

On Repeated Games with General Information Function¹⁾

By S. ZAMIR²⁾

Abstract: For a class of repeated two-person zero-sum games with incomplete information it was proved by AUMANN and MASCHLER that $\lim v_n$ exists, v_n being the value of the game with n repetitions. If the players know at each stage the moves done by both players at all previous stages, AUMANN and MASCHLER could prove that the error term $\delta_n = |v_n - \lim v_n|$ satisfies $\delta_n \leq c/\sqrt{n}$ for some $c > 0$. It was then shown by ZAMIR that this bound is the lowest possible. In this paper it is shown that if previous moves are not always announced, δ_n may be of higher order of magnitude e.g. $\delta_n \geq c/n^{1/3}$ for some $c > 0$. New upper bounds for δ_n are given for two classes of games.

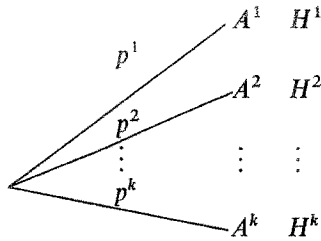
1. Introduction

The class of games considered in this paper are those discussed by AUMANN and MASCHLER [1968]. They can be described as follows:

Let A^1, \dots, A^k ; ($k \geq 2$) be $l \times m$ matrices of real numbers viewed as two-person zero-sum games. Let H^1, \dots, H^k , be $l \times m$ matrices whose elements are symbols (e.g. letters, digits, pairs of digits, etc.). The matrices $\{A^v | v = 1, \dots, k\}$ will be called the *payoff matrices* and their elements will be denoted by a_{ij}^v ; $i = 1, \dots, l$; $j = 1, \dots, m$; $v = 1, \dots, k$. The matrices $\{H^v | v = 1, \dots, k\}$ will be called the *information matrices* and their elements will be denoted by h_{ij}^v . For each p in the simplex $P = \{p = (p^1, \dots, p^k) | p^v \geq 0, v = 1, \dots, k, \sum_1^k p^v = 1\}$ and for each positive integer n , consider the game $\Gamma_n(p)$ played as follows: Chance chooses one of the k games $\{A^v\}$ assigning to A^v the probability p^v . Player I is then informed of chance's choice but player II is not. The game chosen by chance is then played n times. If at a certain stage the player played pure strategies i and j respectively, then player I is credited and player II is debited by a_{ij}^v and *in addition* h_{ij}^v is announced by the referee, where v is the index of the game chosen by chance. Neither the payoffs nor the moves are told explicitly. The only way for the players to learn something about the moves or the payoffs is through the information h_{ij}^v announced after each play (Player I however has the big advantage of knowing v). The payoff in $\Gamma_n(p)$ is defined to be the sum of the payoffs of the n individual plays of the game chosen by chance, divided by n . Both players of course know the rules of $\Gamma_n(p)$ described above. We shall describe the specifications of such a game schematically by:

¹⁾ This research was done in part under NSF Grant No. GP-33431 X.

²⁾ Prof. SHMUEL ZAMIR, Department of Mathematics, University of California, Los Angeles, California.



The main feature of the game $\Gamma_n(p)$ is the following: Player II would like to learn as much as possible concerning which game was chosen by chance since the more he knows about it the higher his profits will usually be (certainly not less). However, the procedure of learning may be costly for him, therefore he will balance learning against profits at each stage of the game.

For the purpose of computing the value of our games we will make use of the minimax theorem which we will read roughly as: "The value of the game is the best that player II (or I) can guarantee if his opponent has to announce his mixed strategy in advance". In other words we will assume that player II knows player I's strategy and suggests a 'best reply' for it. Then we will maximize over all player I's strategies and obtain the value. Now since player I knows the choice of chance v , his move at each stage may depend on v . The strategy of player I together with player II's move, j , induce a probability distribution on the elements $h_{ij}^v; v = 1, \dots, k; i = 1, \dots, l$. Hence given the announced information player II will compute a new conditional distribution on the choice of chance v . Realizing this mechanism of player II to gain knowledge about the choice of chance, the essential notion for classifying strategies of player I becomes that of a non-separating strategy defined as follows:

Definition 1:

A mixed strategy σ of player I in $\Gamma_1(p)$ will be called *non-separating* if for any $j(1 \leq j \leq m)$ the probability distribution induced by σ, j and v on the elements of H^v is the same for all v with $p^v > 0$ (considered as distributions on $H = \{h_{ij}^v | 1 \leq i \leq l, 1 \leq j \leq m, 1 \leq v \leq k\}$).

In other words if at a certain stage player I uses a non-separating strategy then whatever player II does, his conditional probability about the choice of chance after that stage is the same as it was before. This means that player II can gain no knowledge about the choice of chance in a stage in which player I uses a non-separating strategy.

Denote by $NS(p)$ the set of non-separating strategies in $\Gamma_1(p)$ and by $\Delta(p)$ the one stage game $\Gamma_1(p)$ in which the strategy space of player I is restricted to $NS(p)$. Note that it is possible that for some $p \in P, NS(p)$ is the empty set \emptyset in which case $\Delta(p)$ is not defined. However at least when p is an extreme point of P we have $NS(p) \neq \emptyset$. In fact if p is the v th unit vector $(0, \dots, 1, \dots, 0)$ then $\Delta(p)$ is the ordinary two-person zero-sum game A^v .

Let $v_n(p)$ denote the (minimax) value of $\Gamma_n(p)$ and let $u(p)$ denote the value of $A(p)$ ($u(p)$ is defined whenever $NS(p) \neq \emptyset$). Denote by $u^*(p)$ the *concavification* of $u(p)$ i.e. the smallest concave function on P satisfying $u^*(p) \geq u(p)$ whenever $u(p)$ is defined.

Theorem 1 (AUMANN-MASCHLER [1968]):

$$\lim_{n \rightarrow \infty} v_n(p) = u^*(p) \text{ for all } p \in P.$$

The special case in which $h_{ij}^v = (i,j)$ will be referred to as the case of *standard information*. In such games the information announced after each stage is just the moves (i,j) done by the players at that stage. For standard information AUMANN and MASCHLER [1966] proved a stronger version of Theorem 1, namely:

Theorem 1_s:

In the case of standard information

$$v_n(p) = u^*(p) + \delta_n(p) \text{ for } p \in P; \quad n = 1, 2, \dots,$$

where $\delta_n(p) \leq c/\sqrt{n}$ for some $c > 0$.

The error term $\delta_n(p)$ which measures the speed of convergence of $v_n(p)$ is the subject of our paper. In an earlier paper on this subject [ZAMIR, 1971/72] we proved among other things that in the standard information case c/\sqrt{n} is the best bound possible. There are games for which $\delta_n \geq c'/\sqrt{n}$ for $c' > 0$. As for the general case, nothing is said about the error term in Theorem 1. The main question is of course: Does $\delta_n \leq c/\sqrt{n}$ still hold for the non-standard information case? The answer is no: In Section 1 we describe a game in which $c'/n^{1/3} \leq \delta_n \leq c/n^{1/3}$; ($c > 0, c' > 0$). The next natural question is: What is an upper bound for δ_n in the non-standard information case? This question is answered only for the case $k = 2$ (only two possible choices of chance). In Section 3 we shall prove at this case $\delta_n \leq c/n^{1/6}$. If, in addition, we have $NS(p) = \emptyset$ for $0 < p < 1$, a stronger bound is valid, namely $\delta_n \leq c/n^{1/4}$.

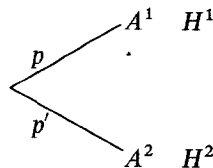
It will be convenient to use in this paper the following terminology:

Definition 2:

Two sequences of non-negative numbers $\{a_n\}$ and $\{b_n\}$ is said to be of the same order of magnitude if there are constants $c_1 > 0$ and $c_2 > 0$ s.t. $c_2 b_n \leq a_n \leq c_1 b_n$; $n = 1, 2, \dots$. This will be denoted by $a_n = O^*(b_n)$ (or $b_n = O^*(a_n)$).

2. A Game with $\delta_n = O^*(1/n^{1/3})$

Consider the game:



where:

$$A^1 = \begin{pmatrix} \alpha & \beta & \gamma \\ 8 & 3 & -1 \\ 8 & -3 & 1 \end{pmatrix} \quad H^1 = \begin{pmatrix} a & c & d \\ b & c & d \end{pmatrix}$$

$$A^2 = \begin{pmatrix} 8 & 2 & -2 \\ 8 & -2 & 2 \end{pmatrix} \quad H^2 = \begin{pmatrix} a & c & d \\ b & c & d \end{pmatrix}$$

(We use the notation x' for $1 - x$ for any $0 \leq x \leq 1$.) Denoting the pure strategies of player II by α, β, γ , we observe that deleting the strategy α and changing H' to standard information matrices we obtain the game with $\delta_n = 0^*(1/\sqrt{n})$ analyzed in our previous paper. The main feature of our present game is that player II does not collect any information about the choice of chance as long as he restricts himself to the strategies β and γ . The only way for him to gain some information is to use his dominated strategy α . In other words player II “has to buy the information”, whenever he wants to know player I’s move he has to pay 8 units. This will cause the enlargement of the order of magnitude of δ_n from $1/n^{1/2}$ to $1/n^{1/3}$.

A general strategy of player I in $\Gamma_1(p)$ is $\sigma = [(s, s'), (t, t')]; 0 \leq s \leq 1, 0 \leq t \leq 1$; which is to say: Play the mixed strategy (s, s') in A^1 and the mixed strategy (t, t') in A^2 . We will write σ in short as $\sigma = (s, t)$. It is easily verified that:

$$NS(0) = NS(1) = \{(s, t) \mid 0 \leq s \leq 1, 0 \leq t \leq 1\}$$

$$NS(p) = \{(s, t) \mid s = t; 0 \leq s \leq 1\} \quad \text{for } pp' \neq 0.$$

(I.e. except for $p = 0$ or $p = 1$, the only non-separating strategies for player I are those in which he plays in the same way in A^1 and A^2 .)

$A(p)$ is therefore the game:

$$\begin{pmatrix} 8 & 3p + 2p' & -p - 2p' \\ 8 & -3p - 2p' & p + 2p' \end{pmatrix}$$

and its value is $u(p) = 0$ for $0 < p < 1$; since also $u(1) = \text{val}(A^1) = 0 = \text{val}(A^2) = u(0)$ we conclude that $u^*(p) = 0$ for $0 \leq p \leq 1$.

Lemma 1:

$$v_{n+1}(p) = \min_{0 \leq \varepsilon \leq 1} \max_{0 \leq s, t \leq 1} \frac{1}{n+1} \{8\varepsilon + (1 - \varepsilon) \min [3p(s - s') + 2p'(t - t'); p(s' - s) + 2p'(t' - t)] + n(1 - \varepsilon)v_n(p) + n\varepsilon[R^a v_n(p^a) + R^b v_n(p^b)]\} \tag{1}$$

for $0 \leq p \leq 1; n = 1, 2, \dots$ where

$$R^a = ps + p't; \quad p^a = ps/R^a$$

$$R^b = ps' + p't'; \quad p^b = ps'/R^b. \tag{2}$$

Proof:

Let $\sigma = (s, t)$ be the part concerning the first play in any strategy of player I in $\Gamma_{n+1}(p)$. Let $\tau = (\varepsilon, x, y)$ where $\varepsilon \geq 0, x \geq 0, y \geq 0, \varepsilon + x + y = 1$ be the

corresponding part for player II. With strategies starting this way the expected payoff in the first play will be:

$$E(H_1) = 8\varepsilon + x[3p(s - s') + 2p'(t - t')] + y[p(s' - s) + 2p'(t' - t)] \\ \geq 8\varepsilon + (1 - \varepsilon) \min [3p(s - s') + 2p'(t - t'); p(s' - s) + 2p'(t' - t)].$$

After the first play the information announced will be:

$$a - \text{with probability } \varepsilon R^a \\ b - \text{with probability } \varepsilon R^b \\ c \text{ or } d - \text{with probability } 1 - \varepsilon,$$

where R^a and R^b given by (2). Given the information $a, b, c,$ or d , the conditional probability (for player II) of the event "Chance has chosen A^1 " is p^a, p^b, p^c or p^d respectively where $p^c = p^d = p$ and p^a and p^b are given by (2). After the first play therefore the players will face a situation equivalent to either

$$\Gamma_n(p) \text{ (with probability } 1 - \varepsilon) \\ \text{or } \Gamma_n(p^a) \text{ (with probability } \varepsilon R^a) \\ \text{or } \Gamma_n(p^b) \text{ (with probability } \varepsilon R^b).$$

By playing optimally in $\Gamma_n(p), \Gamma_n(p^a)$ or $\Gamma_n(p^b)$ player I can guarantee a total expected payoff of $nv_n(p), nv_n(p^a)$ or $nv_n(p^b)$ respectively, for the last n plays, thus:

$$v_{n+1}(p) \geq \min_{0 \leq \varepsilon \leq 1} \max_{0 \leq s, t \leq 1} \frac{1}{n+1} \{8\varepsilon + (1 - \varepsilon) \min [3p(s - s') \\ + 2p'(t - t'); p(s' - s) + 2p'(t' - t)] + n(1 - \varepsilon)v_n(p) \\ + n\varepsilon[R^a v_n(p^a) + R^b v_n(p^b)]\}. \tag{3}$$

On the other hand, for each $\varepsilon, 0 \leq \varepsilon \leq 1$, player II, by playing optimally, can guarantee to pay not more than $\max_{0 \leq s, t \leq 1} \{...\}/(n+1)$ and hence:

$$v_{n+1}(p) \leq \min_{0 \leq \varepsilon \leq 1} \max_{0 \leq s, t \leq 1} \{...\}/(n+1),$$

which together with (3) proves (1).

Lemma 2:

$$v_n(p) \geq pp'/n^{1/3} \tag{4}$$

for $0 \leq p \leq 1, n = 1, 2, \dots$

Proof:

Since (4) is clearly true for $p = 0$ or $p = 1$ we may assume $pp' \neq 0$. It is easily verified that $v_1(p) = \min(p, p') \geq pp'$ in accordance with (4). We proceed by induction: Assume (4) is true for n then apply (1) and restrict s and t by the relation:

$$p(s - s') = p'(t' - t) \text{ i.e. } t = \frac{1}{2} - \frac{p(s - s')}{2p'}, \tag{5}$$

which implies:

$$\min [3p(s - s') + 2p'(t - t'); p(s' - s) + 2p'(t' - t)] = p(s - s'),$$

and also: $p^a = 2ps, p^b = 2ps', R^a = R^b = \frac{1}{2}$. So by (1) we have

$$v_{n+1}(p) \geq \min_{0 \leq \varepsilon \leq 1} \max_{\substack{0 \leq s \leq 1 \\ 0 \leq \frac{1}{2} - \frac{p(s-s')}{2p'} \leq 1}} \frac{1}{n+1} \{8\varepsilon + (1-\varepsilon)p(s-s') + n[(1-\varepsilon)v_n(p) + \frac{\varepsilon}{2}(v_n(2ps) + v_n(2ps'))]\}$$

and by induction hypothesis

$$v_{n+1}(p) \geq \min_{0 \leq \varepsilon \leq 1} \max_{\substack{0 \leq s \leq 1 \\ 0 \leq \frac{1}{2} - \frac{p(s-s')}{2p'} \leq 1}} \frac{1}{n+1} \{8\varepsilon + (1-\varepsilon)pp'n^{2/3} + \varphi(s)\}, \tag{6}$$

where

$$\begin{aligned} \varphi(s) &= (1-\varepsilon)p(s-s') + \frac{\varepsilon n^{2/3}}{2}(2ps(2ps)' + 2ps'(2ps)') \\ &= (1-\varepsilon)p(s-s') + \varepsilon p n^{2/3}(1 - 4ps s' - 2p(s-s')^2) \end{aligned}$$

and since $ss' \leq 1/4$ we have

$$\begin{aligned} \varphi(s) &\geq \psi(s) = (1-\varepsilon)p(s-s') + \varepsilon p n^{2/3}(p' - 2p(s-s')^2) \\ &= \varepsilon p p' n^{2/3} + (1-\varepsilon)p(s-s') - 2\varepsilon p^2 n^{2/3}(s-s')^2 \end{aligned} \tag{7}$$

$\psi(s)$ attains its maximum at s_0 such that

$$(s_0 - s'_0) = \frac{1-\varepsilon}{4p\varepsilon n^{2/3}} \quad \text{and} \quad \psi(s_0) = \varepsilon p p' n^{2/3} + \frac{(1-\varepsilon)^2}{8\varepsilon n^{2/3}}. \tag{8}$$

Consider the two cases:

Case a: s_0 is in the domain of maximization of (6) then by (6) and (8):

$$v_{n+1}(p) \geq \min_{0 \leq \varepsilon \leq 1} \frac{1}{n+1} \left\{ 8\varepsilon + pp'n^{2/3} + \frac{(1-\varepsilon)^2}{8\varepsilon n^{2/3}} \right\}.$$

This minimum is attained at $\varepsilon_0 = 1/\sqrt{64n^{2/3} + 1}$, so

$$\begin{aligned} v_{n+1}(p) &\geq \frac{1}{n+1} \left\{ pp'n^{2/3} + \frac{8}{\sqrt{64n^{2/3} + 1}} + \frac{\sqrt{64n^{2/3} + 1}}{8n^{2/3}} - \frac{1}{4n^{2/3}} \right. \\ &\quad \left. + \frac{1}{8n^{2/3}\sqrt{64n^{2/3} + 1}} \right\} \geq \frac{pp'}{n+1} (n^{2/3} + 1/n^{1/3}) \geq pp'/(n+1)^{1/3}, \end{aligned}$$

which is the desired inequality.

Case b: s_0 is out of the domain of maximization of (6). Let us split this case into three subcases:

Case b₁: $s_0 - s'_0 > 1$ (i.e. $p < (1-\varepsilon)/(4\varepsilon n^{2/3})$) and $p \leq p'$.

In this case choose $s - s' = 1$ and $p(s - s') = p'(t' - t)$. By (6) and (7) we then get:

$$v_{n+1}(p) \geq \min_{\left\{\varepsilon \mid p \leq \frac{1-\varepsilon}{4\varepsilon n^{2/3}}\right\}} \frac{1}{n+1} \{pp'n^{2/3} + p + \varepsilon(8 - p - 2p^2 n^{2/3})\} \geq pp'/(n+1)^{1/3}.$$

The last inequality is a result of a careful and straightforward examination.

Case b₂: $s_0 - s'_0 > 1$ and $p > p'$. In this case we choose $t' - t = 1$ and $p(s - s') = p'(t' - t)$. Again by (6) and (7) we get very similarly:

$$\begin{aligned} v_{n+1}(p) &\geq \min_{\left\{\varepsilon \mid p \leq \frac{1-\varepsilon}{4\varepsilon n^{2/3}}\right\}} \frac{1}{n+1} \{pp'n^{2/3} + p' + \varepsilon(8 - p' - 2p'^2 n^{2/3})\} \\ &\geq \min_{\left\{\varepsilon \mid p' \leq \frac{1-\varepsilon}{4\varepsilon n^{2/3}}\right\}} \frac{1}{n+1} \{\dots\} \geq pp'/(n+1)^{1/3}. \end{aligned}$$

Case b₃: $s_0 - s'_0 \leq 1$ and $t_0 = 1/2 - p(s - s')/2p' < 0$, i.e. $p' < (1 - \varepsilon)/(4\varepsilon n^{2/3}) \leq p$. In this case we again choose $t' - t = 1$ and $p(s - s') = p'(t' - t)$ and get the same expression as in the last step of case b_2 . Hence the proof of Lemma 2 is complete.

Lemma 2 asserts that $1/\sqrt{n}$ which is the upper bound for (the order of magnitude of) δ_n in the standard information case is not an upper bound in the general case. The upper bound is at least $1/n^{1/3}$, and this is the only thing that can be said from our example since in this example δ_n is actually of the order of $1/n^{1/3}$ as it is shown by the next theorem.

Theorem 2:

In the game under consideration,

$$v_n(p) = 0^*(1/n^{1/3}).$$

Proof:

First notice that since $\lim v_n(p) = 0$ we have actually $v_n(p) = \delta_n(p)$. Now in view of Lemma 2 the proof of the theorem will be concluded by proving

$$v_n(p) \leq \alpha \sqrt{pp'}/n^{1/3}, \quad \text{where } \alpha = \sqrt[3]{192}. \tag{9}$$

We shall prove (9) by induction on n . For $n = 1$ we have $v_1(p) = \min(p, p') \leq \sqrt{pp'}$ in accordance with (9). Assume (9) holds for n and let us prove it for $n + 1$. First we notice that by $\min(A, B) \leq (1 - p/4)A + (1 + p/4)B$ we get:

$$\min[3p(s - s') + 2p'(t - t'); p(s' - s) + 2p'(t' - t)] \leq 2pp'(s - t). \tag{10}$$

Combining (10) and (1) we have:

$$\begin{aligned} v_{n+1}(p) &\leq \min_{0 \leq \varepsilon \leq 1} \max_{0 \leq s, t \leq 1} \frac{1}{n+1} \{8\varepsilon + 2(1 - \varepsilon)pp'(s - t) + n(1 - \varepsilon)v_n(p) \\ &\quad + n\varepsilon[R^a v_n(p^a) + R^b v_n(p^b)]\}. \end{aligned}$$

Using the induction hypothesis and noticing that by (2)

$$R^a \sqrt{p^a p^{a'}} + R^b \sqrt{p^b p^{b'}} \leq \sqrt{pp'} \left(1 - \frac{(s - t)^2}{4}\right),$$

we get

$$v_{n+1}(p) \leq \frac{\alpha\sqrt{pp'}n^{2/3}}{n+1} + \min_{0 \leq \varepsilon \leq 1} \max_{0 \leq t, s \leq 1} \frac{1}{n+1} \varphi((s-t)),$$

where

$$\varphi((s-t)) = 8\varepsilon + 2pp'(s-t) - \alpha\varepsilon\sqrt{pp'}n^{2/3} \frac{(s-t)^2}{4}$$

$\varphi((s-t))$ attains its maximum at $(s-t)_0 = 4\sqrt{pp'}/(\alpha\varepsilon n^{2/3})$ and $\varphi((s-t)_0) = 8\varepsilon + 4pp'\sqrt{pp'}/(\alpha\varepsilon n^{2/3})$, therefore,

$$v_{n+1}(p) \leq \frac{\alpha\sqrt{pp'}n^{2/3}}{n+1} + \frac{1}{n+1} \min_{0 \leq \varepsilon \leq 1} \left\{ 8\varepsilon + \frac{4pp'\sqrt{pp'}}{\alpha\varepsilon n^{2/3}} \right\}. \tag{12}$$

The expression $\{\dots\}$ as a function of ε attains its minimum at $\varepsilon_0 = (pp')^{3/4}/(2\alpha)^{1/2}n^{1/3} < 1$, so by (12)

$$\begin{aligned} v_{n+1}(p) &\leq \frac{\alpha\sqrt{pp'}n^{2/3}}{n+1} + \frac{1}{n+1} \frac{16(pp')^{3/4}}{(2\alpha)^{1/2}n^{1/3}} \\ &\leq \frac{\alpha\sqrt{pp'}}{n+1} \left(n^{2/3} + \frac{1}{\sqrt{3}n^{1/3}} \right) < \alpha\sqrt{pp'}/(n+1)^{1/3}, \end{aligned}$$

which completes the proof of Theorem 2.

3. An Upper Bound for δ_n

In this section we discuss the order of magnitude of the error term δ_n when $k = 2$, i.e. chance chooses one of two games A^1 and A^2 with the corresponding information matrix H^1 or H^2 respectively.

Let $H = \{h_{ij}^v \mid 1 \leq i \leq l, 1 \leq j \leq m, v = 1, 2\}$. H may be called the information set and it is the set of all possible pieces of information that can be announced by the referee throughout the game. A k -stage history is a k -tuple of elements of H denoted by $A_k = (h_1, \dots, h_k)$. It is to be interpreted as the list of announcements of the referee for the first k stages of the game. A strategy³⁾ of player I in $\Gamma_n(p)$ is an n -tuple $\sigma_n = (f_1, \dots, f_n)$ where $f_k (1 \leq k \leq n)$ is a function from the set of all $k - 1$ stage histories to the set of mixed strategies (of player I) in $\Gamma_1(p)$. A strategy for player II in $\Gamma_n(p)$ is $\tau_n = (g_1, \dots, g_n)$ defined similarly. Denote by $\tau_k^* = (j_1, \dots, j_k)$ the k -tuple of pure strategies actually played by player II in the first k stages. The probability distribution on all possible τ_k^* is determined by τ . Now, σ_n , τ_k^* and p induce a probability distribution on the set of all $k - 1$ stage histories. Therefore, given σ_n and A_{k-1} , player II can compute the conditional probability of the event ‘‘Chance has chosen A^1 ’’. Denote this probability by p_k , i.e.,

$$p_k = \text{prob} \{ \text{chance has chosen } A^1 \mid \sigma_n, \tau_{k-1}^*, A_{k-1} \}.$$

Notice that $E(p_k) = p$ for $k = 1, 2, \dots$ (E denotes expectation).

³⁾ Essentially a behavioural strategy.

The probability p_k is a state variable in the sense that the situation in $\Gamma_n(p)$ before the k^{th} stage is equivalent to that of a new game $\Gamma_{n-k+1}(p_k)$ with $n - k + 1$ stages and a probability p_k for chance to choose A^1 .

In what follows we shall use the expression “given p_k ”, in symbols $(\dots|p_k)$, as a short form of saying: “Given a strategy σ_n of player I and given that before the k^{th} stage, the history was A_{k-1} , and the sequence of pure strategies played by player II up to that stage was τ_{k-1}^* , such that the conditional probability of the event ‘Chance has chosen A^1 ’ is p_k ”.

We denote as usual the one-stage strategy of player I by $\sigma = (s^1, s^2) = [(s_1^1, \dots, s_{l_1}^1), (s_1^2, \dots, s_{l_2}^2)] \in \Sigma^1 \times \Sigma^2$ where Σ^1 and Σ^2 are the spaces of mixed strategies for player I in A^1 and A^2 respectively. (For the general model as well as for what we are going to prove, player I need not have the same number of pure strategies in all A^v .) The non-separating strategies of player I are then given by:

$$NS(0) = NS(1) = \Sigma^1 \times \Sigma^2, \text{ and}$$

$$NS(p) = \left\{ \sigma \mid \sum_{i=1}^{l_1} s_i^1 \varepsilon_{ijh}^1 - \sum_{i=1}^{l_2} s_i^2 \varepsilon_{ijh}^2 = 0; j = 1, \dots, m; h \in H \right\}$$

for $0 < p < 1$, where:

$$\varepsilon_{ijh}^v = \begin{cases} 1 & h_{ij}^v = h \\ 0 & h_{ij}^v \neq h \end{cases}; \quad v = 1, 2.$$

Notice that $NS(p)$ is the same for any p with $0 < p < 1$.

We are now in the position to state our next result:

Theorem 3:

If $NS(p) = \emptyset$ for $0 < p < 1$ then there exists a positive constant c such that $u_n(p) \leq u^*(p) + c/n^{1/4}$ for $n = 1, 2, \dots, 0 \leq p \leq 1$.

Proof:

Note first that since $u^*(p)$ is concave, it is always true that $u^*(p) \geq pu(1) + p'u(0)$. However when $NS(p) = \emptyset$, $u(p)$ is not defined for $0 < p < 1$ and hence $u^*(p) = pu(1) + p'u(0)$ for $0 \leq p \leq 1$.

Proposition 1:

There is a constant $\alpha > 0$ such that given p_k , player II can guarantee $E(H_k|p_k) \leq u^*(p_k) + \alpha p_k p'_k$. (H_k is the payoff at the k^{th} stage.)

To see this, let τ_1 and τ_0 be two optimal strategies (for player II) in A^1 and A^2 respectively. Using τ_1 player II guarantees: $E(H_k|p_k) \leq u(1) + Bp'_k$ where $B = \max_{i,j,v} |a_{ij}^v|$. Similarly τ_0 guarantees $E(H_k|p_k) \leq u(0) + Bp_k$. Thus mixing τ_1 and τ_0 with probabilities p_k and p'_k guarantees:

$$E(H_k|p_k) \leq p_k u(1) + p'_k u(0) + 2Bp_k p'_k \leq u^*(p_k) + \alpha p_k p'_k$$

where $\alpha = 2B$.

We notice that Proposition 1 is valid (and in fact it was proved) in the general case, without the assumption $NS(p) = \emptyset$.

Proposition 2:

There is a constant $\eta > 0$ such that for any $\sigma \in \Sigma^1 \times \Sigma^2$

$$\max_{h \in H} \left| \sum_i s_i^1 \varepsilon_{ijh}^1 - \sum_i s_i^2 \varepsilon_{ijh}^2 \right| \geq \eta. \tag{13}$$

This follows since the left hand side of the inequality is a non-negative continuous function on the compact set $\Sigma^1 \times \Sigma^2$ and therefore it attains its minimum. This minimum, η , cannot be 0 since this would imply $NS(p) = \emptyset$ for $0 < p < 1$.

For any $\sigma \in \Sigma^1 \times \Sigma^2$ let h_σ be the element in H for which the maximum in (13) is attained. Let j_σ be such that $\varepsilon_{ij_\sigma h_\sigma}^v = 1$ for some i and v . For $0 \leq \varepsilon \leq 1$ define a strategy $\tau(\varepsilon)$ for player II as follows: Choose j_σ with probability ε , and with probability $(1 - \varepsilon)$ play the strategy that guarantees $E(H_k | p_k) \leq u^*(p_k) + \alpha p_k p'_k$. By using $\tau(\varepsilon)$ in the k^{th} stage, player II guarantees:

$$E(H_k | p_k) \leq u^*(p_k) + \varepsilon B + \alpha p_k p'_k, \tag{14}$$

where $B = \max_{i,j,v} |a_{ij}^v|$.

When σ and $\tau(\varepsilon)$ are played, the information h_σ will be announced with a probability $p\{h_\sigma\}$ satisfying:

$$p\{h_\sigma\} \geq \varepsilon \left(p_k \sum_i s_i^1 \varepsilon_{ij_\sigma h_\sigma}^1 + p'_k \sum_i s_i^2 \varepsilon_{ij_\sigma h_\sigma}^2 \right).$$

In this event the new conditional probability will be:

$$p_{k+1}(h_\sigma) = \frac{p_k \sum_i s_i^1 \varepsilon_{ij_\sigma h_\sigma}^1}{p_k \sum_i s_i^1 \varepsilon_{ij_\sigma h_\sigma}^1 + p'_k \sum_i s_i^2 \varepsilon_{ij_\sigma h_\sigma}^2}.$$

Therefore, by (13):

$$E(|p_{k+1} - p_k| | p_k) \geq |p_{k+1}(h_\sigma) - p_k| \cdot p\{h_\sigma\} \geq p_k p'_k \eta \varepsilon. \tag{15}$$

Combining (14) and (15) we have

$$E(H_k | p_k) \leq u^*(p_k) + \varepsilon B + \frac{\alpha E(|p_{k+1} - p_k| | p_k)}{\eta \varepsilon}.$$

Taking $\varepsilon_0 = \sqrt{\alpha/(B\eta)} \sqrt{E(|p_{k+1} - p_k| | p_k)}$, and choosing N such that $k \geq N \Rightarrow \varepsilon_0 \leq 1$. (This is possible since $E(|p_{k+1} - p_k| | p_k) \rightarrow 0$) we get:

$$k \geq N \Rightarrow E(H_k | p_k) \leq u^*(p_k) + 2 \sqrt{\frac{\alpha B}{\eta}} \sqrt{E(|p_{k+1} - p_k| | p_k)}.$$

Let $L = \sum_{k=1}^{N-1} E(H_k)$. Using $E(p_k) = p$, the linearity of $u^*(p)$ and the concavity of the function \sqrt{x} we get:

$$\begin{aligned}
 v_n(p) &\leq \frac{E\left(\sum_i^n H_k\right)}{n} \leq u^*(p) + \frac{L}{n} + 2\sqrt{\frac{\alpha B}{\eta}} \frac{\sum_i^n \sqrt{E(|p_{k+1} - p_k|)}}{n} \\
 &\leq u^*(p) + \frac{L}{n} + 2\sqrt{\frac{\alpha B}{\eta}} \frac{\sqrt{n \sum_i^n E(|p_{k+1} - p_k|)}}{n}
 \end{aligned}$$

Since $E\left(\sum_i^k |p_{k+1} - p_k|\right) \leq \sqrt{n}$, (see, AUMANN & MASCHLER, or ZAMIR), we obtain finally:

$$v_n(p) \leq u^*(p) + c/n^{1/4},$$

where $c = 2\sqrt{\alpha B/\eta} + L$, which concludes the proof of Theorem 3.

Lemma 3:

Let S be a non-empty polyhedron in E^m and let $L_i x = 0, i = 1, \dots, d; x = (x_1, \dots, x_m)$ be hyperplanes in E^m . Define $NS = \{x | x \in S; L_i x = 0, i = 1, \dots, d\}$. If $NS \neq \emptyset$ then there exists a constant $D > 0$ depending only on $L_i, i = 1, \dots, d$ and S such that

$$\max_i |L_i x| \geq D \|x - NS\| \quad \text{for all } x \in S. \tag{16}$$

($\|x - NS\| = \min_{y \in NS} \|x - y\|, \|x\|$ being the Euclidean norm in E^m .)

Proof:

Proposition 1:

If $\{x | Lx = 0\}$ is a hyperplane, $Lx_0 = 0$ and $a \neq x_0$, then for each x in the half line $\{x | x = x_0 + \theta(a - x_0); \theta > 0\}$ we have

$$\frac{|Lx|}{\|x - x_0\|} = \frac{|La|}{\|a - x_0\|}.$$

This is demonstrated by:

$$\frac{|Lx|}{\|x - x_0\|} = \frac{|L(x_0 + \theta(a - x_0))|}{\|\theta(a - x_0)\|} = \frac{L\theta a}{\|\theta(a - x_0)\|} = \frac{|La|}{\|a - x_0\|}.$$

Proposition 2:

If NS is a convex set in $E^m, x_0 \in NS, a \notin NS$ such that $\|a - NS\| = \|a - x_0\|$, then for any x in the half line $\{x | x = x_0 + \theta(a - x_0); \theta > 0\}, \|x - NS\| = \|x - x_0\|$ holds.

To prove this we will show that the hyperplane through x_0 and perpendicular to the half line, separates a from NS . In other words we show that

$$(a - x_0)(y - x_0) \leq 0 \quad \forall y \in NS. \tag{17}$$

Assume that there is $z \in NS$ such that

$$(a - x_0)(z - x_0) > 0, \tag{18}$$

then by the definition of x_0 :

$$\begin{aligned} \|a - x_0\|^2 &< \|a - z\|^2 = [(a - x_0) + (x_0 - z)](a - z) \\ &= \|a - x_0\|^2 + (a - x_0)(x_0 - z) - (a - z)(z - x_0) \\ &< \|a - x_0\|^2 - (a - z)(z - x_0), \end{aligned}$$

hence

$$(a - z)(z - x_0) < 0. \tag{19}$$

From (18) and (19) it follows that there is z^* , $z^* = \theta x_0 + (1 - \theta)z$, $0 < \theta < 1$ such that

$$(a - z^*)(z - x_0) = 0. \tag{20}$$

Since NS is convex it follows that $z^* \in NS$ and

$$\begin{aligned} \|a - z^*\|^2 &= (a - z^*) [(a - z^*) + (1 - \theta)(z - x_0)] \\ &= (a - z^*)(a - x_0) = [(a - x_0) - (1 - \theta)(z - x_0)](a - x_0) \\ &= \|a - x_0\|^2 - (1 - \theta)(a - x_0)(z - x_0) < \|a - x_0\|^2, \end{aligned}$$

which contradicts the definition of x_0 and hence proves Proposition 2.

Let us prove now Lemma 3 by induction on the dimension of S : If $\dim S = 0$ (S is one point) there is nothing to prove. If $\dim S = 1$, i.e., $S = \{x | x = \theta a + (1 - \theta)b; 0 \leq \theta \leq 1\}$ then (16) follows easily from Proposition 1. Assume that the lemma is true for $\dim S \leq k < m$ and it is wrong for $\dim S = k + 1$. Then there is a sequence $\{x_n\}$ of points in S such that:

$$|L_i x_n| < \frac{1}{n} \|x_n - NS\|, \quad i = 1, \dots, d; n = 1, 2, \dots \tag{21}$$

Let y_n be the point in NS nearest to x_n , i.e.:

$$\|x_n - NS\| = \|x_n - y_n\|, \quad n = 1, 2, \dots$$

Define:

$$\begin{aligned} \theta_n &= \max \{ \theta | y_n + \theta(x_n - y_n) \in S \} \\ x'_n &= y_n + \theta_n(x_n - y_n). \end{aligned}$$

Clearly each x'_n is on a face of S and since S has a finite number of faces we can assume (by taking a subsequence) that x'_n are on one face, say π , of S for $n = 1, 2, \dots$

Now by Proposition 1 $|L_i x'_n| / \|x'_n - y_n\| = |L_i x_n| / \|x_n - y_n\|$, and by Proposition 2: $\|x'_n - NS\| = \|x'_n - y_n\|$. Hence by (21):

$$|L_i x'_n| < \frac{1}{n} \|x'_n - NS\|, \quad i = 1, \dots, d; n = 1, 2, \dots \tag{22}$$

Let x_0 be a limit point for a subsequence of $\{x'_n\}$, then since $\|x'_n - NS\|$ is bounded we have: $L_i x_0 = 0$ for $i = 1, \dots, d$; which implies $x_0 \in NS$. On the other hand since the face π is closed we have also $x_0 \in \pi$. So if we define $NS' = NS \cap \pi$ we have $NS' \neq \emptyset$. Now clearly $\|x'_n - NS\| \leq \|x'_n - NS'\|$, hence:

$$|L_i x'_n| < \frac{1}{n} \|x'_n - NS'\|, \quad i = 1, \dots, d; n = 1, 2, \dots; \tag{23}$$

which contradicts the induction hypothesis since π is a polyhedron of dimension k . This completes the proof of Lemma 3.

Theorem 4:

For any game $\Gamma_n(p)$ with $k = 2$, there exists a constant c such that

$$v_n(p) \leq u^*(p) + c/n^{1/6}, \tag{24}$$

for $0 \leq p \leq 1$ and $n = 1, 2, \dots$

Proof:

If $NS(p) = \emptyset$, $0 < p < 1$, then (24) follows from Theorem 3. Assume therefore that $NS(p) = NS \neq \emptyset$ for $0 < p < 1$, then by Lemma 3 there is a constant $D > 0$ such that for any $\sigma \in \Sigma^1 \times \Sigma^2$ we have:

$$\max_h \left| \sum_i s_i^1 \varepsilon_{ijh}^1 - \sum_i s_i^2 \varepsilon_{ijh}^2 \right| \geq D \|\sigma - NS\|, \tag{25}$$

for $0 < p < 1$.

Proposition 1:

Given p_k , player II has a pure strategy j_0 such that when used with probability ε it guarantees $E(|p_{k+1} - p_k| | p_k) \geq \varepsilon p_k p'_k \|\sigma - NS\|$ for any $\sigma \in \Sigma^1 \times \Sigma^2$ played by player I.

This can be shown by very much the same argument used to prove (15) in the proof of Theorem 3 (namely j_0 is the column in which is found h_σ that maximizes the left hand side of (25).)

Proposition 2:

Player II can guarantee:

$$E(H_k | p_k) \leq u^*(p_k) + B \|\sigma - NS\|, \tag{26}$$

where $B = \max_{i,j,v} |a_{ij}^v|$.

In fact he can do this by playing an optimal strategy in $\Delta(p_k)$.

A combination of Proposition 2 and Proposition 1 in the proof of Theorem 3 gives:

Proposition 3:

Given p_k , player II has a strategy τ_* that guarantees $E(H_k | p_k) \leq u^*(p_k) + \min(B \|\sigma - NS\|, \alpha p_k p'_k)$, and hence:

$$E(H_k | p_k) \leq u^*(p_k) + (B\alpha)^{1/2} (p_k p'_k)^{1/2} \|\sigma - NS\|^{1/2}. \tag{27}$$

For $0 \leq \varepsilon \leq 1$ define $\tau_*(\varepsilon) = \varepsilon j_0 \ddagger (1 - \varepsilon) \tau_*$ (j_0 is defined by Proposition 1 and \ddagger is a formal probabilistic mixture). Playing $\tau_*(\varepsilon)$ at stage k player II guarantees:

$$E(H_k | p_k) \leq u^*(p_k) + \varepsilon B + (B\alpha)^{1/2} (p_k p'_k)^{1/2} \|\sigma - NS\|^{1/2},$$

and by Proposition 1:

$$E(H_k | p_k) \leq u^*(p_k) + \varepsilon B + \frac{\bar{c} [E(|p_{k+1} - p_k| | p_k)]^{1/2}}{\varepsilon^{1/2}}, \tag{28}$$

where $\bar{c} = (\alpha B/D)^{1/2}$. Now minimize the right hand side of (28) by choosing $\varepsilon_0 = (\bar{c}/(2B))^{2/3} [E(|p_{k+1} - p_k| | p_k)]^{1/3}$, for $k \geq N$ (where N is such that $k \geq N \Rightarrow \varepsilon_0 \leq 1$), we get:

$$k \geq N \Rightarrow E(H_k | p_k) \leq u^*(p_k) + \bar{D} [E(|p_{k+1} - p_k| | p_k)]^{1/3}, \tag{29}$$

where $\bar{D} = 2B\bar{c}$. Letting $L = \sum_1^{N-1} E(H_k)$ we get by using the concavity of $u^*(p)$ and of $x^{1/3}$:

$$v_n(p) \leq \frac{1}{n} \sum_1^n E(H_k) \leq u^*(p) + \frac{L}{n} + \frac{\bar{D}}{n} \sum_1^n [E(|p_{k+1} - p_k|)]^{1/3}. \tag{30}$$

By HÖLDER inequality and since $E\left(\sum_1^n |p_{k+1} - p_k|\right) \leq n^{1/2}$:

$$\begin{aligned} \sum_1^n [E(|p_{k+1} - p_k|)]^{1/3} &\leq \left[\sum_1^n E(|p_{k+1} - p_k|) \right]^{1/3} n^{2/3} \\ &\leq [n^{1/2}]^{1/3} n^{2/3} = n^{5/6}. \end{aligned}$$

Inserting this in (30) and setting $c = \bar{D} + L$ we obtain finally $v_n(p) \leq u^*(p) + c/n^{1/6}$ which completes the proof of Theorem 4.

4. Some Open Questions

There are several questions still to be answered:

1) What is the upper bound for δ_n when $k > 2$? Quite surprisingly, the generalization of our results for $k = 2$ is not straightforward.

2) What is the least upper bound for the order of magnitude of δ_n ? This is interesting even when $k = 2$ since there is a gap between our bound $1/n^{1/6}$ and the order of magnitude of δ_n in our example which is $1/n^{1/3}$. In fact we conjecture that the least upper bound is lower than $1/n^{1/6}$.

3) Even for $k = 2$ and $NS(p) = \emptyset$ it is not known whether the bound $1/n^{1/4}$ is the least upper bound.

4) In what way does the error term depends on the information matrices of both players? In this work we assumed that the information announced by the referee was the same for both players. More generally this information may be different for each player. AUMANN and MASCHLER [1968] showed that the information to player I (the one who knows the choice of chance) is irrelevant as far as the value of the game is concerned. We feel that it may play an important role in the speed of convergence, namely in δ_n .

References

AUMANN, R. J., and M. MASCHLER: Games Theoretic Aspects of Gradual Disarmament. Report to the U.S.A.C.D.A. (Arms Control and Disarmament Agency, Washington, D.C.) Final report on Contract ACDA/ST-80, prepared by Mathematica, Princeton, New Jersey, Chapter V, June 1966.

AUMANN, R. J., and M. MASCHLER: Repeated Games of Incomplete Information: The Zero-Sum Extensive Case. Report to the U.S.A.C.D.A. Final report on contract ACDA/ST-143, prepared by Mathematica, Princeton, New Jersey, Chapter III, December 1968.

ZAMIR, S.: On the Relation Between Finitely and Infinitely Repeated Games with Incomplete Information. *International Journal of Game Theory* **1**, No. 3, 1971/72.

Received October, 1972 (revised version July, 1973).