{v}(0,q)^* & (q\,\alpha_0 + q'\,\beta_0)^* \\ a_{21}(0,q) & a_{22}(0,q) \end{pmatrix} \end{array} \qquad \text{standard signalling } (ie\ h_{ij} = (i, j)).$$

More precisely Player I (denoted later by P I) is informed upon the chance's move, which chooses the top state with probability p, and P II knows only p. Now as soon as a star is reached, the corresponding entry gives the payoff for the remaining stages (absorbing payoff).

This class of games was studied in a previous paper [*Sorin*, 1984] where it is proved that the minmax exists and equals the value of the one-shot game $\Gamma_1(p, q, \alpha, \beta)$, that we denote by $W_1(p, q, \alpha, \beta)$.

Let us now define two closed convex sets of vector payoffs:

$$H_1 = \{\alpha = (\alpha_1, \alpha_0) \text{ in } \mathbf{R}^2; \alpha_1 \lambda + \alpha_0 \lambda' \geqslant \mathrm{v}(\lambda, 1) \text{ for all } \lambda \in [0, 1]\}$$

$$H_0 = \{\beta = (\beta_1, \beta_0) \text{ in } \mathbf{R}^2; \beta_1 \lambda + \beta_0 \lambda' \geqslant \mathrm{v}(\lambda, 0) \text{ for all } \lambda \in [0, 1]\}$$

and note that H_1 corresponds to the affine majorants of the concave function v (.,1).

We can now state our first result:

Theorem 1

$\bar{\mathrm{v}}(p, q)$ exists on $[0, 1]^2$ and is given by:

$$\bar{\mathrm{v}}(p, q) = \underset{q}{\mathrm{Vex}} \min_{\substack{\alpha \in H_1 \\ \beta \in H_0}} \{\mathrm{maxmin}\, \Gamma(p, q, \alpha, \beta)\}$$

$$= \underset{q}{\mathrm{Vex}} \min_{\substack{\alpha \in H_1 \\ \beta \in H_0}} W_1(p, q, \alpha, \beta).$$

Proof

The proof of the theorem will be divided in two parts, corresponding to conditions i) and ii) of definition 1.C.

2.A. We first show that P II can guarantee this payoff. Recall that if P II can guarantee some function $f(q)$ he can also guarantee Vex $f(q)$. (This basic property for games with incomplete information was proved by *Aumann/Maschler* [see e.g. *Sorin,* 1979, 2.17]).

Thus it is enough to prove that given any (α, β) in $H_1 \times H_0$, P II can guarantee maxmin $\Gamma(p, q, \alpha, \beta)$.

Consider then the following class C of strategies of P II:

(2.1) i) play NS as long as h_n is a regular letter.

ii) if T (resp. B) is announced, play from this stage on optimally in $G(1, q)$ (resp. $G(0, q)$).

iii) if L (resp. R) is announced, approach from this stage on the vector payoff α (resp. β).

Before explaining the meaning of iii) note that ii) is consistent. In fact if $\tilde{m}$ is the stopping time of first appearance of an exceptional letter, then up to stage $\tilde{m}$, P II was playing NS; hence given the regular letters the posterior on his type at stage $\tilde{m} + 1$ is still q.

(2.2) By *Aumann/Maschler*, P II has then a strategy which gives at each stage an expected payoff not more than the value of the revealed game. Now iii) refers to the existence of an "approachability strategy" (due to Blackwell) as soon as $\alpha \in H_1$ (resp. $\beta \in H_0$) [see e.g. *Sorin,* 1979, 2.18]. This property, first proved by *Aumann/Maschler* then extended to the general case by *Kohlberg* can be written as follows:

(2.3) For every α in H_1, and $\epsilon > 0$, there exists τ strategy of P II and N_1 such that:

$$\bar{\gamma}_n^{1,1}(\sigma, \tau) \leqslant \alpha_1 + \epsilon$$

$$\bar{\gamma}_n^{0,1}(\sigma, \tau) \leqslant \alpha_0 + \epsilon$$

for all n greater than N_1 and all strategy σ of P I. (Obviously a similar result holds for $G(p, 0)$ and β in H_0).

It is now easy to see that if P II plays in C, the original game $G(p, q)$ is equivalent to $\Gamma(p, q)$; so that by playing in G an optimal strategy in Γ, P II can get as a payoff in G minmax Γ.

The precise computations are as follows:

Let t in $[0, 1]$ denote an optimal strategy of P II in $\Gamma_1(p, q, \alpha, \beta)$, where t = Prob (play left).

Given $\epsilon > 0$ we define a strategy τ in C where (2.1) i) is now specified as play t i.i.d. up to stage $\underset{\sim}{m}$, where the stopping time $\underset{\sim}{m}$ is defined by:

$$\underset{\sim}{m} = \min\{n \geqslant 1; i_n = \text{Top}\}$$

Given a strategy σ of P I the average expected payoff is $\bar{\gamma}_n^{p,q}(\sigma, \tau)$ but we have:

(2.4) $\bar{\gamma}_n^{p,q}(\sigma, \tau) = p\,\bar{\gamma}_n^{1,q}(\sigma^1, \tau) + p'\,\bar{\gamma}_n^{0,q}(\sigma^0, \tau)$

so that it is enough to majorize $\bar{\gamma}_n^{(1,q)}(\sigma^1, \tau)$ (for example).

Note now that:

$$\text{Prob}_{\sigma^1, \tau}[j_n = \text{left} \mid \underset{\sim}{m} \geqslant n] = t \text{ by (2.1) hence}$$

(2.5) $E_{q, \sigma^1, \tau}[g_n \mid \underset{\sim}{m} \geqslant n] = t\, a_{21}(1, q) + t'\, a_{22}(1, q)$.

Similary (2.1) i) implies

(2.6) $\text{Prob}_{\sigma^1, \tau}[h_n = T \mid \underset{\sim}{m} = n] = t$

$$\text{Prob}_{\sigma^1, \tau}[h_n = L \mid \underset{\sim}{m} = n] = t'\, q$$

$$\text{Prob}_{\sigma^1, \tau}[h_n = R \mid \underset{\sim}{m} = n] = t'\, q'$$

and from (2.1) ii) and (2.2) we have:

$$(2.7)\ E_{q,\sigma^1,\tau}\,[g_n \mid \underset{\sim}{m} < n,\ h_{\underset{\sim}{m}} = T] \leqslant \mathrm{v}\,(1,q)$$

By using (2.1) iii) and (2.3) we now obtain on $\{\underset{\sim}{m} \leqslant n - \max\,\{N_0, N_1\}\}$

$$(2.8)\ E_{q,\sigma^1,\tau}\left[\frac{1}{(n-\underset{\sim}{m})}\ \sum_{\underset{\sim}{m}}^{n} g_k \mid h_{\underset{\sim}{m}} = L\right] \leqslant \alpha_1 + \epsilon$$

From (2.4) – (2.8) we deduce:

$$(2.9)\ n\ \bar{\gamma}^{1,q}\,(\sigma^1,\tau) \leqslant E_{\sigma^1,\tau}\,[\underset{\sim}{m} \wedge n\,(t\,a_{21}\,(1,q) + t'\,a_{22}\,(1,q)) + (n - \underset{\sim}{m} \wedge n)$$
$$(t\,\mathrm{v}\,(1,q) + t'\,(q\,\alpha_1 + q'\,\beta_1))] + n\,\epsilon$$
$$+\, 2\,\|A\|$$
$$\max\,(N_0, N_1)$$

Let $F\,(s^1, s^0, t)$ denote the payoff in $\Gamma_1\ (p, q, \alpha, \beta)$ when P I is using s^1 = Prob (play Top | $\mu = 1$) (resp. s^0 if $\mu = 0$) and P II plays t.

Dual considerations now give from (2.9):

$$(2.10)\ \bar{\gamma}_n^{p,q}\,(\sigma,\tau) \leqslant F\,(s_n^1, s_n^0, t) + \epsilon + 2\,\frac{\|A\|}{n}\,\max\,(N_1, N_0)$$

where $(s_n^1)' = \dfrac{1}{n}\,E_{\sigma^1,\tau}\,(\underset{\sim}{m} \wedge n)$

and similary $(s_n^0)'$ are the normalized mean of the stopping $\underset{\sim}{m}$, up to stage n, under σ^1 (resp. σ^0).

Since t is optimal in $\Gamma_1\ (p, q, \alpha, \beta)$ we obtain finally for $n \geqslant N = \dfrac{2}{\epsilon}\,\|A\|\,.\max\,(N_0, N_1)$: $\bar{\gamma}_n^{p,q}\,(\sigma,\tau) \leqslant W_1\,(p, q, \alpha, \beta) + 2\,\epsilon$.

2\. B. Let us now prove that P I can defend the same payoff.

In order to find a "best reply" of P I (*ie* mainly a distribution of the stopping time $\underset{\sim}{m}$), good for all the possible *posterior* choices for (α, β) we need first the following construction.

Assume (p, q) fixed. Let $\phi\,(\alpha, \beta)$ the set of optimal strategies for P I in $\Gamma_1\ (p, q, \alpha, \beta)$, defined by (s^1, s^0) in $[0, 1]^2$.

We denote by $\psi(s^1, s^0)$ the set of (α, β) in $H_1 \times H_0$ minimizing the absorbing payoff in Γ_1 induced by (s^1, s^0) namely:

$$ps^1 (q\alpha_1 + q'\beta_1) + p's^0 (q\alpha_0 + q'\beta_0).$$

(2.11) Remark that (α, β) belongs to $\psi(s^1, s^0)$ iff α minimizes:

$ps^1 \alpha_1 + p's^0 \alpha_0$ on H_1, and similary for β on H_0.

Hence such an α is a supporting hyperplane to v (., 1) at the posterior $\rho(s^1, s^0) = = \dfrac{ps^1}{ps^1 + p's^0}$.

It is straightforwards to check that the correspondances ϕ and ψ are s.c.s and compact convex valued. It follows that the correspondance $\phi \circ \psi$ has a fixed point that we shall denote by $(\bar{s}^1, \bar{s}^0)$.

The idea of the proof can now be explained as follows: Given τ, strategy of P II, P I plays Bottom untill some stage N after which the posteriors $\tilde{q}_n$, $n \geqslant N$ are essentially constant. For each of these stages n, P I computes z_n which is the non absorbing payoff induced by τ and $(\bar{s}^1, \bar{s}^0)(p, \tilde{q}_N)$ in $\Gamma(p, \tilde{q}_N)$ (note that this quantity is independent of (α, β)). P I plays now Bottom until some stage N_1 where z_n is minimum and plays then once $(\bar{s}_1, \bar{s}_0)(p, \tilde{q}_N)$ at that stage: namely if i_{N_1} = Bottom, P I keeps playing Bottom, otherwise he plays optimally from this stage on in the revealed game. Assuming this strategy for P I a best reply for P II after $\underset{\sim}{m}$ would be to choose $(\alpha, \beta) \in \psi(\bar{s}_1, \bar{s}_0)$. It follows then that P I obtains $W_1(p, \tilde{q}_N, \alpha, \beta)$ for some α, β hence the result.

Let us do now the formal construction.

We are given a strategy τ of P II. Assuming that P I uses the NS strategy $\hat{b}$: play always Bottom, the posteriors $\tilde{q}_n$ are a well defined bounded martingale hence converges in expectation. So let us define N such that

$$(2.12)\; E_{q,\hat{b},\tau}\,[\,|\tilde{q}_n - \tilde{q}_N|\,] \leqslant \epsilon \quad \forall n \geqslant N.$$

Note that this inequality (for a given n) holds true for any σ that coincides with $\hat{b}$ up to stage $n-1$. We shall write $\tilde{S}^1_N$ and $\tilde{S}^0_N$ for the random variables $(\bar{s}^1, \bar{s}^0)(p, \tilde{q}_N)$ (fixed point of $\phi \circ \psi$ at $(p, \tilde{q}_N)$), and $\tilde{\rho}_N$, $\tilde{\lambda}_N$ for the posteriors, given i_N = Top or Bottom, namely:

$$\tilde{\rho}_N = \frac{p\tilde{S}^1_N}{p\tilde{S}^1_N + p\tilde{S}^0_N}, \quad \tilde{\lambda}_N = \frac{P(\tilde{S}^1_N)'}{p(\tilde{S}^1_N)' + p'(\tilde{S}^0_N)'}$$

We now introduce two functions on $[0, 1]^4$ by:

$$f(\tilde{q}, s^1, s^0, t) = t\,[ps^1\, \mathrm{v}\,(1, \tilde{q}) + p' s^0\, \mathrm{v}\,(0, \tilde{q})] + t'\,[ps^1 + p' s^0]$$

$$[q\, \mathrm{v}\,(\rho\,(s^1, s^0), 1) + q'\, \mathrm{v}\,(\rho\,(s^1, s^0), 0)]$$

and

$$\begin{aligned} g\,(\tilde{q}, s^1, s^0, t) = {} & t\,[p\,(s^1)'\, a_{21}\,(1, \tilde{q}) + p'\,(s^0)'\, a_{21}\,(0, \tilde{q})] \\ & + t'\,[p\,(s^1)' \\ & a_{22}\,(1, \tilde{q}) + p'\,(s^0)'\, a_{22}\,(1, \tilde{q})] \end{aligned}$$

(2.13) Note that f and g correspond respectively to the absorbing and non absorbing part of the payoff in $\Gamma_1\,(\tilde{p}, q, \alpha)$ given (s^1, s^0) and t, with moreover (α, β) in $\psi\,(s^1, s^0)$ (see 2.11)

We shall also need the following parameters induced by τ:

$$t_n^i = \mathrm{Prob}_{\hat{b},\tau^i}\,(\{j_n = \mathrm{Left}\} \mid H_N) \quad \text{for } n \geqslant N \text{ and } i = 1, 0$$

$$t_n = \mathrm{Prob}_{q,\hat{b},\tau}\,(\{j_n = \mathrm{Left}\} \mid H_N) \quad \text{hence } t_n = \tilde{q}_N\, t_n^1 + \tilde{q}_N'\, t_n^0.$$

(As for $\tilde{q}_n$ above, these random variables are the same for all strategies σ that coincide with $\hat{b}$ up to stage $n - 1$).

Given $\epsilon > 0$ we define $\tilde{N}_1 \geqslant N$ such that:

$$(2.14)\; g\,(\tilde{q}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_{\tilde{N}_1}) \leqslant g\,(\tilde{q}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_n) + \epsilon,\ \forall\, n \geqslant N.$$

Remark that $\tilde{N}_1$ is H_N mesurable and bounded by some N_1.

The strategy σ of P I is now defined as follows:

- play Bottom up to stage $\tilde{N}_1 - 1$.
- play $(\tilde{S}_N^1, \tilde{S}_N^0)$ at stage $\tilde{N}_1$.

Then: I) if $\{i_{\tilde{N}_1} = \mathrm{Bottom}\}$ play Bottom from stage $\tilde{N}_1 + 1$ on (note that on this event σ coincides with $\hat{b}$).

II) if $\{i_{\tilde{N}_1} = \mathrm{Top}\}$ use an (ϵ) optimal strategy in the revealed game starting at stage $\tilde{N}_1 + 1$ namely:

II a) if $h_{\tilde{N}_1} = T$, play σ_T such that for n larger than some K:

$$(2.15)\, E_{p,q,\sigma_T,\tau}\left(\frac{1}{n}\sum_{\tilde{N}_1+1}^{\tilde{N}_1+1+n} g_m \mid H_{\tilde{N}_1}\right) \geqslant \mathrm{v}\,(1, q_{\tilde{N}_1}) - \epsilon$$

(and similary for B)

II b) if $h_{\tilde{N}_1} = L$, use σ_L such that for $m \geqslant \tilde{N}_1 + 1$:

$$(2.16)\, E_{p,q,\sigma_L,\tau}\,(g_m \mid H_{\tilde{N}_1}) \geqslant \mathrm{v}\,(\tilde{p}_N, 1).$$

(and similary for R).

To see why such strategies exist notice that:

– for II a) v $(p, .)$ is Lipschitz, hence K can be chosen independent of $q_{\tilde{N}_1}$. Now (2.15) follows from (2.3) (or rather its dual) and the fact that after stage $\tilde{N}_1$, the state variables are 1 (for p) and $q_{\tilde{N}_1}$ (for q).

– as for II b) the inequality (2.16) follows from (2.2) and the fact that since P I was playing always bottom up to stage $\tilde{N}_1 - 1$, $P_{\tilde{N}_1 - 1} = p$ hence $P_{\tilde{N}_1} = \tilde{p}_N$.

In order to compute the expected payoff at somme stage $N_1 + m$, conditionnally on H_N, we first study the events generated by $h_{\tilde{N}_1}$.

The conditionnal probabilities on H_N induced by p, q, σ, τ are given by:

$$(2.17)\ \mathrm{Prob}\,(i_{\tilde{N}_1} = \mathrm{Bottom} \mid H_N) = p\,(\tilde{S}_N^1)' + p'\,(\tilde{S}_N^0)'$$

$$(2.18)\ \mathrm{Prob}\,(h_{\tilde{N}_1} = T \mid H_N) = p\,\tilde{S}_N^1 \cdot t_{\tilde{N}_1}$$

$$(2.19)\ \mathrm{Prob}\,(h_{\tilde{N}_1} = L \mid H_N) = (p\,\tilde{S}_N^1 + p'\,\tilde{S}_N^0) \cdot \tilde{q}_N \cdot (t_{\tilde{N}_1}^1)$$

and analogous formulas for B and R.

The last term in (2.19) involves $t_{\tilde{N}_1}^1$ but since the posteriors $\tilde{q}_N$ is essentially constant after stage N, P II is playing almost NS. In fact we have:

$$(2.20)\ E_{q,\hat{b},\tau}(|t_n^i - t_n \| H_N) = \tilde{q}_N | t_n^1 - t_n | + \tilde{q}'_N | t_n^0 - t_n | \leqslant \delta(n+1, N) \text{ with}$$

$$\delta(n+1, N) = E(|\tilde{q}_{n+1} - \tilde{q}_N \| H_N).$$

Now we can use (2.15) and (2.16) to minorize the payoffs after some exceptional letter.

It remains thus to see that after $\{i_{\tilde{N}_i} = \text{Bottom}\}$ the payoff at stage n is given by:

$$(2.21)\ A_n = q_{\tilde{N}}(t_n^1 a_{21}(\tilde{\lambda}_N, 1) + (t_n^1)' a_{22}(\tilde{\lambda}_N, 1))$$

$$+ q'_{\tilde{N}}(t_n^0 a_{21}(\tilde{\lambda}_N, 0) + (t_n^0)' a_{22}(\tilde{\lambda}_N, 0))$$

where we recall that $\tilde{\lambda}$ is the posterior induced by $\tilde{S}_N^1, \tilde{S}_N^0$ on $\{i_{\tilde{N}_1} = \text{Bottom}\}$.

Taking the expectation with respect to H_N in (2.15), (2.16) we obtain, for all $n \geqslant K$, using (2.20).

$$(2.21)\ E_{p,q,\sigma,\tau}\Big(\sum_{N_1+1}^{N_1+1+n} g_m \mid H_N\Big) \geqslant \text{Prob}\,(i_{\tilde{N}_1} = \text{Bottom} \mid H_N) \sum_{N_1+1}^{N_1+1+n} A_m$$

$$+ n\,[\text{Prob}\,(h_{\tilde{N}_1} = T \mid H_N)(\text{v}(1, \tilde{q}_N) - \epsilon)$$

$$+ \text{Prob}\,(h_{\tilde{N}_1} = B \mid H_N)(\text{v}(0, \tilde{q}_N) - \epsilon)$$

$$- \| A \|\, \delta(N_1, N)]$$

$$+ n\,[\text{Prob}\,(h_{\tilde{N}_1} = L \mid H_N)\, \text{v}(\tilde{p}_N, 1)$$

$$+ \text{Prob}\,(h_{\tilde{N}_1} = R \mid H_N)\,.\,\text{v}(\tilde{p}_N, 0)]$$

Using (2.17) (2.20) and (2.21) it follows that:

$$(2.22)\ \text{Prob}\,(i_{\tilde{N}_1} = \text{Bottom} \mid H_N)\,.\,A_m \geqslant g(\tilde{q}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_m) - \| A \|\, \delta(m, N)$$

Denoting by L the left member of (2.21) we get now from (2.22), (2.18) and (2.19) that:

$$(2.23)\ L \geqslant \sum_{N_1+1}^{N_1+1+n} [(g(\tilde{g}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_m) - \|A\|\,\delta(m,N)] + n\,[f(\tilde{q}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_{N_1}) -$$

$$- \epsilon - \|A\|\,\delta(N_1, N)]$$

We use now the definition of $\tilde{N}_1$ (see (2.14)) to minorize the right part by introducing $g(\tilde{q}_N, \tilde{S}_N^1, \tilde{S}_N^0, t_{\tilde{N}_1})$.

Recall also by (2.13) that $f+g$ is the (one shot) payoff in some $\Gamma_1\ (p, \tilde{q}_N, \alpha, \beta)$ where $\tilde{S}_N^1, \tilde{S}_N^0$ is an optimal strategy for P I. (2.23) leads now to:

$$(2.24)\ L \geqslant n \min_{\alpha,\beta} W_1\ (p, \tilde{q}_N, \alpha, \beta) - 2\,n\,\epsilon - \|A\|\,(\sum_{N_1+1}^{N_1+1+n} \delta(m,N) + n\,\delta(N_1, N)).$$

We minorize min W_1 by Vex min W_1 and then take the expectation on both sides to obtain, using (2.12) and Jensen's inequality:

$$E_{p,q,\sigma,\tau}(\sum_{N_1+1}^{N_1+1+n} g_m) \geqslant n \operatorname{Vex} \min_{\alpha,\beta} W_1\ (p, q, \alpha, \beta) - 2\,n\,\epsilon - \|A\|\,(n\,\epsilon + n\,\epsilon)$$

Thus for $n \geqslant (K+1) \vee (N_1+1)$

$$\bar{\gamma}_{n+1+N_1}(\alpha, \tau) \geqslant \operatorname{Vex} \min_{\alpha,\beta}\ (p, q, \alpha, \beta) - 5\,\epsilon\,(\|A\| \vee 1)$$

hence the result.

3 Maxmin

Before stating the proposition we have to introduce some notations and to recall some previous results.

As in Part 2 we can determine a first amount that P I can guarantee by using some *NS* strategy until an exceptional letter is reached and then by playing optimally in the revealed game (if L or R) or by approaching some vector payoff if T or B.

More precisely let:

$$L_1 = \{\gamma = (\gamma_1, \gamma_0) \mid \gamma_1 \lambda + \gamma_0 \lambda' \leqslant \mathrm{v}(1, \lambda) \text{ for all } \lambda \text{ in } [0, 1]\}$$

$$L_0 = \{\delta = (\delta_1, \delta_0) \mid \delta_1 \lambda + \delta_0 \lambda' \leqslant \mathrm{v}(0, \lambda) \text{ for all } \lambda \text{ in } [0, 1]\}$$

Then if P I plays as above, $G(p, q)$ is similar to the infinitely repeated game $\Lambda(p, q, \alpha, \rho)$ which is described by:

$$\overset{q}{\begin{pmatrix} (p\gamma_1 + p'\delta_1)^* & \mathrm{v}(p, 1)^* \\ a_{21}(p, 1) & a_{22}(p, 1) \end{pmatrix}} \quad \overset{q'}{\begin{pmatrix} (p\gamma_0 + p'\delta_0)^* & \mathrm{v}(p, 0)^* \\ a_{21}(p, 0) & a_{22}(p, 0) \end{pmatrix}}.$$

For this class of games, studied in *Sorin* [1985] the maxmin exists. Moreover we obviously have since $\mathrm{v}(\,.\,, \lambda)$ is concave:

(3.1) $p\gamma_1 + p'\delta_1 \leqslant \mathrm{v}(p, 1)$

$p\gamma_0 + p'\delta_0 \leqslant \mathrm{v}(p, 0)$ for all (γ, δ) in $L_1 \times L_0$.

It follows then from *Sorin* [1985] (Part I V, first case) that, if we define: $\Theta = \{\theta;$ Borel positive measure on $[0, 1]$ with total mass less than $1\}$ the maxmin of Λ is given by u with:

$$u(p, q, \alpha, \beta) = \sup_{\theta \in \Theta} \inf_{x, y \in [0,1]^2} q\varphi^1(x, \theta; p, \alpha, \beta) + q'\varphi^0(y, \theta; p, \alpha, \beta)$$

where $\theta(x)$ stands for $\theta([0, x])$ and φ^i is given by:

$$\varphi^i(x, \theta; p, \alpha, \beta) = \int_0^x [t\,\mathrm{v}(p, i) + t'(p\gamma^i + p'\delta^i)]\, d\theta(t)$$

$$+ (1 - \theta(x)) \min\{(xa_{22}(p, i) + x' a_{21}(p, i)), u(p, i, \alpha, \beta)\},$$

$$i = 1, 0.$$

As in Part 2 we can thus state that P I can guarantee:

$$\operatorname*{Cav}_p \max_{\substack{\gamma \in L_1 \\ \delta \in L_0}} u(p, q, \alpha, \beta)$$

Note nevertheless that P I, even without knowing the strategy of P II, obtains some knowledge through the sequence of regular letters: C, D (as long as he plays bottom, of course). This fact (already used in the so called Blackwell strategy) will allow him to get more by changing the approachable vector along the play.

The strategy of P I can then be roughly described as follows: use θ_1 (resp θ_0) if $\mu = 1$ (resp 0) to choose a point x in $[0, 1]$ and play the "Big Match" strategy corresponding to it (*ie* play top if the frequency of right exceeds x [see *Blackwell/Ferguson; Sorin*, 1984, 1985; more generally *Mertens/Neyman*]. As soon as L or R appears, play optimally in the revealed game at p_x (posteriors induced by θ_1 and θ_0). If T or B is announced, approach some γ_x (resp δ_x) in L_1 (resp L_0).

Let us first define:

$\Theta f = \{\theta$ in Θ with finite support$\}$ and given (θ_1, θ_0) in Θf (that we can assume with support included in some set $t_0 = 0 < t_1 < \ldots < t_m = 1$) we introduce on $[0, 1]$: $\theta(t)$, $\rho(t)$ and $p(t)$ by:

(3.2) $\theta(0) = p\,\theta_1(0) + p'\,\theta_0(0)$

$$p(0) = p\,\frac{\theta_1(0)}{\theta(0)} \text{ and } p(t) = p(0) \text{ on } [0, t_1).$$

$$\rho(0) = p\,\frac{\theta_1'(0)}{\theta'(0)} \text{ and } \rho(t) = \rho(0) \text{ on } [0, t_1).$$

and inductively

$$\theta(t) = p\,\theta_1(t) + p'\,\theta_0(t)$$

$$p(t_i) = p\,\frac{\theta_1(t_i) - \theta_1(t_{i-1})}{\theta(t_i) - \theta(t_{i-1})} \text{ and } p(t) = p(t_i) \text{ on } [t_i, t_{i+1})$$

$$\rho(t_i) = p\,\frac{\theta_1'(t_i)}{\theta'(t_i)} \text{ and } \rho(t) = \rho(t_i) \text{ on } [t_i, t_{i+1})$$

Note that if $\theta_j(t)$ is interpreted as the probability of top at t given $\mu = j$, then $p(t_i)$ is the posterior given the first top at t_i and $\rho(t_i)$ given Bottom up to t_i. Obviously the expectation of $p(t_i)$, $\rho(t_i)$ given Bottom up to t_{i-1} is $\rho(t_{i-1})$.

Let $\mathbf{F}$ denote the set of distribution functions corresponding to probabilities on $[0, 1]$.

L_1 (resp. L_0) is the set of measurable mappings from $[0, 1]$ into L_1 (resp L_0). We define now some payoff for (θ_1, θ_0) in Θf^2, $(\bar{\gamma}, \bar{\delta})$ in $\mathsf{L}_1 \times \mathsf{L}_0$, (F, G) in $\mathbf{F}^2$ by:

$$X(p, q; \theta_1, \theta_0; \bar{\gamma}, \bar{\delta}; F, G) = q\, X^1\, (p; \theta_1, \theta_0; \bar{\gamma}, \bar{\delta}; F)$$
$$+ q'\, X^0\, (p; \theta_1, \theta_0; \bar{\gamma}\, \bar{\delta}; G)$$

with

$$X^i\,(p; \theta_1, \theta_0; \bar{\gamma}, \bar{\delta}; F) = \int_0^1 (1 - F(t))\,[t\,\mathrm{v}\,(p\,(t), i)$$
$$+ t'\,[p\,(t)\,\gamma^i\,(t) + p'\,(t)\,\delta^i\,(t)]]\,d\,\theta\,(t)$$
$$+ \int_0^1 (1 - \theta\,(t))\,B\,(t, i)\,d\,F\,(t).$$

Where $B\,(t, i)$ stands for:

$$B\,(t, i) = \min\,\{\min_{0 \leqslant x \leqslant t}\,(x\, a_{22}\,(\rho\,(t), i) + x'\, a_{21}\,(\rho\,(t), i)),\ \mathrm{v}\,(\rho\,(t), i)\}\ i = 1, 0.$$

In order to state the result it remains to introduce:

$$Y(p, q; \theta_1, \theta_0) = \sup_{\substack{\bar{\gamma} \in L_1 \\ \bar{\delta} \in L_0}}\ \inf_{\substack{F \in \mathbf{F} \\ G \in \mathbf{F}}}\ X(p, q; \theta_1, \theta_0; \bar{\gamma}, \bar{\delta}; F, G)$$

$$Z(p, q; \theta_1, \theta_0) = \inf_{\substack{F \in \mathbf{F} \\ G \in \mathbf{F}}}\ \sup_{\substack{\bar{\gamma} \in L_1 \\ \bar{\delta} \in L_0}}\ X(p, q; \theta_1, \theta_0; \bar{\gamma}, \bar{\delta}; F, G)$$

Then we have:

Proposition 2

$Y(p, q; \theta_1, \theta_0) = Z(p, q; \theta_1, \theta_0)$ for all (p, q) in $[0, 1]^2$, (θ_1, θ_0) in Θf^2.

Theorem 3

Maxmin exists on $[0, 1]^2$ and is given by:

$$\underline{\mathrm{v}}\,(p, q) = \sup_{\theta_1, \theta_2 \in \Theta f^2} Y(p, q; \theta_1, \theta_0).$$

3. A. We first prove the proposition.

Remark that X depends upon $\bar{\gamma}$, $\bar{\delta}$ only through their values at the (finitely many) points $\{t_i\}_{i=0,\ldots,m}$ in the support S of θ_1 and θ_0. Hence we can replace L_1 by the convex compact $(L_1 \cap [-\|A\|, \|A\|])^{m+1}$ and similary for L_0. Note now that F is convex and X is affine with respect to $(\bar{\gamma}, \bar{\delta})$ and (F, G). Finally X is obviously continuous with respect to $(\bar{\gamma}, \bar{\delta})$, hence by Sion's minmax theorem $Y = Z$.

We shall use later the fact that a best reply to (F, G) minimizes: $\bar{\gamma}^1(t)(1-F(t))q + \bar{\gamma}^0(t)(1-G(t))q'$ for $t \in S$, with $\bar{\gamma}(t) \in L_1$, and similary for $\bar{\delta}$.

It follows that $\bar{\gamma}(t)$ is a supporting hyperplane for v (1, .) at the point:

$$(3.3)\quad q(t) = q\,\frac{(1-F(t))}{q(1-F(t)) + q'(1-G(t))}$$

The interpretation is the following:

P I is using a "Big match strategy" σ_t blocking at level t, with probability $d\,\theta_i(t)$ if $\mu = i$. Assume that P II plays Right with a frequency increasing form 0 to x, with probability $d\,F(x)$ if $\nu = 1$. If P I plays Top when using σ_t he will deduce that $x \geqslant t$. Thus if L is annonced he will approach $\gamma(t)$ corresponding to $q(t) = \text{Prob}(\nu = 1 \mid x \geqslant t)$.

Let us start now the proof of theorem 3.

3. B. We begin by proving that P II can defend sup $Y(\theta_1, \theta_0)$. The proof relies mainly on two ideas:

· The first one is similar to that used in *Sorin* [1984, lemma 21] or [1985, Lemma 5]. Knowing σ, strategy of P I, P II starts by playing always Left, until reaching the max-min of the probability that P I will play top at this level. From this time on (conditionnally on $\underset{\sim}{m} \geqslant n$ of course) P II will slowly increase his frequency and proceed in the same way. P II will stop at some level x if $\nu = 1$, y if $\nu = 0$. Obviously such a strategy induces with σ a probability $d\,\zeta_i(t)$ of playing Top at level t if $\mu = i$.

Now comes the second step:

P II uses ζ to compute F, G which realizes $Z(p, q; \zeta_1, \zeta_0)$ and plays up to level x (resp. y) according to F (resp. G). If Top is played at level t and L is annonced P II can obtain v $(p(t), 1)$ where $p(\,.\,)$ is induced by p, ζ_1, ζ_0 (see (3.2)). If T appears, P II knows "his" posterior and can get v $(1, q(t))$. This last amount being induced by some (γ, δ) this implies that P II can defend Z, hence the result.

The formal proof is as follows:

$\underset{\sim}{m}$ still denotes the stopping time min $\{n; i_n = \text{Top}\}$ and $S = \{t_r\}_{r=1,\ldots,M}$ is a finite set of points in [0, 1] to be specified later. Given $\epsilon > 0$ and σ strategy of P I we first introduce:

$\tau(0)$: play always left,

$$P_i^*(0) = \text{Prob}_{\sigma^i, \tau(0)}(\underset{\sim}{m} < +\infty) \quad i = 1, 0$$

then n_0 and $P_i(0)$ such that:

$$P_i(0) = \text{Prob}_{\sigma^i, \tau(0)} (\underset{\sim}{m} \leqslant n_0) \geqslant P_i^*(0) - \epsilon \quad i = 1, 0$$

We now define inductively for $r \leqslant M$, given $\tau(r-1)$ and n_{r-1}:
$T(r)$: set of strategies that coincide with $\tau(r-1)$ up to stage n_{r-1} and such that at each following stage Prob (Right) $\leqslant t_r$.

$$P_i^*(r) = \sup_{\tau \in T(r)} \text{Prob}_{\sigma^i, \tau} (\underset{\sim}{m} < +\infty) \quad i = 1, 0,$$

then $\tau(r)$, $n_r \geqslant n_{r-1}$ and $P_i(r)$ such that:

$$P_i(r) = \text{Prob}_{\sigma^i, \tau(r)} (\underset{\sim}{m} \leqslant n_r) \geqslant P_i^*(r) - \epsilon \quad i = 1, 0.$$

Let then ζ_i be the measure in Θf with mass $P_i(r) - P_i(r-1)$ at point t_r, and let F and G in **F** be $\epsilon \| A \|$ optimal strategies for P II in $Z(p, q; \zeta_1, \zeta_0)$ (see Prop. 2) i.e. such that:

$$X(p, q; \zeta_1, \zeta_0; \bar{\gamma}, \bar{\delta}; F, G) \leqslant Y(p, q; \zeta_1, \zeta_0) + \epsilon \| A \|$$

for all $\bar{\gamma}$, $\bar{\delta}$.

We finally introduce $\bar{F}$ atomic probability measure with mass $F(t_r) - F(t_{r-1})$ at point t_r and similary for $\bar{G}$. The strategy of P II can now be described as: P II chooses t_r according to $\bar{F}$ if $\nu = 1$ ($\bar{G}$ if $\nu = 0$). Then he plays $\tau(r)$ up to stage $\underset{\sim}{m}$, and plays optimally in the revealed game after stage $\underset{\sim}{m}$.

In order to compute the expected payoff at some stage n large enough ($\geqslant n_M$) we first study the different events induced by $h_{\underset{\sim}{m}}$.

Recall first that the event $\{\underset{\sim}{m} \leqslant n_r$ and P II plays Right with probability $t > t_r$ at stage $\underset{\sim}{m}\}$ has zero probability.

Moreover by construction we have that Prob $(n \geqslant \underset{\sim}{m} > n_r \mid \tau'(r)) < \epsilon$ for all $n > n_r$ and all $\tau'(r)$ in $T(r)$ that coincides with $\tau(r)$ up to stage n_r. We shall thus neglect the events: $(\{n \geqslant \underset{\sim}{m} > n_r\}$ and $\tau(r)$ is played$\})$.

(3.4) Thus if $\{\underset{\sim}{m} > n\}$ and $\tau(r)$ is played we use the fact that Prob $(\underset{\sim}{m} > n_r \mid \tau(r)$ and $\mu = 1)$ is $1 - \zeta_1(t_r)$ to compute the posteriors Prob $(\mu = 1 \mid \underset{\sim}{m} > n_r$ and $\tau(r)) =$
$= \rho(t_r)$ see (3.2).

(3.5) Now Prob $(\underset{\sim}{m} \in (n_{r-1}, n_r] \mid \tau(s)$ and $\mu = 1) = \zeta_1(t_r) - \zeta_1(t_{r-1})$ for all $s \geqslant r$ and 0 otherwise. It follows that Prob $\{\mu = 1 \mid \underset{\sim}{m} \in (n_{r-1}, n_r], \tau(s)\} = p(t_r)\ \forall\, s \geqslant r$.

On the other hand, even if P I knows the strategy of P II he can only compute Prob $(\nu = 1 \mid \underset{\sim}{m} \in (n_{r-1}, n_r]) =$ Prob $(\nu = 1 \mid \tau(s)$ is played, with $s \geqslant r) = \bar{q}(t_r)$ (see (3.3)) induced by $\bar{F}, \bar{G}$.

Finally if $s(t)$ is defined by $s(t_r) =$ Prob (P II plays Right at stage $\underset{\sim}{m} \mid \underset{\sim}{m} \in (n_{r-1}, n_r]$) then $s(t_r) \leqslant t_r$ a.s..

Hence we have the following complete description:

– if $\underset{\sim}{m} > n$ and $\tau(r)$, the posterior (on μ) is $\rho(t_r)$; then P II can either keep playing with Prob (Right) $\leqslant t_r$ or obtain $\mathrm{v}(\rho(t_r), j)$ if $\nu = j$ (this gives $B(t_r, j)$).

– if $\underset{\sim}{m} \in (n_{r-1}, n_r]$, then $\underset{\sim}{j_m} =$ Right with probability $s(t_r)$. Given T the posteriors are $(1, \bar{q}(t_r))$ and given L they are $(p(t_r), 1)$.

We now obtain, using (3.2) and (3.3) that for N large enough:

$$N \sum_{n_M}^{n_M+N} \gamma_n(\gamma, r) \leqslant \int_0^1 s(t)\,[q\,(1 - \bar{F}(t))\,\mathrm{v}(p(t), 1) + q'\,(1 - \bar{G}(t))$$

$$\mathrm{v}(p(t), 0)]\, d\zeta(t)$$

$$+ \int_0^1 s'(t)\,[q\,(1 - \bar{F}(t)) + q'\,(1 - \bar{G}(t))]\,[p\,\mathrm{v}(1, q(t))$$

$$d\zeta_1(t) + p'\,\mathrm{v}(0, q(t))\, d\zeta_0(t)]$$

$$+ \int_0^1 (1 - \zeta(t))\,[B(t, 1)\, q\, dF(t) + B(t, 0)\, q'\, dG(t)] +$$

$$+ 2\,\|A\|\,\epsilon.$$

By using (3.1), (3.3) and $s(t) \leqslant t$ it follows that there exists $\bar{\gamma}, \bar{\delta}$ in $L_1 \times L_0$ such that:

$$N \sum_{n_M}^{n_M+N} \gamma_n(\sigma, \tau) \leqslant X(p, q; \zeta_1, \zeta_0; \bar{\gamma}, \bar{\delta}; \bar{F}, \bar{G}) + 2\,\|A\|\,\epsilon.$$

It remains to use the fact that $(\bar{F}, \bar{G})$ can be choosen near (F, G) the last being optimal in Z.

In fact we have:

$$| F(t) - \bar{F}(t) | \leqslant \epsilon \text{ on } [0, 1]$$

by defining $t_r = \min \{t; F(t) \geqslant r\,\epsilon\}$

On the other hand given f bounded (by 1 say) and right continuous $| \int_0^1 f(t)\,(d\,F(t) - - d\,\bar{F}(t)) | \leqslant \epsilon$ as soon as $| t_r - t_{r-1} |$ is smaller than some $\eta\,(f)$. It follows then finally that for the corresponding τ:

$$N \sum_{n_M}^{n_M + N} \gamma_n\,(\sigma, \tau) \leqslant X\,(p, q; \zeta_1, \zeta_0; \bar{\gamma}, \bar{\delta}; F, G) + 4\,\| A \|\,\epsilon$$

hence by using Proposition 2 there exists $\bar{N}$ such that $n \geqslant \bar{N}$ implies:

$$\bar{\gamma}_n\,(\sigma, \tau) \leqslant Y\,(p, q; \zeta_1, \zeta_0) + 5\,\| A \|\,\epsilon.$$

this achieves the proof that P II can defend $\underline{v}$.

3. C. This last part, namely that P I can guarantee sup $Y\,(\theta_1, \theta_0)$ is now quite standard [see e.g. *Sorin*, 1984, Prop 26, or *Sorin*, 1985, Lemma 7] but still requires some new notations.

$J\,(t)$ is the infinitely repeated stochastic game with payoff matrix:

$$\begin{pmatrix} -t^* & (1-t)^* \\ t & -(1-t) \end{pmatrix}$$

The value of $J\,(t)$ is 0 and we denote by $\sigma\,(t)$ an ϵ optimal strategy of P I in $J\,(t)$, i.e. such that $n \geqslant N_t$ implies: $\bar{j}_n\,(\sigma, \tau) \geqslant -\epsilon$ where $\bar{j}_n\,(\sigma, \tau)$ is the average expected payoff in $J\,(t)$. (Recall that $J\,(1/(2))$ is the "Big Match" of *Blackwell/Ferguson*):

A sketch of the proof already appears in 3. A. So let (θ_1, θ_0) in Θf realize sup Y up to some ϵ.

We denote by $S = \{t_r \mid t_0 = 0, t_M = 1, r = 0, \ldots M\}$ a finite set including the support of θ_i $(i = 1, 0)$, with $| t_r - t_{r-1} | \leqslant \bar{\eta}$ (to be specified later). Finally $\bar{\gamma}\,(t_r) = = \gamma_r$, $\bar{\delta}\,(t_r) = \delta_r$ are defined as optimal for $Y\,(\theta_1, \theta_0)$ (see proposition 2)

To make the notations simpler we shall write $\sigma\,(r)$ for $\sigma\,(t_r)$ and $d\,\theta_i\,(t_r)$ for $\theta_i\,(\{t_r\})$. We also define $\bar{N} = \max\limits_{t_r \in S} N_{t_r}$.

The strategy σ of P I is now as follows:

Choose r^* in $M \cup \{\alpha\}$ according to the distribution defined by, if $\mu = i$, Prob $(r^* = r) = d\,\theta_i(t_r)$ and Prob $(r^* = \alpha) = 1 - \theta_i(1)$.

Then:

- if $r^* = 0$ play Top at stage 0.
- if $r^* = r \in (0, M]$ play Bottom up to stage m_r and Top at this stage.
- if $r^* = \alpha$ play always Bottom

where the stopping times m_r are defined on each history by:

$$m_1 = \min\ \{m \geqslant 1;\, i_m = \text{Top}\} \text{ induced by } \sigma(1)$$

$$m_r = \min\ \{m \geqslant m_{r-1};\, i_m = \text{Top}\} \text{ induced by } \sigma(r) \text{ from stage } m_{r-1} \text{ on.}$$

Finally after $\underset{\sim}{m}$, P I plays optimally in the revealed game if L or R and approaches γ_{r^*} if T (resp. δ_{r^*} if B).

We shall prove that given $\nu = 1$, P I can guarantee

$$\inf_x\ X^1\ (p;\theta_1,\theta_0;\bar{\gamma},\bar{\delta},F_x)$$

where F_x stands for the Dirac mass at $x \in [0, 1]$.

The result will then follow by using the definition of X and the properties of (θ_1,θ_0) and $(\bar{\gamma},\bar{\delta})$.

Now given n and τ strategy of P II we define:

$$U_r = m_r \wedge n - m_{r-1} \wedge n,\ s_r = \mathbf{I}_{\{j_{m_r} = \text{Right}\}}$$

and

$$\bar{s}_r = \frac{1}{U_r}\ \#\ \{j_m = \text{Right};\, m_{r-1} \wedge n < m \leqslant m_r \wedge n\}$$

Hence P I is using σ_r during U_r stages, $\bar{s}_r$ is the average frequency of Right on this stages and s_r describes the strategy of P II at stage m_r.

Note finally that if $h_{m_r} = L$, the posteriors are $p\,(t_r)$ (as defined by (θ_1,θ_0) in 3.2) and $\rho\,(t_r)$ if h_{m_r} is a regular letter. Thus we obtain:

$$(3.6)\ n\ \bar{\gamma}_n^{p,1}(\sigma,\tau) \geqslant E_{\sigma,\tau}\left(\sum_{r=1}^{\alpha} U_r \sum_{k=0}^{r-1} ((p\, d\,\theta_1(t_k) + p'\, d\,\theta_0(t_k))\, s_k\, \mathrm{v}(p(t_k), 1) + s'_k\,[p\, d\,\theta_1(t_k)\,\gamma^1(k) + p'\, d\,\theta_0(t_k)\,\delta^1(k)]\right)$$

$$+ [p(1-\theta_1(t_{r-1})) + p'(1-\theta_0(t_{r-1}))]\,[a_{22}(\rho(t_{r-1}), 1)\,\bar{s}_r + a_{21}(\rho(t_{r-1}), 1)\,\bar{s}'_r]\Big)$$

Now after stage m_{r-1}, $m \leqslant m_r$ implies that the average of Right is less that some $t_r + \eta$ (otherwise P I plays Top). By the choice of $\sigma(r)$ we thus have:

$$E[U_r \cdot \bar{s}_r] \leqslant N + E(U_r)(t_r + \eta + \epsilon)$$

$$\leqslant N + E(U_r)(t_{r-1} + 2\eta + \epsilon)$$

It follows that the last term in (3.6) is minorized by

$$(1-\theta(t_{r-1}))\,B(t_{r-1}, 1) - [(M+1)N + 2\eta + \epsilon]\,\|A\|.$$

Hence we have, using (3.1)

$$(3.7)\ \bar{n}\,\gamma_n^{p,1}(\sigma) \geqslant E\left(\sum_{r=1}^{\alpha} U_r X^1(t_r)\right) - \|A\|\,((M+1)N + (2\eta+\epsilon)\,n - \Delta)$$

where we write, for short $X^1(t_r) = X^1(p;\theta_1, \theta_0; \bar{\gamma}, \bar{\delta}, F_{t_r})$ and Δ stands for:

$$(3.8)\ \Delta = E \sum_r U_r \sum_k d\,\theta(t_k)(s_k - t_k))$$

It remains to remark that $s_k - t_k = s_k(1-t_k) + (1-s_k)(-t_k)$ is the absorbing payoff in $J(t_k)$, and that this payoff occurs at stage m_r for $\sum_{r+1}^{\alpha} U_k$ stages.

Using the optimality of $\sigma(r)$ we thus have

$$E\left(\sum_{r+1} U_k (s_k - t_k)\right) \geqslant -\epsilon n$$

hence $\Delta \geqslant -\epsilon n(M+1)$.

We get now from (3.7)

$$n\,\bar{\gamma}^{p,1}(\sigma) \geqslant \min_r X^1(t_r) - \|A\|\,((M+1)N + (2\eta+\epsilon)n + \epsilon n(M+1)].$$

A similar result for $\bar{\gamma}^{p,0}$ now implies; $\forall\,\tau$, $\forall\,\epsilon_0 > 0$, there exists $\bar{N}$ and σ such that for $n \geqslant \bar{N}$

$$\bar{\gamma}^{p,q}(\sigma,\tau) \geqslant \inf_{F,G\in\mathbf{F}} X(p,q;\theta_1,\theta_2;\bar{\gamma},\bar{\delta};F,G) - \epsilon_0$$

hence the required result.

4 Concluding Remarks.

4. A. We do not have any statement about the asymptotic behavior of $\mathrm{v}_n(p,q)$: value of the n-time repeated game. Nevertheless it is worthwile to remark that the following recursive formula holds:

$$\begin{aligned}(n+1)\,\mathrm{v}_{n+1}(p,q) = \max_{\substack{0\leqslant s^1\leqslant 1\\ 0\leqslant s^0\leqslant 1}} \min_{\substack{0\leqslant t^1\leqslant 1\\ 0\leqslant t^0\leqslant 1}} &\{p\,q\,s^1 A(1,1)\,t^1 + p'\,q\,s^0\\ &s^0 A(0,1)\,t^1 + p\,q'\,s^1 A(1,0)\,t^0 + p'\,q'\,s^0 A(0,0)\,t^0\\ &+ n\,[q\,t^1 + q'\,t^0]\,[p\,s^1\,\mathrm{v}_n(1,q_l) + p'\,s^0\,\mathrm{v}_n(0,q_l)]\\ &+ n\,[p\,s^1 + p'\,s^0]\,[q\,(t^1)'\,\mathrm{v}_n(p_t,1) + q'\,(t^0)'\,\mathrm{v}_n(p_t,0)]\\ &+ n\,[p\,(s^1)' + p'\,(s^0)']\,([q\,t^1 + q'\,t^0]\,\mathrm{v}_n(p_b,q_l)\\ &+ [q\,(t^1)' + q'\,(t^0)']\,\mathrm{v}_n(p_b,q_r))\}\end{aligned}$$

where

$$x\,A\,y = x\,a_{11}\,y + x\,a_{12}\,y' + x'\,a_{21}\,y + x'\,a_{22}\,y'$$

$$p_t = \frac{p\,s^1}{p\,s^1 + p'\,s^0},\; p_b = \frac{p\,(s^1)'}{p\,(s^1)' + p'\,(s^0)'},\; q_l = \frac{q\,t^1}{q\,t^1 + q'\,t^0},$$

$$q_r = \frac{q\,(t^1)'}{q\,(t^1)' + q'\,(t^0)'}$$

A sketch of the proof is as follows: (s^1, s^0) and (t^1, t^0) define the strategies of both players at the first stage in $G_{n+1}(p, q)$. They induce a payoff at stage 1 and a distribution on the letters. By the minmax theorem, the corresponding posteriors $\tilde{q}(h)$, $\tilde{p}(h)$ on the state space, given some letter h, can be assumed to be common knowledge. Now the players, by playing optimally obtain $n\, \mathrm{v}_n(\tilde{p}(h), \tilde{q}(h))$ for the remaining n stages. The above formula is just:

$$(n+1)\, \bar{\gamma}_{n+1}(\sigma, \tau) = \gamma_1(\sigma, \tau) + n\, E_{\sigma,\tau}\, \mathrm{v}_n(\tilde{p}, \tilde{q})$$

4. B. The analysis in Part II and III shows also that a continuous parameter space is natural in this framework (even with discrete time). This was already the case in *Sorin* [1984] for $\lim \mathrm{v}_n$ and in *Sorin* [1984, 1985] for $\underline{\mathrm{v}}$ and $\bar{\mathrm{v}}$.

In fact we "discretise" the strategies in order to get in an easy way uniform bounds for the convergence but it is clear that there are no "limit problems".

In Part II min $W_1(\alpha, \beta)$ is the value of the game where the strategy space of P II is $\{\text{left} \cup (H_0 \times H_1)\}$.

In part III given θ_1, θ_0 in Θ, $p(t)$ is the Radon-Nidokym derivative of $p\,\theta_1$ with respect to $p\,\theta_1 + p\,\theta_0$.

4. C. The use of "stochastic games" for solving games with incomplete information is introduced first in the "symmetric case" see 1.1 d). The previous analysis seems to prove that this phenomena is more general [another example can be found in *Sorin*, 1984].

It appears then that stochastic games [completely solved now by *Mertens/Neyman*] and games with incomplete information are more and more included in a single topic that could be called: "stochastic games with incomplete information".

References

Aumann, R.J., and *Maschler:* Repeated Games of Incomplete Information. The Zero-Sum Extensive Case, ACDA ST-143. Mathematica, Princeton 1968, 25 – 108.

Blackwell, D., and *T.S. Ferguson:* The Big Match. Ann. of Math. Stat., **39**, 1968, 159 – 163.

Forges, F.: Infinitely Repeated Games of Incomplete Information. Symmetric Case with Random Signals, Int. J. Game Th. **11**, 1982, 203–213.

Kohlberg, E.: Optimal Strategies in Repeated Games with Incomplete Information. Int. J. Game Th., **4**, 1975, 7 – 24.

Kohlberg, E., and *S. Zamir:* Repeated Games of Incomplete Information, the Symmetric Case. Ann. of Stat. **2**, 1974, 1040 – 1041.

Mertens, J.F.: The Value of Two Person Zero Sum Repeated Games: the Extensive Case. Int. J. Game Th., **1**, 1971/72, 217 – 227.

Mertens, J.F., and *A. Neyman:* Stochastic Games. Int. J. Game Th., **10**, 1981, 53 – 66.

Mertens, J.F., and *S. Zamir:* The Value of Two Person Zero Sum Repeated Games with Lack of Information on Both Sides. Int. J. Game Th. **1**, 1971/72, 39 – 64.

–: On a Repeated Game Without a Recursive Structure. Int. J. Game Th., **5**, 1976, 173 – 182.

–: Minmax and Maxmin of Repeated Games with Incomplete Information. Int. J. Game Th., **9**, 1980, 201 – 215.

Sorin, S.: Une Introduction aux Jeux Répétés, à Deux Joueurs à Somme Nulle et à Information Incomplète. Cahiers de Mathématiques Economiques, n° 1, 1979 (English Version, IMSSS, Stanford, TR 312, 1980).

–: "Big Match" with Lack of Information on One Side, Part I. Int. J. Game Th., **13**, 1984, 201 – 255.

–: "Big Match" with Lack of Information on One Side, Part II. Int. J. Game Th., **14**, 1985.

Waternaux, C.: Solution for a Class of Repeated Games without a Recursive Structure. Int. J. Game Th., **12**, 1983, 129 – 160.

Received February 1984
(Revised version February 1985)