Universität Bielefeld/IMW

Working Papers
Institute of Mathematical Economics

Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

Nr. 174

Asymptotically Optimal Strategies in Repeated Zero–Sum Games with Incomplete Information

by

Martin Heuer

May 1989

H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany

0. Introduction

We consider a model of a finitely repeated two–person zero–sum game with incomplete information on both sides that has already been examined by Mertens and Zamir [71/72]. They have determined the limiting value of such a game as the number N of repetitions tends to infinity and they have shown that the difference between the value of the N–stage game and the limit is bounded by $\frac{K}{\sqrt{N}}$. However, their analysis gives no indication to "good" strategies in such a game because it is guided by the following consideration: An equilibrium strategy (which always exists in the finitely repeated case) must guarantee the value of the game even if it is known to the opponent, which implies that he is able to compute posterior probabilities on the other player's types after each stage of the game. Thus the value of the game remains unchanged if both players after each stage are not only told their opponent's choice of action but also the correct posteriors on his types. In such a game the participants must only control their own release of information but not the behaviour of the other player.

In the following we will define a strategy that works without this additional information and guarantees a payoff in the N–stage game that differs from the limiting value by at most $\frac{K}{\sqrt{N}}$. This paper may be regarded as a first version since the model of Mertens and Zamir is not yet covered in full generality: We only consider independent player types.

1. The Model

A finitely repeated two–person zero–sum game with incomplete information on both sides is based on the following data:

- finite sets I and J
  (sets of actions for player 1 and 2)

- finite sets R and S
  (sets of types for players 1 and 2 resp. states of nature)

- for every $(r,s) \in R \times S$ an $I \times J$–matrix $A^{r,s}$
  (payoff matrices)

- probability distributions $p_0 \in \Delta(R)$ and $q_0 \in \Delta(S)$
  (resp. a distribution on $R \times S$ which is the product of its marginals)

- a natural number N
  (number of steps)

The game runs as follows:

1) At stage zero r and s are selected according to p_0 resp. q_0. Both players know both distributions but player 1 is only informed about the choice of $r \in R$ and player 2 about the choice of $s \in S$.

2) At each stage $t = 1,\dots,N$ both players independently pick out parameters $i_t \in I$ resp. $j_t \in J$.

3) Afterwards they learn their opponent's choice of action but not the actual value $A^{r,s}(i_t,j_t)$ of the payoff matrix.

4) Both players have perfect recall, i.e. they can use all information they receive in the course of the game up to stage t for their decision at stage t + 1.

5) Player 1 receives from player 2 the amount $\frac{1}{N}\sum_{t=1}^{N} A^{r,s}(i_t,j_t)$.

Let $H := I \times J$. In the following the identities $h = (i,j)$, $h_t = (i_t,j_t)$ and $h^t = (h_1,\dots,h_t)$ are always tacitly assumed. According to the description above player 1 can make his choice of action at stage t dependent on his type r and the sequence $h^{t-1} = (h_1,\dots,h_{t-1})$ of parameters that have occurred up to stage $t-1$. Thus a strategy of player 1 consists of a sequence $\sigma = (\sigma_1,\dots,\sigma_N)$ of mappings

$$\sigma_t : R \times H^{t-1} \to \Delta(I) \qquad \text{resp. stochastic kernels}$$

$$\sigma_t \mid R \times H^{t-1} \Rightarrow I.$$

Analogously a strategy of player 2 is given by a sequence $\tau = (\tau_1,\dots,\tau_N)$ of mappings

$$\tau_t : S \times H^{t-1} \to \Delta(J) \qquad \text{resp. stochastic kernels}$$

$$\tau_t \mid S \times H^{t-1} \Rightarrow J.$$

The sets of strategies are denoted by

$$\Sigma^N = \prod_{t=1}^{N} \Sigma_t \text{ and } \mathcal{T}^N = \prod_{t=1}^{N} \mathcal{T}_t.$$

Using the strategies $\sigma = (\sigma_1,\dots,\sigma_N)$ and $\tau = (\tau_1,\dots,\tau_N)$ the players generate the probability distribution $\underline{P}^{p_0,q_0}_{(\sigma,\tau)}$ on $R \times S \times H^N$:

$$\underline{P}^{p_0,q_0}_{(\sigma,\tau)}(r,s,h^N) = p_0(r)\, q_0(s) \prod_{t=1}^{N} \sigma_t(r,h^{t-1}; i_t)\, \tau_t(s,h^{t-1}; j_t)$$

The payoff function is naturally defined as the expectation of the average payoffs with respect to this distribution

$$\alpha(\sigma,\tau) = \alpha^N_{p_0,q_0}(\sigma,\tau) = \int_{R\times S\times H^N} \frac{1}{N}\sum_{t=1}^{N} A^{r,s}(h_t)\, d\underline{P}^{p_0q_0}_{(\sigma,\tau)}$$

(r,s,h_t denoting the projection on the corresponding component of the product space).

Hence we have a non–cooperative two–person zero–sum game in the formal form

$$\Gamma^N(p_0,q_0) = (\Sigma^N, \mathcal{T}^N, \alpha^N_{p_0,q_0}).$$

The strategy sets are compact and the payoff function is linear in every component of σ and τ. Thus the value $v_N(p_0,q_0)$ exists.

2. An Example

Define $A(p,q) = \sum_{r,s} p(r)\, q(s)\, A^{r,s}$, $p \in \Delta(R)$, $q \in \Delta(S)$

and let $u(p,q)$ be the value of the matrix game $A(p,q)$. Mertens and Zamir [71/72] have shown that $v(p,q) = \lim_{N \to \infty} v_N(p,q)$ exists and that it is the unique simultaneous solution of the functional equations

$$f = \operatorname{cav}_{I} \min\{u,f\}$$

$$f = \operatorname{vex}_{II} \max\{u,f\}$$

cav_I means concavification w.r.t. the first variable p and vex_{II} means convexification w.r.t. the second variable q.

In the case of lack of information on one side (as usually the side of player 2), i.e. $S = 1$ the functional equations reduce to $v = \operatorname{cav} u$. In this case optimal strategies are known for the infinitely repeated game, and for every $\epsilon > 0$ there exists a number of stages N such that the strategy is ϵ–optimal in the N–stage game. Player 1's strategy involves the concept of type–dependent lotteries which will also be employed in the strategy we are going to construct. (see e.g. Sorin [80]). Player 2's strategy is an application of Blackwell's approaching strategy for games with vector payoffs (Blackwell [56], see e.g. Sorin [80]). A closer look at this strategy will be useful.

$v = \operatorname{cav} u$ is initially defined on $\Delta(R)$ but it may be extended linear homogeneously to $\mathbb{R}_+^R$ preserving concavity. There exists a supporting hyperplane of v at p_0, i.e. there is a vector $\eta_0 \in \mathbb{R}^R$ such that

$$p_0 \eta_0 = v(p_0)$$

$$p\, \eta_0 \geq v(p) \quad \forall\, p \in \Delta(R) .$$

Define $\gamma_T(h^T) = \frac{1}{T} \sum_{t=1}^{T} A(h_t), \ 1 \le T \le N$

with $A(h_t) = (A^1(h_t),\dots,A^R(h_t))$

and a function

$$\text{dist:} \quad \mathbb{R}^R \to Ch(\eta_0) = \{\eta \in \mathbb{R}^R : \eta \le \eta_0)$$

$$x \mapsto \arg \min_{\eta \in Ch(\eta_0)} | x - \eta |$$

Player 2's strategy works like this:

If $\gamma_T(h^T) \notin Ch(\eta_0)$ he computes the probability $p_T^{\prime} \in \Delta(R)$ which is parallel to $\gamma_T(h^T) - \text{dist}\,(\gamma_T(h^T))$ and chooses an optimal strategy $y_T \in \Delta(J)$ of the matrix game $A(p_T^{\prime})$.

If $\gamma_T(h^T) \in Ch(\eta_0)$ he may do anything.

Example 1: (cf. Zamir [74])

$$A^1 = \begin{bmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{bmatrix} \qquad A^2 = \begin{bmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{bmatrix}$$

$$A(p) = \begin{bmatrix} 4p & 4(1-p) & 4p-2 \\ 4p & 4(1-p) & 2-4p \end{bmatrix}$$

There are always pure equilibrium strategies in the NR–game:

Player 1 chooses row 2 if $p \le \frac{1}{2}$

row 1 if $p \ge \frac{1}{2}$

Player 2 chooses column 1 if $p \le \frac{1}{2}$

column 3 if $\frac{1}{4} \le p \le \frac{3}{4}$

column 2 if $p \ge \frac{3}{4}$

Thus the value u(p) of the NR–game A(p) is given by

$$u(p) = \begin{cases} 4p & , \ p \le \frac{1}{4} \\ 2 - 4p & , \ \frac{1}{4} \le p \le \frac{1}{2} \\ 4p - 2 & , \ \frac{1}{2} \le p \le \frac{3}{4} \\ 4(1-p) & , \ p \ge \frac{3}{4} \end{cases}$$

and we find

$$v(p) = \begin{cases} 4p & , \ p \le \frac{1}{4} \\ 1 & , \ \frac{1}{4} \le p \le \frac{3}{4} \\ 4(1-p) & , \ p \ge \frac{3}{4} \end{cases}$$

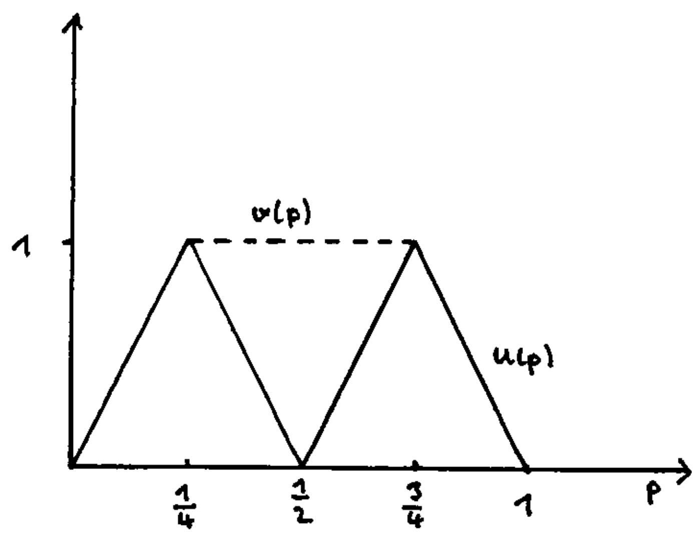

figure 1

Let us now construct an approaching strategy for player 2 in $\Gamma(p_0)$ with $p_0 = (\frac{1}{2},\frac{1}{2})$. First of all η_0 must be determined. The derivative of v (as a function of one variable) at $p = \frac{1}{2}$ is zero, which implies that the gradient of the extended linear homogeneous function v on $\mathbb{R}^2_+$ at $p_0 = (\frac{1}{2},\frac{1}{2})$ contains two equal components. Since $v(p_0) = 1$ we deduce that $\eta_0 = (1,1)$. The approaching strategy is represented by the figure below.

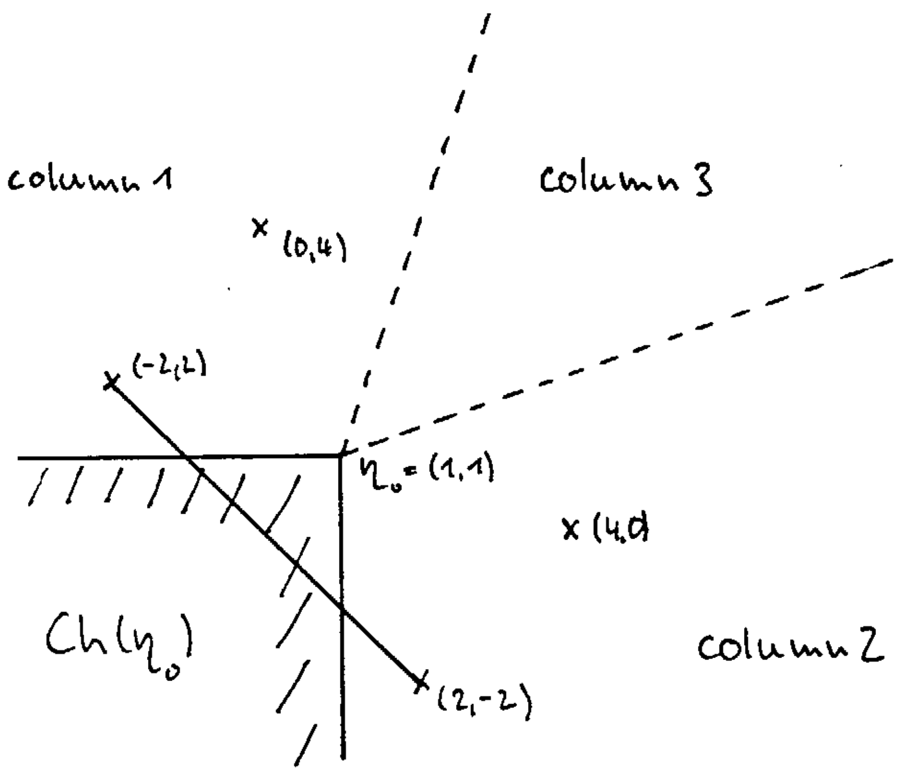

fig. 2

The complement of $Ch(\eta_0)$ is divided into three sectors labeled by the columns player 2 has to select at stage T + 1 if γ_T is contained in the corresponding sector. If $\gamma_T \in$ sector 1 the (T + 1)–stage payoff added to the average is in sector 2 and vice versa, $\gamma_T \in$ sector 3 gives a (T + 1)–stage payoff on the line segment $[(-2,2), (2,-2)]$. It may be plausible from the figure above that this strategy always draws the vector payoff γ_T towards $Ch(\eta_0)$ as the game proceeds. Remark that the approaching strategy works in the finitely and in the infinitely repeated game. We shall now propose an alternative procedure for which the finite horizon is crucial.

Using the approaching strategy player 2 enforces a certain vector payoff given by a supporting hyperplane of the value function. Of course, the vector payoffs he can obtain do not depend on the actual probability p_0. He could choose any supporting hyperplane of v but the one at p_0 gives him the best total payoff given p_0 among all hyperplanes of v. Nevertheless, one can think of a situation that may induce player 2 to change his aim. Suppose by some bad luck the current vector payoff γ_T has moved far away from the vector payoff η_0 he was striving for (which may happen more easily if the equilibrium strategies of the NR–game are mixed unlike in our example). Player 2 might find that no longer favourable to head for η_0 at the present stage but that in order to get close to η_0 after N stages he should choose a different hyperplane for remaining N – T steps. This computation could look like this:

If he starts a new game at time T + 1 consisting of the remaining N – T stages in which he is able to achieve the vector payoff

$$\delta_T = \frac{N}{N-T}\,\eta_0 - \frac{T}{N-T}\,\gamma_T$$

he will in the end get the vector payoff

$$\frac{T}{N}\,\gamma_T + \frac{N-T}{N}\,\delta_T = \eta_0 \,.$$

Of course, δ_T itself will generally not be attainable, i.e. δ_T will not be a supporting hyperplane of v. But in order to keep his potential loss small he could choose the hyperplane with minimal euclidian distance from δ_T.

Let us see how this strategy works in our example. In the first place one must determine the set of supporting hyperplanes V(p) of v at each point p:

$$V(p) = \begin{cases} \begin{bmatrix}0\\4\end{bmatrix} & , \quad 0 < p < \frac{1}{4} \\ Co\left\{\begin{bmatrix}0\\4\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\} & , \quad p = \frac{1}{4} \\ \begin{bmatrix}1\\1\end{bmatrix} & , \quad \frac{1}{4} < p < \frac{3}{4} \\ Co\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}4\\0\end{bmatrix}\right\} & , \quad p = \frac{3}{4} \\ \begin{bmatrix}4\\0\end{bmatrix} & , \quad \frac{3}{4} < p < 1 \end{cases}$$

In addition we define

$$V(0) = \left\{\begin{bmatrix}0\\x\end{bmatrix} : x \geq 4\right\} \qquad V(1) = \left\{\begin{bmatrix}x\\0\end{bmatrix} : x \geq 4\right\}$$

At first, this may seem a bit arbitrary, but it coincides with the general definition of V in section 3 and it will turn out to be useful for the definition of the strategy. The next figure shows the union $V = \bigcup_{p \in \Delta(R)} V(p)$ and it represents the alternative strategy:

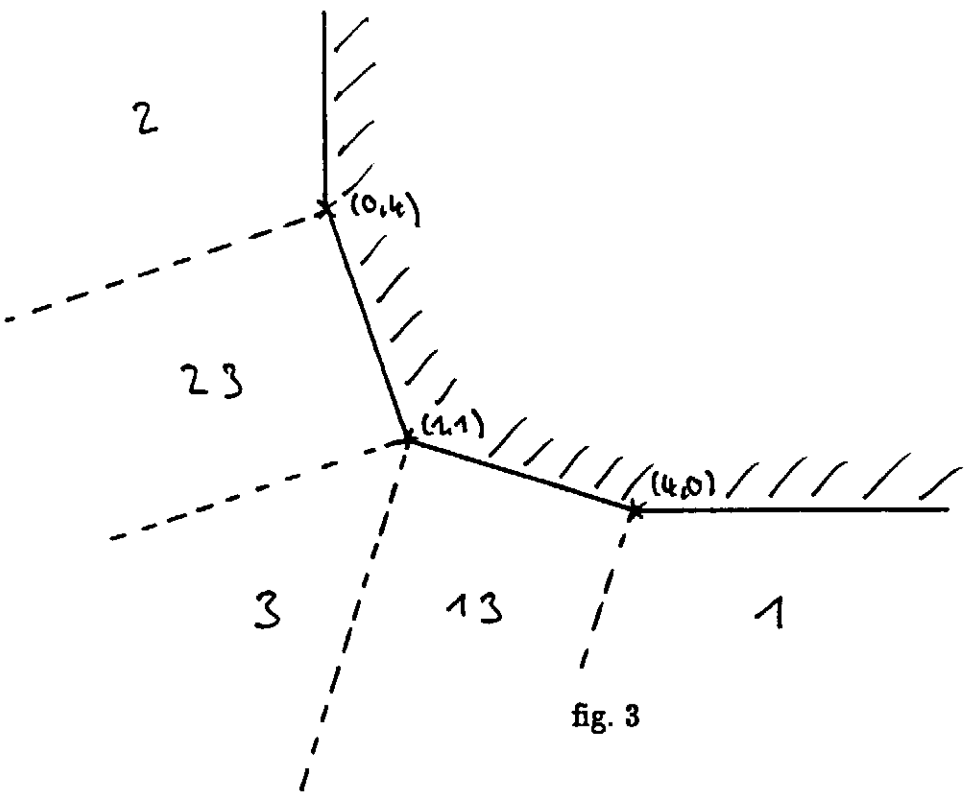

fig. 3

If δ_T is located in the shaded area, player 2 has gained an advantage with regard to his original objective: In order to get η_0 in the end he must in the last $N - T$ stages only achieve an average vector payoff which is greater than one he can obtain in both components. He may be allowed to waste this gain by playing anything as in the approaching strategy. The remainder of $\mathbb{R}^2$ is divided into five sectors each marked by the columns player 2 can choose.

Let e.g. δ_T be situated in sector 23. The element of V with minimal euclidian distance from δ_T is contained in the line segment $\mathrm{Co}\left\{\begin{bmatrix}0\\4\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\}$. This segment is associated to the probability $p = \frac{1}{4}$. So he must choose an optimal strategy of $A(\frac{1}{4})$, i.e. he must play column 2 or 3. A proof that this kind of strategy works as well will be postponed till the main theorem. □

An advantage of the alternative strategy compared with Blackwell's approaching strategy will become recognizable if we think of the general case of lack of information on both sides. Player 2 then does not only have to keep an eye on player 1's vector payoff but he also has to administrate his own private information since he must reckon with player 1 being able to calculate posteriors on S.

Suppose now that the function v(p) we studied in the example is derived from a value function v(p,q) for some fixed q_0. Using Blackwell's strategy it may happen that at stage two player 2 has to act as if $p = 0$ or $p = 1$; i.e. he has to act as if he were in a game with lack of information on the side of player 1 only. Of course these two possibilities (or others in between) generally require a completely different use of player 2's private information. In this way the approximately negligible first stage payoff already widely determines the release of player 2's information, a feature that enables player 1 to cheat his opponent. On the other hand the modification of δ_T and thus of the accompanying probability proceeds slowly throughout the game and gives no opportunity to cheating, as will be shown in the proof of the main theorem.

3. Technical Preliminaries

The results of Mertens and Zamir [71/72] will not be used here, that means we won't refer to the fact that $\lim_{N\to\infty} v_N$ exists and that it is given by the functional equations

$$f = \underset{I}{\text{cav}}\ \min\{u,f\}$$

$$f = \underset{II}{\text{vex}}\ \max\{u,f\}$$

We will presume that there is a unique simultaneous solution to both functional equations called v and that it is continuous if u is continuous, a fact that can be shown purely analytically not making use of the game – theoretical meaning of these equations (Mertens and Zamir [77]).

Lemma 3.1:

The function u is continuous.

Proof:

Let (x^1,y^1) and (x^2,y^2) be pairs of equilibrium strategies in the matrix games $A(p^1,q^1)$ resp. $A(p^2,q^2)$.

W.l.o.g. assume that $u(p^1,q^1) > u(p^2,q^2)$.

$$u(p_1,q_1) - u(p_2,q_2)$$

$$= \sum_{i,j} x_i^1\, y_j^1\, A(p^1,q^1)_{i,j} - \sum_{i,j} x_i^2\, y_i^2\, A(p^2,q^2)_{i,j}$$

$$\leq \sum_{i,j} x_i^1\, y_j^2\, (A(p^1,q^1)_{i,j} - A(p^2,q^2)_{i,j})$$

$$\leq \max_{i,j} \left(\sum_{r,s} p^1(r)\ q^1(s)\ A^{r,s}(i,j) - \sum_{r,s} p^2(r)\, q^2(s)\, A^{r,s}(i,j)\right)$$

$$\leq |p^1 - p^2|\,|q^1 - q^2|\,2M$$

with $M = \max_{i,j,r,s} |A^{r,s}(i,j)|$

□

For any mapping $f : \Delta(R) \times \Delta(S) \to \mathbb{R}$ denote by f_q the function $f_q : \Delta(R) \to \mathbb{R}$, $f_q(p) = f(p,q)$ and with slightly ambiguous notation $f_p : \Delta(S) \to \mathbb{R}$, $f_p(q) = f(p,q)$.

Proposition 3.2:

Let $(p,q) \in \Delta(R) \times \Delta(S)$ with $v(p,q) < u(p,q)$. Then there exist $q^\ell \in \Delta(S)$, $\lambda^\ell \geq 0$, $\ell = 1,\dots,S$ such that

(1) $\sum_\ell \lambda^\ell = 1$, $\sum_\ell \lambda^\ell q^\ell = q$

(2) $v(p,q) = \sum_\ell \lambda^\ell v(p,q^\ell)$

(3) $v(p,q^\ell) = u(p,q^\ell) \qquad \forall \ell = 1,\dots,S$

(4) $v_p < u_p$ on Int Co$\{q^\ell : \ell = 1,\dots,S\}$

Proof:

Due to the second functional equation and the continuity of u v satisfies

$$v(p,q) = \min \left\{ \sum_{\ell=1}^{S} \lambda^\ell \max\{u,v\}(p,q^\ell) : \sum_\ell \lambda^\ell q^\ell = q, \sum_\ell \lambda^\ell = 1, \lambda^\ell \geq 0 \right\}$$

Let $\{\bar{q}^\ell : \ell = 1,\dots,S\}$ be any set of probabilities achieving the minimum:

$$\sum_\ell \lambda^\ell v(p,\bar{q}^\ell) \geq v(p,q) = \sum_\ell \lambda^\ell \max\{u,v\}(p,\bar{q}^\ell)$$

$$\Rightarrow \sum_\ell \lambda^\ell v(p,\bar{q}^\ell) = v(p,q) \text{ and}$$

$$v(p,\bar{q}^\ell) \geq u(p,\bar{q}^\ell)$$

Assume w.l.o.g. that $\lambda^{\ell} > 0 \quad \forall\ \ell$, i.e. q is located in the relative interior of $\bar{C} = Co\{\bar{q}^{\ell} : \ell = 1,...,S\}$. It follows that v_p is affine on $\bar{C}$, i.e. any set of probabilities contained in $\bar{C}$ satisfying (1) for suitable λ^{ℓ} also satisfies (2). Now choose $q^{\ell} \in \bar{C}$ and $\lambda^{\ell} \geq 0$ fulfilling (1) and (2) as well as the next two properties:

$$v(p,q^{\ell}) \geq u(p,q^{\ell})$$

and the volume of $C = Co\{q^{\ell} : \ell = 1,...,S\}$ is minimal among all polyhedra satisfying the three properties above. Obviously (3) and (4) are valid as well.

□

Define

$$W(q) = \{g \in \mathbb{R}^R : p \cdot g \geq v(p,q) \quad \forall\ p \in \Delta(R)\}$$

$$V(q) = \{g \in W(q) : \exists\ p \in \Delta(R) : p \cdot g = v(p,q)\}$$

$$V(p,q) = \{g \in V(q) : p \cdot g = v(p,q)\}$$

W(q) is called the approachable set (of q).

In the case of lack of information on one side the argument q is omitted. In figure 3 of example 1

W is the shaded area

V is the boundary of W

V(p) is a half line for $p = 0,1$

line segment for $p = \frac{1}{4}, \frac{3}{4}$

singleton for all other p's

The next Lemma follows immediately from the definitions of V and W:

Lemma 3.3: Let $g \in V(p,q)$. The p represents a supporting hyperplane of the convex set W(q) at point g.

Proposition 3.4:

Let $p \in \Delta(R)$; $q, q^{\ell} \in \Delta(s)$; $\lambda^{\ell} \geq 0$ or $\ell = 1, \ldots, S$
such that (1),(2) and (4) of proposition 3.2 satisfied. It follows that

$$V(p,q) = \sum_{\ell} \lambda^{\ell} V(p,q^{\ell})$$

Proof:

a) Let $g^{\ell} \in V(p,q^{\ell})$. Then

$$p \cdot (\sum_{\ell} \lambda^{\ell} g^{\ell}) = \sum_{\ell} \lambda^{\ell} p\, g^{\ell} = \sum_{\ell} \lambda^{\ell} v(p,q^{\ell}) = v(p,q)$$

$$p' \cdot (\sum_{\ell} \lambda^{\ell} g^{\ell}) = \sum_{\ell} \lambda^{\ell} p' g^{\ell} \geq \sum_{\ell} \lambda^{\ell} v(p',q^{\ell}) \geq v(p',q) \qquad \forall\, p' \in \Delta(R)$$

$$\Rightarrow \quad \sum_{\ell} \lambda^{\ell} g^{\ell} \in V(p,q)$$

b) Suppose there exists $g \in V(p,q) \setminus \sum_{\ell} \lambda^{\ell} V(p,q^{\ell})$.

Case 1: $p \in \text{Int } \Delta(R)$, i.e. $p(r) > 0 \quad \forall\, r \in R$

In this case there are a vector $x \in \mathbb{R}^R$ and a real number $\epsilon > 0$ such that

$$g \cdot x < g' \cdot x - \epsilon \qquad \forall\, g' \in K_{\epsilon} + \sum_{\ell} \lambda^{\ell} V(p,q^{\ell})$$

with $K_{\epsilon} = \{y \in \mathbb{R}^R : |y| \leq \epsilon\}$.

($K_{\epsilon} + \sum_{\ell} \lambda^{\ell} V(p,q^{\ell})$ is a closed and convex set since it is the sum of closed convex sets).

Let $q_n^{\ell} = \frac{n-1}{n} q^{\ell} + \frac{1}{n} q$, $n \in \mathbb{N}$

Since q_n^{ℓ} is situated in the interior of $\text{Co}\{q_n^{\ell} : \ell = 1, \ldots, S\}$ we deduce that

$v_p < u_p$ on $C_n = \text{Co}\,\{q_n^{\ell} : \ell = 1, \ldots, S\}$ and

⊛ $$v(p,q) = \sum_{\ell} \lambda^{\ell} v(p,q_n^{\ell}) \qquad \forall\, n \in \mathbb{N}$$

Now we don't regard v_q as a function on $\Delta(R)$ but we consider the linear homogeneous extension on $\mathbb{R}_+^R$:

Choose $p_n \in \mathrm{Co}\,\{p,p+x\}$ such that

$$p_n \neq p,\ p_n \in \mathbb{R}_+^R \quad \forall\, n \in N$$

$$\lim_{n\to\infty} p_n = p$$

$$v_{p_n} < u_{p_n} \text{ on } \mathrm{Co}\{q_n^{\ell} : \ell = 1,\dots,S\}$$

It follows that

$$\circledast\circledast \qquad v(p_n,q) = \sum_{\ell} \lambda^{\ell}\, v(p_n,q_n^{\ell}) \qquad \forall\, n \in \mathbb{N}.$$

According to corollary 24.5.1 of Rockafellar [70] there exists $n_0 \in \mathbb{N}$ such that

$$V(p,q_n^{\ell}) \subset V(p,q^{\ell}) + K_{\epsilon} \qquad \forall\, n > n_0,\ \ell = 1,\dots,S$$

Theorem 23.4 of Rockafellar [70] says that the directional derivative $\frac{d}{dx} v_q(p)$ is given by

$$\frac{d}{dx} v_q(p) = \inf\{g' : x \cdot g' \in V(p,q)\}$$

Consequently

$$\frac{d}{dx} v_q(p) = \sum_{\ell} \lambda^{\ell} \frac{d}{dx} v_{q_n^{\ell}}(p) - \epsilon \qquad \forall\, n > n_0.$$

Together with $\circledast$ this implies that for some $n_1 > n_0$

$$v(p_{n_1},p) < \sum_{\ell} \lambda^{\ell}\, v(p,q_{n_1}^{\ell})$$

which yields a contradiction to $\circledast\circledast$.

Case 2: $p(r) = 0$ for at least one $r \in R$ follows from case 1 and theorem 25.6 of Rockafellar [70] .

□

4. Definition of player 2's strategy

In this section a strategy for player 2 is constructed that guarantees a payoff close to $v(p_0,q_0)$ in a sufficiently long game. Several functions resp. random variables are needed to describe player 2's conduct.

q_T denotes the conditional probability on S given the history up to stage T for the strategy τ we are going to define. (Of course, q_T is independent of player 1's strategy.)

For every $q \in \Delta(S)$ define

$$A_T(q): \quad H^N \to \mathbb{R}^R$$

$$h^N \mapsto \Big(\sum_{s\in S} q(s)\, A^{rs}(h_T)\Big)_{r\in R}$$

$$A_T = A_T(q_T)$$

$$\gamma_T = \frac{1}{N+1}\sum_{t=1}^{T} A_t(q_T)$$

$$\gamma_{T+\frac{1}{2}} = \frac{1}{N+1}\sum_{t=1}^{T} A_t(q_{T+1})$$

The next definitions involve the sequence η_T which has to be defined together with the strategy itself stage after stage:

$$\delta_T \quad = \frac{N+1}{N+1-T}\,(\eta_T - \gamma_T)$$

$$\delta_{T+\frac{1}{2}} = \frac{N+1}{N+1-T}\,(\eta_{T+1} - \gamma_{T+\frac{1}{2}})$$

$$\text{dist:} \quad \mathbb{R}^R \times \Delta(S) \longrightarrow W(q)$$

$$(\delta\ ,\ q) \quad \mapsto \arg \min_{g \in W(q)} |\delta - g|$$

Abbreviate $G_T = \text{dist}\ (\delta_T, q_T)$, $G_{T+\frac{1}{2}} = \text{dist}\ (\delta_{T+\frac{1}{2}}, q_{T+1})$

If $G_T \notin W(q_T)$, define

$$p_T' = \frac{G_T - \delta_T}{\sum_r G_T(r) - \delta_T(r)} \quad \in \Delta(R)$$

(Observe that the components of $G_T - \delta_T$ are all positive.)

Since G_T minimizes the euclidean distance between δ_T and $W(q_T)$ and p_T' is parallel to $G_T - \delta_T$ we have $G_T \in V(p_T', q_T)$.

The sequence η_T is defined together with the strategy $\tau = (\tau_1, \ldots, \tau_N)$. It starts with $\eta_0 = \delta_0$. Suppose that everything is defined up to stage T:

If $G_T \in W(q_T)$, i.e. if p_T' is undefined player 2 may choose any distribution on J independent of s.

If $G_T \in W(q_T)$ we distinguish two cases:

Case 1: $\quad v(p_T', q_T) \geq u(p_T', q_T)$

Choose and equilibrium strategy y for the minimizing player in the matrix game $A(p_T', q_T)$ and define

$$\tau_{T+1}(s, h^T; j) = y(j)$$

$$\eta_{T+1} = \eta_T$$

Case 2: $v(p_T^{\cdot},q_T) \geq u(p_T^{\cdot},q_T)$

There exist $q^\ell \in \Delta(s)$, $\lambda^\ell \geq 0$ satisfying (1) – (4) of proposition 3.2 Let y^ℓ be an equilibrium strategy for the minimizing player in the matrix game $A(p_T^{\cdot},q^\ell)$:

$$\tau_{T+1}(s,h^T;j) = \sum_\ell \lambda^\ell \frac{q^\ell(s)}{q_T(s)} y^\ell(j)$$

For every T–stage history h^T there are

$$g^j \in V(p_T^{\cdot},q_{T+1}(j)),\ j \in J$$

such that

$$G_T = \sum_j (\sum_\ell \lambda^\ell y^\ell(j))\, g^j$$

(see Lemma 4.4)

Define

$$\eta_{T+1}(j) = \eta_T + \gamma_{T+\frac{1}{2}}(j) - \gamma_T + \frac{N+1-T}{N+1}(g^j - G_T).$$

Formally the random variables and the strategy are not yet well defined since the definition sometimes permits a choice between several possibilities. One could in these cases always select the lexicographic minimal one among all admissible choices.

Interpretation of the random variables γ_T, η_T, δ_T and $p_T^{\cdot}$:

γ_T reflects an asymmetry in our considerations. Constructing a strategy that guarantees player 2 a certain payoff one must always take into account the possibility that player 1 is able to calculate the posteriors q_T on S after each stage. On the other hand computing posteriors cannot be part of player 2's strategy, he is forced to control player 1's payoff for every possible type $r \in R$. γ_T asymptotically describes the vector payoff that occurred in the game up to stage T from player 2's point of view. For technical reasons to be discussed later we consider a horizon that is one stage longer than the game itself. The updating of γ_T (and δ_T) is done in half steps: γ_T describes the vector payoff up to stage T whereas $\gamma_{T+\frac{1}{2}}$ already includes the posteriors of the following stage.

η_T is the vector payoff player 2 tries to achieve in the course of the game. The first one, η_0, is chosen arbitrarily from the set of supporting hyperplanes of v_{q_0} at p_0. During the game η_T has to be modified from stage to stage according to player 2's information release. The difference

$$\eta_{T+1} - \eta_T = \gamma_{T+\frac{1}{2}} - \gamma_T + \frac{N+1-T}{N+1}(g^j - G_T)$$

depends only on the development of q_T. The first part of the difference, $(\gamma_{T+\frac{1}{2}} - \gamma_T)$, results from the altered evaluation of the payoffs up to stage T. To understand the second part the meaning of δ_T must be explained.

The definition $\delta_T = \frac{N+1}{N+1-T}(\eta_T - \gamma_T)$

can be transformed to $\eta_T = \gamma_T + \frac{N+1-T}{N+1}\delta_T$.

If η_T represents player 2's final objective δ_T can be described as his temporary goal. If after T stages the vector payoff γ_T occurs and if player 2 starts a new game from stage T + 1 on (with probability q_T on his types) achieving the vector payoff δ_T, which is weighted by $\frac{N+1-T}{N+1}$ since we consider an additional stage, then he will get the vector payoff η_T in the end.

Generally δ_T is no realistic objective because it is not contained in the approachable set. It appears sensible to use $G_T = \text{dist}(\delta_T, q_T)$ instead. The difference $G_T - \delta_T$ describes the loss (if $\delta_T \notin W(q_T)$) resp. gain (if $\delta_T \in W(q_T)$) player 2 has attained up to stage T w.r.t. his final objective η_T. The total payoff produced by the strategy will be estimated by this difference. Therefore it is desirable to have δ_N defined which can only be done by considering N + 1 stages.

$p_T^\cdot$ serves as a substitute for the true posterior probability p_T player 2 does not know. If $\delta_T \in V(q_T)$ it can be chosen arbitrarily from the set of supporting hyperplanes at δ_T, if $\delta_T \notin V(q_T)$ it is determined uniquely as the probability vector that is parallel to $G_T - \delta_T$ and satisfies $G_T \in V(p_T^\cdot, q_T)$ as well.

Combining the equations

$$\delta_{T+\frac{1}{2}}(j) = \frac{N+1}{N+1-T}\,(\eta_{T+1}(j) - \gamma_{T+\frac{1}{2}}(j))$$

$$\eta_{T+1}(j) = \eta_T + \gamma_{T+\frac{1}{2}}(j) - \gamma_T + \frac{N+1-T}{N+1}\,(g^j - G_T)$$

it follows that

$$g^j - \delta_{T+\frac{1}{2}}(j) = G_T - \delta_T$$

g^j was defined to be an element of $V(p_T^\cdot, q_{T+1}(j))$. Since $p_T^\cdot$ is parallel to $G_T - \delta_T$ we deduce that $g^j - \delta_{T+\frac{1}{2}}(j)$ defines a supporting hyperplane of $W(q_{T+1}(j))$ at g^j resp. that $g^j = \text{dist}\,(\delta_{T+\frac{1}{2}}(j), q_{T+1}(j)) = G_{T+\frac{1}{2}}(j)$.

The second difference in the definition of $\eta_{T+1}(j)$ now reads $\frac{N+1}{N+1-T}\,(G_{T+\frac{1}{2}}(j) - G_T)$, it regards the development of player 2s "realistic objective" according to his information release. η_{T+1} is designed in such a way that merely the use of his private information does not change his temporary balance, i.e.

Lemma 4.1: $$G_{T+\frac{1}{2}} - \delta_{T+\frac{1}{2}} = G_T - \delta_T$$

Lemma 4.2: $E(\eta_T | h^{T-1}) = \eta_{T-1} \quad \forall\, T=1,\dots,N,\ \sigma \in \Sigma$

τ as defined above

(E denotes the expectation operator w.r.t. $P^{p_0 q_0}_{(\sigma,\tau)}$).

Lemma 4.3: Let p_T be the conditional probability on R given H^T.

p_T and η_T are uncorrelated w.r.t. $P^{p_0 q_0}_{(\sigma,\tau)} \quad \forall\, T=1,\dots,N,\ \sigma \in \Sigma$

Proof:

It is sufficient to show that

$$E(p_T \cdot \eta_T) = E(p_{T-1} \cdot \eta_{T-1})$$

$$E(p_T \cdot \eta_T)$$

$$= E(E(p_T \cdot \eta_T | h^{T-1}))$$

$$= E(E(p_T | h^{T-1}) \cdot E(\eta_T | h^{T-1}))$$

$$= E(p_{T-1} \cdot \eta_{T-1})$$

□

We still have to show that the definition of η_{T+1} in case 2 is admissible. This follows from proposition 3.4 if conditions (1), (2) and (4) are satisfied. (1) is valid due to the martingale property of the posterior probabilities q_T, (2) and (4) are consequences of

Lemma 4.4: $q_{T+1}(j) \in Co\{q^\ell : \ell=1,\dots S\} \quad \forall\, j \in J.$

Proof: (The argument h^T is always omitted)

$$q_{T+1}(j)\,(s)$$

$$= \frac{q_T(s)\ \ \tau_{T+1}(s;j)}{\sum_{s'} q_T(s')\ \tau_{T+1}(s';j)}$$

$$= \frac{q_T(s)\ \sum_{\ell} \lambda^{\ell} \frac{q^{\ell}(s)}{q_T(s)}\, y^{\ell}(j)}{\sum_{s'} \sum_{\ell} \lambda^{\ell}\ q^{\ell}(s')\ y^{\ell}(j)}$$

$$= \frac{\sum_{\ell} \lambda^{\ell}\ y^{\ell}(j)\ q^{\ell}(s)}{\sum_{\ell} \lambda^{\ell}\ y^{\ell}(j)}$$

Thus q_{T+1} is a convex combination of the q^{ℓ}.

□

Lemma 4.5: The conditional distributions on R resp. S are independent given any h^T, $T = 0,\ldots,N$.

5. Main Theorem:

Let σ be an arbitrary strategy for player 1 and τ the strategy for player 2 defined above.

$$\alpha(\sigma,\tau) \leq v(p_0,q_0) + \frac{\sqrt{N}\ \sqrt{R+1}}{N} 2M$$

$$(M = \max_{i,j,r,s} |A^{r,s}(i,j)|)$$

Proof:

At first the difference $\alpha(\sigma,\tau) - v(p_0,q_0)$ must be represented using the difference $G_N - \delta_N$.

E denotes the expectation w.r.t. the marginal distribution of $P^{p_0 q_0}_{(\sigma,\tau)}$ on H^N.

$$\alpha(\sigma,\tau) - v(p_0,q_0)$$

$$= E(p_N \frac{1}{N} \sum_{T=1}^{N} A_T) - p_0\ \eta_0 \qquad \text{(Lemma 4.5)}$$

$$= E(p_N \frac{N+1}{N} \gamma_N) - E(p_N\ \eta_N) \qquad \text{(Lemma 4.2, 4.3)}$$

$$= E(p_N(\frac{N+1}{N} \gamma_N - \eta_N)$$

$$= \frac{1}{N} E(p_N(\eta_N - \delta_N))$$

$$= \frac{1}{N} (E(p_N(\eta_N - G_N)) + E(p_N(G_N - \eta_N)))$$

$$\leq \frac{M}{N} + \frac{1}{N} E(|G_N - \delta_N|)$$

$$\leq \frac{M}{N} + \frac{1}{N} (E((G_N - \delta_N)^2))^{\frac{1}{2}}$$

Next we show by induction that

$$E((G_T - \delta_T)^2) \le \frac{T}{(N+1-T)^2} R M^2$$

T = 0: satisfied by definition

$T \to T+1$:

$$E((G_{T+1} - \delta_{T+1})^2)$$

$$\le E((G_{T+\frac{1}{2}} - \delta_{T+1})^2)$$

$$= E((G_{T+\frac{1}{2}} - \frac{N+1}{N-T}(\eta_{T+1} - \gamma_{T+1}))^2)$$

$$= E((G_{T+\frac{1}{2}} - \frac{N+1}{N-T}(\eta_{T+1} - \gamma_{T+\frac{1}{2}}) + \frac{1}{N-T} A_{T+1})^2)$$

$$= E((G_{T+\frac{1}{2}} - \frac{N+1-T}{N-T} \delta_{T+\frac{1}{2}} + \frac{1}{N-T} A_{T+1})^2)$$

$$= E((\frac{N+1-T}{N-T}(G_{T+\frac{1}{2}} - \delta_{T+\frac{1}{2}}) + \frac{1}{N-T}(A_{T+1} - G_{T+\frac{1}{2}}))^2)$$

$$= E((\frac{N+1-T}{N-T}(G_T - \delta_T) + \frac{1}{N-T}(A_{T+1} - G_{T+\frac{1}{2}}))^2)$$

$$= \left[\frac{N+1-T}{N-T}\right]^2 E((G_T - \delta_T)^2)$$

$$+ \frac{1}{(N-T)^2} E((A_{T+1} - G_{T+\frac{1}{2}})^2)$$

$$+ \frac{N+1-T}{(N-T)^2} E((G_T - \delta_T)(A_{T+1} - G_{T+\frac{1}{2}}))$$

The third term will be estimated separately. In order to show that it is not positive it is sufficient to show that

$$E(p_T^{\,j}\,(A_{T+1} - G_{T+\frac{1}{2}})) \leq 0 \quad \text{resp.}$$

$$E(p_T^{\,j}\,A_{T+1} - v(p_T^{\,j}, q_{T+1})) \leq 0$$

(If $p_T^{\,j}$ is undefined we have $\delta_T \in W(q_T)$ resp. $G_T = \delta_T$ and there is nothing to prove.)

For notational convenience we consider conditional expectations given h^T, omitting the argument h^T during the computation. Let y resp. y^ℓ, λ^ℓ be as in the definition of the strategy and let

$$x(i) \;=\; \sum_r p_T(r)\,\sigma_{T+1}(r, h^T; i).$$

Case 1: $\qquad v(p_T^{\,j}, q_T) \geq u(p_T^{\,j}, q_T)$

$$E(p_T^{\,j}\,A_{T+1} \,|\, h^T)$$

$$= \sum_i x(i) \sum_j y(j)\;\; A(p_T^{\,j}, q_{T+1})\,(i,j)$$

$$= \sum_{i,j} x(i)\,y(j)\;\; A(p_T^{\,j}, q_T)\,(i,j)$$

$$\leq u(p_T^{\,j}, q_T)$$

$$\leq v(p_T^{\,j}, q_T)$$

$$= v(p_T^{\,j}, q_{T+1})$$

Case 2: $v(p_T', q_T) < u(p_T', q_T)$

$$E(p_T' A_{T+1} | h^T)$$

$$= \sum_i x(i) \sum_j \sum_s q_T(s)\ \tau_{T+1}(s;j)\ A(p_T', q_{T+1}(j))\ (i,j)$$

$$= \sum_i x(i) \sum_j \sum_s q_T(s) \sum_\ell \frac{q^\ell(s)}{q_T(s)}\ y^\ell(j)\ A(p_T', q_{T+1}(j))\ (i,j)$$

$$= \sum_i x(i) \sum_j \sum_\ell \lambda^\ell y^\ell(j)\ A(p_T', q_{T+1}(j))\ (i,j)$$

$$= \sum_i x(i) \sum_j \sum_\ell \lambda^\ell y^\ell(j) \sum_{r,s} p_T'(r)\ q_{T+1}(j)(s)\ A^{r,s}(i,j)$$

$$= \sum_i x(i) \sum_j \sum_{r,s} p_T'(r) \sum_\ell \lambda^\ell y^\ell(j)\ q^\ell(s)\ A^{r,s}(i,j)$$

(see proof of Lemma 4.4)

$$= \sum_\ell \lambda^\ell \sum_{i,j} x(i)\ y^\ell(j) \sum_{r,s} p_T'(r)\ q^\ell(s)\ A^{r,s}(i,j)$$

$$= \sum_\ell \lambda^\ell \sum_{i,j} x(i)\ y^\ell(j)\ A(p_T', q^\ell)\ (i,j)$$

$$\leq \sum_\ell \lambda^\ell\ u(p_T', q^\ell)$$

$$= \sum_\ell \lambda^\ell\ v(p_T', q^\ell)$$

$$= v(p_T', q_T)$$

$$= E(v(p_T', q_{T+1}) | h^T)$$

We continue our proof by induction:

$$E((G_{T+1} - \delta_{T+1})^2)$$

$$\leq \left[\frac{N+1-T}{N-T}\right]^2 E((G_T - \delta_T)^2) + \frac{1}{(N-T)^2} E((A_{T+1} - G_{T+\frac{1}{2}})^2)$$

$$\leq \left[\frac{N+1-T}{N-T}\right]^2 \frac{T}{(N+1-T)^2} RM^2 + \frac{1}{(N-T)^2} RM^2$$

$$= \frac{T+1}{(N-T)^2} RM^2$$

Consequently $E((G_N - \delta_N)^2) \leq NRM^2$ and the assertion follows.

□

Example 2: (cf. Mertens and Zamir [71/72])

$$A^{11} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{bmatrix} \qquad A^{12} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A^{21} = \begin{bmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad A^{22} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$A(p,q) = \begin{bmatrix} p-q & q-p & p-q & q-p \\ 1-p-q & p+q-1 & p+q-1 & 1-p-q \end{bmatrix}$$

The value of the NR–game u(p,q) is given by figure 4

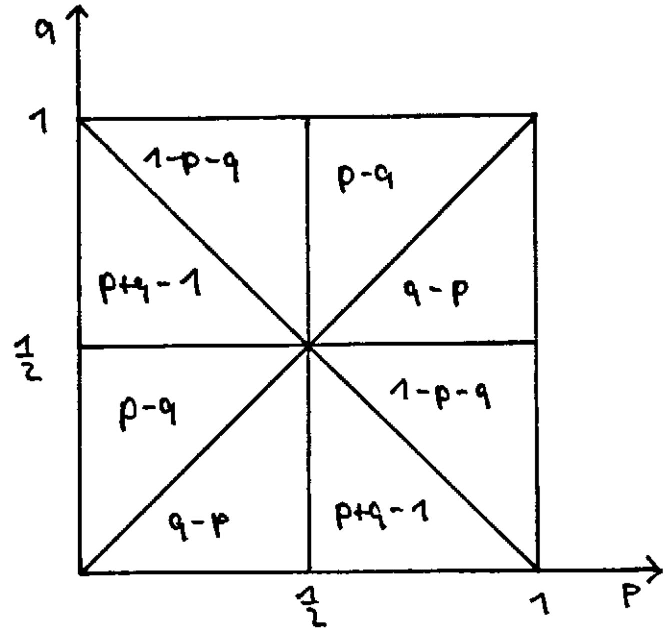

fig. 4

and v(p,q) is represented in figure 5

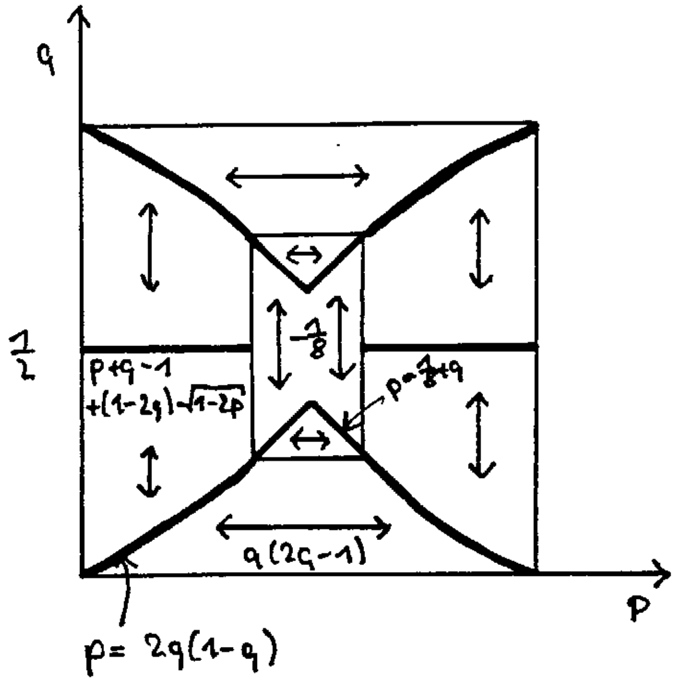

fig. 5

The thick lines show where $v = u$ and the arrows denote the directions of affinity of v. (According to Mertens and Zamir). Explicit definition of v:

$0 \le q \le \frac{1}{4}$:

$$v(p,q) = \begin{cases} p+q-1+(1-2q)\sqrt{1-2p}, & p \le 2q(1-q) \\ q(2q-1), & 2q(1-q) \le p \le 1-2q(1-q) \\ q-p+(1-2q)\sqrt{2p-1}, & p \ge 1-2q(1-q) \end{cases}$$

$\frac{1}{4} \le q \le \frac{1}{2}$:

$$v(p,q) = \begin{cases} p+q-1+(1-2q)\sqrt{1-2p}, & p \le \frac{3}{8} \\ -\frac{1}{8}, & \frac{3}{8} \le p \le \frac{5}{8} \\ q-p+(1-2q)\sqrt{2p-1}, & p \ge \frac{5}{8} \end{cases}$$

$q \ge \frac{1}{2}$: v is symmetric w.r.t. $q = \frac{1}{2}$ (and $p = \frac{1}{2}$).

In order to determine the approachable sets, we must compute the derivative of v w.r.t. the first variable.

$0 \le q \le \frac{1}{4}$:

$$\frac{d}{dp} v(p,q) = \begin{cases} 1 - \frac{1-2q}{\sqrt{1-2p}}, & p < 2q(1-q) \\ 0, & 2q(1-q) < p < 1-2q(1-q) \\ -1 + \frac{1-2q}{\sqrt{2p-1}}, & p > 1-2q(1-q) \end{cases}$$

$\frac{1}{4} \le q \le \frac{1}{2}$:

$$\frac{d}{dp} v(p,q) = \begin{cases} 1 - \frac{1-2q}{\sqrt{1-2p}}, & p < \frac{3}{8} \\ 0, & \frac{3}{8} < p < \frac{5}{8} \\ -1 + \frac{1-2q}{\sqrt{2p-1}}, & p > \frac{5}{8} \end{cases}$$

If v_q is differentiable at p, V(p,q) is a singleton, $V(p,q) = \{(g_1,g_2)\}$ and $g = (g_1,g_2)$ is determined by the equations:

$$g_1 - g_2 = \frac{d}{dp} v(p,q)$$

$$p\, g_1 + (1-p)\, g_2 = v(p,q)$$

$0 \le q \le \frac{1}{4}$:

$$V(p,q) = \begin{cases} \left[q-(1-2q)\frac{p}{\sqrt{1-2p}},\ q-1+(1-2q)\frac{1-p}{\sqrt{1-2p}} \right], & p < 2q(1-q) \\ \left[q(2q-1),\ q(2q-1) \right] & , 2q(1-q) < p < 1-2q(1-q) \\ \left[-q+(1-2q)\frac{1-p}{\sqrt{2p-1}}, 1-q-(1-2q)\frac{p}{\sqrt{2p-1}} \right], & p > 1-2q(1-q) \end{cases}$$

$\frac{1}{4} \le q \le \frac{1}{2}$:

$$V(p,q) = \begin{cases} \left[q-(1-2q)\frac{p}{\sqrt{1-2p}},\ q-1+(1-2q)\frac{1-p}{\sqrt{1-2p}} \right], & p < \frac{3}{8} \\ \left[-\frac{1}{8}, -\frac{1}{8} \right] & , \frac{3}{8} < p < \frac{5}{8} \\ \left[-q+(1-2q)\frac{1-p}{\sqrt{2p-1}}, 1-q-(1-2q)\frac{p}{\sqrt{2p-1}} \right], & p > \frac{5}{8} \end{cases}$$

The next figure represents the approachable sets for q = 0, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{1}{2}$.

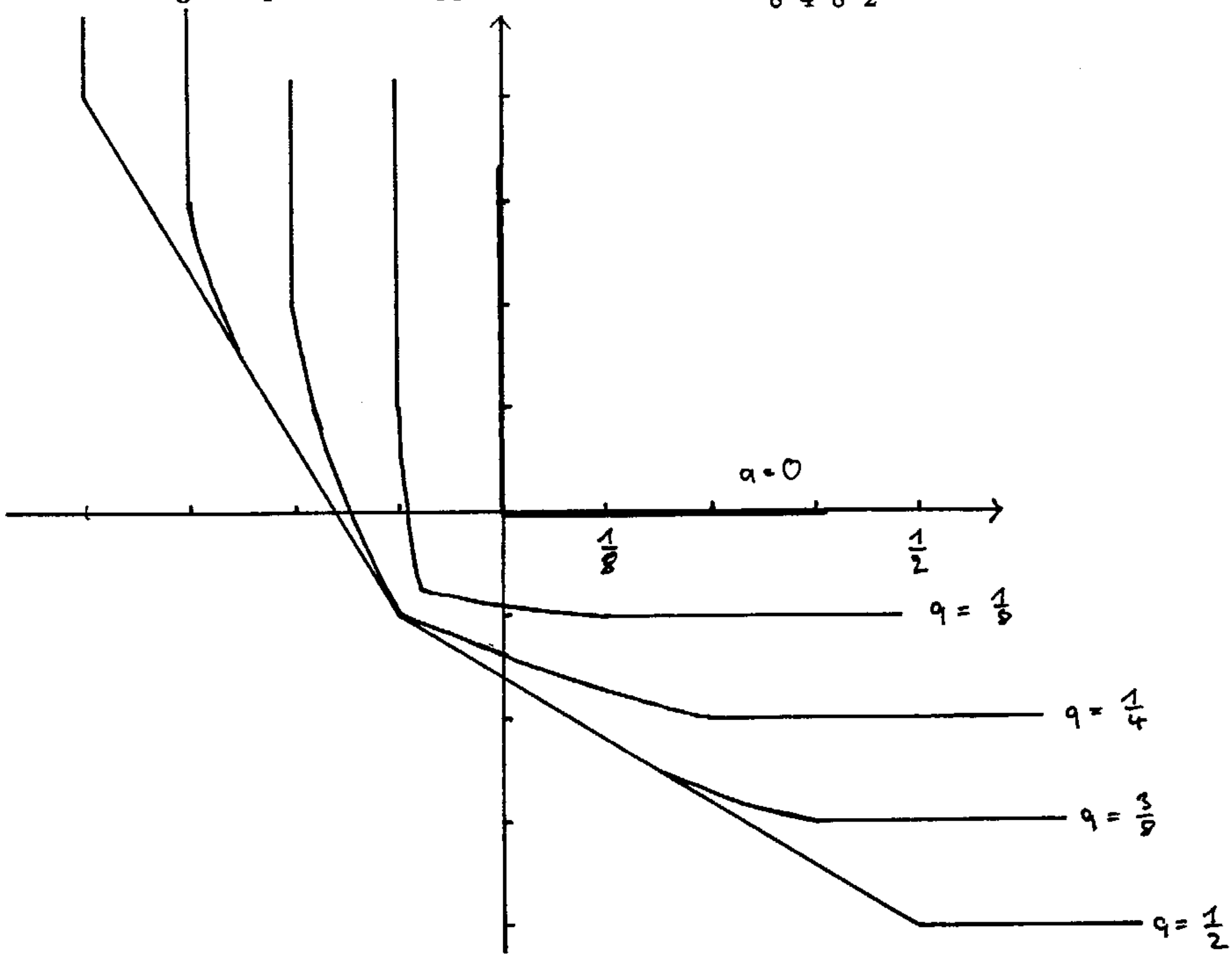

fig. 6

□

6. The Recursive Game

The random variable $p_T^{\prime}$ describes the amount of information player 2 can deduce from player 1's behaviour without running the risk of being cheated. He uses $p_T^{\prime}$ in the computation of his strategy because the rules of the game don't enable him to calculate the actual posterior probabilities on player 1's types. We shall now see how player 2's strategy works if by some change of the rules of the game he is able to employ p_T instead of $p_T^{\prime}$.

The description of the original model in section 1 applies except for point 3). In addition to the actions they also learn the posteriors on the opponent's types. A strategy for player 1 is now given by a sequence of stochastic kernels $\sigma = (\sigma_1, \ldots, \sigma_N)$

$$\sigma_T \mid R \times H^{T-1} \times \Delta(S)^{T-1} \Rightarrow I.$$

(analogously for player 2).

In the definition of player 2's strategy $p_T^{\prime}$ is now replaced by p_T, thus making his actions dependent on the posterior probability of the foregoing stage only. Consequently $p_T^{\prime}$ no longer determines G_T, but G_T is selected out of $V(p_T, q_T)$.

The random variables γ_T, δ_T and η_T become superfluous.

Theorem 6.1: Let σ be an arbitrary strategy for player 1 and τ the strategy for player 2 defined above.

$$\alpha(\sigma,\tau) \leq v(p_0, q_0) + \frac{M}{\sqrt{N}} \sum_r \sqrt{p_0(r)(1-p_0(r))}$$

Proof:

$$\alpha(\sigma,\tau)$$

$$= E(p_N \frac{1}{N} \sum_{T=1}^{N} A_T)$$

$$= \frac{1}{N} E(p_N(\sum_{T=1}^{N} (A_T - G_{T-\frac{1}{2}}) + \sum_{T=1}^{N} G_{T-\frac{1}{2}}))$$

$$= \frac{1}{N} E(\sum_{T=1}^{N} p_T(A_T - G_{T-\frac{1}{2}}) + \sum_{T=1}^{N} p_T G_{T-\frac{1}{2}})$$

$$\leq \frac{1}{N} E(\sum_{T=1}^{N} (p_T - p_{T-1})(A_T - G_{T-\frac{1}{2}}) + \sum_{T=1}^{N} p_{T-1} G_{T-1})$$

(The inequality is due to the change of the first term, see proof of the main theorem. As for the second term note that p_T and $G_{T-\frac{1}{2}}$ are conditionally independent given h^{T-1}.)

$$= \frac{1}{N} E(\sum_{T=1}^{N} (p_T - p_{T-1})A_T + \sum_{T=0}^{N-1} p_T G_T)$$

$$\leq \frac{M}{N} \sum_r E(\sum_{T=1}^{N} |p_T(r) - p_{T-1}(r)|) + \frac{1}{N} E(\sum_{T=0}^{N-1} v(p_T,q_T))$$

$$E(\sum_{T=0}^{N-1} v(p_T,q_T))$$

$$\leq v(p_0,q_0) + E(\sum_{T=1}^{N-1} v(p_{T-1},q_T))$$

(due to I-concavity of v)

$$= v(p_0,q_0) + E(\sum_{T=1}^{N-1} v(p_{T-1},q_{T-1}))$$

(see definition of player 2's strategy and Lemma 4.4: player 2 only uses his information at stage T if $v_{p_{T-1}}$ is locally affine at q_{T-1}, Lemma 4.4. shows that the T–stage posteriors are located within this region).

$$= 2\, v(p_0,q_0) + E\Big(\sum_{T=1}^{N-2} v(p_T,q_T)\Big)$$

$$= \cdots$$

$$\vdots$$

$$= N\, v(p_0,q_0)$$

Using the martingale properties of the sequence p_T one can show that

$$E\Big(\sum_{T=1}^{N} |p_T(r) - p_{T-1}(r)|\Big) \le \sqrt{N}\,\sqrt{p_0(r)(1-p_0(r))}$$

(cf. Mertens and Zamir [71/72])

and the result follows.

□

Remark: If player 1 uses his private information in the same way as placer 2 does (especially if he employs the analog strategy), the only inequality in the estimation of the second term becomes an equality, i.e.

$$\frac{1}{N} E\Big(\sum_{T=0}^{N-1} v(p_T,q_T)\Big) = v(p_0,q_0).$$

The proof shows that the players can only achieve a payoff exceeding the limiting value $v(p_0,q_0)$ if they make use of their private information.

References

[56] Blackwell, D.: An Analogue of the Minimax Theorem for Vector Payoffs, Pacific Journal of Mathematics, Vol. 6, 1–8.

[71/72] Mertens, J.–F. and S. Zamir: The Value of Two–Person Zero–Sum Repeated Games with Lack of Information on Both sides, International Journal of Game Theory 1, 36–64.

[77] Mertens, J.–F. and S. Zamir: A Duality Theorem on a Pair of Simultaneous Functional Equations, Journal of Mathematical Analysis and Applications 60, 550–558.

[70] Rockafellar, R.T.: Convex Analysis, Princeton University Press.

[80] Sorin, S.: An Introduction to Two–Person Zero–Sum Repeated Games with Incomplete Information, Technical Report No. 312, Institute for Mathematical Studies in the Social Sciences, Stanford University.

[74] Zamir, S.: Convexity and Repeated Games, Analyse Convexe et ses Applications, J.–P.Aubin (ed.), Lecture Notes in Economics and Mathematical Systems 102, Springer Verlag, 216–231.