Chapter 5

REPEATED GAMES OF INCOMPLETE INFORMATION: ZERO-SUM

SHMUEL ZAMIR

The Hebrew University of Jerusalem

Contents

1. Introduction

This chapter and the next apply the framework of repeated games, developed in the previous chapter, to games of incomplete information. The aim of this combination is to analyze the strategic aspects of information: When and at what rate to reveal information? When and how should information be concealed? What resources should be allocated to acquiring information? Repeated games provide the natural paradigm for dealing with these dynamic aspects of information. The repetitions of the game serve as a signaling mechanism which is the channel through which information is transmitted from one period to another.

It may be appropriate at this point to clarify the relation of repeated incomplete information games to stochastic games, treated in a forthcoming volume of this Handbook. Both are dynamic models in which payoffs at each stage are determined by the state of nature (and the player's moves). However, in stochastic games the state of nature changes in time but is common knowledge to all players, while *in repeated games of incomplete information the state of nature is fixed but not known to all players*. What changes in time is each player's knowledge about the other players' past actions, which affects his beliefs about the (fixed) state of nature. But it should be mentioned that it is possible to provide a general model which has both stochastic games and repeated games of incomplete information as special cases [see Mertens, Sorin and Zamir (1993, ch. IV), henceforth MSZ].

An important feature of the analysis is that when treating any specific game one has to consider a whole *family of games*, parameterized by the prior distribution on the states of nature. This is so because the state of information, which is basically part of the initial data of the game, changes during the play of the repeated game.

Most of the work on repeated games of incomplete information was done for two-person, zero-sum games, which is also the scope of this chapter. This is not only because it is the simplest and most natural case to start with, but also because it captures the main problems and aspects of strategic transmission of information, which can therefore be studied "isolated" from the phenomena of cooperation, punishments, incentives, etc. Furthermore, the theory of non-zero-sum repeated games of incomplete information makes extensive use of the notion of punishment, which is based on the minmax value borrowed from the zero-sum case.

1.1. Illustrative examples

Before starting our formal representation let us look at a few examples illustrating some of the main issues of strategic aspects of information. We start with an example, studied very extensively by Aumann and Maschler (1966, 1967):

Example 1.1. Imagine two players I (the maximizer) and II (the minimizer) playing repeatedly a zero-sum game given by a 2×2 payoff matrix. This matrix is chosen (once and for all) at the beginning to be either G^1 or G^2 where

$$G^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Player I is told which game was chosen but player II is not; he only knows that it is either G^1 or G^2 with equal probabilities and that player I knows which one is it. After the matrix is chosen the players repeatedly do the following: player I chooses a row, player II chooses a column (simultaneously). A referee announces these moves and records the resulting payoff (according to the matrix chosen at the beginning). He *does not announce the payoffs* (though player I of course knows them, knowing the moves and the true matrix). The game consists of n such stages and we assume that n is very large. At the end of the nth stage player I receives from player II the total payoff recorded by the referee divided by n (to get an average payoff per stage in order to be able to compare games of different length).

So player I has the advantage of knowing the real state (the real payoff matrix). How should he play in this game?

A first possibility is to choose the dominating move in each state: always play Top if the game is G^1 and always play Bottom if it is G^2. Assuming that I announces this strategy (which we may as well assume by the minmax theorem), it is a *completely revealing* strategy since player II will find out which matrix has been chosen by observing whether player I is playing Top or Bottom. Having found this, he will then choose the appropriate column (Right in G^1 and Left in G^2) to pay only 0 from then on. Thus, a completely revealing strategy yields the informed player an average payoff of almost 0 (that is a total payoff of at most 1, at the first stage after which the matrix is revealed, thus an average of at most $1/n$).

Another possible strategy for player I is to play *completely non-revealing*, that is, ignoring his private information and playing as if he, just like player II, does not know the matrix chosen. This situation is equivalent to repeatedly playing the average matrix game:

$$D = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

This game (which may also be called the non-revealing game) has a value $1/4$ and player I can guarantee this value (in the original game) by always playing Top and Bottom with equal probabilities ($1/2$ each) *independently of what the true matrix is*.

So, strangely enough, in this specific example the informed player is better off not using his information than using it. As we shall see below, the completely non-revealing strategy is in fact the (asymptotically) optimal strategy for player I; in very long games he cannot guarantee significantly more than $1/4$ per stage.

Example 1.2. The second example has the same description as the first one except that the two possible matrices are now

$$G^1 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Following the line of discussion of the previous example, if player I uses his dominating move at each state, Bottom in G^1 and Top in G^2, he will guarantee a payoff of 0 at each stage and again this will be a completely revealing strategy. Unlike in the previous example, here this strategy is an optimal strategy for the informed player. This is readily seen without even checking other strategies: 0 is the highest payoff in both matrices. Just for comparison, the completely non-revealing strategy would yield the value of the non-revealing game

$$D = \begin{pmatrix} -1/2 & 0 \\ 0 & -1/2 \end{pmatrix},$$

which is $-1/4$.

Example 1.3. Consider again a game with the same description as the previous two examples, this time with the two possible matrices given by

$$G^1 = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}.$$

Playing the dominant rows (Top in G_1 and Bottom in G_2) is again a completely revealing strategy which leads to a long-run average payoff of (almost) 0 (the value of each of the matrices is 0). Playing completely non-revealing leads to the non-revealing game

$$D = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix},$$

which has a value 0. So, both completely revealing and completely non-revealing strategies yield the informed player an average payoff of 0. Is there still another, more clever, way of using the information to guarantee more than 0? If there is, it must be a strategy which *partially reveals the information*. In fact such a strategy exists. Here is how an average payoff of 1 per stage can be guaranteed by the informed player.

Player I prepares two non-symmetric coins, both with sides (T, B). In coin C^1 the corresponding probabilities are (3/4, 1/4) while in coin C^2 they are (1/4, 3/4). Then he plays the following strategy: if the true matrix is G^k, $k = 1, 2$, flip coin C^k. Whichever coin was used, if the outcome is T play Top in all stages, if it is B, play Bottom in all stages.

To see what this strategy does, let us assume that even player II knows it. He will then, right after the first stage, know whether the outcome of the coin was T or B (by observing whether player I played Top or Bottom). He will not know which coin was flipped (since he does not know the state k). However, he can update his beliefs about the probability of each matrix in view of a given outcome of the coin. Using Bayes' formula we find

$$P(G^1 \mid T) = 3/4 \quad \text{and} \quad P(G^1 \mid B) = 1/4.$$

Given that player I is playing Top, the payoffs will be either according to the line (4, 0, 2) (this with probability 3/4) or according to the line (0, 4, −2) (and this with probability 1/4). The expected payoffs given Top are therefore (depending on the move of player II)

$$(3/4)(4, 0, 2) + (1/4)(0, 4, -2) = (3, 1, 1).$$

Similarly, given that player I is playing Bottom, the expected payoffs are

$$(1/4)(4, 0, -2) + (3/4)(0, 4, 2) = (1, 3, 1).$$

We conclude that in any event, and no matter what player II does, the conditional expected payoff is at least 1 per stage. Therefore the expected average payoff for player I is at least 1. We shall see below that this is the most player I can guarantee in this game. So the optimal strategy of the informed player in this example is partially to reveal his information.

Let us have a closer look at this strategy. In what sense is it partially revealing? Player I, when being in G^1, *will more likely* (namely with probability 3/4) play Top and when being in G^2, he will more likely play Bottom. Therefore when Top is played it becomes more likely that the matrix is G^1,

while when Bottom is played it becomes more likely that it is G^2. Generally, player I is giving player II information in the right direction, but it is not definite; player II will adjust his *beliefs* about the true matrix from (1/2, 1/2) to either (3/4, 1/4) or (1/4, 3/4) and with probability 3/4 this adjustment will be *in the right direction*, increasing the subjective probability for the true game. This idea of changing a player's beliefs by giving him a signal which is partially correlated with the true state is undoubtedly the heart of the theory of games with incomplete information.

There is one point we wish to add about the notion of revealing. In all three examples we discussed the informed player was *revealing* information when- ever he was *using* it. However, in principle, and in fact in the general model which will be presented below, these are two distinct concepts. Using informa- tion means to play differently in two different informational states; for instance, in our first example player I was using his information when he was playing Top in G^1 and Bottom in G^2. Revealing information is changing the beliefs of the uninformed player. Clearly, when the move of the informed player is observed by the uninformed player – which we shall later call *the full monitoring case* – the two concepts are two expressions of the same thing; the only way to play non-revealing is to play the same way, independently of one's information, i.e. not to use the information. This was the case in all our examples. More generally, the move of the informed player need not be observable. The uninformed player receives some *signal* which is correlated in an arbitrary way with the move of the informed player. It may then well be that in order to play non-revealing, a player *has to use his information*. Similarly, he may be revealing his information by not using it.[1]

2. A general model

A repeated game of incomplete information consists of the following elements:
- A finite set I, the set of players.
- A finite set K, the set of states of nature.
- A probability distribution p on K, the *prior* probability distribution on the states of nature.
- For each $i \in I$, a partition K^i of K, the initial information of player i.
- For each $i \in I$ and $k \in K$, a finite set S_k^i, which is the same for all k in the same partition element of K^i. This is the set of moves available to player I at state k. By taking the Cartesian product $\Pi_k S_k^i$ we may assume, without loss of generality, that the sets of moves S^i are state independent. Let $S = \Pi_i S^i$.

---

[1]For examples, see MSZ (1993, ch. V, section 3.b).

- For each $k \in K$, a payoff function $G^k: S \to R^I$. That is $G^k(s)$ is the vector of payoffs to the players when they play moves s and the state is k.
- For each $i \in I$, a finite alphabet A^i, the set of *signals* to player i. Let $A = \Pi_i A^i$ and $\mathcal{A} = \Delta(A)$ be the set of probability distributions on A.
- A transition probability Q from $K \times S$ to $\mathcal{A}$, the signaling probability distribution [we use the notation $Q^k(s)$ for the image of (k, s)].

On the basis of these elements the repeated game (or supergame) is played in stages as follows. At stage 0 a chance move chooses an element $k \in K$ according to the probability distribution p. Each player i is informed of the element of K^i containing the chosen k. Then at each stage m ($m = 1, 2, \ldots$), each player i chooses $s_m^i \in S^i$, a vector of signals $a \in A$ is chosen according to the probability distribution $Q^k(s)$ and a^i is communicated to player i. This is the signal to player i at stage m.

Notice that, as mentioned in the Introduction, the state k is chosen at stage 0 and remains fixed for the rest of the game. This is in contrast to stochastic games (to be discussed in a forthcoming volume of this Handbook) in which the state may change along the play. Also note that the payoff g_m at stage m is not explicitly announced to the players. In general, on the basis of the signals he receives, a player will be able to deduce only partial information about his payoffs.

2.1. Classification

Games of incomplete information are usually classified according to the nature of the three important elements of the model, namely players and payoffs, prior information, and signaling structure.

(1) *Players and payoffs.* Here we have the usual categories of two-person and n-person games. Within the two-person games one has the zero-sum games treated in this chapter and the non-zero-sum games treated in Chapter 6 of this Handbook.

(2) *Prior information.* Within two-person games the main classification is games with incomplete information on *one side*, versus incomplete information on *two sides*. In the first class are games in which one player knows the state chosen at stage 0 (i.e. his prior information partition consists of the singletons in K) while the other player gets no direct information at all about it (i.e. his prior information partition consists of one element $\{K\}$). More general prior information may sometimes be reduced to this case, for example if one of the player's partition is a refinement of the other's partition, and the signaling distribution Q, as a function of k, is measurable with respect to the coarser partition.

(3) *Signaling structure.* The simplest and most manageable signaling struc-

ture is that of *full monitoring*. This is the case in which the moves at each stage are the only observed signals by all players, that is $A^i = S$ for all i, and for all k and all s, $Q^k(s)$ is a probability with mass 1 at $(s, \ldots, s)$. The next level of generality is that of *state independent signals*. This is the case in which $Q^k(s)$ is constant in k, and consequently the signals do not reveal any direct information about the state but only about the moves. Hence the only way for a player to get information about k is by deducing it from other players' moves, about which he learns something through the signals he receives. There is no established classification beyond that, although two other special classes will be treated separately: the case in which the signals are the same for all players and include full monitoring (and possibly more information), and the case in which the signal either fully reveals the state to all players or is totally non-informative.

3. Incomplete information on one side

In this section we consider repeated two-person, zero-sum games in which only one player knows the actual state of nature. These games were first studied by Aumann, Maschler and Stearns, who proved the main results. Later contributions are due to Kohlberg, Mertens and Zamir.

Since in this chapter we consider only two-person, zero-sum games it is convenient to slightly modify our notation for this case by referring to the two players as player I (the maximizer) and player II (the minimizer). Their sets of pure actions (or moves) are S and T, respectively, and their corresponding mixed moves are $X = \Delta(S)$ and $Y = \Delta(T)$. The payoff matrix (to I) at state $k \in K$ is denoted by G^k with elements G^k_{st}. The notation for general signaling will be introduced later.

3.1. General properties

We shall consider the games $\Gamma_n(p)$, $\Gamma_\lambda(p)$ and $\Gamma_\infty(p)$ which are defined with the appropriate valuation of the payoffs sequence $(g_m)^\infty_{m=1} = (G^k(s_m))^\infty_{m=1}$. Before defining and analyzing these we shall first establish some general properties common to a large family of games with incomplete information on one side.

The games considered in this section will all be zero-sum, two person games of the following form: chance chooses an element k from the set K of states (games) according to some $p \in \Delta(K)$. Player I is informed which k was chosen but player II is not. Players I and II then, simultaneously, choose $\sigma^k \in \Sigma$ and $\tau \in \mathcal{T}$, respectively, and finally $G^k(\sigma^k, \tau)$ is paid to player I by player II. The

sets Σ and $\mathcal{T}$ are convex sets of strategies, and the payoff functions $G^k(\sigma^k, \tau)$ are bilinear and uniformly bounded on $\Sigma \times \mathcal{T}$. We may think of k as the *type* of player I which is some private information known only to him and could attain various values in K. This is thus a game of *incomplete information on one side*, on the side of player II.

Even though the strategies in Σ and $\mathcal{T}$ will usually be strategies in some repeated game (finite or infinite), it is useful at this point to consider the above described game as a one-shot game in strategic form in which the strategies are $\sigma \in \Sigma^K$ and $\tau \in \mathcal{T}$, respectively, and the payoff function is $G^P(\sigma, \tau) = \Sigma_k p^k G^k(\sigma^k, \tau)$. Denote this game by $\Gamma(p)$.

Theorem 3.1. *The functions* $\bar{w}(p) = \inf_\tau \sup_\sigma G^P(\sigma, \tau)$ *and* $\underline{w}(p) = \sup_\sigma \inf_\tau G^P(\sigma, \tau)$ *are concave.*

Proof. The argument is the same for both functions. We will show it for $\bar{w}(p)$. Let $p = (p_e)_{e \in E}$ be finitely many points in $\Delta(K)$, and let $\alpha = (\alpha_e)_{e \in E}$ be a point in $\Delta(E)$ such that $\Sigma_{e \in E} \alpha_e p_e = p$. We claim that $\bar{w}(p) \geqslant \Sigma_{e \in E} \alpha_e \bar{w}(p_e)$. To see this, consider the following two-stage game: chance chooses $e \in E$ according to the probability distribution α, then $k \in K$ is chosen according to p_e, the players then choose $\sigma^k \in \Sigma$ and $\tau \in \mathcal{T}$, respectively, and the payoff is $G^k(\sigma^k, \tau)$. We consider two versions, in both of which player I is informed of everything (both e and k) while player II may or may not be informed of the outcome e (in any case he is not informed of the value of k).

Now, if player II is informed of the outcome e the situation following the first lottery is equivalent to $\Gamma(p_e)$. Thus, the $\inf_\tau \sup_\sigma$ for the game in which player II is informed of the outcome of the first stage is $\Sigma_{e \in E} \alpha_e \bar{w}(p_e)$. This game is more favorable to II than the game in which he is not informed of the value of e, which is equivalent to $\Gamma(\Sigma_e \alpha_e p_e) = \Gamma(p)$. Therefore we have:

$$\bar{w}(p) \geqslant \sum_{e \in E} \alpha_e \bar{w}(p_e). \quad \square$$

Remark. Although this theorem is formulated for games with incomplete information on one side it has an important consequence for games with incomplete information on both sides. This is because we did not assume anything about the strategy set of player II. In a situation of incomplete information on both sides, when a pair of types, one for each player, is chosen at random and each player is informed of his type only, we can still think of player II as being "uninformed" (of the type of I) but with strategies consisting of choosing an action after observing a chance move (the chance move choosing his type). When doing this, we can use Theorem 3.1 to obtain the concavity of $\bar{w}(p)$ and $\underline{w}(p)$ in games with incomplete information on both

sides, when p (the joint probability distribution on the pairs of types) is restricted to the subset of the simplex where player I's conditional probability on the state k, given his own type, is fixed.

The concavity of $\underline{w}(p)$ can also be proved constructively by means of the following useful proposition, which we shall refer to as the *splitting procedure*.

Proposition 3.2. *Let p and $(p_e)_{e \in E}$ be finitely many points in $\Delta(K)$, and let $\alpha = (\alpha_e)_{e \in E}$ be a point in $\Delta(E)$ such that $\Sigma_{e \in E} \alpha_e p_e = p$. Then there are vectors $(\mu^k)_{k \in K}$ in $\Delta(E)$ such that the probability distribution P on $K \times E$ obtained by the composition of p and $(\mu^k)_{k \in K}$ (that is $k \in K$ is chosen according to p and then $e \in E$ is chosen according to μ^k) satisfies*

$$P(\cdot \mid e) = p_e \quad and \quad P(e) = \alpha_e , \quad for \ all \ e \in E .$$

Proof. In fact, if $p^k = 0$, μ^k can be chosen arbitrarily in $\Delta(E)$. If $p^k > 0$, μ_k is given by $\mu^k(e) = \alpha_e p_e^k / p^k$. $\square$

Let player I use the above described lottery and then guarantee $\underline{w}(p_e)$ (up to ϵ). In this way he guarantees $\Sigma_e \alpha_e \underline{w}(p_e)$, even if player II were informed of the outcome of the lottery. So $\underline{w}(p)$ is certainly not smaller than that. Consequently the function $\underline{w}(p)$ is concave.

The idea of splitting is the following. Recall that the informed player, I, knows the state k while the uninformed player, II, knows only the probability distribution p according to which it was chosen. Player I can design a state dependent lottery so that if player II observes only the outcome e of the lottery, his conditional distribution (i.e. his new "beliefs") on the states will be p_e. Let us illustrate this using Example 1.3. At $p = (1/2, 1/2)$ player I wants to "split" the beliefs of II to become $p_1 = (3/4, 1/4)$ or $p_2 = (1/4, 3/4)$ (note that $p = 1/2 p_1 + 1/2 p_2$.) He does this by the state dependent lottery on $\{T, B\}$: $\mu^1 = (3/4, 1/4)$ and $\mu^2 = (1/4, 3/4)$.

Another general property worth mentioning is the Lipschitz property of all functions of interest (such as the value functions of the discounted game, the finitely repeated game, etc.), in particular $\bar{w}(p)$. This follows from the uniform boundedness of the payoffs, and hence is valid for all repeated games discussed in this chapter.

Theorem 3.3. *The function $\bar{w}(p)$ is Lipschitz with constant C (the bound on the absolute value of the payoffs).*

Proof. Indeed, the payoff functions of two games $\Gamma(p_1)$ and $\Gamma(p_2)$ differ by at most $C\|p_1 - p_2\|_1$. $\square$

Let us turn now to the special structure of the repeated game. Given the

basic data $(K, p, (G^k)_{k \in K}, A, B, Q)$ (here A and B are the signal sets of I and II, respectively), any play of the game yields a payoff sequence $(g_m)_{m=1}^{\infty} = (G^k(s_m t_m))_{m=1}^{\infty}$. On the basis of various valuations of the payoff sequence, we shall consider the following games (as usual, E denotes expectation with respect to the probability induced by p, Q, and the strategies).

The *n*-stage game, $\Gamma_n(p)$, is the game in which the payoff is $\bar{\gamma}_n = E(\bar{g}_n) = E((1/n) \sum_{m=1}^{n} g_m)$. Its value is denoted by $v_\lambda(p)$.

The λ-discounted game, $\Gamma_\lambda(p)$ (for $\lambda \in (0, 1]$), is the game in which the payoff is $E(\sum_{m=1}^{\infty} \lambda(1 - \lambda)^{m-1} g_m)$. Its value is denoted by $v_\lambda(p)$.

The values $v_n(p)$ and $v_\lambda(p)$ clearly exist and are Lipschitz by Theorem 3.3.

As in the previous section, the infinite game $\Gamma_\infty(p)$ is the game in which the payoff is some limit of $\bar{g}_n$ such as lim sup, lim inf or, more generally, any Banach limit $\mathscr{L}$. It turns out that the results in this chapter are independent of the particular limit function chosen as a payoff. The definition of the value for $\Gamma_\infty(p)$ is based on a notion of guaranteeing.

Definition 3.4. (i) Player I can *guarantee* α if

$$\forall \epsilon > 0, \exists \sigma_\epsilon, \exists N_\epsilon, \text{ such that } \bar{\gamma}_n(\sigma_\epsilon, \tau) \geq \alpha - \epsilon, \forall \tau, \forall n \geq N_\epsilon .$$

(ii) Player II can *defend* α if

$$\forall \epsilon > 0, \forall \sigma, \exists \tau, \exists N, \text{ such that } \bar{\gamma}_n(\sigma, \tau) \leq \alpha + \epsilon, \forall n \geq N .$$

$\underline{v}(p)$ is the *maxmin* of $\Gamma_\infty(p)$ if it can be guaranteed by player I and can be defended by player II. In this case a strategy σ_ϵ associated with $\underline{v}(p)$ is called ϵ-*optimal*. The *minmax* $\bar{v}(p)$ and ϵ-optimal strategies for player II are defined in a dual way. A strategy is optimal if it is ϵ-optimal for all ϵ.

The game $\Gamma_\infty(p)$ has a *value* $v_\infty(p)$ iff $\underline{v}(p) = \bar{v}(p) = v_\infty(p)$. It follows readily from these definitions that:

Proposition 3.5. *If $\Gamma_\infty(p)$ has a value $v_\infty(p)$, then both $\lim_{n \to \infty} v_n(p)$ and $\lim_{\lambda \to 0} v_\lambda(p)$ exist and they are both equal to $v_\infty(p)$. An ϵ-optimal strategy in $\Gamma_\infty(p)$ is an ϵ-optimal strategy in all $\Gamma_n(p)$ with sufficiently large n and in all $\Gamma_\lambda(p)$ with sufficiently small λ.*

By the same argument used in Theorem 3.1 or by using the splitting procedure of Proposition 3.2 we have:

Proposition 3.6. *In any version of the repeated game $(\Gamma_n(p), \Gamma_\lambda(p)$ or $\Gamma_\infty(p))$, if player I can guarantee $f(p)$ then he can also guarantee* Cav $f(p)$.

Here Cav f is the (pointwise) smallest concave function g on $\Delta(K)$ satisfying $g(p) \geq f(p)$, $\forall p \in \Delta(K)$. We now have:

Theorem 3.7. $v_n(p)$ *and* $v_\lambda(p)$ *converge uniformly (as* $n \to \infty$ *and* $\lambda \to 0$, *respectively) to the same limit which can be defended by player II in* $\Gamma_\infty(p)$.

Proof. Let τ_n be an ϵ-optimal strategy of player II in $\Gamma_n(p)$ with $\epsilon = 1/n$ and let $v_{n_i}(p)$ converge to $\liminf_{n \to \infty} v_n(p)$. Now let player II play τ_{n_i} for n_{i+1} times $(i = 1, 2...)$ – thus for $n_i \cdot n_{i+1}$ periods – before increasing i by 1. By this strategy player II guarantees $\liminf_{n \to \infty} v_n(p)$. Since player II certainly cannot guarantee less than $\limsup_{n \to \infty} v_n(p)$, it follows that $v_n(p)$ converges (uniformly by Theorem 3.3).

As for the convergence of v_λ, since clearly player II cannot guarantee less than $\limsup_{\lambda \to 0} v_\lambda(p)$, the above described strategy of player II proves that

$$\limsup_{\lambda \to 0} v_\lambda(p) \leq \lim_{n \to \infty} v_n(p).$$

To complete the proof we shall prove that $\lim_{n \to \infty} v_n(p) \leq \liminf_{\lambda \to 0} v_\lambda(p)$ by showing that $\lim_{n \to \infty} v_n(p) \leq v_\lambda(p)$ for any $\lambda > 0$. In fact, given $\lambda > 0$ let τ_λ be an optimal strategy of player II in the λ-discounted game and consider the following strategy (for player II): start playing τ_λ and at each stage restart τ_λ with probability λ and with probability $(1 - \lambda)$ continue playing the previously started τ_λ. With this strategy, for any $\epsilon > 0$, we have $E(\bar{g}_n) \leq v_\lambda + \epsilon$ for all n sufficiently large (compared with $1/\lambda$). It follows that $\lim_{n \to \infty} v_n(p) \leq v_\lambda(p)$. $\square$

Remark. If we interpret the discounted game as a repeated game with a probability λ of stopping after each stage, then the convergence of v_λ can be generalized as follows. Let $a = \{a_n\}_{n=1}^{\infty}$ be a probability distribution, with finite expectation, of T – the stopping time of the game – and let $v_a(p)$ be the value of this repeated game. If $\{a^l\}_{l=1}^{\infty}$ is a sequence of such distributions with mean going to infinity, then $\lim_{l \to \infty} v_{a^l}(p) = \mathrm{Cav}\, u(p)$.

3.2. Full monitoring

The first model we consider is that of incomplete information on one side and with full monitoring. This is the case when the moves of the players at each stage are observed by both of them and hence they serve as the (only) device for transmitting information about the state of nature. The repeated game with the data $(K, p, S, T, (G^k)_{k \in K})$ is denoted by $\Gamma(p)$ and is played as follows.

At stage 0 a chance move chooses $k \in K$ with probability distribution $p \in \Delta(K)$, i.e. p^k is the probability of k. The result is told to player I, the row chooser, but not to the column chooser, player II who knows only the initial probability distribution p.

At stage $m = 1, 2, \ldots$, player I chooses $s_m \in S$ and II chooses, simultaneously and independently, $t_m \in T$ and then (s_m, t_m) is announced (i.e. it becomes common knowledge).

Actually $\Gamma(p)$ is not a completely defined game since the payoff is not yet specified. This will be done later; according to the specific form of the payoff, we will be speaking of $\Gamma_n(p)$ (the n-stage game), $\Gamma_\lambda(p)$ (the discounted game) or $\Gamma_\infty(p)$ (the infinitely repeated game).

The main feature of these games is that the informed player's moves will typically depend on (among other things) his information (i.e. on the value of k). Since these moves are observed by the uninformed player, they serve as a channel which can transfer information about the state k. This must be taken into account by player I when choosing his strategy. In Example 1.1., for instance, playing the move $s = 1$ if $k = 1$ and $s = 2$ if $k = 2$ is a dominant strategy as far as the single-stage payoff is concerned. However, such behavior will reveal the value of k to player II and by that enable him to reduce the payoffs to 0 in all subsequent stages. This is of course very disadvantageous in the long run and player I would be better off even by simply ignoring his information. In fact, playing the mixed move $(1/2, 1/2)$ at each stage independently of the value of k guarantees an expected payoff of at least $1/4$ per stage. We shall see that this is indeed the best he can do in the long run.

3.2.1. Posterior probabilities and nonrevealing strategies

For $n = 1, 2, \ldots$, let $H_n^{II} = [S \times T]^{n-1}$ be the set of possible histories for player II at stage n (that is, an element $h_n \in H_n^{II}$ is a sequence $(s_1, t_1; s_2, t_2; \ldots; s_{n-1}, t_{n-1})$ of the moves of the two players in the first $n - 1$ stages of the game). Similarly, H_∞^{II} denotes the set of all infinite histories (i.e. plays) in the game. The set of all histories is $H^{II} = \bigcup_{n>1} H_n^{II}$. Let $\mathcal{H}_n^{II}$ be the σ-algebra on H_∞^{II} generated by the cylinders above H_n^{II} and let $\mathcal{H}_\infty^{II} = \bigvee_{n>1} \mathcal{H}_n^{II}$.

A pure strategy for player I in the supergame $\Gamma(p)$ is $\sigma = (\sigma_1, \sigma_2, \ldots)$, where for each n, σ_n is a mapping from $K \times \mathcal{H}_n^{II}$ to S. Mixed strategies are, as usual, probability distributions over pure strategies. However, since $\Gamma(p)$ is a game of perfect recall, we may (by Aumann's generalization of Kuhn's Theorem; see Aumann (1964)) equivalently consider only behavior strategies that are sequences of mappings from $K \times \mathcal{H}_n^{II}$ to X or equivalently from $\mathcal{H}_n^{II}$ to X^K. Similarly, a behavior strategy of player II is a sequence of mappings from $\mathcal{H}_n^{II}$ (since he does not know the value of k) to Y. Unless otherwise specified the word "strategy" will stand for behavior strategy. A strategy of player I is denoted by σ and one of player II by τ.

Any strategies σ and τ of players I and II, respectively, and $p \in \Delta(K)$ induce a joint probability distribution on states and histories – formally, a probability distribution on the measurable space $(K \times H_\infty^{II}, 2^K \otimes \mathcal{H}_\infty^{II})$. This will be our

basic probability space and we will simply write P or E for probability or expectation when no confusion can arise.

Let $p_1 \equiv p$ and for $n \geq 2$ define

$$p_n^k = \mathrm{P}(k \mid \mathcal{H}_n^{\mathrm{II}}) \quad \forall k \in K .$$

These random variables on $\mathcal{H}_n^{\mathrm{II}}$ have a clear interpretation: p_n is player II's *posterior probability distribution* on K at stage n given the history of moves up to that stage. These posterior probabilities turn out to be the natural *state variable* of the game and therefore play a central role in our analysis.

Observe first that the sequence $(p_n)_{n=1}^{\infty}$ is a $(\mathcal{H}_n^{\mathrm{II}})_{n=1}^{\infty}$ martingale, being a sequence of conditional probabilities with respect to an increasing sequence of σ-algebras, i.e.

$$\mathrm{E}(p_{n+1} \mid \mathcal{H}_n^{\mathrm{II}}) = p_n \quad \forall n = 1, 2, \ldots .$$

In particular, $\mathrm{E}(p_n) = p$ $\forall n$. Furthermore, since this martingale is uniformly bounded, we have the following bound on its L_1 variation (derived directly from the martingale property and the Cauchy–Schwartz inequality):

Proposition 3.8.

$$\frac{1}{n} \sum_{m=1}^{n} \mathrm{E}\|p_{m+1} - p_m\| \leq \sum_k \sqrt{\frac{p^k(1-p^k)}{n}} .$$

Note that $\Sigma_k \sqrt{p^k(1-p^k)} \leq \sqrt{\#K - 1}$ since the left-hand side is maximized for $p^k = 1/(\#K)$ for all k. Intuitively, Proposition 3.8 means that in "most of the stages" p_{m+1} cannot be very different from p_m.

The explicit expression of p_m is obtained inductively by Bayes' formula: given a strategy σ of player I, for any stage n and any history $h_n \in H_n^{\mathrm{II}}$, let $\sigma(h_n) = (x_n^k)_{k \in K}$ denote the vector of mixed moves of player I at that stage. That is, he uses the mixed move $x_n^k = (x_n^k(s))_{s \in S} \in X = \Delta(S)$ in the game G^k. Given $p_n(h_n) = p_n$, let $\bar{x}_n = \Sigma_{k \in K} p_n^k x_n^k$ be the (conditional) average mixed move of player I at stage n. The (conditional) probability distribution of p_{n+1} can now be written as follows: $\forall s \in S$ such that $\bar{x}_n(s) > 0$ and $\forall k \in K$,

$$p_{n+1}^k(s) = \mathrm{P}(k \mid \mathcal{H}_n^{\mathrm{II}}, s_n = s) = \frac{p_n^k x_n^k(s)}{\bar{x}_n(s)} . \tag{5.1}$$

It follows that if $x_n^k = \bar{x}_n$ whenever $p_n^k > 0$, then $p_{n+1} = p_n$, that is:

Proposition 3.9. *Given any player II's history h_n, the posterior probabilities do not change at stage n if player I's mixed move at that stage is independent of k over all values of k for which $p_n^k > 0$.*

In such a case we shall say that player I plays *non-revealing* at stage n and, motivated by that, we define the corresponding set

$$NR = \{x \in X^K \mid x^k = x^{k'} \quad \forall k, k' \in K\} \ .$$

We see here, because of the full monitoring assumption, that not revealing the information is equivalent to not using the information. But then the outcome of the initial chance move (choosing k) is not needed during the game. This lottery can also be made at the end, just to compute the payoff.

Definition 3.10. For $p \in \Delta(K)$ the non-revealing game at p, denoted by $D(p)$, is the (one-stage) two-person, zero-sum game with payoff matrix

$$D(p) = \sum_{k \in K} p^k G^k \ .$$

Let $u(p)$ denote the value of $D(p)$. Clearly, u is a continuous function on $\Delta(K)$ (it is, furthermore, Lipschitz with constant $C = \max_{k,s,t} |G_{st}^k|$).

So if player I uses NR moves at all stages, the posterior probabilities remain constant. Hence the (conditional) payoff at each stage can be computed from $D(p)$. In particular, by playing an optimal strategy in $D(p)$ player I can guarantee an expected payoff of $u(p)$ at each stage. Thus we have:

Proposition 3.11. *Player I can guarantee $u(p)$ in $\Gamma_n(p)$, in $\Gamma_\lambda(p)$, and in $\Gamma_\infty(p)$ by playing i.i.d. an optimal strategy in $D(p)$.*

Combined with Proposition 3.6 this yields:

Corollary 3.12. *The previous proposition holds if we replace $u(p)$ by Cav $u(p)$.*

Given a strategy σ of player I, let $\sigma_n = (x_n^k)_{k \in K}$ be "the strategy at state n" [see MSZ (1993, ch. IV, section 1.6)]. Its average (over K) is the random variable $\bar{\sigma}_n = E(\sigma_n \mid \mathcal{H}_n^{II}) = \Sigma_k p_n^k \sigma_n^k$. Note that $\bar{\sigma}_n \in NR$.

A crucial element in the theory is the following intuitive property. If the σ_n^k are close (i.e. all near $\bar{\sigma}_n$), p_{n+1} will be close to p_n. In fact a much more precise relation is valid; namely, if the distance between two points in a simplex $[\Delta(S)$

or $\Delta(K)$] is defined as the L_1 norm of their difference, then the expectations of these two distances are equal. Formally,

Proposition 3.13.   *For any strategies σ and τ of the two players*

$$\mathrm{E}(\|\sigma_n - \bar{\sigma}_n\| \mid \mathcal{H}_n^{\mathrm{II}}) = \mathrm{E}(\|p_{n+1} - p_n\| \mid \mathcal{H}_n^{\mathrm{II}}) \ .$$

This is directly verified using expression (5.1) for p_{n+1} in terms of σ_n.

Next we observe that the distance between payoffs is bounded by the distance between the corresponding strategies. In fact, given σ and τ let $\rho_n(\sigma, \tau) = \mathrm{E}(g_n \mid \mathcal{H}_n^{\mathrm{II}})$, and define $\tilde{\sigma}(n)$ to be the same as the strategy σ except for stage n where $\tilde{\sigma}_n(n) = \bar{\sigma}_n$, we then have:

Proposition 3.14.   *For any σ and τ,*

$$|\rho_n(\sigma, \tau) - \rho_n(\tilde{\sigma}(n), \tau)| \leq C\mathrm{E}(\|\sigma_n - \bar{\sigma}_n\| \mid \mathcal{H}_n^{\mathrm{II}}) \ .$$

Proof.   In fact, since p_n is the same under σ and under $\tilde{\sigma}(n)$, we have (for any ω in H_∞^{II}):

$$|\rho_n(\sigma, \tau) - \rho_n(\tilde{\sigma}(n), \tau)| (\omega) \leq C \sum_k p_n^k(\omega)\|\sigma_n^k - \bar{\sigma}_n\|$$

$$= C\mathrm{E}(\|\sigma_n - \bar{\sigma}_n\| \mid \mathcal{H}_n^{\mathrm{II}})(\omega) \ . \quad \square$$

3.2.2. The limit values $\lim v_n(p)$ *and* $v_\infty(p)$

The main consequence of the bound derived so far is:

Proposition 3.15.   *For all $p \in \Delta(K)$ and all n,*

$$v_n(p) \leq \mathrm{Cav}\, u(p) + \frac{C}{\sqrt{n}} \sum_k \sqrt{p^k(1 - p^k)} \ .$$

Proof.   Making use of the minmax theorem, it is sufficient to prove that for any strategy σ of player I in $\Gamma_n(p)$, there exists a strategy τ of player II such that

$$E_{\sigma,\tau}(\bar{g}_n) \leq \mathrm{Cav}\, u(p) + \frac{C}{\sqrt{n}} \sum_k \sqrt{p^k(1 - p^k)} \ .$$

Given σ let τ be the following strategy of player II: at stage m; $m = 1, \ldots, n$, compute p_m and play a mixed move τ_m which is optimal in $D(p_m)$.

By Proposition 3.14 and Proposition 3.13, for $m = 1, \ldots, n$:

$$\rho_m(\sigma, \tau) \leqslant \rho_m(\tilde{\sigma}(m), \tau) + CE(\| p_{m+1} - p_m \| \mid \mathcal{H}_m^{\mathrm{II}}) \, .$$

Now

$$\rho_m(\tilde{\sigma}_m, \tau) = \sum_k p_m^k \, \bar{\sigma}_m G^k \tau_m \, ,$$

with $\bar{\sigma}_m \in NR$ and τ_m optimal in $D(p_m)$. Hence

$$\rho_m(\tilde{\sigma}(m), \tau) \leqslant u(p_m) \leqslant \mathrm{Cav}\, u(p_m) \, ,$$

which yields

$$\rho_m(\sigma, \tau) \leqslant \mathrm{Cav}\, u(p_m) + CE(\| p_{m+1} - p_m \| \mid \mathcal{H}_m^{\mathrm{II}}) \, .$$

Averaging on $m = 1, \ldots, n$ and over all possible histories $\omega \in H_n^{\mathrm{II}}$ we obtain [using $E(\mathrm{Cav}\, u(p_m(\omega))) \leqslant \mathrm{Cav}\, u(p)$ by Jensen's inequality]:

$$E_{\sigma, \tau}(\bar{g}_n) \leqslant \mathrm{Cav}\, u(p) + \frac{C}{n} \sum_{m=1}^n E\| p_{m+1} - p_m \| \, .$$

The claimed inequality now follows from Proposition 3.8. $\quad \square$

Combining Proposition 3.15 with Corollary 3.12 we obtain [Aumann and Maschler (1967)]:

Theorem 3.16. *For all $p \in \Delta(K)$, $\lim_{n \to \infty} v_n(p)$ exists and equals $\mathrm{Cav}\, u(p)$. Furthermore, the speed of convergence is bounded by*

$$0 \leqslant v_n(p) - \mathrm{Cav}\, u(p) \leqslant \frac{C}{\sqrt{n}} \sum_n \sqrt{p^k(1 - p^k)} \, .$$

The strategy in Proposition 3.15 yields also:

Corollary 3.17. $\lim_{\lambda \to 0} v_\lambda(p)$ *exists and equals $\mathrm{Cav}\, u(p)$ and the speed of convergence satisfies*

$$0 \leqslant v_\lambda(p) - \mathrm{Cav}\, u(p) \leqslant C \sqrt{\frac{\lambda}{2 - \lambda}} \sum_k \sqrt{p^k(1 - p^k)} \, .$$

This follows using

$$\sum_{m=1}^\infty \lambda(1 - \lambda)^{m-1} E\| p_{m+1}^k - p_m^k \| \leqslant \sqrt{\frac{\lambda}{2 - \lambda}} \sqrt{p^k(1 - p^k)} \, ,$$

which is a consequence of the Cauchy–Schwartz inequality and Proposition 3.8.

Combining now Corollary 3.12, Theorem 3.16 and Theorem 3.7 establishes the existence of the value of the infinite game $\Gamma_\infty(p)$ [Aumann, Maschler and Stearns (1968)]:

Theorem 3.18.   *For all* $p \in \Delta(K)$ *the value* $v_\infty(p)$ *of* $\Gamma_\infty(p)$ *exists and equals* Cav $u(p)$.

3.2.3. Recursive formula for $v_n(p)$

The convergence of $v_n(p)$ is actually a monotone convergence. This follows from the following recursive formula for $v_n(p)$. Recall that $x = (x^k)_{k \in K} \in [\Delta(S)]^K$ is a one-stage strategy of player I (i.e. he plays the mixed move x^k in game G^k), then we have

$$v_{n+1}(p) = \frac{1}{n+1} \max_x \left\{ \min_t \sum_k p^k x^k G^k_{\cdot,t} + n \sum_{s \in S} \bar{x}_s v_n(p_s) \right\},$$

where $\bar{x} = \Sigma_k\, p^k x^k$ and for each s in S for which $\bar{x}_s > 0$, p_s is the probability on K given by $p^k_s = p^k x^k_s / \bar{x}_s$, and $G^k_{\cdot,t}$ denotes the t-th column of the matrix G^k.

By this recursive formula it can be proved inductively, using the concavity of $v_n(p)$, that:

Proposition 3.19.   *For all* $p \in P$, *the sequence* $v_n(p)$ *is monotonically decreasing.*

The above recursive formula and the monotonicity are valid much more generally than in the full monitoring case under consideration. They hold (with the appropriate notation) in any signalling structure in which the signal received by player I includes the signal received by player II. However, when this condition is not satisfied, $v_n(p)$ may not be monotone. In fact, Lehrer (1987) has exhibited an example of a game with incomplete information on one side in which $v_1 \geq v_2 < v_3$.

3.2.4. Approachability strategy

Corollary 3.12 provides an explicit simple optimal strategy for player I in $\Gamma_\infty(p)$ which is played as follows. Express p as a convex combination $p = \Sigma_{e \in E} \alpha_e p_e$ of points $(p_e)_{e \in E}$ in $\Delta(E)$ such that Cav $u(p) = \Sigma_{e \in E} \alpha_e u(p_e)$. Perform the appropriate lottery described in Proposition 3.2 to choose $e \in E$, and then play i.i.d. at each stage an optimal strategy of the non-revealing game $D(p_e)$ (with the chosen e.)

On the other hand, the optimal strategy for player II provided by Theorem 3.7 is far from easy to compute. We now describe a simple optimal strategy for player II, making use of Blackwell's approachability theory for vector payoff games.

At any stage $(n + 1)$, given the history $h_{n+1} = (s_1, t_1, \ldots, s_n, t_n)$, player II can compute $\bar{g}_n^k = \frac{1}{n} \Sigma_{m=1}^n G_{s_m t_m}^k$, which is what his average payoff would be up to that stage if the state was k. Since the prior distribution on the states is p, his expected average payoff is $\langle p, \bar{g}_n \rangle = \Sigma_{k \in K} p^k \bar{g}_n^k$. We shall show that player II can play in such a way that, with probability one, the quantity $\langle p, \bar{g}_n \rangle$ will be arbitrarily close to Cav $u(p)$, for n sufficiently large.

Having focused our attention on the vector of averages $\bar{g}_n = (\bar{g}_n^k)_{k \in K}$, it is natural to consider the game with vector payoffs in the Euclidean space $\mathbb{R}^K$. So when moves (s, t) are played, the resulting vector payoff is $G_{st} \in \mathbb{R}^K$.

Consider the game $\Gamma_\infty(p)$. Let $u(p)$ be its NR-value function, and let $l = (l^k)_{k \in K}$ be the vector of intercepts of a supporting hyperplane to Cav u at p (see Figure 1); that is,

$$\text{Cav } u(p) = \langle l, p \rangle = \sum_k l^k p^k \quad \text{and} \quad u(q) \leqslant \langle l, q \rangle \quad \text{for all } q \text{ in } \Delta(K) .$$

If player II can play so that the average vector payoff $\bar{g}_n$ will *approach*[2] the "corner set" $C = \{x \in \mathbb{R}^K \mid x \leqslant l\}$, it will mean that $\forall \epsilon \geqslant 0$, $\langle p, \bar{g}_n \rangle \leqslant \langle l, p \rangle + \epsilon = $ Cav $u(p) + \epsilon$, both in expectation and with probability one, for n sufficiently large. This is precisely the optimal strategy we are looking for.
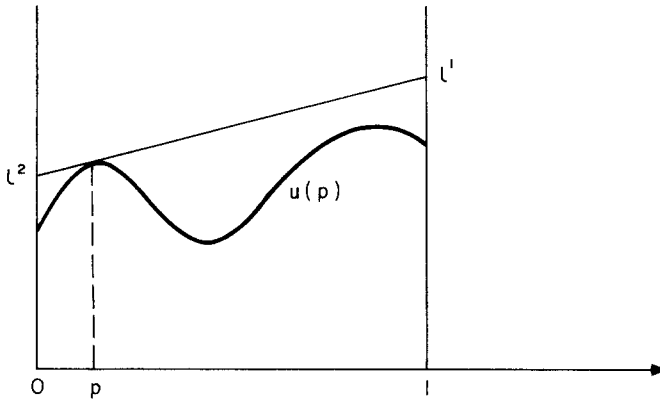

Figure 1

[2]That is, for any strategy of player I, the distance $d(\bar{g}_n, C)$ tends to 0 with probability 1. See Blackwell (1956).

For any mixed move $y \in Y$ of player II let $Q(y) = \text{Co}\{\Sigma_t \, G_{st} y_t \mid s \in S\}$, where Co A denotes the convex hull of A. (This is the set in which lies the expected vector payoff when y is played.) A sufficient condition for the approachability of C by player II [Blackwell (1956)] is that for any $g \not\in C$, he has a mixed move y such that if c is the closest point to g in C, then the hyperplane H orthogonal to $[cg]$ through c separates g from $Q(y)$ (see Figure 2).
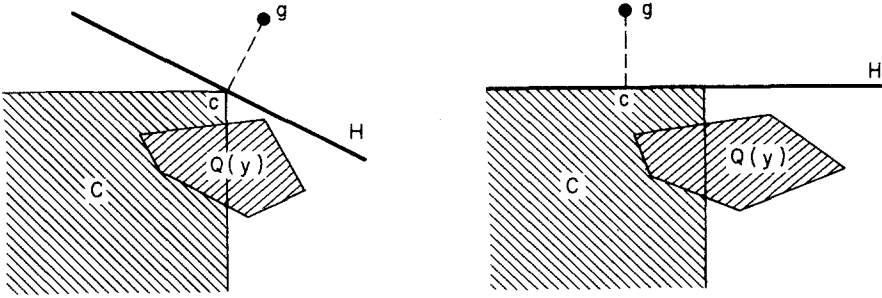

Figure 2

To verify this condition in our case let $q \in \Delta(K)$ be the unit vector in the direction $(g - c)$ and let $H = \{x \in \mathbb{R}^K \mid \langle q, x \rangle = \langle q, c \rangle\}$. Note that since C is a corner set, $q \geqslant 0$ and therefore $\langle q, c \rangle \geqslant \langle q, c' \rangle$ for all points c' in C, in particular $\langle q, c \rangle \geqslant \langle q, l \rangle$. Since $c \in C$, we also have $c \leqslant l$, which implies $\langle q, c \rangle \leqslant \langle q, l \rangle$; hence $l \in H$ and C lay in the half space defined by H, which can be written as $H_1 = \{x \in \mathbb{R}^K \mid \langle q, x \rangle \leqslant \langle q, l \rangle\}$. Now, by playing optimally in the non-revealing game $D(q)$, player II guarantees

$$\sum_{k \in K} q^k \sigma G^k y \leqslant u(q) \quad \forall \sigma \in \Delta(S) .$$

This means that for any mixed strategy σ of player I, the vector payoff $x = (\sigma G^k y)_{k \in K}$ satisfies $\langle q, x \rangle \leqslant u(q) \leqslant \langle q, l \rangle$, i.e. $x \in H_1$, establishing the approachability of C.

The optimal strategy of player II in $\Gamma_\infty(p)$ can now be summarized as follows:

(1) Choose $l \in \mathbb{R}^K$ such that $\langle p, x \rangle = \langle p, l \rangle$ is the supporting hyperplane to the graph of Cav u at p.

(2) Define the corner set $C = \{x \in \mathbb{R}^K \mid x \leqslant l\}$, and at each stage n compute the average vector payoff $\bar{g}_n$ up to that stage.

(3) At stage $(n + 1)$, $n = 1, 2, \ldots$, if $\bar{g}_n \in C$, play arbitrarily. If $\bar{g}_n \not\in C$, let $c \in C$ be the closest point to $\bar{g}_n$ in C, compute $q = (\bar{g}_n - c)/\|\bar{g}_n - c\| \in \Delta(K)$ and play an optimal mixed move in $D(q)$.

3.2.5. The examples revisited

Let us look again at the examples we discussed in the Introduction in view of the general results.

In Example 1.1, $D(p)$ is the matrix game:

$$p\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (1-p)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix},$$

and its value is $u(p) = p(1-p)$. Since this is a concave function, Cav $u(p) = u(p) = p(1-p)$ and we have (see Figure 3)

$$\lim_{n\to\infty} v_n(p) = \lim_{\lambda\to 0} v_\lambda(p) = v_\infty(p) = p(1-p).$$

Example 1.1:
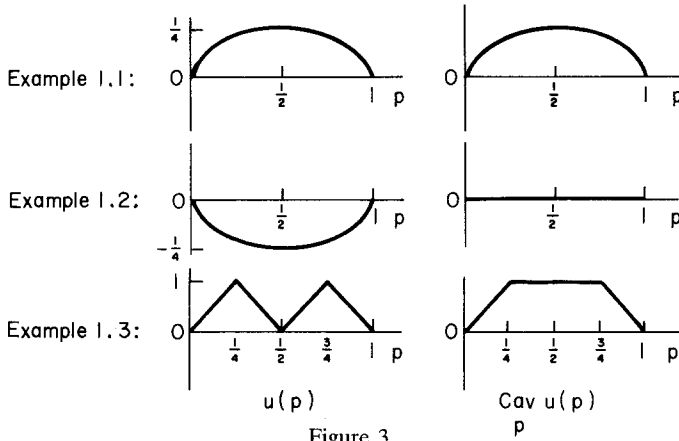
Example 1.2:

Example 1.3:

$u(p)$ Cav $u(p)$

Figure 3

In particular, for $p = 1/2$ this limit is $1/4$. So asymptotically the value is that of the game in which no player is informed about the value of k. In other words, the informed player has an advantage only in games of finite length. This advantage may be measured by $v_n(p) - v_\infty(p)$. By Theorem 3.16 this is bounded by

$$v_n(p) - p(1-p) \leq \frac{2\sqrt{p(1-p)}}{\sqrt{n}} \leq \frac{1}{\sqrt{n}}.$$

It turns out that for this specific game this bound can be improved and the actual speed of convergence is [see Zamir (1971–72)]

$$v_n(p) - p(1-p) = O\left(\frac{\ln n}{n}\right).$$

In Example 1.2,

$$D(p) = p\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + (1-p)\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -p & 0 \\ 0 & -(1-p) \end{pmatrix},$$

and its value is $u(p) = -p(1-p)$. Since this is a convex function and its concavification is the constant function 0, Cav $u(p) = 0$ $\forall p \in [0, 1]$ and we have (see Figure 3)

$$\lim_{n \to \infty} v_n(p) = \lim_{\lambda \to 0} v_\lambda(p) = v_\infty(p) = 0.$$

For $p = 1/2$ (as we had in our example), the value is 0.

In Example 1.3,

$$D(p) = p\begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} + (1-p)\begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 4p & 4-4p & 4p-2 \\ 4p & 4-4p & 2-4p \end{pmatrix},$$

and its value is (see Figure 3)

$$u(p) = \begin{cases} 4p, & 0 \le p \le 1/4, \\ 2-4p, & 1/4 \le p \le 1/2, \\ 4p-2, & 1/2 \le p \le 3/4, \\ 4-4p, & 3/4 \le p \le 1. \end{cases}$$

Therefore

$$\lim_{n \to \infty} v_n(p) = \lim_{\lambda \to 0} v_\lambda(p) = v_\infty(p) = \text{Cav } u(p),$$

where Cav $u(p)$ is (see Figure 3)

$$\text{Cav } u(p) = \begin{cases} 4p, & 0 \le p \le 1/4, \\ 1, & 1/4 \le p \le 3/4, \\ 4-4p, & 3/4 \le p \le 1. \end{cases}$$

For $p = 1/2$ (as we had in our example), the value is 1.

Remark. To all results so far there are of course corresponding **dual results** for the case in which the informed player is player II (while **player I is** uninformed). In particular the dual to Theorem 3.18 is:

Theorem 3.20. *In the game in which player II is informed and player I is not, for all $p \in \Delta(k)$ the value $v_\infty(p)$ of $\Gamma_\infty(p)$ exists and equals* Vex $u(p)$.

Here Vex $u(p)$ is the maximal convex function pointwise majorized by $u(p)$.

3.3. The general case

The main results so far, specifically the existence of $\lim v_n$ and of v_∞, extend to the general case without full monitoring, so we no longer assume that the moves are announced after each stage but rather that some individual message is transmitted to each of the players. This model was also treated by Aumann, Maschler and Stearns (1968), who proved the main result about the existence and the formula of $v_\infty(p)$. The generalization of the strategy for the uninformed player, using Blackwell approachability, is due to Kohlberg (1975a). Although the analysis follows the lines developed for the case of full monitoring, the mathematical details require several new ideas. These will be only outlined in this section [for the detailed proofs see, for example, MSZ (1993, Ch. V).

Recalling the general model, we add a signaling structure – two finite sets of signals A and B and transition probability Q from $K \times S \times T$ to $\Delta(A) \times \Delta(B)$. We denote by Q^k_{st} the probability distribution at (k, s, t). The repeated game is played as in the previous model with the following modification. At each stage n, $n \geqslant 1$, instead of announcing the moves (s_n, t_n), the signal a_n is announced to player I and b_n is announced to player II, where (a_n, b_n) is chosen according to the distribution $Q^k_{s_n t_n}$. It turns out that the value of $\Phi_\infty(p)$ does not depend on the signaling structure to the informed player, so by abuse of notation we denote the marginal of Q on B also by $Q^k_{s_n t_n}$.

The generalization of the notion of non-revealing utilizes the property that when a non-revealing strategy is played by player I at a certain stage, the conditional probability on K does not change at that stage. This is equivalent to:

Definition 3.21. $x \in X^K$ is said to be *non-revealing* at $p \in \Delta(K)$ if, for each move $t \in T$ of player II, the distribution of b (induced by t and x^k) in the kth state is the same for all k for which $p^k > 0$.

We denote by $NR(p)$ the set of non-revealing strategies in $\Gamma_1(p)$. For $p \in \Delta(K)$ let $K(p) \subseteq K$ denote the support of p. Then

$$NR(p) = \{x \in X^K \mid x^k Q^k = x^{k'} Q^{k'} \quad \forall (k, k') \in K(p) \times K(p)\}$$

Note that $NR(p) \subseteq NR(\tilde{p})$ whenever $K(p) \supseteq K(\tilde{p})$. Therefore $NR(p)$ is a "step set-valued function" on $\Delta(K)$ with possible discontinuities at the intersections of two (or more) facets. $NR(p)$ may be empty for some $p \in \Delta(K)$; however, if p is an extreme point of $\Delta(K)$, then $NR(p) = X^K$. This is intuitive – an extreme point of $\Delta(K)$ corresponds to a situation of *complete information*, where k is known to both players and hence every strategy of **I** is non-revealing since there is nothing to reveal.

The non-revealing game (*NR*-game), denoted by $D(p)$, is the (one-stage) two-person, zero-sum game in which player I's strategy set is $NR(p)$, player II's strategy set is $\Delta(T)$ and for $x = (x^k)_{k \in K} \in NR(p)$ and $y \in \Delta(T)$ the payoff is $\Sigma_{k \in K} p^k x^k G^k y$.

We denote by $u(p)$ the value of $D(p)$ and refer to it as the *NR*-value. If $NR(p) = \emptyset$ [hence $D(p)$ is undefined] we define $u(p) = -\infty$. Since $u(p)$ is finite at least on the extreme points, it follows that Cav $u(p)$ is well defined [and Lipschitz on $\Delta(K)$ with constant C].

Theorem 3.18 can now be proved for the general signaling case with this Cav $u(p)$. The proof that player I can guarantee $v(p)$ in $\Gamma_\infty(p)$ is the same as in the full monitoring case, that is, by applying an appropriate "splitting" followed by a non-revealing strategy. The major difficulty is in generalizing the optimal strategy of the uninformed player. The problem is that the above described optimal strategy for player II is based on the *statistics* $\bar{g}_n = (1/n) \Sigma_{m=1}^n g_n$. This is the vector whose kth coordinate is $(1/n) \Sigma_{m=1}^n G^k_{s_m t_m}$ which is *observable* by player II in the full monitoring case since he observes the moves (s_m, t_m). In the general case $\bar{g}_n$ is not observable by player II. Another optimal strategy is to be provided which is based only on the history $h_n = (b_1, \ldots, b_n)$ available to player II at each stage.

For any signal $b \in B$, and any move $t \in T$ at any stage n, let ρ_n^{tb} be the proportion of stages, up to stage n, in which b was received by player II following a move t, out of all stages in which move t was played, i.e.

$$\rho_n^{tb} = \frac{\#\{m \mid m \leqslant n, b_m = b, t_m = t\}}{\#\{m \mid m \leqslant n, t_m = t\}} \,.$$

The vector $\rho_n = (\rho_n^{tb})_{t \in T, b \in B}$, which is observable by player II after each stage n, is the basis for his strategy. There is also a vector payoff ξ_n which plays the role of the non-observable g_n. We do not define it formally here; it is, roughly speaking, the worst vector payoff which is compatible (up to a small deviation δ) with the observed vectors $\rho_1, \ldots, \rho_n$. To this vector payoff one applies Blackwell's approachability theory. The definition of ξ_n and the strategy of player II are such that [for the details see Kohlberg (1975a, 1975b) or MSZ (1993, ch. V)]:

- The ξ-payoff, i.e. $\langle p, \xi \rangle$, will be as close as we wish to Cav $u(p)$.
- The actual unobserved payoff will not exceed the observed ξ-payoff by more than an arbitrarily small ϵ.
- Player II plays each mixed move in a large block of stages so that, using (an appropriate version of) the strong law of large numbers, both the signals distribution and the (unobserved) payoffs are close to their means.

4. Incomplete information on two sides

The case of incomplete information on two sides is that in which each of the two players initially has only partial information about the state of nature, represented by a general partition of K. We denote these partitions by K^I and K^{II}. (The case in which one of the partitions is $\{\{1\}, \{2\}, \dots, \{\#K\}\}$ is the case of incomplete information on one side treated in the previous section.) By common terminology, the elements of K^I and K^{II} are called the *types* of players I and II, respectively. The initial probability p can then be thought of as a joint prior probability distribution on the pairs of types.

A special case is that in which the types of the two players are *independent*, i.e. there exist two probability vectors q^I and q^{II} on the elements of K^I and K^{II}, respectively, such that

$$p(\kappa_j^I \cap \kappa_l^{II}) = q_j^I q_l^{II} \quad \forall \kappa_j^I \in K^I \quad \text{and} \quad \kappa_l^{II} \in K^{II} .$$

No general results are available for the whole class of these games. Most of this section is devoted to the special case in which Q^k is independent of k. This will be called the case of *state independent signaling*. That is, the information gained at each stage does not depend on the state of nature and it is determined completely by the players' moves at that stage. We omit the index k and denote the signaling mechanism by one transition probability from $S \times T$ to $A \times B$.

4.1. Minmax and maxmin

Let $\mathcal{K}^I$ and $\mathcal{K}^{II}$ be the σ-fields generated by K^I and K^{II}, respectively. A one-stage strategy $x = (x^k)_{k \in K}$ of player I in X^K is *non-revealing* if it is $\mathcal{K}^I$ measurable and $\Sigma_{s \in S} x^k(s) Q_{s,t}(b)$ is independent of k for all t in T and b in B. In words, for each column of Q, the marginal probability distribution on B induced on the letters of that column is independent of the state of nature k. The set of non-revealing one-stage strategies of player I is denoted by NR^I. The set of non-revealing one-stage strategies of player II is defined in a dual way and is denoted by NR^{II}. These sets are obviously non-empty; they contain, for instance, the strategies constant on K. Denote by $D(p)$ the one-stage game in which players I and II are restricted to strategies in NR^I and NR^{II}, respectively. Let $u(p)$ be the value of $D(p)$.

Remarks. (i) The above definition of non-revealing strategy differs formally from Definition 3.21 in that there we required the induced distribution on B

(resp. on A) to be constant in k only on $K(p)$, while here we require it over all of K. However, it is easily seen that in this case the two definitions lead to the same $u(p)$. Since all results are formulated in terms of $u(p)$, we prefer to use here the above introduced definitions which have the advantage of making NR^{I} and NR^{II} independent of p.

(ii) Note that $u(p)$ is continuous in p on the simplex $\Delta(K)$ of prior probabilities.

We need now to generalize the notion of concavity and convexity:

A function on $\Delta(K)$ is said to be *concave with respect to* I (abbreviated w.r.t. I) if for every $p = (p^k)_{k \in K}$ it has a concave restriction on the subset $\Pi^{I}(p)$ defined by

$$\Pi^{I}(p) = \left\{ (\alpha^k p^k)_{k \in K} \mid \alpha^k \geq 0 \quad \forall k, \sum_k \alpha^k p^k = 1 \text{ and } (\alpha^k)_{k \in K} \right.$$

$$\left. \text{is } \mathcal{K}^{I}\text{-measurable} \right\}.$$

Interpretation: Given the prior probability distribution p on K and given any one stage strategy of player I (which is hence $\mathcal{K}^{I}$-measurable), the conditional probability distribution on K given the move of player I is an element of $\Pi^{I}(p)$. In other words, when updating the distribution on K in view of observations on player I's moves only (knowing his strategy), the range of the posterior distribution is $\Pi^{I}(p)$.

A function on $\Delta(K)$ is said to be *convex with respect to* II (abbreviated w.r.t. II) if for every $p = (p^k)_{k \in K}$ it has a convex restriction on the subset $\Pi^{II}(p)$ defined by

$$\Pi^{II}(p) = \left\{ (\beta^k p^k)_{k \in K} \mid \beta^k \geq 0 \quad \forall k, \sum_k \beta^k p^k = 1 \text{ and } (\beta^k)_{k \in K} \right.$$

$$\left. \text{is } \mathcal{K}^{II}\text{-measurable} \right\}.$$

Note that for any p in $\Delta(K)$ both $\Pi^{I}(p)$ and $\Pi^{II}(p)$ are convex and compact subsets of $\Delta(K)$ containing p, which justify the above definitions of concavity w.r.t. I and convexity w.r.t. II.

In the independent case it is more convenient to work not with p in $\Pi = \Delta(K)$ but rather with the product probability $(q^{I}, q^{II}) \in \Delta^{I} \times \Delta^{II}$, where Δ^{I} and Δ^{II} are the simplices of probability distributions on the types of player I (i.e. the elements of K^{I}) and of player II, respectively. In this case,

$$\Pi^{I}(q^{I}, q^{II}) = \{q^{I}\} \times \Delta^{II} \quad \text{and} \quad \Pi^{II}(q^{I}, q^{II}) = \Delta^{I} \times \{q^{II}\}.$$

Thus, concavity w.r.t. I means simply concavity in the first variable q^I (for any value of q^{II}), and similarly for convexity w.r.t. II.

Given any function g on $\Delta(K)$, the *concavification* of g w.r.t. I (denoted by $\mathrm{Cav}_I\, g$) is the (pointwise) minimal function which is concave w.r.t. I and is greater than or equal to g on $\Delta(K)$. Similarly, the *convexification* of g w.r.t. II (denoted by $\mathrm{Vex}_{II}\, g$) is the (pointwise) minimal function which is convex w.r.t. II and is less than or equal to g on $\Delta(K)$.

Remark. Note that in the special case of incomplete information on one side ($K^I = \{\{1\}, \{2\}, \ldots, \{\#K\}\}$ and $K^{II} = \{K\}$), $\mathrm{Cav}_I\, g$ is the usual $\mathrm{Cav}\, g$ and $\mathrm{Vex}_{II}\, g$ is g.

Theorem 4.1. *The minmax of $\Gamma_\infty(p)$ exists and is given by*

$$\bar{v}(p) = \mathrm{Vex}_{II}\, \mathrm{Cav}_I\, u(p) .$$

Similarly, $\mathrm{Cav}_I\, \mathrm{Vex}_{II}\, u(p)$ *is the maxmin of* $\Gamma_\infty(p)$.

Proof. The heuristic arguments of the proof are as follows.[3] Proving that the minmax of $\Gamma_\infty(p)$ is $\mathrm{Vex}_{II}\, \mathrm{Cav}_I\, u(p)$ consists of two parts.

Part (i): Player II can guarantee $\mathrm{Vex}_{II}\, \mathrm{Cav}_I\, u(p)$. If player II ignores his private information (i.e. K^{II}) and if for each $\kappa^I \in K^I$ let $q^{\kappa^I} = \Sigma_{k \in \kappa^I}\, p^k$ and take as payoffs $A^{\kappa^I} = (1/q^{\kappa^I}) \Sigma_{k \in \kappa^I}\, p^k G^k$ (keeping the same distribution on signals), we obtain a game $\Gamma(q)$ with incomplete information on one side, with K^I as the set of states of nature, with initial probability distribution q on it, and player I informed. In this game, denoting the value of the non-revealing game by $w(q)$, player II can guarantee $\mathrm{Cav}\, w(q)$. Now by our construction $w(q) = u(p)$ and $\mathrm{Cav}\, w(q) = \mathrm{Cav}_I\, u(p)$. Finally, by the dual of Proposition 3.6, player II can also guarantee the Vex_{II} of this function, namely $\mathrm{Vex}_{II}\, \mathrm{Cav}_I\, u(p)$ (by applying the appropriate splitting procedure established in Proposition 3.2).

Part (ii): Player I can defend $\mathrm{Vex}_{II}\, \mathrm{Cav}_I\, u(p)$. Any pair of strategies σ and τ of the two players induces a martingale of posterior probability distributions $\{p_n\}_{n=1}^\infty$, converging with probability one and hence having a bounded total variation: $\mathrm{E}_{\sigma,\tau}[\Sigma_{k \in K} \Sigma_{n=1}^\infty (p_n^k - p_{n-1}^k)^2]$. Given any strategy τ of player II, define σ_0 as a non-revealing strategy of player I inducing a martingale with N-stage variation ϵ-close to the supremum, over all his non-revealing

strategies, of the total variation. Denoting by NR_∞^I the set of all non-revealing strategies of player I, this means that σ_0 and N are defined by

$$E_{\sigma_0,\tau}\left[\sum_{k\in K}\sum_{n=1}^{N}(p_n^k - p_{n-1}^k)^2\right] > \sup_{\sigma\in NR_\infty^I}E_{\sigma,\tau}\left[\sum_{k\in K}\sum_{n=1}^{\infty}(p_n^k - p_{n-1}^k)^2\right] - \epsilon .$$

By playing this σ_0 against τ up to stage N, player I "exhausts" almost all the variation of the martingale, i.e. player II will be playing "practically non-revealing" from that stage on. Thus, the situation is almost that of incomplete information on one side in which player I is informed and he can then guarantee $\text{Cav}_I\, u(p_N)$ (where p_N is the posterior probability at stage N). Finally, since up to stage N player I is playing non-revealing, we have $p_N \in \Pi^{II}(p)$ and $E(p_N) = p$ implying that the expected average payoff to player I is at least $E(\text{Cav}_I\, u(p_N)) \geq \text{Vex}_{II}\,\text{Cav}_I\, u(p)$. $\square$

It should be noted that the formal proof of the above outlined arguments is quite intricate and non-trivial mainly because in a general signaling structure "exhausting" the information from the other player's strategy usually involves revealing the player's own information. Another general difficulty in all proofs involving the posterior probabilities p_m of a certain player is that they have to be assumed computable by the other player as well, which is usually not the case when there is general signaling. The way to overcome these difficulties is the following. Assume that we want to prove that player I can guarantee a certain payoff level. We perturb the game to make it slightly · more dis-advantageous to him. This perturbation consists of not giving player I his signal according to Q unless he buys it for an amount C. Furthermore, he is restricted to use this option of buying information exactly with probability $\delta > 0$ while with probability $(1 - \delta)$ he gets no information whatsoever. Whenever he does receive non-trivial information, his signal is completely known to player II. This implies that $\mathcal{H}_m^{II} \supset \mathcal{H}_m^{I}$ and hence p_m, the posterior distribution of player I, is also computable by player II. If, despite the disadvantageous modi-fications, player I can guarantee a certain amount (for sufficiently small δ) then he can also certainly guarantee it in the original game.

A corollary of Theorem 4.1 is that the infinite game $\Gamma_\infty(p)$ has a value if and only if

$$\text{Cav}_I\,\text{Vex}_{II}\, u(p) = \text{Vex}_{II}\,\text{Cav}_I\, u(p) .$$

An example of a game without a value is the following game with independent types and full monitoring in which there are two types of each player with the payoff matrices given by

$$G^{11} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix}, \qquad G^{12} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$G^{21} = \begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad G^{22} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

That is, the set of states is $K = \{11, 12, 21, 22\}$ and the partitions of initial information are $K^{\mathrm{I}} = \{11, 12\}\{21, 22\}$ and $K^{\mathrm{II}} = \{11, 21\}\{12, 22\}$. If, for instance, the initial probability distributions on types are $q^{\mathrm{I}} = q^{\mathrm{II}} = (1/2, 1/2)$, then[4] $\underline{v}_\infty(1/2, 1/2) = \mathrm{Cav}_{\mathrm{I}} \mathrm{Vex}_{\mathrm{II}}\, u(1/2, 1/2) = -1/4$ and $\bar{v}_\infty(1/2, 1/2) = \mathrm{Vex}_{\mathrm{II}} \mathrm{Cav}_{\mathrm{I}}\, u(1/2, 1/2) = 0$.

4.2. The asymptotic value $\lim_{n\to\infty} v_n(p)$

The non-existence of a value for infinite games with incomplete information on both sides is a very important feature of these games which, among other things, exemplifies the difference between repeated games with incomplete information and stochastic games, in which the value always exists. Given this result, the next natural question is that of the existence of the asymptotic value $\lim_{n\to\infty} v(p)$. Here the result is positive [Mertens and Zamir (1971–72)]:

Theorem 4.2. $v(p) = \lim_{n\to\infty} v_n(p)$ *exists for all* $p \in \Delta(K)$ *and is the unique solution to the following set of functional equations*:
  (1) $f(p) = \mathrm{Vex}_{\mathrm{II}} \max\{u(p), f(p)\}$;
  (2) $f(p) = \mathrm{Cav}_{\mathrm{I}} \min\{u(p), f(p)\}$.

Proof. To outline the main arguments let $\underline{v}(p)$ and $\bar{v}(p)$ be, respectively, the lim inf and lim sup of $\{v_n\}_{n=1}^\infty$. Both functions are Lipschitz, $\underline{v}$ is concave w.r.t. I and $\bar{v}$ is convex w.r.t. II. For any strategy of player II consider the following response strategy of player I (actually a sequence of strategies one for each finite game): play optimally in $D(p_m)$ at stage m as long as $u(p_m) \geqslant \underline{v}(p_m)$. As soon as $u(p_m) < \underline{v}(p_m)$ play optimally in the remaining subgame (p_m is the posterior probability distribution on K at stage m). This strategy guarantees player I an expected average payoff arbitrarily close to the maximum of $u(p)$ and $\underline{v}(p)$, for n large enough, proving that:

  (1') $\underline{v}(p) \geqslant \mathrm{Vex}_{\mathrm{II}} \max\{u(p), \underline{v}(p)\}$,

and similarly:

  (2') $\bar{v}(p) \leqslant \mathrm{Cav}_{\mathrm{I}} \min\{u(p), \bar{v}(p)\}$.

[4]For the detailed computations see Mertens and Zamir (1971–72).

Actually, this argument shows that if player I can guarantee $f(p)$ in $\Gamma_n(p)$ for large enough n, he can also guarantee $\mathrm{Vex}_{\mathrm{II}} \max\{u(p), f(p)\}$.

Now, since $\bar{v}(p)$ is convex w.r.t. II, it follows from (2') that

(2'') $\bar{v}(p) \leqslant \mathrm{Vex}_{\mathrm{II}} \mathrm{Cav}_{\mathrm{I}} \min\{u(p), \bar{v}(p)\}$.

Next, when a player plays an optimal strategy in $D(p_m)$ at stage m, his expected payoff at that stage differs from $u(p_m)$ by at most a constant times $|p_{m+1} - p_m|$. Combining this with Proposition 3.8 one shows that any function $f(p)$ satisfying $f(p) \leqslant \mathrm{Vex}_{\mathrm{II}} \mathrm{Cav}_{\mathrm{I}} \min\{u(p), f(p)\}$ must satisfy

$$(f(p) - v_n(p))^+ \leqslant R \frac{\Sigma_k \sqrt{p^k(1-p^k)}}{n} \,, \tag{5.2}$$

for some constant R. In particular, letting $n \to \infty$, this implies $f(p) \leqslant \underline{v}$. It follows now from (2'') that $\bar{v}(p) \leqslant \underline{v}(p)$ and hence $v_n(p)$ converge to, say, $v(p)$ with the speed of convergence of $1/\sqrt{n}$. The limit is the smallest solution to

$$f(p) \geqslant \mathrm{Cav}_{\mathrm{I}} \mathrm{Vex}_{\mathrm{II}} \max\{u(p), f(p)\} \,,$$

and the largest solution to

$$f(p) \leqslant \mathrm{Vex}_{\mathrm{II}} \mathrm{Cav}_{\mathrm{I}} \min\{u(p), f(p)\} \,.$$

It is then the only simultaneous solution to both. Finally, since $v(p)$ is both concave w.r.t. I and convex w.r.t. II, it must also satisfy (1) and (2), and is the only solution to this system. $\square$

The above outline can be made a precise proof for the case of full monitoring. For the general signaling case, one has to use a sequence of δ-perturbations of the game. This provides the same results as far as the functional equations are concerned but with different bound on the speed of convergence for $v_n(p)$, namely [see MSZ (1993)]

$$|v_n(p) - v(p)| \leqslant \tilde{C} \frac{[\Sigma_{k \in K} \sqrt{p^k(1-p^k)}]^{2/3}}{\sqrt[3]{n}} \,, \tag{5.3}$$

for some constant $\tilde{C}$ which depends only on the game.

4.3. Existence and uniqueness of the solution of the functional equations

The pair of dual equations (1) and (2) that determine $v(p)$ are of interest and

can be analyzed without reference to game theoretic context and techniques. This was in fact done [see Mertens and Zamir (1977b), Sorin (1984b)] and the results can be summarized as follows:

Denote by $\mathscr{C}(\Delta)$ the space of all continuous functions on the simplex Δ, and by U the subset of $\mathscr{C}(\Delta)$ consisting of those functions that are "*u*-functions": values of $D(p)$, for some two-person, zero-sum game with incomplete information $\Gamma(p)$ with full monitoring. Denote by φ the mapping from U to $\mathscr{C}(\Delta)$ defined by $\varphi(u) = v = \lim v_n$ [using Theorem 4.2, this mapping is well defined since $\lim v_n$ is the same for all games $\Gamma(p)$ having the same *u*-function]. Let $\mathscr{C}(\Delta)$ be endowed with the topology of uniform convergence.

Proposition 4.3. (a) *U is a vector lattice[5] and a vector algebra[6] which contains all the affine functions.*

(b) *U is dense in $\mathscr{C}(\Delta)$.*

Proposition 4.4. *The mapping $\varphi: U \to \mathscr{C}(\Delta)$ has a unique continuous extension $\varphi: \mathscr{C}(\Delta) \to \mathscr{C}(\Delta)$. This extension is monotone and Lipschitz with constant 1 [or non-expansive, i.e. $\|\varphi(f) - \varphi(g)\| \leq \|f - g\|$].*

Theorem 4.5. *Consider the following functional inequalities and equations in which u, f and g denote arbitrary functions on the simplex Δ:*

(α) $f \geq \mathrm{Cav}_I \, \mathrm{Vex}_{II} \, \max\{u, f\}$;

(β) $f \leq \mathrm{Vex}_{II} \, \mathrm{Cav}_I \, \min\{u, f\}$;

(α') $g = \mathrm{Vex}_{II} \, \max\{u, g\}$;

(β') $g = \mathrm{Cav}_I \, \min\{u, g\}$.

There exists a monotone non-expansive mapping $\varphi: C(\Delta) \to C(\Delta)$ such that, for any $u \in \mathscr{C}(\Delta)$:

(i) *$\varphi(u)$ is the smallest f satisfying (α) and the largest f satisfying (β), and thus in particular it is the only solution f of the system (α)–(β).*

(ii) *$\varphi(u)$ is also the only solution g of the system (α')–(β').*

Theorem 4.6 [An approximation procedure for $\varphi(u)$]. *Define $\underline{v}_0 = -\infty$, $\bar{v}_0 = +\infty$, and for $n = 1, 2, \ldots$ let $\underline{v}_{n+1} = \mathrm{Cav}_I \, \mathrm{Vex}_{II} \, \max\{u, \underline{v}_n\}$ and $\bar{v}_{n+1} = \mathrm{Vex}_{II} \, \mathrm{Cav}_I \, \min\{u, \bar{v}_n\}$. Then $\{\underline{v}_n\}_{n=1}^{\infty}$ is monotonically increasing, $\{\bar{v}_n\}_{n=1}^{\infty}$ is monotonically decreasing and both sequences converge uniformly to $\varphi(u)$.*

Note that $\underline{v}_1$ (resp. $\bar{v}_1$) is the maxmin (resp. minmax) of $\Gamma_\infty(p)$ if $u(p)$ is the value of $D(p)$.

---

[5]That is, an ordered vector space V such that the maximum and the minimum of two elements of V exist (in V).

[6]That is, the product of two elements of U is in U.

4.4. The speed of convergence of $v_n(p)$

As mentioned in previous sections, the proofs for the convergence of $v_n(p)$ yield as a byproduct a bound for the speed of convergence: $1/\sqrt{n}$ for the full monitoring case [inequality (5.2)] and $1/\sqrt[3]{n}$ for the general signaling case [inequality (5.3)]. It turns out that these bounds are the best possible. In fact, games with these orders of speed of convergence can be found in the special case of incomplete information on one side.

Example 4.7. Consider the following game in which $k = \{1, 2\}$, player I is informed of the value of k, with full monitoring and payoff matrices:

$$G^1 = \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix},$$

and the prior probability distribution on K is $(p, 1-p)$.

For this game it is easily verified that $v_\infty(p) \equiv 0$. More precisely, we have [see Zamir (1971–72)]

$$\frac{p(1-p)}{\sqrt{n}} \leq v_n(p) \leq \frac{\sqrt{p(1-p)}}{\sqrt{n}}, \tag{5.4}$$

for all n and for all $p \in [0, 1]$.

Remaining in the framework of the previous example we change the payoffs and signals to obtain:

Example 4.8. Let $K = \{1, 2\}$. The payoff matrices G^1 and G^2 and the signaling matrices Q^1 and Q^2 (to player II) are given by

$$G^1 = \begin{pmatrix} 8 & 3 & -1 \\ 8 & -3 & 1 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 8 & 2 & -2 \\ 8 & -2 & 2 \end{pmatrix}$$

and

$$Q^1 = Q^2 = \begin{pmatrix} a & c & d \\ b & c & d \end{pmatrix}.$$

Here the signals (to player II) are deterministic (e.g. if top and middle are played the signal is c, etc.). Observe that deleting the left strategy of player II and changing the signaling matrices to provide full monitoring we obtain the game in Example 4.7. It is easily verified that the differences between the examples do not affect the value and in the game of this example we still have $v_\infty(p) \equiv 0$. However, in the present example player II is not informed of the

last move of his opponent unless he chooses his left strategy which is strictly dominated in terms of payoffs. In other words, player II has to pay 8 units whenever he wants to observe his opponent's move. Since observing the moves of the informed player is his only way to collect information about the state k, it is not surprising that his learning process will be slower and more costly than in Example 4.7. This yields a slower rate of convergence of v_n to v_∞, [see Zamir (1973a)]

$$\frac{p(1-p)}{\sqrt[3]{n}} \leqslant v_n(p) \leqslant \frac{\alpha\sqrt{p(1-p)}}{\sqrt[3]{n}}$$

for some positive constant α, for all n and for all $p \in [0, 1]$.

The speed of convergence of $v_n(p)$ can also be of lower order, such as $(\ln n)/n$, $1/n$. There are some partial results for classification of games according to those speeds [see Zamir (1971–72, 1973a)].

The special role of the normal distribution

One of the interesting, and still quite puzzling results in the study of the speed of convergence of $v_n(p)$ is the appearance of the normal distribution. Consider again the game in Example 4.7. It follows from inequality (5.4) that for any $0 < p < 1$, $\sqrt{n}v_n(p)$ is bounded between $p(1-p)$ and $\sqrt{p(1-p)}$. A natural question is then: Does this sequence converge? If it does, the limit is the coefficient of the leading term (i.e. $1/\sqrt{n}$) in the expansion of $v_n(p) - v_\infty(p)$ in fractional powers of n [recall that $v_\infty(p) \equiv 0$]. The sequence does turn out to converge and the limit is the well-known standard normal distribution function:

Theorem 4.9. *For all* $p \in [0, 1]$,

$$\lim_{n \to \infty} \sqrt{n}\, v_n(p) = \phi(p)\,,$$

where

$$\phi(p) = \frac{1}{\sqrt{2\pi}}\, e^{-(1/2)x_p^2} \quad \text{and} \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_p} e^{-(1/2)x^2}\, dx = p\,.$$

In words: the limit of $\sqrt{n}\, v_n(p)$ *is the standard normal density function evaluated at its p-quantile.*

The proof is rather technical [see Mertens and Zamir (1976b)] and does not give the intuition behind the result. It is based on a general result about the

variation of martingales in $[0, 1]$ [Mertens and Zamir (1977a)]. Let $\mathscr{X}_p^n = \{X_m\}_{m=1}^n$ denote an n-martingale bounded in $[0, 1]$ with $E(X_1) = p$, and let $V(\mathscr{X}_p^n)$ denote its L_1 variation, i.e.

$$V(\mathscr{X}_p^n) = \sum_{m=1}^{n-1} E(|X_{m+1} - X_m|) .$$

Then we have

Theorem 4.10. (The L_1 variation of a bounded martingale).

$$\lim_{n \to \infty} \sup_{\mathscr{X}_p^n} \left[\frac{1}{\sqrt{n}} V(\mathscr{X}_p^n) \right] = \phi(p) .$$

It turns out that this is not an isolated incident for this specific example but rather part of a general phenomenon. Consider a game with incomplete information on one side and two states of nature, each with a 2×2 payoff matrix. If the error term is of the order of $1/\sqrt{n}$, then $\sqrt{n}[v_n(p) - v_\infty(p)]$ tends (as $n \to \infty$) to an appropriately scaled normal density function [Mertens and Zamir (1990)]. This result was recently further generalized by De Meyer (1989) to any (finite) number of states of nature and any (finite) number of strategies for each player.

5. Incomplete information on two sides: The symmetric case

In the general situation of incomplete information on two sides, the case of state independent signaling treated in the previous section is the case with the most complete analysis. In this section we consider a special case in which signaling may be state dependent but it is *symmetric* in the sense that at each stage both players get the same signal.

Formally, we are given a finite collection of $S \times T$ payoff matrices $\{G^k\}_{k \in K}$ with initial probability p in $\Pi = \Delta(K)$, and the players have no initial information about the true state k except the prior distribution p. We denote by A the finite set of signals and by A^k the signaling matrix for state k. Given k and a pair of moves (s, t), a signal a is announced to both players according to a given probability distribution A_{st}^k on A. Assuming perfect recall means in this framework that for all k and k' in K, $s \neq s'$ or $t \neq t'$ implies that A_{st}^k and $A_{s't'}^{k'}$ have disjoint support.

Denoting the above described infinite game by $\Gamma_\infty(p)$ the result is:

Theorem 5.1. $\Gamma_\infty(p)$ *has a value.*

Proof. To see the idea of the proof consider first the special case in which the signals are deterministic – the support of $A_{s,t}^k$ consists of a single element of A (which will also be denoted by A_{st}^k). Define the set of non-revealing moves:

$$NR = \{(s, t) \in S \times T \mid A_{st}^k = A_{st}^{k'} \quad \forall k, k' \in K\} .$$

That is, a non-revealing move is one which gives no additional information about the state k and hence after a non-revealing move, the players face the same (infinite) game as the one they faced before that move. Whenever a move $(s, t) \notin NR$ is played and a certain signal a is announced, a non-empty subset of K is eliminated from the set of possible states, namely all states k for which $A_{st}^k \neq a$. The resulting situation is a game having the same data as the original one but with K replaced by a proper subset of itself, and the prior probability distribution on this smaller set is the normalization of its marginal according to p. Now if we prove our theorem by induction on $\#K$, then by the induction hypothesis the game resulting from a move not in NR has a value which can be guaranteed by both players from that stage on. In other words, using stochastic games terminology, the result of such a move is an *absorbing* state with payoff equal to that value.

Writing this formally, for each move (s, t) and for each signal a let $K_{st}(a) = \{k \in K \mid A_{st}^k = a\}$ and $p_{st}(a) = \Sigma_{k' \in K_{st}(a)} p^{k'}$. Let p_a be the probability distribution on $K_{st}(a)$ given by $p_a^k = p^k/p_{st}(a)$. Finally, denote by $v_{st}(a)$ the value of the game obtained from $\Gamma_\infty(p)$ when replacing K by $K_{st}(a)$ and p by p_a. The game $\Gamma_\infty(p)$ is equivalent to an $S \times T$ game with absorbing states in which the payoffs are given by (x^* indicates an absorbing state with payoff x):

$$\tilde{G}_{st} = \begin{cases} \Sigma_{k \in K} p^k G_{st}^k & \text{if } (s, t) \in NR , \\ (\Sigma_{a \in A} p_{st}(a) v_{st}(a))^* & \text{otherwise .} \end{cases}$$

Since this game (like any finite stochastic game) has a value, the original game $\Gamma_\infty(p)$ also has a value, completing the inductive step of the proof. $\square$

Remark. It is worth noting that historically the reduction of symmetric games of incomplete information to games with absorbing states was done before the latter were known to have a value [see Kohlberg and Zamir (1974)]. In fact, this focused attention on games with absorbing states and on the particular example of *The big match* treated by Blackwell and Ferguson (1968). The general solution of these games by Kohlberg (1974) then led to the solution of general stochastic games [Bewley and Kohlberg (1976a, 1976b, 1978), Mertens and Neyman (1981)].

In the general signaling case a "revealing" signal need not eliminate elements of K as impossible but rather it leads to a new (posterior) probability

distribution $p_1 \neq p$ on K. The value function is then a continuous function on $\pi = \Delta(K)$ and its existence is proved by induction on the dimension of this simplex [see Forges (1982) and MSZ (1993)].

6. Games with no signals

We consider here a class of games which was introduced by Mertens and Zamir (1976a) under the name "repeated games without a recursive structure". These games consist again of a finite collection of $S \times T$ payoff matrices G^k, $k \in K$, with an initial probability distribution p on K. No player is informed of the initial state. The signals are defined by a family of matrices A^k with deterministic entries (the extension to random signals is simple). Moreover, we assume that in each matrix A^k there are only two possible signals; either both players receive a "white" (totally uninformative) signal (0) or the game is completely revealed to both players. We can thus assume in the second case that the payoff is absorbing and equal to the value of the revealed game from this time on. It is then enough to define the strategies on the "white" histories; hence the name "game with no signals". Note that unlike the games considered in the previous section, the signal 0 does not include the moves of the players. By Dalkey's theorem [Dalkey (1953)], each player may be assumed to remember his own move, and hence the "white" signal is actually asymmetric information.

For a typical simple example of such games consider a game with two states, $\#S = \#T = 2$ and signaling matrices given by

$$A^1 = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix}.$$

To see the special feature of these games assume that in our example the prior probability distribution is $(1/2, 1/2)$, both players play at the first stage the mixed move $(1/2, 1/2)$ which results in (*top, right*) and a white signal (an event of probability $3/4$). Consequently, the posterior probability distribution of player I is $(1/3, 2/3)$ while that of player II is $(2/3, 1/3)$. The "state variable" of the problem can no longer be just a probability distribution on K. A larger and actually unbounded dimensional space is needed, and hence the name "games without a recursive structure". The above mentioned example was introduced and solved in Mertens and Zamir (1976a). The general case (i.e. general K and general size A^k) was solved by Waternaux (1983a, 1983b).

The analysis of these games brought about a new tool. The minmax and the maxmin of the game Γ_∞ are equal, respectively, to the values of two auxiliary one-shot games $\overline{G}$ and $\underline{G}$ in strategic form. The pure strategies in each of these

games mimic strategies in Γ_∞, and the payoffs are defined to be the correspond-ing asymptotic payoffs in Γ_∞ to the strategies which are mimicked. Fix p and write Γ for $\Gamma_\infty(p)$, and $\overline{v}(\Gamma)$ and $\underline{v}(\Gamma)$ for its minmax and maxmin, respec-tively.

Formally, define the (pure) strategy sets $\overline{X}$ and $\overline{Y}$, in the one-shot game $\overline{G}$, by

$$\overline{X} = \bigcup_{S' \subset S} \Delta(S') \times \mathbb{N}^{S \backslash S'} \times S' ,$$

$$\overline{Y} = \bigcup_{T' \subset T} \Delta(T') \times \mathbb{N}^{T \backslash T'} ,$$

where $\mathbb{N}$ is the set of positive integers. Given x in $\overline{X}$ (resp. y in $\overline{Y}$) we denote the corresponding subset S' by S^x, the first component by α^x, the second by c^x and the third by s^x.

The heuristic representation of the strategies in $\overline{G}$ is given by the following strategies in Γ_∞. For x in $\overline{X}$, player I plays i.i.d. the mixed move $\alpha^x \in \Delta(S^x)$ except for $c(x) = \Sigma_s c_s^x$ exceptional moves; each move s which is not part of the mixed move α^x (i.e. $s \notin S^x$) is played c_s^x times uniformly distributed before some large state N_0. From stage N_0 on, player I uses the (pure) move s^x.

A strategy $y \in \overline{Y}$ of player II has a similar meaning with the difference that, after stage N_0, he continues playing i.i.d. his mixed move (with no exceptional moves).

Note that these (behavioral) strategies are specified only for uninformative histories of the type $0...0$. As soon as a signal other than 0 appears, both players know the true payoff matrix G^k and the payoff stream in the super-game is assumed to be "absorbed" at $v(G^k)$ from that stage on.

The payoffs in $\overline{G}$ when x and y are used is defined as the asymptotic payoff corresponding to these strategies in the finite game [for formal definitions see Waternaux (1983a, 1983b) or MSZ (1993)].

Note that the players are not symmetric in $\overline{G}$ since this game is designed to provide the upper value $\overline{v}(\Gamma)$. In a dual way we define the strategy sets $\underline{X}$ and $\underline{Y}$ for the game $\underline{G}$ which provides the lower value $\underline{v}(\Gamma)$.

We first have:

Proposition 6.1. *The game $\overline{G}$ has a value $v(\overline{G})$ and both players have ϵ-optimal strategies.*

Theorem 6.2. (i) *$\overline{v}(\Gamma)$ exists and equals $v(\overline{G})$.*

(ii) *Player II has an ϵ-optimal strategy which is a finite mixture of i.i.d. sequences, each of which is associated with a finite number of exceptional moves, uniformly distributed before some stage N_0.*

(iii) *Dual results hold for $\underline{v}(\Gamma)$.*

It follows that the game under consideration has a value iff $v(\overline{G}) = v(\underline{G})$, which is generally not the case [examples of games where $v(\overline{G}) \neq v(\underline{G})$, were exhibited by Mertens and Zamir (1976a))]. In view of this, one is led to study $\lim v_n$ and $\lim v_\lambda$ [Sorin (1989)]. Sorin's approach is similar to that adopted in the study of Γ_∞; that is, using more manageable auxiliary "approximating games" as follows. For each L in $\mathbb{N}$ we construct a game G_L. The heuristic interpretation of G_L is Γ_n played in L large blocks, during each of which both players use stationary strategies, except for some singular moves. The strategy sets in G_L are $\overline{X}^L$ and $\overline{Y}^L$, where

$$\bar{X} = \bigcup_{S' \subset S} \Delta(S') \times \mathbb{N}^{S \setminus S'} ,$$

$$\bar{Y} = \bigcup_{T' \subset T} \Delta(T') \times \mathbb{N}^{T \setminus T'} .$$

Again, the payoff to a pair of strategies is defined as the corresponding asymptotic average payoff (as the block size tends to ∞). Then we have first:

Proposition 6.3. G_L has a value w_L and both players have ϵ-optimal strategies.

Theorem 6.4. $\lim_{n \to \infty} v_n$ and $\lim_{L \to \infty} w_L$ exist and coincide.

Then, a similar construction gives:

Theorem 6.5. $\lim v_\lambda$ exists and $\lim_{\lambda \to 0} v_\lambda = \lim_{L \to \infty} w_L$.

7. A game with state dependent signaling

For games with incomplete information on two sides, the general results so far are mainly those described in Section 4. In that section we considered the special case in which the signals provided to the players after each stage do not depend on the state k (but only on the player's moves). When the signals depend also on the states, we have results only for two special cases: the symmetric case (Section 5), and "games with no signals" (Section 6).

In this section we briefly introduce another game with state dependent signals which was studied by Sorin (1985b). This work illustrates an example of a game at the forefront of the research in games with incomplete information. It is not only strongly related to stochastic games (as were the games studied in the previous two sections), but it involves what may be called *stochastic games with incomplete information*.

Consider the class of games with lack of information on both sides (and state

dependent signaling) given by the following data: $K = \{0, 1\}^2 = L \times M$ [we write $k = (l, m)$], and the probability on K is the product $p \otimes q$ of its marginals.

At stage 0, player I is informed about l and player II about m. The payoffs are defined by 2×2 payoffs matrices G^{lm}, and the signaling matrices are given by

$$A^{11} = \begin{pmatrix} T & L \\ c & d \end{pmatrix}, \qquad A^{10} = \begin{pmatrix} T & R \\ c & d \end{pmatrix},$$

$$A^{01} = \begin{pmatrix} B & L \\ c & d \end{pmatrix}, \qquad A^{00} = \begin{pmatrix} B & R \\ c & d \end{pmatrix}.$$

The special features of this information structure to be noted are:

(a) The signals include the moves.

(b) As soon as player I plays Top, the "type" of one of the players is revealed: l if player II played Left at that stage, m if he played Right.

Denoting this game by $\Gamma(p, q)$, we note first that as soon as $p^1 p^0 q^1 q^0 = 0$, it is reduced to a game with incomplete information on one side (treated in Section 2). In particular it has a value $v(p, q)$.

Sorin has given explicit expressions for the minmax and maxmin of these games, which will not be given here. We just mention that these also rely on a family of auxiliary games which are of the form:

$$G^1 = \begin{pmatrix} a_{11}^* & a_{12}^* \\ a_{21} & a_{22} \end{pmatrix}, \qquad G^0 = \begin{pmatrix} b_{11}^* & b_{12}^* \\ b_{21} & b_{22} \end{pmatrix}, \qquad p = (p^1, p^0).$$

That is, the auxiliary games are repeated games of incomplete information on one side in which the games G^1 and G^0 are stochastic games with absorbing states (more specifically, "Big match" type games). In fact, when studying the game under consideration Sorin found the minmax and the maxmin of this family of games and of the dual family in which the absorbing states are in the columns, i.e. in the control of the uninformed player:

$$G^1 = \begin{pmatrix} a_{11}^* & a_{12} \\ a_{21}^* & a_{22} \end{pmatrix}, \qquad G^0 = \begin{pmatrix} b_{11}^* & b_{12} \\ b_{21}^* & b_{22} \end{pmatrix}, \qquad p = (p^1, p^0).$$

The analysis of these games, which is beyond the scope of this review, is rather deep and involves new ideas and tools developed specifically for this purpose.

8. Miscellaneous results

In this section we mention some interesting results which somehow remained isolated and were not followed by further research.

8.1. Discounted repeated games with incomplete information

Mayberry (1967) studied a game with incomplete information on one side, full monitoring, and λ-discounted payoff. Specifically, he considered the game in Example 1.1. Denoting this game by $\Gamma_\lambda(p)$ and its value by $v_\lambda(p)$, he first derived the following formula:

$$v_\lambda(p) = \max_{s,t}\{\lambda \min(ps, p't') + (1 - \lambda)(\bar{s}v_\lambda(ps/\bar{s}) + \bar{s}'v_\lambda(ps'/\bar{s}'))\} . \quad (5.5)$$

Here, for $x \in [0, 1]$, x' stands for $(1 - x)$, and $(s, t) \in X^2$ is the pair of mixed moves used by player I in one stage [that is, play $(s, 1 - s)$ in game G^1 and $(t, 1 - t)$ in game G^2], and $\bar{s} = ps + p't$.

Using this formula and the concavity of v_λ, it can be proved that the value of v_λ at any rational $p = n/m \le 1/2$ is given in terms of v_λ at some other rational numbers q with smaller denominator.

By differentiating (5.5) we obtain (letting $v_\lambda' = dv_\lambda/dp$):

$$v_\lambda'(p) = (1 - \lambda)(1 - p/p')v_\lambda'(p/p') - (1 - \lambda)v_\lambda(p/p') . \quad (5.6)$$

From this it follows (using the symmetry of v_λ) that for $2/3 < \lambda < 1$, the function has a left derivative and a right derivative at $p = 1/2$, *but they are not equal*.

By induction on the denominator, one can then prove that for any rational p, the sequence of derivatives obtained by repeated use of equation (5.6) leads to an expression for $v_\lambda'(p)$ in terms of $v_\lambda'(0)$, $v_\lambda'(1)$ and $v_\lambda'(1/2)$.

Combining the last two results we conclude, for $2/3 < \lambda < 1$, that although v_λ is concave, it has discontinuous derivatives at every rational point.

8.2. Sequential games

Sequential games with incomplete information were first studied by Ponssard in a series of papers [Ponssard (1975a, 1975b, 1976)] and also by Ponssard and Zamir (1973), Ponssard and Sorin (1980a, 1980b) and Sorin (1979). The basic model is the following. The players' type sets are $K = \{1, \ldots, L\}$ for player I and $R = \{1, \ldots, M\}$ for player II. For each pair of types (k, r) the corresponding payoff matrix is $G^{kr} = (G^{kr}_{st})_{s \in S, t \in T}$. For each $p \in P = \Delta(K)$ and $q \in Q = \Delta(R)$, the n-stage sequential game $\Gamma_n(p, q)$ is played as follows:
- At stage 0, a chance move chooses independently k according to p and r according to q. Player I is informed of (his type) k and player II is informed of r.

- At stage m $(m = 1, \ldots, n)$, knowing $h_m = (s_1, t_1, \ldots, s_{m-1}, t_{m-1})$, the history of the moves up to that stage, player I chooses s_m in S. This is told to player II who then, knowing $h'_m = (s_1, t_1, \ldots, s_{m-1}, t_{m-1}, s_m)$, chooses t_m.
- At the end of n stages player II pays player I the amount $(1/n) \sum_{m=1}^n G^{kr}_{s_m t_m}$.

Let $v_n(p, q)$ denote the value of $\Gamma_n(p, q)$ and $v(p, q) = \lim_{n \to \infty} v_n(p, q)$. Clearly, by normalizing the strategies of player II at each stage, this is shown to be a special case of the simultaneous repeated games discussed in Section 3, in which the payoff matrices are of size $|S| \times |T|^{|S|}$ (a move of player II is an element of T depending on the choice of player I at that stage). However, it turns out that stronger results hold for this case because of its special structure.

The (behavior) strategies σ_n and τ_n of $\Gamma_n(p, q)$ are defined in the natural way as sequences of mappings from the player's type and available history to the set of his mixed moves [$\Delta(S)$ and $\Delta(T)$, respectively]. The non-revealing game is again the one-shot sequential game with payoff matrix $G(p, q) = \sum_{k,r} p^k q^r G^{k,r}$, and its value is therefore

$$u(p, q) = \max_s \min_t \sum_{k,r} p^k q^r G^{kr}.$$

8.2.1. Incomplete information on one side

For incomplete information on the side of player II (the minimizer and the second to move), it was proved by Ponssard and Zamir (1973) that:

Proposition 8.1. *For all $p \in P$, $v_1(p) = \mathrm{Cav}_p u(p)$.*

Using the monotonicity of $v_n(p)$ (Proposition 3.19), one has:

Corollary 8.2. $v_n(p) = \mathrm{Cav}_p u(p)$, *for all n and all $p \in P$. Consequently* $\lim_{n \to \infty} v_1(p) = \mathrm{Cav}_p u(p)$.

8.2.2. Incomplete information on two sides

In this case one can prove a recursive formula for $v_n(p, q)$ which is much simpler than the corresponding formula for the general simultaneous move game:

$$v_{n+1}(p, q) = \frac{1}{n+1} \mathrm{Cav}_p \max_s \mathrm{Vex}_q \min_t \left\{ \sum_{k,r} p^k q^r G^{kr}_{st} + n v_n(p, q) \right\}.$$

Using this, it was proved by Sorin (1979) that for all p and all q the sequence $v_n(p, q)$ is increasing (and therefore it converges), the speed of convergence is

bounded by

$$0 \le v(p, q) - v_n(p, q) \le \frac{C}{n}$$

for some positive constant C, and that this is the best bound.

8.3. A game with incomplete information played by "non-Bayesian players"

Megiddo (1980) considered a game with incomplete information on one side in which there is no given prior on the states of nature. More specifically, the uninformed player II knows only the set of his moves (columns) and is told his payoff at each stage. Megiddo provided an algorithm to construct an optimal strategy for the uninformed "non-Bayesian" player. Basically the algorithm considered a dense grid of games with a given number of columns, and tested statistically the performance of each strategy which is optimal in one of these games.

Looking carefully at the problem, it turns out that this result can be derived as a consequence of the general results in Section 2 along the following lines [Mertens (1987)].

(a) Assume first that the unknown payoff matrix is an element of a finite set $(G^k)_{k \in K}$ of matrices having the same set J of columns and any (finite or infinite) number of rows.

(b) Since player II is told his payoff at each stage, any non-revealing strategy $\sigma \in NR(p)$ yields the same distribution of payoffs, and a fortiori the same expected payoff, in all games G^k in the support of p, for all columns $j \in J$. It follows that $u(p)$ is constant in the interior of each facet of the simplex $\Delta(K)$. Since u is upper-semicontinuous, this implies that Cav u *is linear in p* on $\Delta(K)$.

(c) Since Cav $u(p)$ is linear, player II has a strategy (in Γ_∞) which guarantees $v(G^k)$ if the true state is k, for all k. In fact, any optimal strategy $\tau(p)$ of player II at some interior point p has this property [otherwise player I could obtain against $\tau(p)$ strictly more than $v(G^{k_0})$ at some state k_0 and, by playing optimally at each other state, he could get strictly more than $\Sigma_k p^k v(G^k) =$ Cav $u(p)$, contradicting the optimality of $\tau(p)$].

(d) These results are valid not only for a finite state set K but also for a countable K. In particular, if we consider the countable set $\mathcal{G}$ of all finite matrices with J columns and rational entries, it follows that if the true game is in $\mathcal{G}$, then player II has a strategy τ^r which guarantees its value. To extend this to any real entries, we perform the following approximation procedure.

(e) For any $\epsilon > 0$ let τ_ϵ^r be the strategy of player II which consists of playing τ^r while "rationalizing" the histories as follows: if the announced payoff (at

some stage) is α, replace it by a rational number $r(\alpha) \geq \alpha$ such that $r(\alpha) - \alpha < \epsilon$. Clearly, for any play of the game induced by τ_ϵ^r there is a $G^\epsilon \in \mathcal{G}$ with $\|G^\epsilon - G\| < \epsilon$ such that if it was the true game instead of G, it would have induced the same play when player II is using τ^r, and hence the expected payoff would be at most $v(G^\epsilon)$, which is at most $v(G) + \epsilon$.

We conclude that for any $\epsilon > 0$ player II has a strategy τ_ϵ^r which guarantees $v(G) + \epsilon$.

(f) Finally, choose a sequence $\{\epsilon_n\}_{n=1}^\infty$ decreasing to 0 and play successively $\tau_{\epsilon_n}^r$ in large blocks with appropriately increasing sizes so that the resulting strategy guarantees $v(G) + \epsilon_n$ for all n and hence it guarantees $v(G)$. □

The main idea of this argument is that the announcement of the payoffs induces the linearity of Cav u, which in turn implies the existence of a strategy for the uninformed player which is uniformly optimal for all prior distributions on K. This is the sense in which the player is non-Bayesian, since he does not need any prior in order to play his optimal strategy.

8.4. A stochastic game with signals

Ferguson, Shapley and Weber (1970) considered the following game which was the first treated example of a stochastic game with incomplete information.

We are given two states of nature with the following payoff matrices.:

$$G^1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad G^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The transition probability from state 1 to state 2 is a constant $(1 - \pi) \in (0, 1)$, independent of the moves. The reverse transition, from state 2 to state 1, takes place if and only if player I plays Bottom. Player I knows everything while player II is told only the times of the transition from 2 to 1.

Let us consider Γ_λ, the discounted game starting from $k = 1$, and write v_λ for its value. It can be shown that

$$v_\lambda = \frac{[1 - (1 - \lambda)^k][1 - \pi(1 - \lambda)] - \lambda\{[1 - 2[\pi(1 - \lambda)]^k\}}{[1 - (1 - \lambda)^{k+1}][1 - \pi(1 - \lambda)] + 2(1 - \lambda)^{k+1}\pi^k\lambda} .$$

Letting

$$v_0 = \lim_{\lambda \to 0} v_\lambda = \frac{r(1 - \pi) - (1 - 2\pi^r)}{(r + 1)(1 - \pi) + 2\pi^r}$$

where r is the positive integer satisfying $\pi^{r-1} > 1/2$, and $\pi^r \leq 1/2$, one can then find optimal strategies σ^* and τ^* for the two players such that for each ϵ,

each player (with his optimal strategy), can ϵ-guarantee v_0 in all Γ_λ with sufficiently small λ [for details see MSZ (1993)].

Bibliography

Aumann, R.J. (1964) 'Mixed and behavior strategies in infinite extensive games', in: M. Dresher et al., eds., *Advances in game theory*, Annals of Mathematical Studies, Vol. 52: Princeton: Princeton University Press, pp. 627–650.

Aumann, R.J. (1985) 'Repeated games', in: G.R. Feiwel, ed., *Issues in contemporary micro-economics and welfare*. New York: Macmillan, pp. 109–242.

Aumann, R.J. and M. Maschler (1966) *Game-theoretic aspects of gradual disarmament*, Reports to the U.S. Arms Control and Disarmament Agency, ST-80, Chapter V, pp. 1–55.

Aumann, R.J. and M. Maschler (1967) *Repeated games with incomplete information: A survey of recent results*, Reports to the U.S. Arms Control and Disarmament Agency, ST-116, Chapter III, pp. 287–403.

Aumann, R.J. and M. Maschler (1968) *Repeated games of incomplete information: The zero-sum extensive case*, Reports to the U.S. Arms Control and Disarmament Agency ST-143, Chapter II, pp. 37–116.

Aumann, R.J., M. Maschler and R. Stearns (1968) *Repeated games of incomplete information: An approach to the nonzero sum case*, Reports to the U.S. Arms Control and Disarmament Agency, ST-143, Chapter IV, pp. 117–216.

Bewley, T. and E. Kohlberg (1976a) 'The asymptotic theory of stochastic games', *Mathematics of Operations Research*, 1: 197–208.

Bewley, T. and E. Kohlberg (1976b) 'The asymptotic solution of a recursion equation occurring in stochastic games', *Mathematics of Operations Research*, 1: 321–336.

Bewley, T. and E. Kohlberg (1978) 'On stochastic games with stationary optimal strategies', *Mathematics of Operations Research*, 3: 104–125.

Blackwell, D. (1956) 'An analog of the minmax theorem for vector payoffs', *Pacific Journal of Mathematics*, 65: 1–8.

Blackwell, D. and T.S. Ferguson (1968) 'The big match', *Annals of Mathematical Statistics*, 39: 159–163.

Dalkey, N. (1953) 'Equivalence of information patterns and essentially determinate games', in: H.W. Kuhn and A.W. Tucker, eds., *Contributions to the theory of games*, Vol. II, Annals of Mathematics Study, 28: Princeton: Princeton University Press, pp. 217–243.

Ferguson, T.S., L.S. Shapley and R. Weber (1970) 'A stochastic game with incomplete information'. Mimeograph.

Forges, F. (1982) 'Infinitely repeated games of incomplete information: symmetric case with random signals', *International Journal of Game Theory*, 11: 203–213.

Harsanyi, J.C. (1967–68) 'Games of incomplete information played by Bayesian players. Parts I, II, III', *Management Science*, 14: 159–182, 320–334, 486–502.

Kohlberg, E. (1974) 'Repeated games with absorbing states', *Annals of Statistics*, 2: 724–738.

Kohlberg, E. (1975a) 'Optimal strategies in repeated games with incomplete information', *International Journal of Game Theory*, 4: 7–24.

Kohlberg, E. (1975b) 'The information revealed in infinitely-repeated games of incomplete information', *International Journal of Game Theory*, 4: 57–59.

Kohlberg, E. and S. Zamir (1974) 'Repeated games of incomplete information: the symmetric case', *Annals of Statistics*, 2: 1040–1041.

Lehrer, E. (1987) 'A note on the monotonicity of v_n', *Economic Letters*, 23: 341–342.

Mayberry, J.-P. (1967) 'Discounted repeated games with incomplete information', *Mathematica*, ST-116, Ch. V, pp. 435–461.

Megiddo, N. (1980) 'On repeated games with incomplete information played by non-Bayesian players', *International Journal of Game Theory*, 9: 157–167.

Mertens, J.-F. (1972) 'The value of two-person zero-sum repeated games: The extensive case', *International Journal of Game Theory*, 1: 217–225.

Mertens, J.-F. (1973) 'A note on "the value of two-person zero-sum repeated games: The extensive case"', *International Journal of Game Theory*, 9: 189–190.

Mertens, J.-F. (1982) 'Repeated games: An overview of the zero-sum case', in: W. Hildenbrand, ed., *Advances in economic theory*, Cambridge: Cambridge University Press.

Mertens, J.-F. (1986) 'The minmax theorem for u.s.c.-l.s.c. payoff functions', *International Journal of Game Theory*, 10(2): 53–56.

Mertens, J.-F. (1987) 'Repeated games', *Proceedings of the International Congress of Mathematicians* (Berkeley), 1986, American Mathematical Society. pp. 1528–1577.

Mertens, J.-F. and A. Neyman (1981) 'Stochastic games', *International Journal of Game Theory*, 10(2): 53–56.

Mertens, J.-F. and A. Neyman (1982) 'Stochastic games have a value', *Proceedings of the National Academy of Sciences of the U.S.A.*, 79: 2145–2146.

Mertens, J.-F. and S. Zamir (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, J.-F. and S. Zamir (1976a) 'On a repeated game without a recursive structure', *International Journal of Game Theory*, 5: 173–182.

Mertens, J.-F. and S. Zamir (1976b) 'The normal distribution and repeated games', *International Journal of Game Theory*, 5: 187–197.

Mertens, J.-F. and S. Zamir (1977a) 'The maximal variation of a bounded martingale', *Israel Journal of Mathematics*, 27: 252–276.

Mertens, J.-F. and S. Zamir (1977b) 'A duality theorem on a pair of simultaneous functional equations', *Journal of Mathematical Analysis and Application*, 60: 550–558.

Mertens, J.-F. and S. Zamir (1980) 'Minmax and Maxmin of repeated games with incomplete information', *International Journal of Game Theory*, 9: 201–215.

Mertens, J.-F. and S. Zamir (1981) 'Incomplete information games with transcendental values', *Mathematics of Operations Research*, 6: 313–318.

Mertens, J.-F. and S. Zamir (1985) 'Formulation of Bayesian analysis for games with incomplete information', *International Journal of Game Theory*, 14: 1–29.

Mertens, J.-F. and S. Zamir (1990) 'Incomplete information games and the normal distribution', mimeograph.

Mertens, J.-F., S. Sorin and S. Zamir (1993) *Repeated games*, forthcoming.

De Meyer, B. (1989) 'Repeated games and multidimensional normal distribution', CORE Discussion Paper 89, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Ponssard, J.-P. (1975a) 'Zero-sum games with "almost" perfect information', *Management Science*, 21: 794–805.

Ponssard, J.-P. (1975b) 'A note on the L-P formulation of zero-sum sequential games with incomplete information', *International Journal of Game Theory*, 4: 1–5

Ponssard, J.-P. (1976) 'On the subject of nonoptimal play in zero-sum extensive games: The trap phenomenon', *International Journal of Game Theory*, 5: 107–115.

Ponssard, J.-P. and S. Sorin (1980a) 'The LP formulation of finite zero-sum games with incomplete information', *International Journal of Game Theory*, 9: 99–105.

Ponssard, J.-P. and S. Sorin (1980b) 'Some results on zero-sum games with incomplete information: The dependent case', *International Journal of Game Theory*, 9: 233–245.

Ponssard, J.-P. and S. Sorin (1982) 'Optimal behavioral strategies in zero-sum·games with almost perfect information', *Mathematics of Operations Research*, 7: 14–31.

Ponssard, J.-P. and S. Zamir (1973) 'Zero-sum sequential games with incomplete information', *International Journal of Game Theory*, 2: 99–107.

Sorin, S. (1979) 'A note on the value of zero-sum sequential repeated games with incomplete information', *International Journal of Game Theory*, 8: 217–223.

Sorin, S. (1980) 'An introduction to two-person zero-sum repeated games with incomplete information', IMSS-Economics, Stanford University, TR 312. French version in *Cahiers du Groupe de Mathématiques Economiques*, 1: Paris (1979).

Sorin, S. (1984a) '"Big match" with lack of information on one side (part I)', *International Journal of Game Theory*, 13: 201–255.

Sorin, S. (1984b) 'On a pair of simultaneous functional equations', *Journal of Mathematical Analysis and Applications*, 98(1): 296–303.

Sorin, S. (1985a) '"Big match" with lack of information on one side (part II)', *International Journal of Game Theory*, 14: 173–204.

Sorin, S. (1985b) 'On a repeated game with state dependent signaling matrices', *International Journal of Game Theory*, 14: 249–272.

Sorin, S. (1986) 'On repeated games with complete information', *Mathematics of Operations Research*, 11: 147–160.

Sorin, S. (1989) 'On repeated games without a recursive structure: Existence of lim v_n', *International Journal of Game Theory*, 18: 45–55.

Sorin, S. and S. Zamir (1985) 'A 2-person game with lack of information on $1\frac{1}{2}$ sides', *Mathematics of Operations Research*, 10: 17–23.

Sorin, S. and S. Zamir (1991) '"Big match" with lack of information on one side (III)', in: T.E.S. Raghavan, T.S. Ferguson, T. Parthasarthy and O.J. Vrieze, eds., *Stochastic games and related topics in honor of Professor L.S. Shapley*. Kluwer Academic Publishers, pp. 101–112.

Stearns, R.E. (1967) 'A formal information concept for games with incomplete information', *Mathematica*, ST-116, Ch. IV, pp. 405–433.

Waternaux, C. (1983a) 'Solution for a class of games without recursive structure', *International Journal of Game Theory*, 12: 129–160.

Waternaux, C. (1983b) 'Minmax and maxmin of repeated games without a recursive structure', Core Discussion Paper 8313, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, to appear in *International Journal of Game Theory*.

Zamir, S. (1971–72) 'On the relation between finitely and infinitely repeated games with incomplete information', *International Journal of Game Theory*, 1: 179–198.

Zamir, S. (1973a) 'On repeated games with general information function', *International Journal of Game Theory*, 2: 215–229.

Zamir, S. (1973b) 'On the notion of value for games with infinitely many stages', *Annals of Statistics*, 1: 791–796.