International Journal of Game Theory (1992) 20: 377–392

Asymptotically Optimal Strategies in Repeated Games with Incomplete Information

M. Heuer[1]

Abstract: It is known that for repeated zero-sum games with incomplete information the limit of the values of the N-stage game exists as N tends to infinity. In this paper strategies are constructed that guarantee in the N-stage game the limit of values up to an error term $\frac{K}{\sqrt{N}}$.

0 Introduction

In the model considered here two players have to play a matrix game repeatedly. The payoff matrix is determined once and forever by a chance move at the beginning of the game. Both players receive partial information about the outcome of the chance move, during the game they only learn the opponent's choice of action.

There are two main results concerning the n-stage value $v_n(p)$ of a game with prior distribution p on the set of possible payoff matrices:

- $\lim_{n\to\infty} v_n$ exists and satisfies a pair of functional equations [Mertens and Zamir 1971/72]
- the pair of functional equations has a unique solution [Sorin 1984].

The latter result is now used to construct strategies that asymptotically guarantee the limit of values. This allows to recover the first result which had been derived without exhibiting "good" strategies.

1 The Model

A finitely repeated two-person zero-sum game with incomplete information on both sides is given by the following data:

- finite sets R and S
  (sets of types of player 1 and 2)

[1] Dr. Martin Heuer, Frankenweg 8, 4800 Bielefeld 11

0020-7276/92/4/377-392 $ 2.50

- for every pair $(r, s) \in R \times S$ a $m \times n$-matrix $A^{r,s}$
  (payoff matrices)
- a probability distribution p on $R \times S$ with $p(r) > 0$, $p(s) > 0 \ \forall\ r, s$
- a natural number N
  (the number of stages)

The game is played according to the rules below:

- At stage zero a pair of types (r, s) is selected according to p. Both players know p, but they only learn their own type.
- At each stage $t = 1, \ldots, N$ the players independently and simultaneously choose parameters

  $$i_t \in I = \{1, \ldots, m\}, \quad j_t \in J = \{1, \ldots, n\}.$$

- Afterwards they are told the opponent's choice of action.
- Both players have perfect recall.
- After stage N player 1 receives from player 2 the amount $\frac{1}{N} \sum_{t=1}^{N} A^{r,s}(i_t, j_t)$.

The product of the action spaces is abbreviated by $H = I \times J$. In the following the identities

$$h = (i, j), \quad h_t = (i_t, j_t) \quad \text{and} \quad h^t = (h_1, \ldots, h_t)$$

are always tacitly assumed.

A behavior strategy of player 1 consists of a sequence

$$\sigma = (\sigma_1, \ldots, \sigma_N)$$

of stochastic kernels

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

(i.e. $\sigma_t(r, h^{t-1}; \cdot)$ is a distribution on I for every $(r, h^{t-1}) \in R \times H^{t-1}$.)

Analogously a strategy of player 2 is given by a sequence

$$\tau = (\tau_1, \ldots, \tau_N)$$

of stochastic kernels

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

Using the strategies $\sigma = (\sigma_1, \ldots, \sigma_N)$ and $\tau = (\tau_1, \ldots, \tau_N)$ the players generate a probability distribution $P^p_{\sigma,\tau}$ on the set $R \times S \times H^N$:

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

The payoff as a function on the product of the strategy spaces $\alpha_n^p(\sigma, \tau)$ is naturally defined as the expectation of the average payoffs with respect to this distribution.

Mertens and Zamir [1971/72] describe the chance move and the information the players receive about it in a different manner which at first sight seems to be more general. They consider an arbitrary finite set K instead of the product set $R \times S$ and represent the player's information by two partitions of K. Defining the type sets R and S to consist of the elements of the partition of players 1 and 2 and assigning to every pair of types the probability of their intersection, an equivalent random experiment with underlying product space is obtained.

The definition of I-concavity resp. II-convexity (definition 1.1) is a reformulation of the corresponding definition by Mertens and Zamir [1971/72] for the case of product spaces. Some notation first:
For any $p \in \Delta(R \times S)$, $p^R(p^S)$ denotes the marginal distribution of p on $R(S)$ and $p^{R|S}(p^{R|S})$ is the conditional distribution on S given R (on R given S). If $p^R > 0$ holds true, then $p^{S|R}$ is uniquely defined. Otherwise the symbol $p^{S|R}$ denotes a specific version of the conditional probability.

If q is a probability on R let $q \otimes p^{S|R}$ be the probability distribution on $R \times S$ defined by

$$q \otimes p^{S|R}(r, s) = q(r) \cdot p(s|r).$$

Especially we have

$$p = p^R \otimes p^{S|R} = p^S \otimes p^{R|S}.$$

Definition 1.1: A function φ: $\Delta(R \times S) \to \mathbb{R}$ is called I-concave iff the function

$$q \mapsto \varphi(q \otimes \textstyle\prod)$$

defined on $\Delta(R)$ is concave for every stochastic kernel $\prod | R \Rightarrow S$.
A function φ: $\Delta(R \times S) \to \mathbb{R}$ is called II-convex iff the function

$$q \mapsto \varphi(q \otimes \textstyle\prod)$$

defined on $\Delta(S)$ is convex for every stochastic kernel $\prod | S \Rightarrow R$.

If φ: $\Delta(R \times S)$ is I-concave (II-convex) this implies that the restriction of φ on the set of product measures $\varphi|_{\Delta(R) \times \Delta(S)}$ is concave in the first (convex in the second) component in the ordinary sense.

Let $A(p) = \sum_{r,s} p(r,s) A^{r,s}$ be the so called non-revealing game and define a function u: $\Delta(R \times S) \to \mathbb{R}$

$$p \mapsto \operatorname{val} A(p)$$

where $\operatorname{val} A(p)$ denotes the min-max value of the matrix game $A(p)$. The function u is continuous so that the pair of functional equations

$$\varphi = \underset{\mathrm{I}}{\mathrm{cav}} \min\{u, \varphi\}$$

$$\varphi = \underset{\mathrm{II}}{\mathrm{vex}} \max\{u, \varphi\}$$

has a unique simultaneous solution (Mertens and Zamir [1977], Sorin [1984]). ($\underset{\mathrm{I}}{\mathrm{cav}}\,\varphi$ is the minimal I-concave function that is greater or equal to φ and $\underset{\mathrm{II}}{\mathrm{vex}}\,\varphi$ is the maximal II-convex function that is less or equal to φ.)

Let v be the solution of the functional equations. Of course v is I-concave and II-convex.

2 Player 2's Strategy

Let us assume for a moment that the players do not only learn their opponent's choice of action after each stage but also the corresponding posterior probability $p_T(h^T) = P^p_{\sigma,\tau}(\cdot \mid h^T)$ on $R \times S$ after each stage T (which they can only partially compute according to the original rules of the game). In this case it may seem reasonable to employ an optimal strategy of the non-revealing matrix game $A(p_T)$ at stage $T+1$. Nevertheless there is one case where this conduct is clearly not optimal. Suppose that $u(p_T) > v(p_T)$ holds true and let player 2 play optimally in $A(p_T)$. Player 1 can guarantee the payoff $u(p_T)$ by playing a nonrevealing optimal strategy as well, which by assumption is more than he should get. In this case player 2 has to make his strategy dependent on his type. He does this by performing a so called type dependent lottery (cf. Aumann and Maschler [1966]). The posteriors the lottery produces are supporting points of the function v. They are provided by proposition 2.1.

Proposition 2.1: If $v(p) < u(p)$ is satisfied for some $p \in \Delta(R \times S)$, then there exist distributions $q^\ell \in \Delta(S)$ and real numbers λ^ℓ such that the conditions (1)–(4) hold true (abbreviating $p^\ell = q^\ell \otimes p^{R|S}$):

$$\sum_\ell \lambda^\ell q^\ell = p^S, \quad \sum_\ell \lambda^\ell = 1, \quad \lambda^\ell \geq 0 \;\forall \ell \tag{1}$$

$$v(p) = \sum_\ell \lambda^\ell v(p^\ell) \tag{2}$$

$$v(p^\ell) = u(p^\ell) \;\forall \ell \tag{3}$$

$$v(q \otimes p^{R|S}) < u(q \otimes p^{R|S}) \text{ for all } q \text{ in the relative interior of } \mathrm{Conv}\{q^\ell\}. \tag{4}$$

Proof: Due to the functional equation

$$\varphi = \underset{\mathrm{I}}{\mathrm{cav}} \max\{u, \varphi\}$$

the epigraph of the mapping $q \mapsto v(q \otimes p^{S|R})$ is the convex hull of the closed epigraph of the continuous function

$$q \mapsto \max\{u(q \otimes p^{S|R}), v(q \otimes p^{S|R})\}$$

defined on $\Delta(S)$. Since the condition $v(p) < u(p)$ is assumed we also have $v(p) < \max\{u(p), v(p)\}$. It follows that the point $(p^S, v(p))$ is contained in a face of the epigraph of $q \mapsto v(q \otimes p^{S|R})$, i.e. it can be represented as a convex combination of points $(q^\ell, \max\{u(p^\ell), v(p^\ell)\})$. Choosing among the tupels $\{q^\ell\}$ appropriate for this purpose the one that puts up the minimal polyhedron $\mathrm{Conv}\{q^\ell\}$, a representation satisfying (1)–(4) is found.

So we could develop our idea of a good strategy for player 2 as follows:

- Play optimally in $A(p_T)$ if $u(p_T) \leq v(p_T)$ holds true.
- Perform a type dependent lottery using the supporting points p^ℓ of v if $u(p_T) > v(p_T)$ holds true (a precise definition of this follows later on).

Nevertheless there is still a serious obstacle. Not knowing his opponent's strategy player 2 is not able to compute posteriors. But it is easy to check that he is able to compute conditional posterior probabilities $p_T(s|r)$ after each stage T since they only depend on his own strategy. Moreover he computes a unique version of $p_T^{S|R}$ since we assumed that p^R is strictly positive. (From now on $p_T^{S|R}$ will always denote player 2's version of the conditional posterior probability.) So he must find a substitute $\bar{p}_T^R$ for the missing marginal probability p_T^R. To this end we have to consider supergradients of v.

Define

$$W(p) = \{x \in \mathbb{R}^R : p \cdot x = v(p),\ q \cdot x \geq v(q \otimes p^{S|R})\ \forall\, q \in \Delta(R)\}$$

Of course the definition is only unique if $p^{S|R}$ is uniquely determined. Since it will only be applied to posteriors p_T no ambiguity arises. Suppose now that in a game $v_N(p)$ player 2 is able to confine the payoff every type r of player 1 would receive by $\eta_0(r)$ for any $\eta_0 \in W(p)$. (The proof of the main theorem will show that this is almost the case.) Suppose further that player 2 has observed a sequence h^T of actions during the first T stages and that he has generated the conditional posterior $p_T^{S|R}$. He knows that, provided player 1's type is r, he has gathered the payoff

$$\sum_{t=1}^{T} \sum_{S} p_T(s|r, h^T) A^{r,s}(h_T)$$

or, abbreviating

$$B_T(h^T) = \Big(\sum_{S} p_T(s|r, h^T) A^{r,s}(h_T)\Big)_{r \in R},$$

he has received the vector payoff

$$\sum_{t=1}^{T} B_t(h^t).$$

Comparing this quantity with his original aim η_0 he finds that in order to end up with η_0 he must achieve the vector payoff

$$\frac{N}{N-T}\eta_0 - \frac{1}{N-T}\sum_{t=1}^{T} B_t$$

during the remaining $N-1$ stages. For technical reasons it is desirable to have the above difference defined for $T=N$, therefore we consider $N+1$ instead of N stages. Furthermore it will turn out that player 2's information release has to be taken into account in the shape of the sequence η_t, $t=0, \frac{1}{2}, 1, \frac{3}{2}, \ldots, N$, which will be introduced in the sequel. Therefore we finally define the sequence δ_T,

$$\delta_T: H^N \to \mathbb{R}^R, \quad T=0, \ldots, N$$

by

$$\delta_T = \frac{N+1}{N+1-T}\eta_0 - \frac{1}{N+1-T}\sum_{t=1}^{T} B_t + \sum_{t=0}^{T-1} \frac{N-t}{N+1-T}\left(\eta_{t+\frac{1}{2}} - \eta_t\right).$$

Maybe the following recursive definition of δ_t looks a bit more attractive:

$$\delta_0 = \eta_0$$

$$\delta_{T+1} = \delta_T + \frac{1}{N-T}(\delta_T - B_{T+1}) + \left(\eta_{T+\frac{1}{2}} - \eta_T\right), \quad T=0, \ldots, N-1.$$

Anyway δ_T represents the vector payoff player 2 has to achieve from stage $T+1$ on in order to get the total payoff $v(p_0)$ in the end. According to our assumption this is only a realistic aim if δ_T is a supergradient of the function $q \mapsto v(q \otimes p_T^{S|R})$. If this condition is not satisfied we must distinguish two cases which can be described by the set

$$U(p) = \{x \in \mathbb{R}^R : q \cdot x \geq v(q \otimes p^{S|R}) \; \forall \, q \in \Delta(R)\}.$$

As in the case of $W(p)$ the definition is only unique if $p^{S|R}$ is uniquely determined. $U(p)$ is closed, convex, bounded from below and unbounded from above. Geometrically it can be described as the comprehensive hull from above of the union $\bigcup_{q \in \Delta(R)} W(q \otimes p^{S|R})$ which will be abbreviated by $V(p)$.

If $\delta_T \in U(p_T)$ holds true there is a supergradient $\xi \in V(p_T)$ satisfying $\xi \leq \delta_T$, i.e. player 2 is able to reduce player 1's payoff further than he has to. He has gained a temporary advantage. The strategy to be defined will allow him to waste it by doing anything.

If $\delta_T \notin U(p_T)$ holds true there is no supergradient $\xi \in V(p)$ with $\xi \leq \delta_T$, player 2 has suffered a loss in terms of vector payoffs. In this case he tries to get the supergradient

$$\eta_T = \underset{\xi \in U(p_T)}{\operatorname{argmin}} |\xi - \delta_T|$$

with minimal euclidean distance from δ_T. Since $U(p_T)$ is convex, η_T is always uniquely defined and if $\delta_T \notin U(p_T)$ holds, η_T is contained in the boundary of $U(p_T)$, i.e. $\eta_T \in V(p_T)$ is valid. Define

$$\bar{p}_T^R = \frac{1}{\sum\limits_r \eta_T(r) - \delta_T(r)} (\eta_T - \delta_T), \quad T = 1, \ldots, N$$

$$\bar{p}_0^R = p^R.$$

(In the case $\delta_T \in U(p_T)$ $\bar{p}_T^R$ is undefined.)

$\bar{p}_T^R$ is defined as a normal vector of the set $U(p_T)$ at the point δ_T but it also satisfies

Lemma 2.2: $\eta_T \in W(\bar{p}_T^R \otimes p_T^{S|R})$

Proof: By definition of η_T we have

$$\bar{p}_T^R \cdot \eta_T \leq \bar{p}_T^R \cdot \xi \;\forall\; \xi \in U(p_T)$$

and because of the supergradient property of η_T

$$\bar{p}_T^R \cdot \eta_T \geq v(\bar{p}_T^R \otimes p_T^{S|R}).$$

Choosing $\xi \in W(\bar{p}_T^R \otimes p_T^{S|R})$ we find

$$v(\bar{p}_T^R \otimes p_T^{S|R}) \leq \bar{p}_T^R \cdot \eta_T \leq \bar{p}_T^R \cdot \xi = v(\bar{p}_T^R \otimes p_T^{S|R}).$$

Consequently

$$\eta_T \in W(\bar{p}_T^R \otimes p^{S|R})$$

holds true as well.

Trying to enforce the vector payoff η_T (instead of δ_T) player 2 will act as if the probability $\bar{p}_T^R$ associated with η_T were the actual marginal of the posterior probability p_T, i.e. his conduct is base on the probability

$$\bar{p}_T = \bar{p}_T^R \otimes p_T^{S|R}.$$

Remark that δ_T depends on the sequence (η_t) only up to stage $T-1$ so that no circular definition occurs. The half steps of the sequence (η_t) are defined together with the strategy itself.

Suppose now that everything is defined up to stage T and a history h^T has occurred. (The argument h^T will be omitted.)

Case 1: $\delta_T \in U(p_T)$

Define $\quad \bar{\tau}_{T+1}(s;j) = \dfrac{1}{|J|} \;\forall j \in J, s \in S$

and $\qquad \eta_{T+\frac{1}{2}} = \eta_T.$

Case 2: $\delta_T \notin U(p_T)$

Subcase 2a: $v(\bar{p}_T) \geq u(\bar{p}_T)$

Choose an optimal strategy $y \in \Delta(J)$ for the minimizing player in the matrix game $A(\bar{p}_T)$ and define

$$\bar{\tau}_{T+1}(s;j) = y_j \;\forall j \in J, s \in S$$

$$\eta_{T+\frac{1}{2}} = \eta_T$$

Subcase 2b: $v(\bar{p}_T) < u(\bar{p}_T)$

There exist $q^\ell \in \Delta(S)$, $\lambda^\ell \in \mathbb{R}$ such that the conditions (1) to (4) of proposition 2.1 hold true. Let y^ℓ be an optimal strategy for the minimizing player in the matrix game $A(p^\ell)$. Player 2's strategy is given by

$$\bar{\tau}_{T+1}(s;j) = \begin{cases} \sum_\ell \lambda^\ell \dfrac{q^\ell(s)}{\bar{p}_T(s)} y_j^\ell & \text{if } \bar{p}_T(s) > 0 \\ \sum_\ell \lambda^\ell y_j^\ell & \text{if } \bar{p}_T(s) = 0 \end{cases}.$$

In this case also η_T has to be adjusted:

$$\eta_{T+\frac{1}{2}}(j) = \xi^j,$$

the ξ^j being provided by the following theorem:

Theorem 2.3: If the conditions of subcase 2b are satisfied, there exist supergradients $\xi^j \in V(p_{T+1}(j))$ satisfying

$$E\left(\eta_{T+\frac{1}{2}}(j)(r) \mid r\right) = \eta_T(r).$$

Proof: Straightforward computation yields that the probability

$$\bar{p}_{T+\frac{1}{2}}(j) = \bar{p}_T^R \otimes p_{T+1}^{S\mid R}(j)$$

is a convex combination of the p^ℓ,

$$\bar{p}_{T+\frac{1}{2}}(j) = \sum_\ell \mu_j^\ell p^\ell$$

with

$$\mu_j^\ell = \frac{\lambda^k y_j^k}{\sum_k \lambda^k y_j^k};$$

i.e. $\bar{p}_{T+\frac{1}{2}}(j)$ is contained in the convex hull $\mathrm{Conv}\{p^\ell\}$. Consequently it is sufficient to prove

Theorem 2.4: Let $p^j \in \Delta(R \times S)$, $q^j \in \Delta(S)$ and $p^j = q^j \otimes p^{R\mid S}$, $j \in J$ such that the subsequent conditions are satisfied:

$$\sum_j \mu^j p^j = p, \ \sum_j \mu^j = 1, \ \mu^j \geq 0 \quad j \in J \tag{1}$$

$$v(p) = \sum_j \mu^j v(p^j) \tag{2}$$

$$v(p^j) \leq u(p^j) \quad j \in J \tag{3a}$$

$$v(q \otimes p^{R\mid S}) < (q \otimes p^{R\mid S}) \text{ for all } q \text{ in the relative interior of } \mathrm{Conv}\{q^j; j \in J\} \tag{4}$$

Define a stochastic kernel $(p^j)^{S\mid R}$ by

$$p^j(s\mid r)=\begin{cases}\dfrac{\dfrac{q^j(s)}{p(s)}\Pi(s\mid r)}{\sum\limits_{s':p(s')>0}\dfrac{q^j(s')}{p(s')}p(s'\mid r)+\sum\limits_{s':p(s')=0}p(s'\mid r)} & \text{if } p(s)>0\\[2ex] \dfrac{p(s\mid r)}{\sum\limits_{s':p(s')>0}\dfrac{q^j(s')}{p(s')}p(s'\mid r)+\sum\limits_{s':p(s')=0}p(s'\mid r)} & \text{if } p(s)=0.\end{cases}$$

For any supergradient $\xi\in W(p^R, p^{S\mid R})$ there exists a tupel $\xi^j\in W((p^j)^R, (p^j)^{S\mid R})$ that represents ξ as follows:

$$\xi(r)=\sum_j \mu^j\left(\sum_{s:p(s)>0}\frac{q^j(s)}{p(s)}p(s\mid r)+\sum_{s:p(s)=0}p(s\mid r)\right)\xi^j(r)\ \forall\ r\in R \tag{5}$$

Remark: $(p^j)^{S\mid R}$ is a version of the corresponding conditional probability of p^j. If $p^j(r)>0\ \forall\ r\in R$ holds, its explicit definition is not necessary. In this case the representation of ξ reduces to

$$\xi=\sum_j \mu^j\frac{p^j(r)}{p(r)}\xi^j(r)$$

and if p is a product measure we have

$$\xi=\sum_j \mu^j\xi^j.$$

Proof: First of all it is sufficient to consider distributions p with strictly positive marginals p^R. In order to see this apply the assertion to

$$p_n=\frac{n-1}{n}p+\frac{1}{n}(x\otimes p^{S\mid R}),$$

$$q_n^j=\frac{n-1}{n}q^j+\frac{1}{n}(x\otimes p^{S\mid R})^S$$

instead of p and q^ℓ and check that

$$\lim_{n\to\infty}\frac{p_n^j(r)}{p_n(r)}=1\quad \text{if } p(r)=0$$

and that

$$\lim_{n\to\infty}p_n^j(s\mid r)=p^j(s\mid r)$$

hold true for any strictly positive probability $x \in \Delta(R)$.

Secondly we have to consider distributions p^j satisfying

$$v(p^j) < u(p^j) \ \forall j \in J \tag{3 b}$$

instead of (3 a) because due to (4) any tupel $\{p^j, j \in J\}$ satisfying (1), (2), (3 a), (4) can be approximated by tupels satisfying (1), (2), (3 b), (4).

Define

$$W^*(p) = \left\{\xi \in \mathbb{R}^R : \exists\, \xi^j \in W(p^j) : \xi(r) = \sum_j \mu^j \frac{p^j(r)}{p(r)} \xi^j(r) \ \forall\, r \in R\right\}$$

$W^*(p)$ is closed, convex and comprehensive from above. Let us now assume that (5) does not hold true, i.e. there is a supergradient $\bar{\xi} \in W(p) \setminus W^*(p)$. Then there is a separating hyperplane between $\bar{\xi}$ and $W^*(p)$, i.e. is a normal vector $x \in \Delta(R)$ and a real number $\delta > 0$ such that

$$x \cdot \bar{\xi} < x \cdot \xi^* - \delta \ \ \forall\, \xi^* \in W^*(p)$$

is valid.

Comparing the values of the function v at the points p, p^j with those slightly shifted in the direction x, x^j with

$$x^j(r) = x(r) \frac{p^j(r)}{p(r)}$$

a contradiction will be produced.

For $\varepsilon > 0$ define

$$p_\varepsilon(r, s) = (p^R + \varepsilon x)(r) p(s \mid r)$$

$$p_\varepsilon^j(r, s) = (p^R + \varepsilon x)(r) \frac{p^j(r)}{p(r)} p^j(s \mid r)$$

Of course p_ε and p_ε^j are no longer probability distributions. This could be mended by dividing by the sum of their components. But it is more convenient to consider the linear homogeneous extension of v on the positive orthant $(\mathbb{R}_+)^{|R| \cdot |S|}$. I-concavity and II-convexity are preserved and supergradients of v^{II} are also supergradients of the linear homogeneous extension. Especially the notion "conditional probability" makes sense for p_ε and p_ε^j since e.g. the quotient $\dfrac{p_\varepsilon(r, s)}{p_\varepsilon(s)}$ is independent of the normalization.

Choosing ε small enough we can assume that $v < u$ holds on $\mathrm{Conv}\{p_\varepsilon^\ell\}$. By definition we have

$$(p_\varepsilon^j)^{S \mid R} = (p^j)^{S \mid R}$$

and straightforward computation yields

$$(p_\varepsilon^j)^{R|S} = (p^j)^{R|S}.$$

Let $\xi^j \in W(p^j)$ describe the directional derivative of the function $p \mapsto v(q \otimes (p^j)^{S|R})$ at the point $(p^j)^R$ in the direction x^j, i.e. let ξ^j satisfy

$$\lim_{\varepsilon \to 0} \frac{v(p^j) + \varepsilon x^j \cdot \xi^j - v(p_\varepsilon^j)}{\varepsilon} = 0.$$

For the vector $\xi^* \in W^*(p)$ defined by

$$\xi^*(r) = \sum_j \mu^j \frac{p^j(r)}{p(r)} \xi^j(r), \quad r \in R$$

we can deduce that

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \frac{v(p) + \varepsilon x \cdot \xi - v(p_\varepsilon)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\sum_j \mu^j v(p^j) + \varepsilon \sum_j \mu^j x^j \cdot \xi^j - \sum_j \mu^j v(p_\varepsilon^j)}{\varepsilon} \\
&= \sum_j \mu^j \lim_{\varepsilon \to 0} \frac{v(p^j) + \varepsilon x^j \xi^j - v(p_\varepsilon^j)}{\varepsilon} \\
&= 0
\end{aligned}$$

Consequently $x \cdot \xi$ is the derivative of $q \mapsto v(q \otimes p^{S|R})$ at p^R in the direction x. On the other hand theorem 2.3.4 of Rockafellar [1970] implies that the directional derivative is given by

$$\min \{x \cdot \xi' : \xi' \in W(p)\}.$$

But x was determined such that the minimum is not attained by any $\xi' \in W^*(p)$. Consequently $\xi^* \in W(p) \setminus W^*(p)$ cannot exist and (5) is proved by contradiction.

q.e.d.

3 Main Theorem

Theorem 3.1: Let σ be an arbitrary strategy for player 1 in $\Gamma_N(p)$ and let $\bar{\tau}$ be the strategy of player 2 defined in section 2. Then the total payoff is bounded by

$$\alpha_N^p(\sigma, \bar{\tau}) \leq v(p) + 2\left(\sqrt{\frac{|R|}{N}} + \frac{1}{N}\right) M$$

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

Proof: It is straightforward to check that the total payoff can be represented by the vector payoff:

$$\alpha_N^p(\sigma, \bar{\tau}) = \frac{1}{N} E\left(p_N^R \cdot \sum_{T=1}^{N} B_T\right)$$

Now we can estimate the difference $\alpha_N^p(\sigma, \tau) - v(p)$ by the difference $\eta_N - \delta_N$ or rather the euclidean distance between δ_N and the set $U(p_N)$:

$$\begin{aligned}
&\alpha_N^p(\sigma, \tau) - v(p) \\
&= \frac{1}{N} E\left(p_N^R \cdot \sum_{T=1}^{N} B_T\right) - p^R \cdot \eta_0 \\
&= \frac{1}{N} E\left(p_N^R \left(\sum_{T=1}^{N} B_T - N\eta_0\right)\right) \\
&= \frac{1}{N} E\left(p_N^R \left(\sum_{T=1}^{N} B_T - \sum_{T=0}^{N-1} (N-T)\left(\eta_{T+\frac{1}{2}} - \eta_T\right) - (N+1)\eta_0 + \eta_0\right)\right) \\
&\qquad\qquad\qquad\qquad \text{(due to theorem 2.3)} \\
&= \frac{1}{N} E(p_N^R(\eta_0 - \delta_N)) \\
&= \frac{1}{N} E(p_N^R(\eta_0 - \eta_N)) + \frac{1}{N} E(p_N^R(\eta_N - \delta_N)) \\
&\leq 2\frac{M}{N} + \frac{1}{N} E(|\eta_N - \delta_N|) \\
&\leq 2\frac{M}{N} + \frac{1}{N} E((\eta_N - \delta_N)^2)^{\frac{1}{2}}
\end{aligned}$$

The quantity $E((\eta_T - \delta_T)^2)$ is estimated inductively from stage to stage. We claim that

$$E((\eta_T - \delta_T)^2) \leq \frac{T}{(N+1-T)^2} |R| 4M^2$$

holds true for $T = 0, \ldots, N$. For $T = 0$ the assertion is satisfied by definition. Assuming that it is valid for $T < N$, $E((\eta_{T+1} - \delta_{T+1})^2)$ is estimated as follows:

$$\begin{aligned}
&E((\eta_{T+1} - \delta_{T+1})^2) \\
&\leq E\left(\left(\eta_{T+\frac{1}{2}} - \delta_{T+1}\right)^2\right)
\end{aligned}$$

$$=E\left(\left(\eta_T-\frac{N+1-T}{N-T}\delta_T+\frac{1}{N-T}B_{T+1}\right)^2\right)$$

$$=E\left(\left(\frac{N+1-T}{N-T}(\eta_T-\delta_T)+\frac{1}{N-T}(B_{T+1}-\eta_T)\right)^2\right)$$

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

$$+\frac{1}{(N-T)^2}E((B_{T+1}-\eta_T)^2)$$

$$+2\frac{N+1-T}{(N-T)^2}E((\eta_T-\delta_T)(B_{T+1}-\eta_T))$$

$$\leq\left[\frac{N+1-T}{N-T}\right]^2\frac{T}{(N+1-T)^2}|R|\,4M^2+\frac{1}{(N-T)^2}|R|\,4M^2$$

$$=\frac{T+1}{(N-T)^2}|R|\cdot 4M^2$$

The last inequality requires further explanation:
The first term is estimated by inductive hypothesis, the estimation of the second one is trivial. It remains to show that

$$E((\eta_T-\delta_T)(B_{T+1}-\eta_T))\leq 0$$

or that

$$E((\eta_T-\delta_T)(B_{T+1}-\eta_T)\mid h^T, i)\leq 0 \;\;\forall\,(h^T, i)\in H^T\times I$$

We examine the different cases according to the definition of player 2's strategy (omitting the argument h^T):

Case 1: $\delta_T\in U(p_T^{S|R})$

Obviously $\eta_T=\delta_T$ holds true by definition.

Case 2: $\delta_T\notin U(p_T^{S|R})$

According to the definition of $\bar{p}_T^R$ the inequality

$$E((\eta_T-\delta_T)(B_{T+1}-\eta_T)\mid h^T, i)\leq 0$$

is equivalent to

$$E(\bar{p}_T^R(B_{T+1}-\eta_T)\mid h^T, i)\leq 0$$

or

$$\bar{p}_T^R\cdot E(B_{T+1}\mid h^T, i)\leq v(\bar{p}_T)$$

Subcase 2a: $v(\bar{p}_T)\geq u(\bar{p}_T)$

Player 1 does not make use of his private information so that $p_{T+1}^{S|R} = p_T^{S|R}$.

$$
\begin{aligned}
&\bar{p}_T^R \cdot E(B_{T+1} \mid h^T, i) \\
&= \sum_r \bar{p}_T(r) \sum_s p_T(s \mid r) \sum_j y_j A^{r,s}(i,j) \\
&= \sum_j y_j A(\bar{p}_T)(i,j) \\
&\leq u(\bar{p}_T) \qquad \text{(since } y \text{ is optimal in } A(\bar{p}_T)) \\
&\leq v(\bar{p}_T)
\end{aligned}
$$

Subcase 2b: $v(\bar{p}_T) < u(\bar{p}_T)$

$$
\begin{aligned}
&\bar{p}_T^R \cdot E(B_{T+1} \mid h^T, i) \\
&= \sum_r \bar{p}_T(r) \sum_j \sum_{s'} p_T(s') \tau_{T+1}(s';j) \sum_s p_{T+1}(j)(s \mid r) A^{r,s}(i,j) \\
&= \sum_r \bar{p}_T(r) \sum_j \sum_s p_T(s \mid r) \tau_{T+1}(s;j) A^{r,s}(i,j) \\
&= \sum_j \sum_{r,s} \bar{p}_T(r,s) \tau_{T+1}(s;j) A^{r,s}(i,j) \\
&= \sum_j \sum_{r,s} \bar{p}_T(r,s) \sum_\ell \lambda^\ell \frac{q^\ell(s)}{\bar{p}_T(s)} y_j^\ell A^{r,s}(i,j) \\
&= \sum_j \sum_{r,s} \sum_\ell \lambda^\ell p^\ell(r,s) y_j^\ell A^{r,s}(i,j) \\
&= \sum_\ell \lambda^\ell \sum_j y_j^\ell A(p^\ell)(i,j) \\
&\leq \sum_\ell \lambda^\ell u(p^\ell) \\
&= \sum_\ell \lambda^\ell v(p^\ell) \\
&= v(\bar{p}_T)
\end{aligned}
$$

The proof by induction is complete. Applying the assertion for $T = N$ we obtain

$$E((\eta_N - \delta_N)^2) \leq N \cdot |R| \, 4M^2$$

and the result follows. q.e.d.

References

Aumann RJ, Maschler M (1966) Game Theoretic Aspects of Gradual Disarmament, Mathematica, Chapter V.

Aumann RJ, Maschler M (1967) Repeated Games with Incomplete Information: A Survey of Recent Results. Mathematica, St-116, Chapter III, 287–403.

Blackwell D (1956) An Analogue of the Minimax Theorem for Vector Payoffs, Pacific Journal of Mathematics, 65, 1–8.

Heuer M (1990): The Role of Vector Payoffs in Repeated Zero-Sum Games with Incomplete Information, Doctoral Dissertation in Mathematics, IMW, Universität Bielefeld.

Kuhn HW (1953) Extensive Games and the Problem of Information, Contributions to the Theory of Games, Princeton.
Mertens JF, Zamir S (1971/72) The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides, International Journal of Game Theory 1, 39–64.
Mertens JF, Zamir S (1977) A Duality Theorem on a Pair of Simultaneous Functional Equations, Journal of Mathematical Analysis and Application, 60, 550–558.
Mertens JF, Zamir S (1980) Minmax and Maxmin of Repeated Games with Incomplete Information, International Journal of Game Theory 9, 201–215.
Rockafellar, RT (1970) Convex Analysis, Princeton University Press.
Sorin S (1984) On a Pair of Simultaneous Functional Equations, Journal of Mathematical Analysis and Applications, 98 (1), 296–303.
Stearns RE (1967) A Formal Information Concept for Games with Incomplete Information. In Mathematica (1967), St–116, Chapter IV, 405–433.

Received February 1991
Revised version December 1991