TOPICS IN NONCOOPERATIVE GAME THEORY

by

Shmuel Zamir

Hebrew University of Jerusalem, Israel

Table of Contents

Chapter 1 - Minmax and Equilibria

Chapter 2 - Games in Extensive Form

Chapter 3 - Multistage Games

Chapter 4 - Modeling Incomplete Information

Chapter 5 - Repeated Games with Incomplete Information (I)

Chapter 6 - Repeated Games with Incomplete Information (II)

Chapter 1

MINMAX AND EQUILIBRIA

The first and simplest game theoretical model we shall discuss is meant to describe the following interactive decision situation: Two decision makers, called player I and player II have to choose an action each. Player I chooses an element x of a set X while player II chooses an element y of a set Y. The choices are done simultaneously and the chosen x and y determine a certain money (or utility) transfer between the players. This motivates the following.

Definition 1.1 A *two-person zero-sum game* is an ordered triple (X, Y, h) where X and Y are sets and h is a real-valued function defined on the product set $X \times Y$.

Remarks: 1) The sets X and Y will be referred to as the *strategy sets* of player I and II, respectively. The function h is called the *pay-off function.* For $x \in X$ and $y \in Y$, $h(x, y)$ is interpreted as the amount of money that player II pays player I if I chooses x and II chooses y. In view of this interpretation player I will also be called the *maximizer* and player II the *minimizer.*

2) Two aspects of the interpretation should be emphasized for future reference: (i) The strategy choices are done *simultaneously* and *independently*, e.g. each player hands his choice to a referee who then announces (x, y) and executes the pay-off. (ii) The data of the game, namely (X, Y, h) are 'publicly' known to both players, what we shall later call a *common knowledge.*

Throughout this lecture, unless we specify otherwise, we shall say for convenience 'a game' instead of 'a two-person zero-sum game'.

Example 1.2 The special case in which X and Y are finite will be called a *finite game* or a *matrix game.* In this case, the function h is described as a *pay-off matrix* A whose rows names are labelled by the elements of X (usually denoted as $M = \{1,\ldots,m\}$) and the columns by the elements of Y (denoted as $N = \{1,\ldots,n\}$). Examples of matrix games are:

(i)

i \ j	1	2	3	4
1	-2	2	-1	4
2	1	6	1	2
3	9	-8	-3	-1

(ii)

i \ j	1	2
1	1	-1
2	-1	1

.

Definition 1.3 The *upper value* of the game $G = (X, Y, h)$ is $\bar{v} \in [-\infty,\infty]$ defined by: $\bar{v} = \inf_{y \in Y} \sup_{x \in X} h(x, y)$. The *lower value* is $\underline{v} = \sup_{x \in X} \inf_{y \in Y} h(x, y)$.

It readily follows that for any game G :

$$\underline{v} = \sup_{x \in X} \inf_{y \in Y} h(x, y) \leq \inf_{y \in Y} \sup_{x \in X} h(x, y) = \bar{v} \quad . \tag{1.1}$$

Strategies $x^* \in X$ and $y^* \in Y$ for which $\underline{v} = \inf_{y \in Y} h(x^*, y)$ and $\bar{v} = \sup_{x \in X} h(x, y^*)$ are called *minmax strategies* of the respective players. That is, if player I has a minmax strategy then the sup may be replaced by max (similarly for player II). In example (i): $\underline{v} = \max_i \min_j a_{ij} = 1$ (a minmax strategy $i = 2$)

$\bar{v} = \min_j \max_i a_{ij} = 1$ (a minmax strategy $j = 3$) .

In example (ii): $\underline{v} = -1$ (both $i = 1$ and $i = 2$ are minmax strategies);
$\bar{v} = 1$ (both $j = 1$ and $j = 2$ are minmax strategies).

Definition 1.3 A game $G = (X, Y, h)$ is said to have a *value* (or a *minmax value*) if both players have minmax strategies and:

$$\sup_{x \in X} \inf_{y \in Y} h(x, y) = \inf_{y \in Y} \sup_{x \in X} h(x, y) = v \quad . \tag{1.2}$$

v is called the value of the game and the minmax strategies are then also called *optimal strategies*. So the game (i) has a value $v = 1$ with $i = 2$; $j = 3$ as optimal strategies while game (ii) does not have a value.

Definition 1.4 A pair of strategies $(x_0, y_0) \in X \times Y$ is called a *saddle-point* of the game $G = (X, Y, h)$ if

$$h(x, y_0) \leq h(x_0, y_0) \leq h(x_0, y) \quad \forall \quad (x, y) \in X \times Y \quad .$$

In game (i): (2,3) is a saddle-point with a corresponding pay-off $h(2, 3) = 1$ (which is the value of the game).

In game (ii): There is no saddle-point.

The relation between the notions of the minmax value and the saddle-point is formulated in the following lemma whose proof is rather simple and will be omitted.

Lemma 1.5 A game $G = (X, Y, h)$ has a value if and only if it has a saddle-point. In such a case:

(i) The value of the game is the pay-off corresponding to the saddle-point.

(ii) Any pair of optimal strategies is a saddle-point and any saddle-point consists of a pair of optimal strategies.

In view of non-existence of the value for matrix games such as game (ii), it is self-suggested that a player can sometimes do better by choosing his strategy randomly. For instance, if in (ii) plyaer I chooses his two strategies each with probability ½ his expected pay-off will be 0 independently of what player II does (compared to his security level $\underline{v} = -1$). This motivates the following definition:

<u>Definition 1.6</u> The *mixed extension* of a matrix game $G_0 = (M, N, A)$ is the game $G = (X, Y, H)$ where

$$X = \{x \in R^m | x_i \geq 0, \quad \forall\ i \in M ; \quad \sum_{i \in M} x_i = 1\}$$

$$Y = \{y \in R^n | y_j \geq 0 , \quad \forall\ j \in N ; \quad \sum_{j \in N} y_j = 1\}$$

$$H(x, y) = xA\tilde{y} \qquad (x \text{ and } y \text{ are rows, } \tilde{y} \text{ is the transposition of } y).$$

In other words, the strategy sets in G are the $(m - 1)$ and $(n - 1)$ dimensional simplices of probability distributions on M and N, respectively. The pay-off function H is just the expectation of the random pay-off a_{ij}. The extreme points of X (or Y) can be identified with the strategy set M (or N) in G_0. They are therefore termed *pure strategies* compared to *mixed strategies*, which is the name for general elements of X and Y.

<u>Example.</u> The mixed extension of the matrix game (ii) is $G = (X; Y\ H)$ where:

$$X = \{x = (x, 1 - x) | 0 \leq x \leq 1\} ; \qquad Y = \{y = (y, 1 - y) | 0 \leq y \leq 1\}$$

$$H(x, y) = xA\tilde{y} = 4xy - 2x - 2y + 1 \quad .$$

Unlike the original game this game has a value 0 and optimal strategies which are $(\frac{1}{2}, \frac{1}{2})$ for both players.

<u>Theorem 1.7</u> (The Minmax Theorem, J. Von-Neumann, 1928). *The mixed extension of a (finite) matrix game has a value.*

<u>Proof</u>. Let $A = (a_{ij})$ be an $m \times n$ matrix game in which the (pure) strategy sets are $M = \{1,\ldots,m\}$, $N = \{1,\ldots,n\}$. In view of Lemma 1.5 it is enough to prove the existence of a saddle-point for the mixed extension, i.e. the existence of $x^* \in X$ and $y^* \in Y$ s.t. $\forall\ x \in X$ and $\forall\ y \in Y$:

$$H(i, y^*) \leq H(x^*, y^*) \leq H(x^*, j) \tag{1.3}$$

Here i and j stand for e^i - the i-th unit vector in X - and e^j the j-th unit vector in Y - respectively, i.e. $\forall\ x \in X \quad \forall\ y \in Y$:

$$H(i, y) = e^i A\tilde{y} = \sum_{j \in N} a_{ij} y_j$$

$$H(x, j) = xA\tilde{e}^j = \sum_{i \in M} a_{ij}\, x_i \quad .$$

Consider now the product space $S = X \times Y$ and define $f: S \to R^{m+n}$ as follows: For $s = (x, y) \in S$, $f(s) = (f_1(s),\ldots,f_m(s);\ f^1(s),\ldots,f^n(s))$ where:

$$f_i(s) = \max(H(i, y) - H(x, y), 0) \; ; \; \forall \; i \in M$$

$$f^j(s) = \max(H(x, y) - H(x, j), 0) \; ; \; \forall \; j \in N \qquad (1.4)$$

Define a mapping $F: S \to S$ by: For $s = (x, y) \in S$,

$F(s) = (F_1(s), \ldots, F_m(s); \; F^1(s), \ldots, F^n(s))$ where

$$F_i(s) = \frac{x_i + f_i(s)}{1 + \sum_{\ell \in M} f_\ell(s)} \; , \; \forall \; i \in M$$

$$F^j(s) = \frac{y_j + f^j(s)}{1 + \sum_{\ell \in N} f^\ell(s)} \; , \; \forall \; j \in N \qquad (1.5)$$

S is a convex compact set (in R^{m+n}). H and f are continuous functions and therefore F is continuous. It follows by Brouwer's fixed point theorem that there exists $s^* = (x^*, y^*) \in S$ s.t. $F(s^*) = s^*$. By (1.5) this implies

$$f_i(s^*) = x_i^* \sum_{\ell \in M} f_\ell(s^*) \; ; \; \forall \; i \in M$$

$$f^j(s^*) = y_j^* \sum_{\ell \in N} f^\ell(s^*) \; ; \; \forall \; j \in M \qquad (1.6)$$

Claim. There exists $i \in M$ s.t. $x_i^* > 0$ and $f_i(s^*) = 0$.

Assume this is not true. Using the definition of f we would have that

$x_i^* > 0$ implies $f_i(s^*) > 0$ i.e. $H(i, y^*) > H(x^*, y^*)$. Thus:

$$\sum_{\{i | x_i^* > 0\}} x_i^* H(i, y^*) > \sum_{\{i | x_i^* > 0\}} x_i^* H(x^*, y^*) \; ,$$

which implies $\sum_{i \in M} x_i^* H(i, y^*) > H(x^*, y^*) \sum_{i \in M} x_i^*$, a contradiction since both sides equal $H(x^*, y^*)$. It follows from this claim that $\sum_{\ell \in M} f_\ell(s^*) = 0$ and since $f_\ell(s^*) \geq 0$ (by definition of f) it follows from (1.4) that $H(i, y^*) \leq H(x^*, y^*)$ $\forall \; i \in M$ which is one part of (1.3). The second part is proved in the same way showing that (x^*, y^*) is an equilibrium point and thus (by lemma 1.5) it is also a pair of optimal strategies and $H(x^*, y^*)$ is the value of the game.

Q.E.D.

Remark. The proof of the Minmax theorem given here is due to John Nash. Of the many other proofs of the theorem, at least two should be mentioned: the one using the duality theorem in linear programming, and the one using a separating hyperplane argument. Actually, the Minmax theorem is equivalent to the duality theorem in linear programming.

Extensions

The Minmax theorem was extended to apply for games far more general than mixed extensions of finite matrix games. Let us mention here two important results. The first result is that of Sion (1958) which proved the theorem for a game (X, Y, h) under rather weak properties imposed on X, Y, and h.

Theorem 1.8 (M. Sion) Let $G = (X, Y, h)$ be a game in which X and Y are convex topological spaces of which one is compact. h is an extended real-valued function defined on $X \times Y$ and satisfying the following condition: For every real c, the sets $\{y | h(x_0, y) \leq c\}$ and $\{x | h(x, y_0) \geq c\}$ are closed and convex for every $(x_0, y_0) \in X \times Y$. Then

$$\sup_{x \in X} \inf_{y \in Y} h(x, y) = \inf_{y \in Y} \sup_{x \in X} h(x, y) .$$

If X (respectively, Y) is compact then sup (respectively, inf) may be replaced by max (respectively, min).

The second result to be mentioned is in the direction of extending the range of the pay-off function h: assuming that h is not necessarily a real-valued function but rather has values in some ordered field F. That is, a commutative field with a subset P of positive elements which is closed under addition and multiplication and for any $x \in F$ either $x \in P$ or $x = 0$ or $-x \in P$. The order in F is then defined in the natural way: $a > b$ iff $a - b \in P$, etc.

Theorem 1.9 Let $A = (a_{ij})$ be an $m \times n$ matrix with elements a_{ij} in an ordered field F. Then there exists a unique element v of F and there exist $x_1,\dots,x_m$ and $y_1,\dots,y_n$ in F s.t. $x_i \geq 0 \ \forall \ i \in M$, $y_j \geq 0 \ \forall \ j \in N$;

$\sum_{i \in M} x_i = \sum_{j \in M} y_j = 1$ and

$$\sum_{j \in N} y_j a_{ij} \leq v \quad \forall \ i \in M , \qquad \sum_{i \in M} x_i a_{ij} \geq v \quad \forall \ j \in N .$$

The proof follows from the fact that a solution of an L.P. problem e.g. by the simplex method can be carried out in any ordered field. For real closed F the result follows from the standard minmax theorem using Tarski's principle.

Non-zero Sum Games

We end our first chapter by mentioning briefly a possible extension of our model of two-person zero-sum games to more players and to pay-offs not necessarily adding up to 0.

Definition 1.10 A *non-cooperative n-person game* in *strategic form* is an ordered

2n-tuple: $G = (X_1,\dots,X_n;\ h_1,\dots,h_n)$ where $X_1,\dots,X_n$ are sets and for each i, $1 \le i \le n$, h_i is a real-valued function defined on $X = X_1 \times,\dots,\times X_n$.

Interpretation. $N = \{1,\dots,n\}$ is the set of players, for each $i \in N$, X_i is the strategy set of player i and h_i is his pay-off function.

Remark. Our model of two-person zero-sum game (X, Y, h) is the special case in which $N = \{1, 2\}$; $X_1 = X$; $X_2 = Y$; $h_1 = h$; $h_2 = -h$.

Now two concepts were used in the two-person 0-sum case: the solution of *minmax* and that of *equilibrium*. Each of these concepts lends itself to a natural extension to the more general case. To do that let us introduce some notations. Given a game $(X_1,\dots,X_n,\ h_1,\dots,h_n)$ we let $X = \underset{i \in N}{\times} X_i$ and $\forall\ i \in N$, $X_{-i} = \underset{j \ne i}{\times} X_j$. Each $x_i \in X_i$ and $x_{-i} \in X_{-i}$ determines an element of X which is denoted by (x_i, x_{-i}) . For $x \in X$ and $\hat{x}_i \in X_i$ we denote by $(x|\hat{x}_i)$ the element of X obtained from x by replacing the i-th coordinate x_i by $\hat{x}_i$.

Definition 1.11 The *Minmax value* of player i in the game $G = (X_1,\dots,X_n,\ h_1,\dots,h_n)$ is denoted by v_i and defined by

$v_i = \sup_{x_i \in X_i} \inf_{x_{-i} \in X_{-i}} h_i(x_i, x_{-i})$. A strategy $\bar{x}_i$ which satisfies

$v_i = \inf_{x_{-i} \in X_{-i}} h_i(\bar{x}_i, x_{-i})$ is called a *minmax strategy* of player i . (Thus, if player i has a minmax strategy, the sup may be replaced by max.).

Definition 1.12 A strategy n-tuple $x^* \in X$ is called a Nash Equilibrium Point (N.E.P.) if for each $i \in N$:

$$h_i(x^*|x_i) \le h_i(x^*) \quad \forall\ x_i \in X_i \quad .$$

As the name suggests, the concept of Equilibrium was introduced by John Nash in 1950 who proved its existence for mixed extensions of finite (strategy sets) games. The proof is almost identical to the one we gave here for the two-person 0-sum case. Here again the result was generalized by considerably weakening of the conditions on the strategy sets and the pay-off functions (see, for instance Glicksberg, 1952).

It should be emphasized, however, that Lemma 1.5 is no longer true for the general case. Even if there are only two players, then generally the case is that a pair of *minmax strategies is not an E.P.* and vice versa: *a strategy in an E.P. is not a minmax strategy*. An easy example which demonstrates this diversion of the two concepts is the following two-person non-zero sum game:

$$\begin{pmatrix} 2,\ 2 & 4,\ 1 \\ 4,\ 1 & -2,\ 2 \end{pmatrix} \quad .$$

That is, each player has two strategies and the pay-off functions are given by the 2×2 matrix whose entries are ordered pairs (a_{ij}, b_{ij}) where a_{ij} is the pay-off for player I and b_{ij} for player II.

The following observations are easily verified.

(1) The minmax values are:

for player I, $v_1 = 2\frac{1}{2}$ with minmax strategy $(\frac{3}{4}, \frac{1}{4})$

for player II, $v_2 = \frac{3}{2}$ with minmax strategy $(\frac{1}{2}, \frac{1}{2})$.

(2) The unique N.E.P. is $(\frac{1}{2}, \frac{1}{2})$ for player I and $(\frac{3}{4}, \frac{1}{4})$ for player II corresponding to the pay-offs $(\frac{5}{2}, \frac{3}{2})$., So, although the equilibrium pay-offs are equal to the minmax payoffs, the equilibrium strategies are not minmax strategies, and vice versa. In other words, by playing $(\frac{1}{2}, \frac{1}{2})$ in equilibrium, player I *does not guarantee* $\frac{5}{2}$ which he can guarantee by playing $(\frac{3}{4}, \frac{1}{4})$. However, if both players will play minmax to guarantee the pay-offs $(\frac{5}{2}, \frac{3}{2})$, this will not be in equilibrium, each of them can improve his pay-off by a unilateral deviation.

Remark 1.13 It should be noted that there is no analogue of Theorem 1.9 for the N.E.P. in the non-zero sum case. In other words, a finite game with pay-offs in a certain ordered field may not have a N.E.P. in that field. To see that, consider a three-person game in which player I chooses one of two rows, player II chooses one of two columns and player III chooses one of two pay-off matrices:

$$A = \begin{pmatrix} 0, 3, 1 & 1, 0, 0 \\ 1, 0, 0 & 0, 1, 1 \end{pmatrix} \quad \text{or} \quad B = \begin{pmatrix} 2, 0, 5 & 0, 2, 0 \\ 0, -1, 0 & 1, 0, 0 \end{pmatrix} .$$

It can be shown that this game has a *unique* N.E.P. in which players I, II and II use the mixed strategies $(x, 1 - x)$; $(y, 1 - y)$ and $(z, 1 - z)$, respectively, where

$$x = \frac{9 + \sqrt{24}}{19} ; \quad y = \frac{7 - \sqrt{24}}{25} ; \quad z = \frac{12 - \sqrt{24}}{15} .$$

Hence, the game does not have a N.E.P. within the ordered field of rational numbers.

Glicksberg, I. (1952). A further generalization of the Kakutani fixed point theorem with application to Nash Equilibrium points. Proc. Amer. Math. Society, 38, 170-174.

Nash, J.F. (1950). Equilibrium points in n-person games. Proc. National Academy of Sciences, USA, 36, 48-49.

Von-Neumann, J. (1928). Zur theorie der gesellschaftesspiele. Mathematische Annalen, 100, 295-320.

Von-Neumann, J. and O. Morgenstern (1944, 1947). Theory of Games and Economic Behaviour. Princeton University Press: Princeton.

Weyl, H. (1950). Elementary proof of a minmax theorem due to Von-Neumann. Contributions to the theory of games I. Ann. Mathe. Studies, no. 24, 19-25, Princeton University Press: Princeton.

Chapter 2

GAMES IN EXTENSIVE FORM

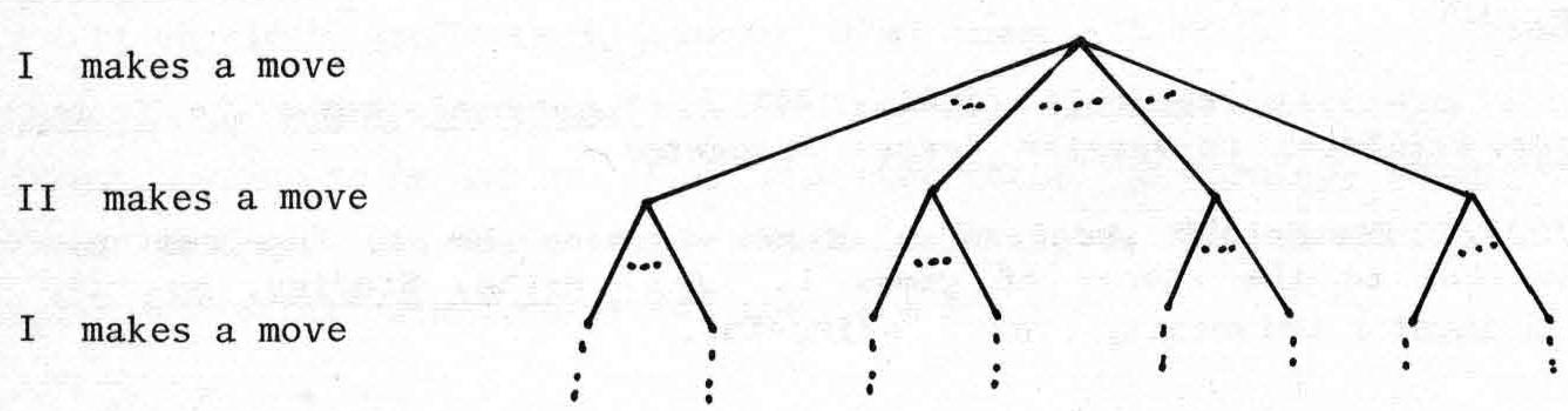
So far we know only one way to describe a game, namely the *strategic form.* Let us try to describe the game of chess in this way. That is, we look for an ordered 4-tuple $(S_I, S_{II}, h_I, h_{II})$, where by convention I and II are the white and black players, respectively. S_I, S_{II} are their respective strategy sets and h_I, h_{II} are the pay-off functions. This game has only three outcomes: W (white wins), B (black wins, and D (draw). It is natural to have pay-offs 1 for I and -1 for II when the outcome is W ; -1 for I and 1 for II when the outcome is B; and 0 for both players when the outcome is D . This makes chess a zero-sum game. But, *what are the strategy sets?* A strategy in chess (for I or II) is a *complete instruction book* for the player which instructs him in choosing his move in any possible situation in the game, where by 'situation' we mean here a complete *history* of the play which led to that decision point.

One readily observes that:

1. The rules of chess allow only a *finite* number of moves (though very large) for each player, thus:
2. Both S_I and S_{II} are *finite but astronomically large.* Therefore:
3. By the minmax theorem we can conclude that the game of chess has a value and each player has an optimal *mixed* strategy which guarantees this value.

This description of chess looks quite artificial and not very appealing. Our strategic form model for chess suppressed its dynamic structure and condensed all decision-making into one stage. The strategies are extremely complex objects and non-manageable in any practical sense: even more so are the mixed strategies.

Is there a more appealing way to describe the game of chess? Yes, there is the natural way of describing the evolution of the play using the notions of graph theory:



and so on until terminal points denoted by W , B or D are reached.

Such a description of a game is called an extensive form game. In its simplest version it is defined formally as follows:

<u>Definition 2.1</u> A finite two-person zero-sum game in *extensive form* is an

ordered collection $\Gamma = (X, X_I, X_{II}, X_T, x_0, f, h)$ where:

1) X is a finite set (the set of positions);

2) The sets X_I (decision positions of I), X_{II} (decision positions of II), and X_T (terminal positions) form a partition of X into disjoint sets.

3) x_0 (the initial position) is a point in $X_I \cup X_{II}$.

4) f (the immediate predecessor mapping) is a mapping from $X - \{x_0\}$ onto $X - X_T$ s.t. for any $x \in X$ there is an integer $n \geq 0$ satisfying $f^n(x) = x_0$.

5) h (pay-off function for player I) is a real-valued function defined on X_T.

An extensive form game is also called a *game tree*.

Remark. At a later stage, as we generalize our model, we shall refer to the games defined in Definition 2.1 as *extensive games with perfect information*. For the moment, since these are the only extensive games we have, we prefer to use a simple name.

A (pure) *strategy* of player I in Γ is a complete decision rule for him, i.e. a mapping s which maps each $x \in X_I$ to an alternative available for him at x, i.e. an element of the set $A(x) = \{y \in X | f(y) = x\}$. Denote by S_I the set of all pure strategies of I. S_{II} is derived similarly.

A *play* (or a *path*) in the game Γ is a finite sequence $p = (x_0, \ldots, x_n)$ of points in X s.t. $f(x_k) = x_{k-1}$ $\forall$ $k \geq 1$, and $x_n \in X_T$.

It is easily seen that a pair of strategies $s \in S_I$ and $t \in S_{II}$ determine uniquely a play $P(s, t) = (x_0, \ldots, x_n)$ and thus a pay-off $H(s, t) = h(x_n)$.

As long as we are interested merely in the strategies used by the players and the resulting pay-offs, any game in extensive form Γ is equivalent to the game in strategic form $\tilde{\Gamma} = (S_I, S_{II}, H)$ with the above-derived S_I, S_{II} and H. However it is important to notice the following.

1) Different extensive form games may have the same equivalent strategic form.

2) Not any finite strategic form game is obtainable from some extensive form game. For example, the matrix game $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is not equivalent to any extensive form game as defined in Definition 2.1.

The most important feature of this structure is:

<u>Definition 2.2</u> Given a game $\Gamma = (X, X_I, X_{II}, X_T, x_0, f, h)$ and any $x \in X - X_T$, the *subgame starting at* x is the game $\Gamma_x = (\tilde{X}, \tilde{X}_I, \tilde{X}_{II}, \tilde{X}_T, \tilde{f}, \tilde{h})$ where:

1) $\tilde{X} = \{y \in X \mid \text{there is } n \geq 0 \text{ s.t. } f^n(y) = x\}$.

2) $\tilde{X}_I = \tilde{X} \cap X_I$; $X_{II} = \tilde{X} \cap X_{II}$; $\tilde{X}_T = \tilde{X} \cap X_T$.

3) $\tilde{x}_0 = x$; $\tilde{f}$ is the restriction of f to $\tilde{X} - \{x\}$ and $\tilde{h}$ is the restriction of h to $\tilde{X}_T$.

This special structure of the game tree lends itself to a *dynamic programming* approach to determine the value of and the optimal strategies of the game by *forward or backward* induction, using the finiteness of the tree. The first result of this approach is:

Thereom 2.3 (Zermelo) Any finite zero-sum two-person game in extensive form has a value and each player has a (pure) optimal strategy.

This can be considered as the first important result in game theory, proved by Zermelo in 1912 for chess. The proof, which is a standard induction argument (on the maximal length of the game), is valid for any game given by Definition 2.1.

Note that in addition to the more appropriate description by a game tree we have here a result stronger than the one provided by the minmax theorem, namely the existence of *pure optimal* strategies. In other words, given an extensive form game, its reduction to a strategic form *itself* (rather than its mixed extension) has a minmax value.

n-person Non-zero Sum Games

Definition 2.1 has a straightforward extension to n-person non-zero sum games in extensive form. Any such game has a reduction to an equivalent strategic form game. The induction proof of Zermelo's theorem can be repeated to yield:

Theorem 2.4 Any finite n-person game in extensive form has a Nash equilibrium point (in pure strategies).

Two properties of the extensive games discussed so far were very crucial for the proof of Theorem 2.4, namely:

1. The game tree is finite.
2. The collection of positions succeeding a certain position x is a *subgame* (Definition 2.2).

In the rest of this lecture we discuss the generalizations of the model obtained by abandoning these properties.

Infinite Extensive Form Games

Infinite games in extensive form were discussed first by Gale and Stewart who considered zero-sum two person games similar to those of Definition 2.1 but with infinite length. To simplify the model, let us consider a very simple pay-off function which attains the values 1 (I wins) or -1 (II wins) only. One then obtains what is called a *win-lose* game defined as follows.

Definition 2.5 A win-lose game Γ is an ordered collection $(X, X_I, X_{II}, x_0, f, S, S_I, S_{II})$ where:

1) X is an infinite set (the set of positions);
2) X_I, X_{II} is a partition of X.

3) $x_0 \in X$.

4) f maps $X - \{x_0\}$ onto X s.t. for any $x \in X$ there is an integer $n \geq 0$ satisfying $f^n(x) = x_0$.

5) S is the set of infinite sequences $s = (s_0, s_1, \ldots)$ of elements of X satisfying $s_0 = x_0$ and $s_i = f(s_{i+1})$ for all $i \geq 0$. An element of S is called a *play*.

6) S_I (winning set for I) and S_{II} (winning set for II) form a partition of S .

Example 2.6 The two players alternate in choosing 0 or 1 . A play can then be identified with a point in $S = [0, 1]$ (i.e. the binary expansion of ..). S_I and S_{II} are two disjoint subsets s.t. $S_I \cup S_{II} = [0, 1]$.

The notion of a (pure) strategy is exactly as in the finite case, namely: A strategy of player I (respectively, II) is a function σ (respectively, τ) with domain X_I (respectively, X_{II}) satisfying $\sigma(x) \in f^{-1}(x)$ (respectively, $\tau(x) \in f^{-1}(x)$). We denote the players' strategy sets by $\Sigma_I^{(\Gamma)}$ and $\Sigma_{II}^{(\Gamma)}$. Any pair of strategies (σ, τ) , $\sigma \in \Sigma_I^{(\Gamma)}$, $\tau \in \Sigma_{II}^{(\Gamma)}$ determines in an obvious way a play $s \in S$ which we therefore write as (σ, τ) .

A strategy σ of player I is a *winning strategy* if $(\sigma, \tau) \in S_I$ for all $\tau \in \Sigma_{II}(\Gamma)$. A winning strategy for II is defined similarly. To say that the game has a minmax value is equivalent to:

Definition 2.7 A game $\Gamma = (X, X_I, X_{II}, x_0, f, S, S_I, S_{II})$ is *determined* if one of the players has a winning strategy.

An extension of Zermelo's Theorem 2.3 for infinite games would say that any such Γ is determined. However, this turned out to be false.

Theorem 2.8 (Gale and Stewart) There is an infinite game $\Gamma(X, X_I, X_{II}, x_0, f, S, S_I\ S_{II})$ which is not determined.

The proof is by constructing a counterexample of the type of Example 2.6. The construction is based on the observation that, roughly speaking, the strategy sets of the players are "very big" namely $2^{\aleph_0}$. Consequently, given any strategy of one of the players, the other can force $2^{\aleph_0}$ different plays (which is also the cardinality of the set S of all possible plays) . This enables construction of two disjoint sets of plays A and B such that: given any strategy of II , player I can force an outcome in A ; and given any strategy of I player II can force an outcome in B .

In view of this negative result, the natural question is: What interesting families of games can be proved to be determined? To put that more formally, let us introduce a topology on S . Actually, there is a natural one, namely the topology in which the basic open sets are those of the form $\{s \mid \rho_n(s) = \rho_n(s_0)\}$ for some $s_0 \in S$ and for some integer $n > 0$, where ρ_n denotes the projection operator on the first n coordinate space. It is a matter of straightforward verification to prove that this is a Housdorff topology for S in which S is totally disconnected.

A game $\Gamma = (X, X_I, X_{II}, x_0, f, S_I, S_{II})$ is said to be open, closed, G_δ etc. according to whether S_I is open, closed, G_δ etc.

Theorem 2.9 (Gale and Stewart) If S_I belongs to the Boolean algebra generated by the open sets then Γ is determined.

An important consequence of this result is that any game with continuous pay-off function h has a minmax value and the players have optimal strategies.

Theorem 2.10 (Wolfe) Any win-lose game is determined if one player's winning set is G_δ.

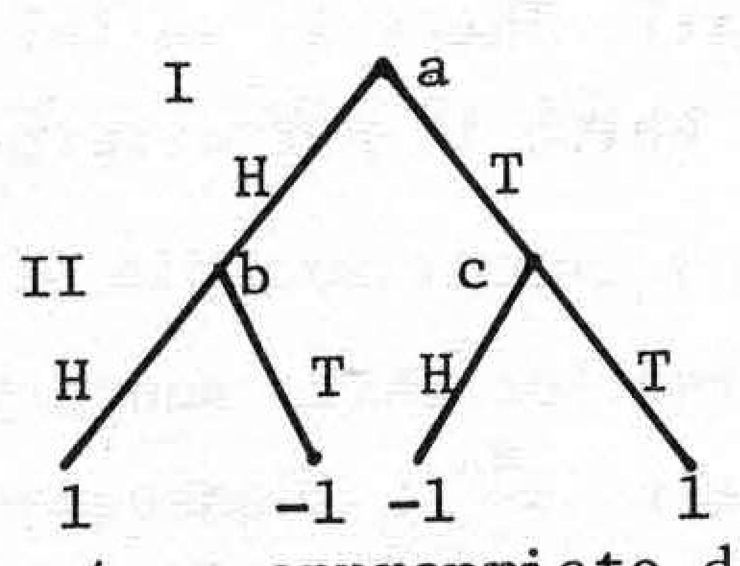
The problem of determinacy of games in which S_I is any Borel set was a long-standing difficult problem which was finally proved by D. Martin in 1975.

Theorem 2.11 (D. Martin) Any Borel game is determined.

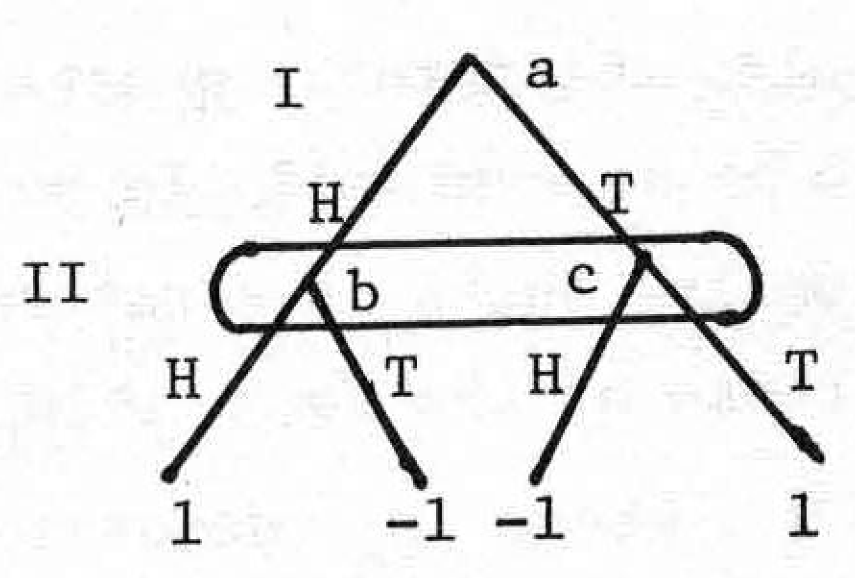
The consequence of this result is that any game with a measurable pay-off function h has a minmax value. However, in contrast to the continuous pay-off case, the players may not be able to guarantee the value but rather only ε-guarantee it for any $\varepsilon > 0$.

Games with Imperfect Information

Let us look now at the second property — the subgame property' used in the proof of Zermelo's theorem (and Martin's theorem). Consider the game of 'matching pennies'. Two players, I and II choose simultaneously H or T. If they both choose the same thing II pays I one dollar, otherwise I pays II one dollar. Can this game be described in extensive form? The obvious candidate for a game tree is:



It is readily seen that this is not an appropriate description of the game unless we add more structure to it: player II cannot distinguish between positions b and c. This means in particular that he cannot choose T in b and H in c (as he would certainly like to do). We indicate this by saying that b, c is an *information set* of player II and describe it by:



In other words, the right notion of decision point of a player is not a node in the game tree but rather a set of nodes which are indistinguishable for him. One immediately realizes that this game does not have a value. In fact, it is equivalent to the matrix game $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ which has no value (in pure strategies). This shows already that Theorem 2.3 cannot be extended to extensive form finite games with additional structure of information sets. The failure of the inductive proof is quite transparent: the part of the tree succeeding node b (or c) *is not a subgame.*

Unfortunately, the formal definition of this intuitively simple notion is quite complicated. This is so because one has to express the fact that a player cannot distinguish between two nodes in the same information set. This means, for instance, that he must be (from his point of view) in the 'same stage of the game'. Also, we allow chance moves in the game tree. This makes the pay-offs random variables whose expectations are all by convention the utilities of the corresponding players.

Definition 2.12 *Extensive form game* Γ *of* n-*players* consists of the following elements:

(1) A set $N = \{1, 2, \ldots, n\}$ of players;

(2) A finite connected graph G with no cycles called the *game tree.*

(3) A distinguished node of the tree x_0 called *the first move.* A node of degree one, different from x_0, is called a *terminal node.* The set of terminal nodes is denoted by T.

(4) The set X of non-terminal nodes is called the set of *moves* and is partitioned into $n + 1$ sets $X^0, X^1, \ldots, X^n$. Elements of X^i are called *moves of player* i, while elements of X^0 are called *chance moves.*

(5) For each node in X^0 there is a probability distribution on the branches out of it with positive probability to each one of them.

(6) For each $i \in N$, there is a partition of X^i into $U^i_1, \ldots, U^i_{k_i}$, called the information sets of player i, such that for each $j \in \{1, \ldots, k_i\}$,

 (i) There is a 1 - 1 correspondence between the sets of outgoing branches of any two nodes in U^i_j.

 (ii) Any path from x_0 to a terminal node (i.e. a *play*) can cross U^i_j at most once.

(7) For each terminal node $t \in T$ attached an n-dimensional real vector $h(t) = (h^1(t), \ldots, h^n(t))$ called the pay-off vector at t.

If all information sets are singletons the game is called *a game with perfect information.* Thus the game in Definition 2.1 is a finite game with perfect information and no-chance moves, while the games of Gale and Stewart are infinite games of this kind.

A *pure strategy* of player i is a k_i-tuple $\sigma^i = (\sigma^i(U^i_j))_{j=1}^{k_i}$ where $\sigma^i(U^i_j)$ is an element of the set of alternatives available to player i in his information set U^i_j.

Denote by S^i the set of pure strategies of player i and let $S = S^1 \times \ldots \times S^n$.

Given an n-tuple of strategies $s = (s^1,\dots,s^n) \in S$ the *expected pay-off* to player i is defined as

$$H^i(s) = \sum_{t \in T} P_s(t)\, h^i(t)$$

where $P_s(t)$ is the probability that $t \in T$ will be reached when s is played.

Any finite n-person game in normal form can be reduced to a strategic form game $(S^1,\dots,S^n, H^1,\dots,H^n)$. If the extensive form we started with was a game of perfect information, by Zermelo's proof it will have an N.E.P. in pure strategies. This result is no longer true for imperfect information games as the game of matching pennies already shows. For these games we have, by Nash's result, the existence of N.E.P. for the mixed extension.

Behaviour Strategies

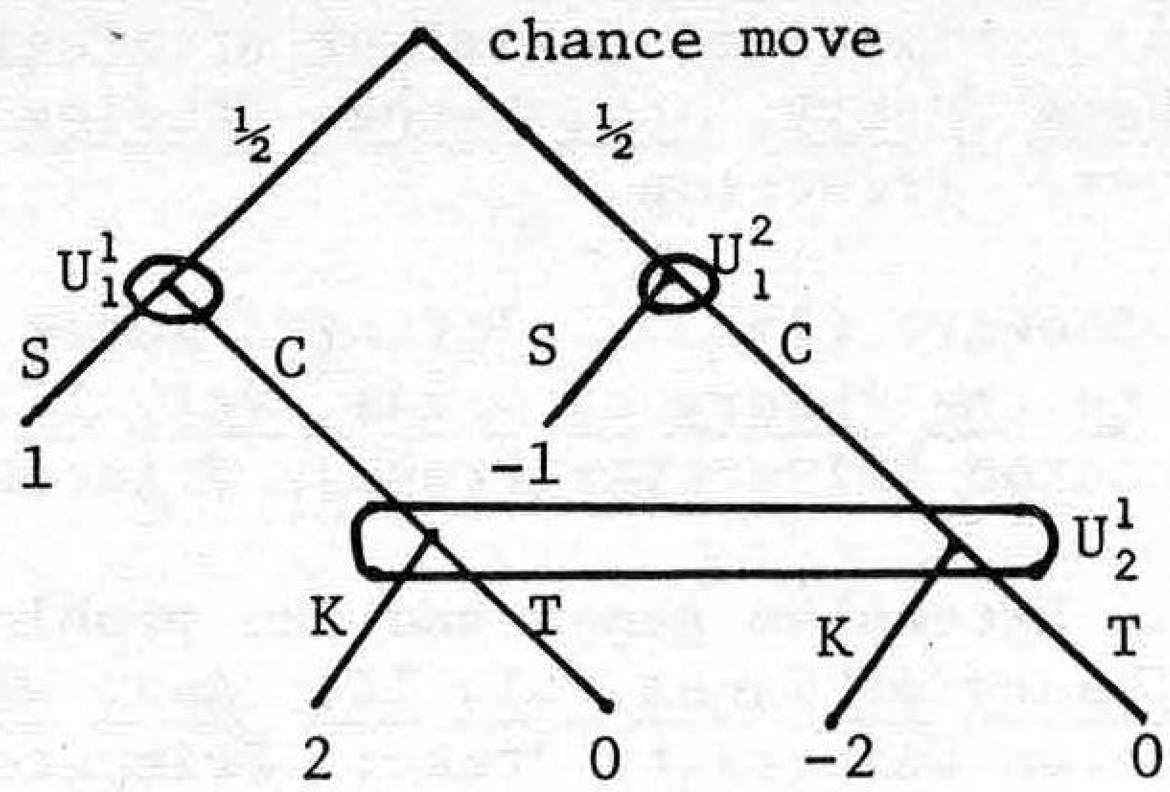
In a game in extensive form, a mixed strategy means a single randomization at the beginning of the game after which a certain pure strategy is followed, i.e., a *deterministic* choice of an alternative at each information set. Another way for a player to randomize his choice is to randomize on his possible alternatives at each information set, and to do these randomizations independently in his various information sets.

Definition 2.13 A *behaviour strategy*, b^i, of player i in an extensive form game Γ is a k_i-tuple $b^i = (b^i(U^i_j))_{j=1}^{k_i}$ where $b^i(U^i_j)$ is a probability distribution over the set of alternatives at the information set U^i_j. Denote by B^i the set of behaviour strategies of player i, and denote by Σ^i his set of mixed strategies (i.e., probability distributions on S^i).

Beside its intuitive appeal for extensive form, the behaviour strategies set is usually much smaller than the mixed strategies set. For instance, consider a game in which a certain player i has three information sets with two alternatives in each. Then $|S^i| = 8$ and therefore Σ^i is seven-dimensional simplex. On the other hand, a behaviour strategy is determined by three probabilities in $[0, 1]$ and thus B^i is a three-dimensional cube.

In what circumstances can we work with B^i instead of Σ^i ?

First observe that any $b^i \in B^i$ generates in a natural way a probability distribution on S^i, i.e. a mixed strategy $x^i \in \Sigma^i$. This x^i *leads to the same pay-offs as* b^i regardless of the strategies chosen by the other players. In this sense we may say that x^i is *strategically equivalent* to b^i. Denote this mapping by $\varphi: B^i \to \Sigma^i$. If φ is 'onto' (i.e., $\varphi(B^i) = \Sigma^i$) then any $\sigma^i \in \Sigma^i$ could be replaced by a behaviour strategy which is strategically equivalent to it, namely any $b^i \in \varphi^{-1}(\sigma^i)$. However, φ may not be 'onto' as can be seen in the following example.



With the obvious notation player I has four pure strategies: {(S, K); (S, T); (C, K); (C, T)} . It is easily seen that if we consider the mixed strategy $\sigma^i = \frac{1}{2}(S, T) + \frac{1}{2}(C, K)$, (which happened to be the optimal strategy of player I), then there is no $b^i \in B^i$ s.t. $\varphi(b^i) = \sigma^i$. The reason for that is also quite transparent: in σ^i , the choices in the two information sets are highly correlated. This correlation cannot be produced by appropriate choices of the probability distributions comprising the behaviour strategies since player I, when in U_2^1 , *does not remember* his move in U_1^1 .

<u>Definition 2.14</u> An extensive form game Γ is said to be a *game with perfect recall* if each player at each move remembers what he knew in previous moves and what choices he made at those moves.

<u>Remark:</u> There is no difficulty in writing this formally at the cost of introducing some more notations which we prefer to avoid here.

<u>Theorem 2.15</u> (Kuhn, 1953) Let Γ be an extensive form game in which player i has perfect recall. Then, for each mixed strategy $\sigma^i \in \Sigma^i$, there is a behaviour strategy $b^i \in B^i$ which is strategically equivalent to σ^i , i.e. for each $j \in N$ and $\sigma \in \Sigma$, $H^j(\sigma) = H^j(\sigma | b^i)$, where $(\sigma | b^i)$ is the n-tuple σ in which σ^i is replaced by b^i .

<u>Corollary 2.16</u> Any (finite) game Γ in extensive form has a N.E.P. in behaviour strategies.

Aumann (1964) generalized Kuhn's theorem to infinite games with perfect recall, i.e. both the length of the game and the number of alternatives at each move may be infinite.

Aumann, R.J. (1964). Mixed and behaviour strategies in infinite extensive games. Adv. Game Theory, Ann. Mathe. Studies 52, 627-650. Princeton University Press: Princeton.

Gale, D. and F.M. Stewart (1953). Infinite games with perfect information. Contributions to the Theory of Games, Vol. II. Ann. Mathe. Studies 28, 245-266. Princeton University Press: Princeton.

Kuhn, H.W. (1953). Extensive games and the problem of information. Contributions to the Theory of Games Vol. II. Ann. Mathe. Studies. 28, 193-216. Princeton University Press: Princeton.

Martin, D.A. (1975). Borel determinacy. Ann. Mathe. 102, 363-371.

Wolfe, P. (1955). The strict determinateness of certain infinite games. Pacific J. Math. 5, 891-897.

Zermelo, E. (1912). Über eine anwendung der mengenlelre auf die theorie des Schachspiels. Proc. Fifth Int. Cong. Math., Cambridge, Vol. II, 501-504.

Chapter 3

MULTISTAGE GAMES

The Notion of Super-Game

Multiperson decision situations for which we attempt to provide game theoretical models, are very seldom one-time affairs, but rather repeated over and over again. One may therefore gain additional insight about various phenomena by studying not merely the static one-shot games but also some *multi-stage* or a *repeated* game. These models seem to be the correct paradigm for studying phenomena such as communication, retaliation, flow of information, etc.

Consider the following two-person non-zero sum game known as the "Prisoner's Dilemma":

$$\begin{array}{c} \\ G \\ C \end{array} \begin{array}{c} \begin{array}{cc} G & \quad C \end{array} \\ \begin{pmatrix} 1, 1 & 5, 0 \\ 0, 5 & 4, 4 \end{pmatrix} \end{array} .$$

The only N.E.P. in this game is (G, G) yielding a pay-off of (1, 1) which is dramatically inferior to (4, 4) from the point of view of both players. This is especially disturbing if the game is *played many times by the same players*, since one would expect some 'silent understanding' between the players and the emergence of the cooperative outcome (4, 4) at least in some of the repetitions. Can we provide a model that predicts this phenomenon?

The first attempt is to consider, say, a 1,000-times repeated prisoner's dilemma played by the same players. One easily sees that the only N.E.P. in this game is again such that each player plays G in all stages independently of what the other player does. So this is not the appropriate model we are trying to find. A moment of reflection reveals the reason. The presence of a last stage which is recognized as such by both players, aside from being unrealistic, creates unnatural end effects which propagate themselves backwards and distort the entire analysis. This suggests that a game "without an end" may be more appropriate. Without bothering much about details, let us show the following.

<u>Proposition 3.1</u> In the infinitely repeated Prisoner's dilemma there is an N.E.P. with the cooperative pay-offs (4, 4) as an "average" pay-off for the players.

<u>*Proof*</u>. Consider the following strategy, σ , for a player. Play C in the first stage and keep on playing C as long as the other player continues playing C . As soon as he plays G , play G following that stage on.

Clearly, if both players play σ the pay-off sequence for both of them will be:

$$h^1(\sigma, \sigma) = h^2(\sigma, \sigma) = (4, 4, 4, \ldots) \quad .$$

If a player, say player I, uses $\hat{\sigma} \neq \sigma$, while the other player is using σ he will play G for the first time, say at stage k . His pay-off sequence will be at most (coordinate-wise): $(4,\ldots,4,5,1,1,\ldots)$ with 5 as the k-th coordinate.

By any reasonable definition of 'average pay-off' such as Cesaro limit, Abel limit, or any Banach limit, the value of $(4, 4,\ldots)$ is 4 and that of $(4,\ldots,4,5,1,1,\ldots)$ is 1 . Thus (σ, σ) is in fact a N.E.P. with 'pay-offs' (4, 4).

Q.E.D.

So, what the players could get in the one-shot game, by signing a *binding agreement* to play (C, C) can be *self-enforced* as an N.E.P. in the *super-game* (i.e., the infinitely repeated game). Many other pay-offs can be reached in the one-shot game via binding agreements. For instance, the expected pay-offs (2,2) by signing an agreement to draw a lottery (controlled by the 'authorities') to choose C or G with respective probabilit es $\frac{1}{3}$ and $\frac{2}{3}$. Whatever the outcome is both players are committed to play it. Their expected pay-off is $\frac{1}{3}(4, 4) + \frac{2}{3}(1, 1) = (2, 2)$. Can this also be sustained by an N.E.P. in the super-game? The answer is 'yes' and the N.E. strategies are the following (again, the same for both).

Play repeatedly C, G, G, C, G, G,... so long as the other player is following the same pattern. As soon as the other player deviates from this preseribed pattern, play G from there on. It is clear that when both players follow this strategy each will have the pay-off sequence $(4,1,1,4,1,1,\ldots)$ that is, worth 2 by any reasonable definition. Any unilateral deviation of one of the players will yield him a payoff sequence with at most 1 from one stage on.

The general ideas should be clear by now, and we move quickly towards the general theorem.

Consider an n-person game in strategic form $\Gamma = (S^1,\ldots,S^n, h^1,\ldots,h^n)$.

<u>Definition 3.2</u> A *correlated strategy* of a coalition $T \subset N$ is a probability distribution on $\prod_{i \in T} S^i$.

When correlated strategies are used the set of expected vector-payoffs is the convex hull of the vector payoffs attainable by pure strategies. We denote this set by C and refer to it as the set of *correlated pay-offs*.

<u>Definition 3.3</u> The individual rationality level of player i is r_i defined by:

$$r_i = \min_{\tau} \max_{\sigma} H^i(\sigma, \tau)$$

where σ ranges over the (mixed) strategies of i and τ ranges over all correlated strategies of $N \setminus \{i\}$.

A pay-off vector $(\alpha_1,\ldots,\alpha_n)$ is said to be *individually rational* if $\alpha_i \geq r_i$

for all $i \in N$.

Remark 3.4 Note that $\min_\tau \max_\sigma H^i(\sigma, \tau) = \max_\sigma \min_\tau H^i(\sigma, \tau)$ but this is so because $N \setminus \{i\}$ are allowed to use correlated strategies. It is not true if τ ranges over mixed strategies of $N \setminus \{i\}$. As an example, consider the three-person game in which I chooses a row, II chooses a column and III chooses the matrix. The pay-offs for III are

$$\begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} -3 & 0 \\ 0 & -1 \end{pmatrix}.$$

If we denote by $(x, 1 - x)$, $(y, 1 - y)$ and $(z, 1 - z)$ the mixed strategies of the three players, respectively, then for player III:

$$\min_{x,y} \max_z H(x,y,z) = \min_{x,y} \max (-xy - 3(1 - x)(1 - y), -3xy - (1 - x)(1 - y)) = -1$$

while

$$\max_z \min_{x,y} H(x,y,z) \leq \min(-z - 3(1 - z), -3z - (1 - z)) \leq -2 .$$

Theorem 3.5 (the "folk theorem') The pay-off vectors to Nash equilibrium points in the super-game Γ^* are the individually rational correlated pay-offs $C_R = \{x \in C | x^i \geq r^i \ \forall \ i \in N\}$.

This theorem has been known for about 20 years but has not been published and its authorship is obscure although it is probably to be credited mainly to Aumann and Shapley. The idea of the proof is the one that can be read in our example: prescribe the right pattern of correlated moves to approach the desired point in C_R . As soon as player i deviates $N \setminus \{i\}$ switch to the *punishment* strategy, i.e., the correlated strategy that keeps his pay-off to r_i .

We shall not go through the formal definitions of the super-game Γ here. Later we shall discuss the point of the definitions of pay-offs, which is an issue of general importance to all infinite stage games.

Stochastic Games

Stochastic games are multistage games in which the game played at each stage changes randomly. The following short review of the subject will be confined to two-person zero-sum stochastic games with finitely many states and finitely many alternatives in each for both players.

The first model and result is due to Shapley (1953). There is a finite set of *states* $S = \{1,\ldots,S\}$ and additional state $s = 0$ which is the 'game is over' (by abuse of notation, S denotes both the set and its cardinality as does I, J, etc.). At state $s \in S$ each player has a finite number of possible actions: $i=1,\ldots,I_s$ for player 1 and $j=1,\ldots,J_s$ for player 2. We may assume w.l.g. that $I_s = I$ and $J_s = J$ for all $s \in S$, and thus associate to each $s \in S$ an $I \times J$ pay-off matrix A^s (from 2 to 1). For $i \in I$, $j \in J$ and $s \in S$ there is a ***transition*** *probabilities* vector $P^s_{ij} = (P^{st}_{ij})_{t=0,1,\ldots,s}$.

The stochastic game is played in stages: at each stage the game is in some state $s \in S \cup \{0\}$. If $s \neq 0$, player 1 chooses $i \in I$, player 2 chooses $j \in J$. Then (i, j) is announced, player 2 pays player 1 a^s_{ij}, the referee chooses the new state according to the probability vector P^s_{ij} and informs the players about the new state asking them to play the next stage (unless the new state is 0).

Stochastic games generalize Markov decision processes in that Markov decision processes may be viewed as stochastic games in which one of the players has only one action in each state.

The most crucial element in Shapley's first model was:

Assumption 3.6 $\quad \lambda = \min_{i,j,s} P^{so}_{ij} > 0$.

Due to this assumption, expected total pay-offs are bounded and the existence of value and optimal strategies could be derived from general minmax theorems. However, we shall use an alternative approach used by Shapley and proved to be very fruitful in more general models. This is basically the dynamic programming approach.

Special Case: If $P^{so}_{ij} = \lambda$ for all i,j, and $s \in S$ we have a λ-*discounted* game: we may forget about the state o and normalize the probability vectors on S (i.e. divide by $(1 - \lambda)$). The game then has denumberably many stages and a pay-off stream $x = (x_1, x_2, \ldots)$ is *evaluated* by $\sum_{k=1}^{\infty} \lambda^{k-1} x_k$.

For the sake of simplicity of notations, we shall derive Shapley's results for this special case.

A ***history*** prior to stage n is $[(i_k, j_k, s_k)]_{k=1}^{n-1}$, where i_k, j_k are the actions chosen at stage k and s_k was the state at stage k. Denote by H_n the set of all possible such histories.

Definition 3.7

(i) A *behaviour strategy* of player 1 is a sequence $\sigma = (\sigma_n)_{n=1}^{\infty}$ where $\sigma_n : H_n \times S \to I^*$, and I^* is the simplex of probability distributions on I.

(ii) If the mappings σ_n are to I (i.e. the extreme points of I^*), then σ is a *pure strategy*.

(iii) If the σ_n's are independent of H_n, then σ is called a *stationary strategy*.

Strategies for player 2 are defined similarly. We denote by Σ_1, Σ_2 the sets of (behaviour) strategies of the two players. Given an *initial state* $s \in S$, any pair of strategies $(\sigma, \tau) \in \Sigma_1 \times \Sigma_2$ determines a probability distribution on pay-off streams. The expected evaluation of these vector streams defines a pay-off function $H_s: \Sigma_1 \times \Sigma_2 \to R$. Denote by $\Gamma_s(\lambda)$ the two-person zero-sum game $(\Sigma_1; \Sigma_2; H_s)$ and let $\Gamma = (\Gamma_1(\lambda),\ldots,\Gamma_S(\lambda))$.

Given the pay-off matrices $A = (A^1,\ldots,A^S)$ and any $x \in R^S$ define $G(x) = (G_1(x),\ldots,G_S(x))$ where $G_s(x)$ is the $I \times J$ matrix defined by:

$$(G_s(x))_{ij} = a^s_{ij} + \sum_{t=1}^{S} p^{st}_{ij} x_t .$$

Denoting by 'val' the value operator, and $\text{val } G = (\text{val } G_1,\ldots,\text{val } G_S)$ we have:

Theorem 3.8 The stochastic games $\Gamma(\lambda) = (\Gamma_1(\lambda),\ldots,\Gamma_S(\lambda))$ have a value $V(\lambda) = (V_1(\lambda),\ldots,V_s(\lambda))$ which is the unique solution of the equation

$$x = \text{val } G((1 - \lambda)x) . \tag{3.1}$$

Proof. Observe first that with respect to the norm $|x| = \max_s |x_s|$ we have for any $x,y \in R^S$:

$$|\text{val } G(x) - \text{val } G(y)| \leq |x - y| .$$

From this it follows that the function of x , $\text{val } G((1 - \lambda)x)$ is a *contraction* and thus has a unique fixed point which is a solution for (3.1).

Next, if we denote by $\Gamma^n(\lambda) = (\Gamma^n_1(\lambda),\ldots,\Gamma^n_S(\lambda))$ the stochastic game with n-stages and its values by $V^n = (V^n_1,\ldots,V^n_S)$ we readily see that:

$$V^n = \text{val } G((1 - \lambda)V^{n-1}) ; \quad n = 1,2,\ldots ,$$

with $V^o = (0,\ldots,0)$. Therefore $\lim_{n\to\infty} V^n = x$, the only solution of (3.1).

Finally, for any $\varepsilon > 0$ both players can guarantee $\lim_{n\to\infty} V^n$ up to an ε by playing optimally in Γ^N for some N large enough (remember that $\lambda > 0$ and hence the contribution to the pay-off of stages $n > N$ is less than ε if N is large enough).

Q.E.D.

As for the optimal strategies, given any S-tuple of mixed strategies $x = (x^1,\ldots,x^S)$ of player 1 (or 2), in the one-stage game (i.e., elements of I^* or J^*). We identify x with the *stationary behaviour strategy* which consists of playing the mixed strategy x^s whenever the state is s .

Theorem 3.9 If for each $s \in S$, x^s is an optimal strategy in the matrix $G_s((1 - \lambda)V(\lambda))$, then $x = (x^1,\ldots,x^S)$ is a stationary optimal strategy in the λ-discount game $\Gamma(\lambda)$.

Proof. For each n let $\hat{\Gamma}^n(\lambda)$ be the same game as $\Gamma^n(\lambda)$, except that when stage n is reached and the state is s , the pay-off is according to $G_s((1 - \lambda)V(\lambda))$

instead of A^s . One checks then easily that: (i) By using the stationary strategy x each player guarantees $V(\lambda)$ in $\hat{\Gamma}^n(\lambda)$. (ii) The difference in pay-offs between $\hat{\Gamma}^n(\lambda)$ and $\Gamma^n(\lambda)$ is arbitrarily small if n is large enough. Hence, for any $\varepsilon > 0$ choosing n large enough, x guarantees $V^n(\lambda) \pm \varepsilon$ in $\Gamma^n(\lambda)$ and hence $V(\lambda) \pm 2\varepsilon$ in $\Gamma(\lambda)$. Since this is true for any $\varepsilon > 0$, the result follows.

Q.E.D.

The great importance of Shapley's work is not only in formulating the first model and opening a new field of research, but also in using the *dynamic programming* approach and the *contraction mapping* which proved to be very useful tools in most of the research that followed.

However, as soon as assumption 3.6 was to be relaxed, that is, away from the λ-discount game, a lot of mathematical ingenuity and depth was needed. We are able to mention here only part of the important results.

Gillette (1957), Hoffman and Karp (1966) and Stern (1975) looked for conditions under which the undiscounted infinite stage game (to be defined later) has a min max value. Such a condition was, for instance, that for any pair of strategies used by the players, the resulting Markov chain is ergodic. An example in which this condition is not satisfied was studied by Blackwell and Ferguson (1968) under the name of "the Big Match." Their result was generalized by Kohlberg (1968) to 'games with absorbing states.' The most important breakthrough was done by Bewley and Kohlberg (1976), and finally Mertens and Neyman (1981) answered the long-standing difficult problem by proving that any stochastic game has a value.

Bewley and Kohlberg (B.K. hereafter) studied the asymptotics of stochastic games in two directions:

(i) Considering the λ-discount game $\Gamma(\lambda)$ and letting λ tend to 0 .

(ii) Considering the undiscounted $(\lambda = 0)$ n-stage game Γ^n and letting n to to ∞ .

From Shapley's result we know that for any $\lambda > 0$ the λ-discount game $\Gamma(\lambda)$ has a value $V(\lambda)$. If we think of λ as the probability of stopping the game after each stage then the expected number of stages is $1/\lambda$ and then $\lambda V(\lambda)$ can be interpreted as 'a value per stage'. B.K. proved:

Theorem 3.10 $\lim_{\lambda\to 0} \lambda V(\lambda)$ exists.

Considering now the limit value of the undiscounted truncated game Γ^n , note first that there is no problem of existence for the value V^n of Γ^n . In order to compare games of different lengths one looks at the 'value per stage' V^n/n . B.K.'s next results were:

Theorem 3.11 $\lim_{n\to\infty} V^n/n$ exists.

Theorem 3.12 $\lim_{\lambda\to 0} \lambda V(\lambda) = \lim_{n\to\infty} V^n/n$.

Actually, B.K. managed to find the expansion of the value $V(\lambda)$ and the optimal strategies in fractional powers of λ for an interval $0 < \lambda \leq \lambda_0$. Similarly, they found an approximate expansion of V^n in powers of n. More precisely, they proved:

Theorem 3.13 There exists an integer M such that:

(i) There exists $\lambda_0 > 0$ such that the following expansion holds for $0 < \lambda \leq \lambda_0$:

$$V(\lambda) = a_M \lambda^{-1} + a_{M-1}\lambda^{-(M-1)/M} + a_{M-2}\lambda^{-(M-2)/M} + \dots .$$

(ii) There exists a stationary strategy for player 1 described by vectors:

$$x_s(\lambda) = x_{0s} + x_{1s}\lambda^{1/M} + x_{2s}\lambda^{2/M} + \cdots ,$$

where $x_s(\lambda)$ is a probability vector in I^*, and $x_{ks} \in R^I$ for all k, and there exists $\lambda_{00} > 0$ such that for each $0 < \lambda \leq \lambda_{00}$, the stationary strategy $x(\lambda) = (x_1(\lambda),\dots,x_S(\lambda))$ is optimal in $\Gamma(\lambda)$. The above works similarly for player 2.

(iii) There is an expansion of the form

$$w_s^n = a_s^M n + a_s^{M-1} n^{(M-1)/M} + ,\dots, + a_s^1 n^{1/M} ,$$

such that $|V_s^n - w_s^n| < C \log (n + 1)$ for some constant C.

To prove these results, B.K. adopted an algebraic approach rather than analytic. Their impressive proofs are based on the following main steps:

Step 1: Consider the ordered field F of real Puiseux series, i.e. series of the form $\sum_{k=-\infty}^{K} a_k \theta^{k/M}$, where M is a positive integer, K is any integer and a_k are real numbers. Addition and multiplications are defined in the natural way and order is defined by: $\sum_{k=-\infty}^{K} a_k \theta^{k/M} > 0$ if and only if $a_N > 0$ where N is the largest integer k s.t. $a_k \neq 0$.

Step 2: If the fundamental limit discount equation (3.1) (with θ^{-1} replacing λ),

$$x = \text{val } G((1 - \theta^{-1})x) , \tag{3.2}$$

has a solution in F, then for small enough λ, substitution of λ^{-1} for θ gives $V(\lambda)$. In other words, this is then the desired expansion. Thus the problem is reduced to prove that (3.2) has a solution in F.

Step 3: As it was noted in our first lecture, the minmax theorem is true in any ordered field (Weyl, 1950), thus val G is defined for any matrix G with entries in an arbitrary ordered field, F in our case. Furthermore, an equation of the type $y = \text{val } G$ may be expressed as an *elementary formula* over F, i.e. an expression constructed in a finite number of steps from *atomic formulae* ($p > 0$ or $p = 0$, where p is a polynomial with integer coefficients, in one or more variables) by means of conjunction ($\wedge$), disjunction ($\vee$), negation ($\sim$) and quantifiers of the form $\exists x$, $\forall x$.

The statement, "there exists a solution x in F to $x = \text{val } G((1 - \theta^{-1})x)$", can be expressed as an *elementary sentence in* F i.e. an elementary formula in which *all* variables are quantified by $\exists$ or $\forall$.

Step 4 (Tarski's Principle): An elementary sentence which is valid over one real closed field is valid over every real closed field. (An ordered field F is *real closed* if it has no ordered algebraic extention.)

Step 5: By Shapley's result, the elementary sentence stating, 'there is a solution in H to (3.2)', is valid over the real closed field of the real numbers.

Step 6: F is a real closed field, therefore by Steps 4 and 5 (3.2) has a solution in F.

For the n-stage values V^n the same real closed field of real Puiseux series is used with θ representing the function n.

Remark Parts (i) and (ii) of theorem 3.13 apply for non-zero sum n-person games as well. The result is then the existence of pay-off vectors $v(\lambda)$ and strategy vectors $x(\lambda)$ each of which has a convergent expansion in fractional power of λ such that in some neighborhood of $\lambda = 0$, $x(\lambda)$ is a N.E. point in the λ-discounted game with corresponding pay-offs $v(\lambda)$ (see Mertens, 1982).

The Value of the Infinite Game

In the asymptotic approach of Bewley-Kohlberg one considers the *limit of value*, either the limit of $\lambda V(\lambda)$ as $\lambda \to 0$ or the limit of V^n as $n \to \infty$. Another natural approach to study the *very long undiscounted game* is to look at *the value of the limit*, i.e. the value of the undiscounted infinite stage game Γ_∞. The strategies in Γ_∞ are defined as in definition 3.7. However, there is a technical difficulty in defining an appropriate pay-off function. This difficulty which is *common to all undiscounted infinite state games* (not necessarily stochastic) can be overcome by one of two ways:

(i) By defining the evaluation of a pay-off stream $x = (x_1, x_2 \ldots)$ as lim inf, lim sup or more generally any Banach limit of the n-stage averages $\rho_n = \frac{1}{n} \sum_{k=1}^{n} x_k$.

(ii) By avoiding the definition of pay-off function and defining directly the notion of value.

We shall adopt usually the second alterantive. More precisely, we define:

Definition 3.14 An undiscounted infinite stage game Γ_∞ is said to have a value v if $\forall\ \varepsilon > 0$ there is a strategy $\hat{\sigma}$ of player 1 and $\hat{\tau}$ of player 2 and an integer $N > 0$ s.t.

$$\rho_n(\hat{\sigma}, \tau) \geq v - \varepsilon \ ; \qquad \forall\ n > N \quad \forall\ \tau$$

$$\rho_n(\hat{\tau}, \sigma) \leq v + \varepsilon \ ; \qquad \forall\ n > N \quad \forall\ \sigma$$

where $\rho_n(\sigma, \tau)$ is the expected n-stage average pay-off when σ and τ are used. This implies in particular that player 1 can guarantee that lim inf ρ_n will be as close as he wishes to v and player 2 can guarantee that lim sup ρ_n will be as close as he wishes to v. We shall use the following terminology: $\hat{\sigma}$ (as well as $\hat{\tau}$) ε-*guarantee* v in Γ_∞.

Remark 3.15 Note that if Γ_∞ has a value v and if we denote by v_n the (average per stage) value of Γ_n then $\lim_n v_n$ exists and is equal to v.

The problem of existence of a value for a general undiscounted stochastic game Γ_∞ was an open problem for many years, in spite of many attempts to solve it. It was finally solved in 1981 by Mertens and Meyman who used the B.K. asymptotic theory to prove:

Theorem 3.16 The infinite game has a value which equals the asymptotic values:

$$v = \text{val}(\Gamma_\infty) = \lim_{\lambda\to 0} \lambda V(\lambda) = \lim_{n\to\infty} V^n/n \quad .$$

A rough description of the strategy of player 1 which guarantees lim inf $\rho_n \geq v - \varepsilon$ looks as follows. At stage k player 1 computes a number $\lambda_k \in (0, 1]$ and plays optimally in the λ_k-discounted game (according to the state he is in). $\lambda_k = \lambda(\xi_k)$ where λ: $[1, \infty) \to (0, 1]$ is an appropriately designed continuous decreasing function and ξ_k is a statistic updated as follows:

$$\xi_{k+1} = \text{Max}[C, \xi_k + x_k - \lambda_k V(\lambda_k) + 4\varepsilon] \ ,$$

where $C \geq 1$ is a sufficiently large constant. So roughly speaking, ξ_k is the excess of the actual pay-offs $x_1 + x_2 + \cdots + x_k$ over the intended pay-offs $\lambda_1 V(\lambda_1) + \lambda_2 V(\lambda_2) + \cdots + \lambda_k V(\lambda_k)$. The higher ξ_k becomes the lower λ_k is, which means that he plays for lower discount rates, i.e. with more importance attached to later stages compared to the present one.

Remark 3.17 Mertens and Neyman's result holds for a class of stochastic games much wider than that treated by B.K. One does not have to make any finiteness assumptions, neither on the state space nor on the action sets, provided the following conditions hold:

(i) Pay-offs are uniformly bounded.

(ii) The value $V(\lambda)$ of the λ-discounted games exists.

(iii) $\forall\ \varepsilon < 1$ there exists a sequence λ_i decreasing to 0 such that $\lambda_{i+1} \geq \varepsilon\lambda_i\ \ \forall\ i$ and $\Sigma||V(\lambda_{i+1}) - V(\lambda_i)|| < \infty$.

It is a consequence of B.K.'s results that these conditions are always satisfied in the finite case treated there.

Aumann, R.J. (1981). Survey of repeated games. In R.J. Aumann, et al. Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Wissenschaftsverlag, Manheim, Wien, Zurich.

Bewley, T. and E. Kohlberg (1976a). The asymptotic theory of stochastic games. Math. Oper. Res. 1, 197-208.

Bewley, T. and E. Kohlberg (1976b). The asymptotic solution of a recursion equation occuring in stochastic games. Math. Oper. Res. 1, 321-336.

Blackwell, D. (1956). An analog of the minmax theorem for vector pay-offs. Pacific J. Math. 6, 1-8.

Blackwell, D. and T.S. Ferguson (1968). The big match. Ann. Math. Statist. 39, 159-163.

Gillette, D. (1957). Stochastic games with zero-stop probabilities. Contributions to the Theory of Games, Vol. III (Ann. Mathe. Studies, No.39,). Princeton University, NJ., 179-187.

Hoffman, A.J. and R.M. Karp (1966). On nonterminating stochastic games. Management Sci. 12, 359-370.

Kohlberg, E. (1974). Repeated games with absorbing states. Ann. Statist. 2, 724-738.

Mertens, J.-F. (1971-72). Repeated games: an overview of the zero-sum case. Advance Economic Theory, W. Hildenbrand (ed.). Cambridge University Press: Cambridge, 175-182.

Mertens, J.-F. and A. Neyman (1982). Stochastic games. Internat. J. Game Theory 10, 53-66.

Shapley, L. (1953). Stochastic games. Proc. Nat. Acad Sci U.S.A. 39, 1095-1100.

Stern, Martin A. (1975). On stochastic games with limiting average pay-off. Doctoral dissertation in mathematics, University of Illinois.

Chapter 4

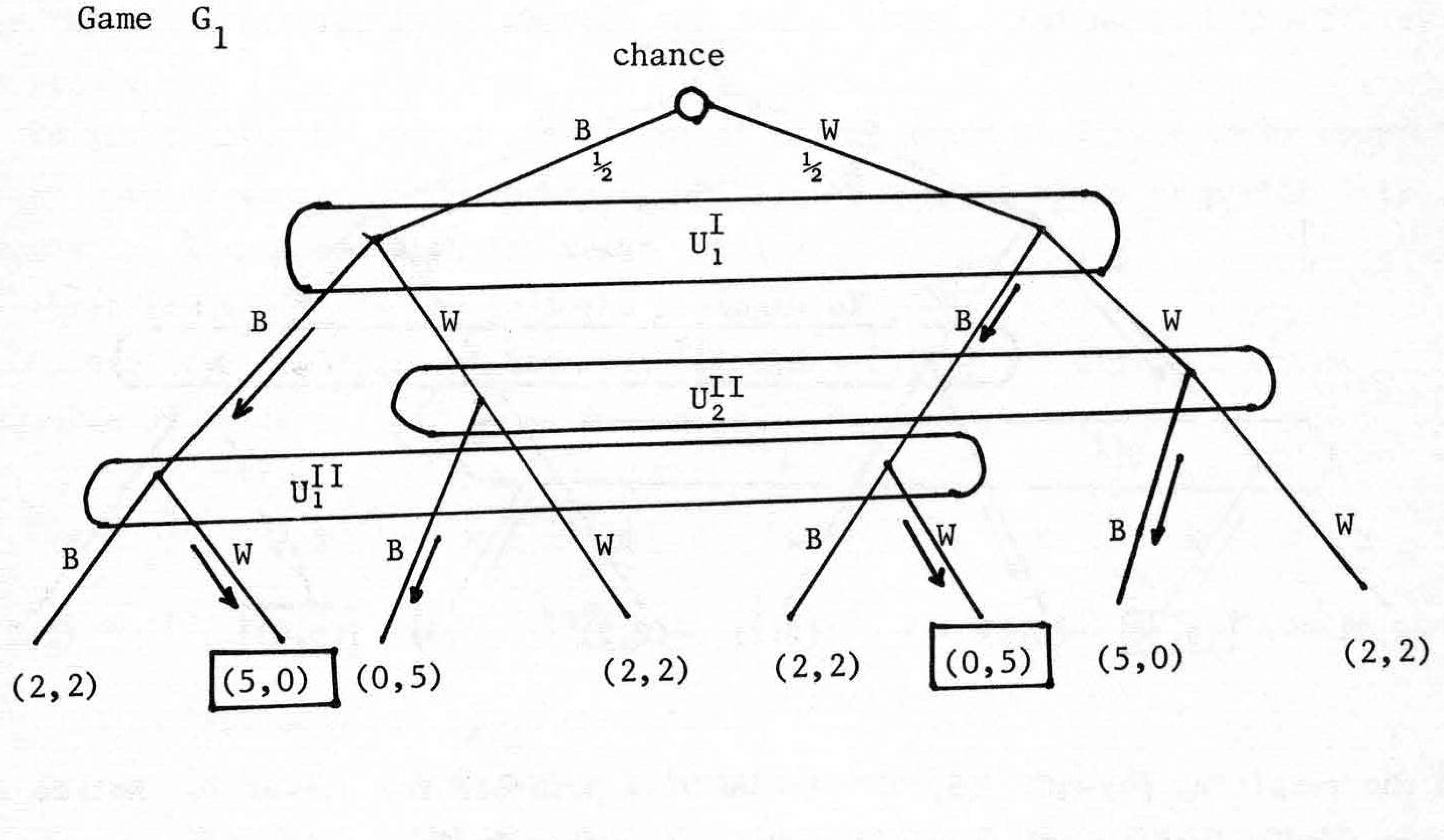
MODELING INCOMPLETE INFORMATION

In all models we discussed so far there was an implicit but very crucial underlying assumption: the description of the game and all the data involved in this description is *known to all players*. In particular each player knows the strategy sets and the pay-off functions. On the other hand we know that this is not a very realistic assumption: players are often uncertain even about their own pay-off function and their available actions, and even more so about those of the other players. Can we model such situations in which *players are uncertain as to what game they are playing?*

Example 4.1 The state of nature is chosen by a chance move to be B (black) or W (white) with probability ½ for each possibility. Players I and II are engaged in the following situation. Player I has to choose B or W. Hearing that, player II also chooses B or W, if they both choose the same thing they receive 2 each. If one chooses B and the other W, the one choosing the real state of nature receives 5 and the other player receives 0.

Case (i). Both players do not know the real state of nature. This is the game:

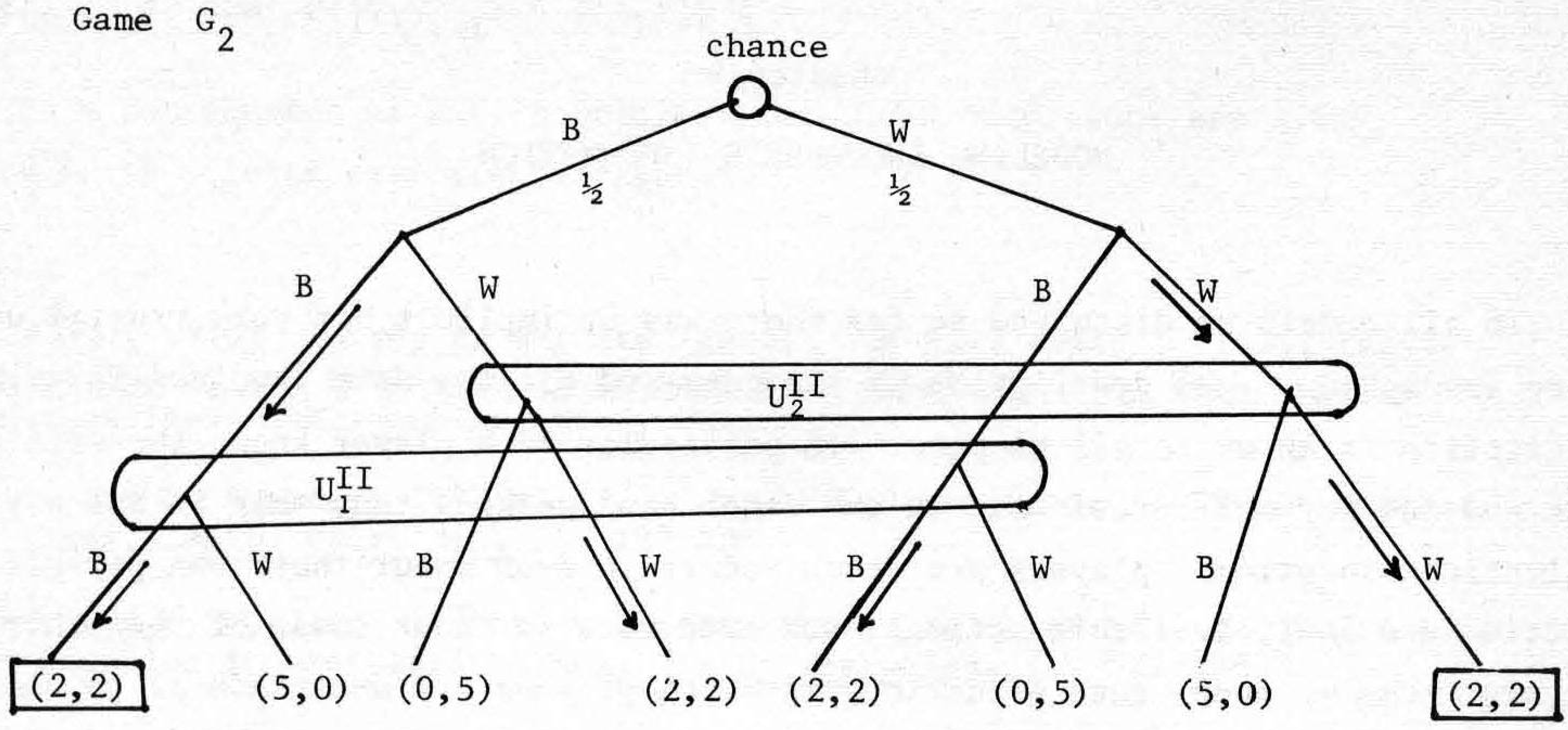
Game G_1



and with a unique N.E. pay-off $\frac{1}{2}(5, 0) + \frac{1}{2}(0, 5) = (2\frac{1}{2}, 2\frac{1}{2})$.

Case (ii). Player I knows the state of nature while player II does not, even though he is aware of the fact that I knows. The game is then:

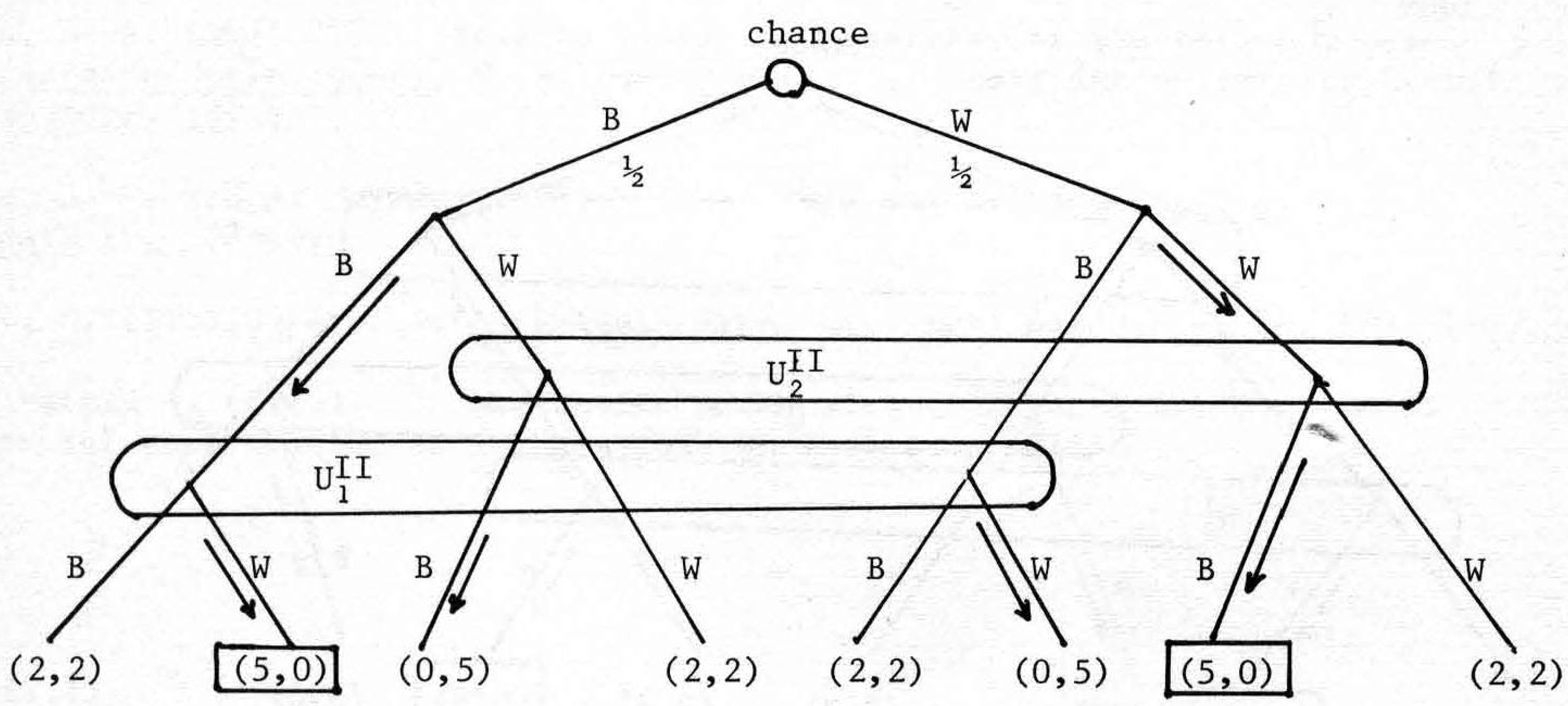
Game G_2



with a unique N.E. pay-off (2, 2) .

Here we already see a tricky thing about information; *additional information may be disadvantageous*. A moment of reflection shows that the problem of Player I is that *player II knows that he knows*. In fact if I could get his information without player II suspecting as much, we would get:

Case (iii). Player I (and only he) knows the state of nature and player II "thinks" that he does not know. The pattern of behaviour will be:



with the resulting pay-off (5, 0) -- the best possible for player I. Notice that we were careful not to call the last case a 'game'. In fact this is not a game. This is a situation in which player I knows that he is playing G_2 while II thinks that he is playing G_1 . More typically player II may not be sure whether player I knows the state of nature. In such a situation he is not sure whether he is playing G_1 or G_2 . And what about player I? What does he think about player II's beliefs concerning the real game he is playing? And what does player II think about this? The problem is getting more and more complicated and the question is how to treat it.

To fix ideas we consider a situation of incomplete information involving a set of players $N = \{1,\dots,n\}$, the members of which are uncertain about the parameters of the game they are playing which may be any element of some set S (we may think of a point of S as a full listing of the strategy sets and the pay-off functions). We shall refer to S as the set of *States of Nature*.

As we saw already in our example, a full description of the situation should include the beliefs (i.e. subjective probability distribution) of each player on S . These may be called the first level beliefs. Then we have to include what each player believes about the other player's beliefs on S . These are the second level beliefs. Then we have what a player believes are the second level beliefs of the others (i.e. what he thinks that they think that he thinks...) and so on. We are led to an infinite hierarchy of beliefs which seems inavoidable and hardly manageable.

In an attempt to overcome this difficulty, Harsanyi (1967-68) introduced the concept of *type*. A type of a player is an entity that summarizes all parameters and beliefs (of all levels) of that player. The game starts by a chance move that selects the type of each player. Of course each player knows his own type and has some beliefs (as part of his type) on the types of the other players.

The concept of type proved to be very useful but its formal derivation from the more basic notions of beliefs, beliefs on beliefs, etc. was done only some 12 years later (Böge and Eisele, 1979; Mertens and Zamir, 1985). Let us look briefly at this result.

We start with the set S of states of nature which we assume to be compact. For any compact space X we denote by $\Pi(X)$ the compact space of probability measures on X endowed with the weak* typology.

First level beliefs are just the elements of $\Pi(S)$.

Second level beliefs are elements of $\Pi(S \times [\Pi(S)]^n)$, etc. We define a sequence of spaces $\{Y_k\}_{k=0}^{\infty}$ as follows:

$Y_0 = S$ and for $k=1,2,\dots$

$Y_k = \{y_k \in Y_{k-1} \times [\Pi(Y_{k-1})]^n$ s.t. if t^i denotes the projection on the i-th copy of $[\,]^n$, then:

(a) $\forall\ i$, the marginal distribution of $t^i(y_k)$ on Y_{k-2} is $t^i(y_{k-1})$.

(b) $\forall\ i$, the marginal distribution of $t^i(y_k)$ on the i-th copy of $\Pi(Y_{k-2})$ is a unit mass at $t^i(y_{k-1})\}$

Conditions (a) and (b) are coherency conditions saying that each player knows his own beliefs and any event whose probability can be computed according to beliefs of two different levels, will have the same probability in both levels.

Now let $\mathcal{Y}$ be the projective limit of $\{Y_k\}_{k=0}^{\infty}$. $\mathcal{Y}$ is a well-defined compact space if S is compact. Let T^i be the projection of $\mathcal{Y}$ on player i's coordinates.

$$\mathcal{Y} = S \times T^1 \times \ldots \times T^n \quad . \tag{4.1}$$

The set T^i can be called the set of *types* of player i. Clearly all T^i's are copies of the same set T. An element $t^i \in T^i$ defines uniquely a probability distribution on $\mathcal{Y}$ i.e. on $S \times T^1 \times \ldots \times T^N$. By properties (a) and (b), the marginal distribution of t^i on T^i is a unit mass on $\{t^i\}$. This is a formal expression of the fact that *each player knows his own type*. Therefore:

$$T^i = \Pi(S \times \underset{j \neq i}{X} T^j) \quad . \tag{4.2}$$

Equations (4.1) and (4.2) give the structure of what we call *the universal beliefs* (BL) *space* $\mathcal{Y}$ generated by S and n. A point $y = (s, t^1, t^2, \ldots, t^n)$ of $\mathcal{Y}$ may also be called a *state of the world* (compared to *state of nature* which is an element of S). A state of the world thus consists of a state of nature and an n-tuple of types, one for each player. A type of a player which can also be called *the state of mind* of the player is just a joint probability distribution on the states of nature and the types of the other players.

Beliefs Subspaces

As the name indicates, the universal beliefs space is a very big space. It contains all possible configurations of hierarchy of beliefs. Often the uncertainty of players is confined to a small subset of $\mathcal{Y}$.

Definition 4.2 A *Beliefs subspace* (BL subspace) is a closed subset C of $\mathcal{Y}$ s.t. if $y = (s, t^1, \ldots, t^n) \in C$ then $t^i(C) = 1 \ \forall \ i$.

This is the notion of *common knowledge*, first defined formally by Aumann (1976): Every player knows that the state of the world is in C, he knows that everybody knows that the state is in C, he knows that everybody knows that everybody knows that the state is in C, etc.

Example 4.3 Players $N = \{I, II\}$; $C = \{y_1, y_2, y_3, y_4\}$ where:

$$y_1 = \{s_{11}; (\tfrac{2}{5}, \tfrac{3}{5}, 0, 0); (\tfrac{1}{3}, 0, \tfrac{2}{3}, 0)\}$$

$$y_2 = \{s_{12}; (\tfrac{2}{5}, \tfrac{3}{5}, 0, 0); (0, \tfrac{3}{4}, 0, \tfrac{1}{4})\}$$

$$y_3 = \{s_{21}; (0, 0, \tfrac{4}{5}, \tfrac{1}{5}); (\tfrac{1}{3}, 0, \tfrac{2}{3}, 0)\}$$

$$y_4 = \{s_{22}; (0, 0, \tfrac{4}{5}, \tfrac{1}{5}); (0, \tfrac{3}{4}, 0, \tfrac{1}{4})\}$$

In this BL subspace there are two types of player I:

$$I_1 = (\tfrac{2}{5}, \tfrac{3}{5}, 0, 0); \quad I_2 = (0, 0, \tfrac{4}{5}, \tfrac{1}{5}) ,$$

and two types of player II:

$$II_1 = (\tfrac{1}{3}, 0, \tfrac{2}{3}, 0); \quad II_2 = (0, \tfrac{3}{4}, 0, \tfrac{1}{4}) .$$

The mutual beliefs of each player on the other player's types are:

I on II:

	II_1	II_2
I_1	$\frac{2}{5}$	$\frac{3}{5}$
I_2	$\frac{4}{5}$	$\frac{1}{5}$

II on I:

	II_1	II_2
I_1	$\frac{1}{3}$	$\frac{3}{4}$
I_2	$\frac{2}{3}$	$\frac{1}{4}$

This is equivalent to the situation in which the pair of types is chosen according to the following probability distribution on the product of the type sets:

	II_1	II_2
I_1	$\frac{2}{10}$	$\frac{3}{10}$
I_2	$\frac{4}{10}$	$\frac{1}{10}$

.

Then each player is told his type from which he derives his subjective probability as "the conditional probability on the types of the other player given my own type."

When such a prior on the BL subspace exists it is called a *consistent* BL subspace.

<u>Example 4.4</u> $N = \{I, II\}$; $C = \{y_1, y_2, y_3, y_4\}$

$$y_1 = \{s_{11}; (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0); (\tfrac{1}{3}, 0, \tfrac{2}{3}, 0)\}$$

$$y_2 = \{s_{12}; (\tfrac{1}{2}, \tfrac{1}{2}, 0, 0); (0, \tfrac{1}{5}, 0, \tfrac{4}{5})\}$$

$$y_3 = \{s_{21}; (0, 0, \tfrac{1}{4}, \tfrac{3}{4}); (\tfrac{1}{3}, 0, \tfrac{2}{3}, 0)\}$$

$$y_4 = \{s_{22}; (0, 0, \tfrac{1}{4}, \tfrac{3}{4}); (0, \tfrac{1}{5}, 0, \tfrac{4}{5})\} .$$

I on II:

	II_1	II_2
I_1	$\frac{1}{2}$	$\frac{1}{2}$
I_2	$\frac{1}{4}$	$\frac{3}{4}$

II on I:

	II_1	II_2
I_1	$\frac{1}{3}$	$\frac{1}{5}$
I_2	$\frac{2}{3}$	$\frac{4}{5}$

No prior on $\{I_1, I_2\} \times \{II_1 \times II_2\}$ can give these as conditionals which means that this is an *inconsistent case*.

To define formally the notion of consistency we need some notation. If C is a BL subspace and $y = (s, t^1, \ldots, t^n) \in C$ we denote player I's type, t^i, in y (which is a probability measure on C) by t^i_y.

Definition 4.5 A BL subspace C is *consistent* if there exists a probability measure P on C s.t. $\forall\ i \in N$:

$$P = \int_C t^i_y \, dP \quad . \tag{4.3}$$

We will also say that this $P \in \Pi(C)$ is consistent. Any $y \in C$ is a consistent state of the world with respect to P .

With the appropriate measurability structure on C and on $\Pi(C)$, let $F(t^i)$ be the sub σ-field of measurable sets of $\Pi(C)$ generated by the projection t^i. Then:

Theorem 4.7 If y is a consistent state of the world w.s.t. a consistent P with finite support, then P (and in particular its support - the BL subspace containing y) is uniquely determined and is common knowledge.

In other words, each player, with his information only, can answer the question: Is the state of the world consistent? If the state is in fact consistent all players will know that and compute correctly the same BL subspace and the prior on it.

The way for player i to find the BL subspace, which he believes contains the state of the world y , is rather straightforward. In $y = (s, t^1, \ldots, t^n)$ player i knows t^i. He finds $C^i_{y,1} = \mathrm{Supp}(t^i_y)$ (i.e. support of t^i_y) and then inductively:

$$C^i_{y,k+1} = C^i_{y,k} \cup [\bigcup_{\tilde{y} \in C^i_{y,K}} \bigcup_j \mathrm{Supp}(t^j_{\tilde{y}})] \qquad k=1,2,\ldots \quad .$$

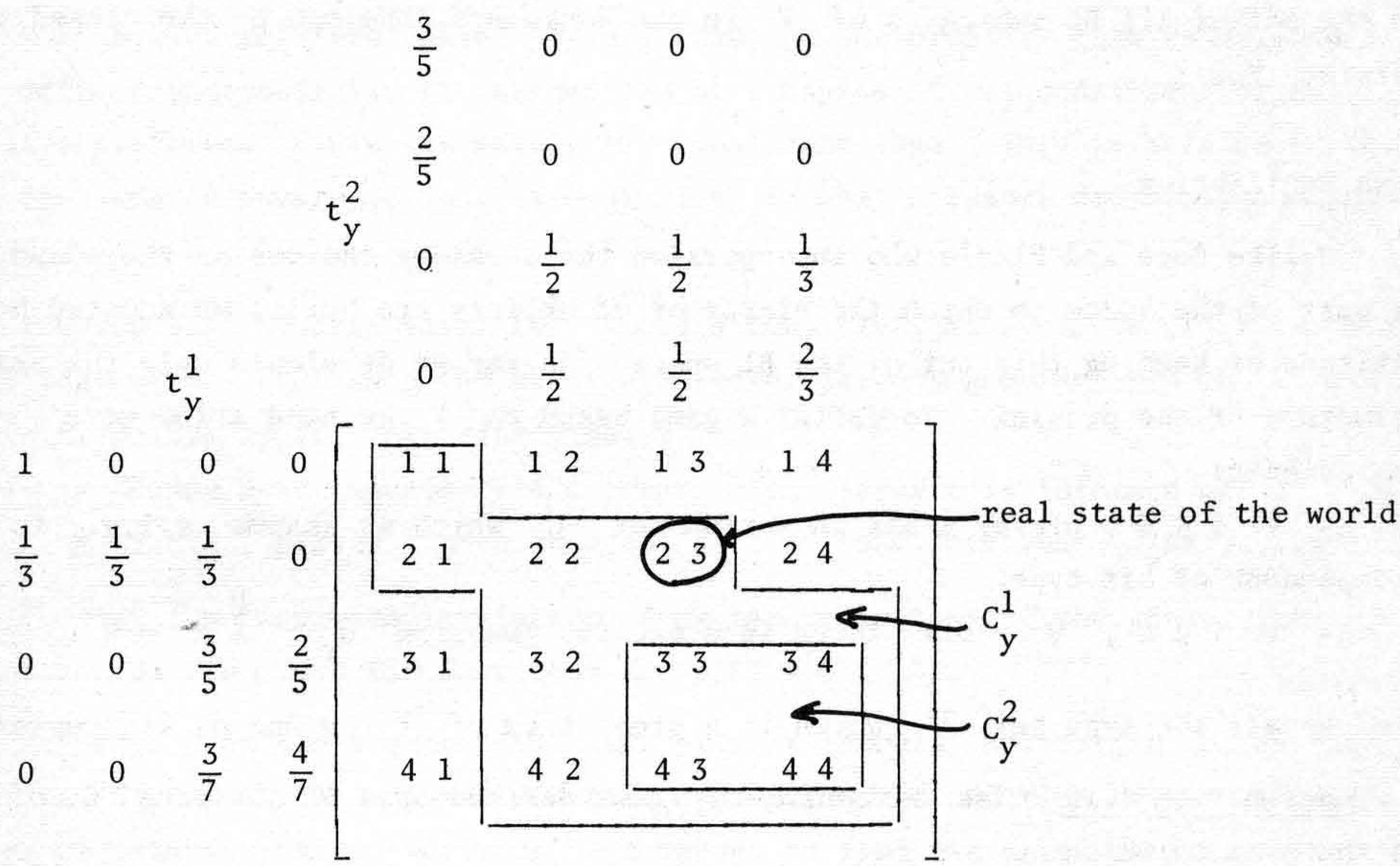
We have $C^i_{y,1} \subset C^i_{y,2} \subset \ldots$ and if C (the support of P) is finite we get a limit set C^i_y. Theorem 4.6 asserts that if y is consistent then C^i_y is the same for all i . Denoting this by C , it is the minimal BL subspaces containing the real state of the world *according to the beliefs of every player*.

The fact that the prior P on C can be computed correctly by each player follows from the consistency of P which implies:

If $P(z) > 0$ and $y \in \mathrm{Supp}(t^i_z)$ then $\dfrac{P(y)}{P(z)} = \dfrac{t^i_z(y)}{t^i_z(z)} > 0$. From this it follows by proceeding inductively on sets converging to C^i_y, that for any y and i either $P(C^i_y) = 0$ or $P(\cdot|C^i_y)$ is uniquely determined by C .

So in a consistent state of the world, players cannot draw wrong conclusions concerning the consistency. Can this happen in an inconsistent state of the world? It turns out that if y is inconsistent then player i may think wrongly that y is consistent only if $y \notin \mathrm{Supp}(t^i_y)$. Otherwise he concludes correctly that y is not consistent. For instance, in Example 4.4 in any state $y \in C$ both players will recognize correctly that the state is not consistent. On the other hand, look at the following example.

Example 4.8 Consider the following BL subspace consisting of 16 states and 4 types for each of the two players. We arrange the 16 states in a matrix as a product of the two type sets (ij means player I is of type i and II of type j). Being interested only by the beliefs structure we omit from y the state of nature s and write next to each row the corresponding type of player I which is a probability distribution on the columns (types of player II). We do similarly for player II.



If the state of the world is $y = 23$ it is inconsistent. Also, for this state $y \in t^1_y$ but $y \notin t^2_y$ so we expect player I to get to the correct conclusion which may not be the case for player II. In fact, player I will compute $C^1_y = \{11, 21, 22, 23, 32, 33, 34, 42, 43, 44\}$ but no consistent probability distribution on it, so he will conclude that the state is inconsistent. On the other hand, player II will compute $C^2_y = \{33, 34, 42, 44\}$ with the consistent distribution $P = (1/4, 1/6, 1/4, 1/3)$ on it. So he may wrongly conclude that the state is consistent.

Approximation of a BL subspace by a Finite BL subspace

As it can be easily seen, even if we start with a finite set S, both $\mathcal{Y}$ and 'most' of its BL subspaces are sets of high cardinality. On the other hand, most of the work on games with incomplete information assumes finitely many possible states of the world. To make this discrepancy slightly less disturbing we have the following theorem which we bring without its proof which is technically quite complicated (see Mertens and Zamir, Theorem 3.1).

Theorem 4.9 For any BL subspace C of $\mathcal{Y}$ and any finite open cover $\mathcal{O}$ of $\mathcal{Y}$, there is a finite BL subspace C^* of $\mathcal{Y}$ s.t.

(i) $C \subset \cup\{O \in \mathcal{O} | O \cap C^* \neq \phi\}$

(ii) $C^* \subset \cup\{O \in \mathcal{O} | O \cap C \neq \phi\}$.

In other words this theorem states that *the finite BL subspaces of* $\mathcal{Y}$ *are dense* in the set of all BL subspaces of $\mathcal{Y}$ in the *Hansdorff topology* on the closed subsets of $\mathcal{Y}$.

Nash Equilibrium

Unlike Böge and Eisele who incorporated the strategy choices of the players as part of the space on which the hierarchy of beliefs are built, we adopted here the attitude of keeping this out of the BL space. So far we developed only the beliefs structure of the problem. To define a game based on $\mathcal{Y}$ we need a few more ingredients:

- $\forall\ i \in N$, player i has an action set A^i which we assume w.l.g. to be independent of his type.
- $\forall\ i \in N$, $\forall\ y \in \mathcal{Y}$ there is a *utility function* $u^i_y : \underset{j=1}{\overset{n}{X}} A^j \to R$.

Recall the type set $\mathcal{T}^i$ which is a projection of $\mathcal{Y}$ on one of its coordinates.

Definition 4.10 The *vector pay-off* game defined on a BL subspace C of $\mathcal{Y}$ is the game in which:

- the set of players is $N = \{1,2,\ldots,n\}$;
- the (pure) strategy set Σ^i of player i is the set of mappings $\sigma^i: C \to A^i$ which is $\mathcal{T}^i$-measurable;
- the 'pay-off' to player i resulting from an n-tuple of strategies $\sigma = (\sigma^1,\ldots,\sigma^n)$ is a vector $u_i = (u_{t^i})_{t^i \in \mathcal{T}^i}$, where $u_{t^i}(\sigma) = \int u^i_y(\sigma(\tilde{y}))\, dt^i(\tilde{y})$, with the interpretation that type t^i is paid $u_{t^i}(\sigma)$.

We note that u_{t^i} is $\mathcal{T}^i$-measurable as it should be. Although this is not a game in the usual sense, the concept of N.E. can be defined in the usual way, namely: $\sigma = (\sigma^1,\ldots,\sigma^n)$ is N.E. if:

$$\forall\ i \in N,\ \forall\ t^i \in \mathcal{T}^i,\ \forall\ \tilde{\sigma}^i \in \Sigma^i, \quad u_{t^i}(\sigma) \geq u_{t^i}(\sigma|\tilde{\sigma}^i),$$

where as usual $(\sigma|\tilde{\sigma}^i)$ is the n-tuple σ in which the i-th component is replaced by $\tilde{\sigma}^i$. This is also called a *Bayesian Equilibrium*.

When C is a finite BL subspace this game is an n-person game in which the pay-off for player i is a vector of dimension equal to $|T^i|$, the number of types of player i. This is actually the game studied by Harsanyi. We can, in this case, make this an ordinary n-person game in which the pay-off to player i is $\bar{u}_i = \sum_{t^i \in T^i} \gamma_{t^i} u_{t^i}$ where $\forall\ t^i \in T^i$, γ_{t^i} is a strictly positive constant. Clearly, *independently of the constant* γ_{t^i} we choose, this game has the same N.E. points as the above vector pay-off game. Aumann and Maschler (1967) suggested γ_{t^i} s.t. $\sum_{t^i \in T^i} \gamma_{t^i} = 1$ to treat the inconsistent case.

Notice that both the vector pay-off game and the ordinary game we defined are well defined independently of whether the BL subspace C is consistent or not. Harsanyi preferred to discuss mainly the consistent case. This is because in that case the game in strategic form is equivalent to what Harsanyi calls a "game in standard form."

<u>Theorem 4.11</u> (Harsanyi): Let C be a consistent BL subspace of $\mathcal{Y}$ with a consistent prior P. Then the strategic vector pay-off game defined on C has the same N.E. points as the following game:

- A chance move chooses $y \in C$, then each player i is informed of t^i_y.
- $\forall\ i \in N$, player i then chooses $a^i \in A^i$ and receives $u^i_y(a^1,\ldots,a^n)$.

Proof. The proof readily follows from the definition of the games, the definition of N.E., and the fact that $\mathrm{Supp}(P) = C$.

Harsanyi made the argument that the players 'should' believe in P as the prior on C. We by no means claim that here. The introduction of P is just a matter of mathematical convenience. It serves to find the original N.E. points naturally defined by C via subjective probabilities. Furthermore, since by Theorem 4.7, P is common knowledge (in the consistent case), the above-described 'game in standard form' is also common knowledge which gives even more justification for using it in analyzing the situation of incomplete information.

Aumann, R.J. (1976). Agreeing to disagree. <u>Ann. Statist.</u> 4, 1236-1239.

Aumann, R.J. and M. Maschler (1967). Repeated games with incomplete information: a survey of recent results. <u>Mathematica</u> ST-116, Ch.III, 287-403.

Boge, W. and Th. Eisele (1979). On situations of Bayesian games. <u>Internat. J. Game Theory</u> 8, 193-215.

Harsanyi, J.C. (1967, 1968). Games with incomplete players played by Bayesian players. Parts I, II, III. <u>Management Sci.</u> 14 (3,5,7).

Mertens, J.-F. and S. Zamir (1985). Formulation of Bayesian analysis for games with incomplete information. <u>Internat. J. Game Theory</u> 14, 1-29.

Chapter 5

REPEATED GAMES WITH INCOMPLETE INFORMATION (I)

One of the most interesting and important aspects of incomplete information situations is the strategic use of information: When and how to reveal information? When and how much to invest in collecting new information? How does information flow between players? and so on. Clearly the right setting to deal with these issues is that obtained by combining the Bayesian games of the last chapter with repeated games. This is what we plan to do here. Most of the research done so far in this direction was in the consistent 0-sum two-person games. This is because problems of information appear already in this case.

Incomplete Information on One Side

The first and simplest model of repeated games with incomplete information was presented and studied by Aumann and Maschler in 1966. In their model the state of nature was presented by a pay-off matrix chosen at random and known to one player only:

- At stage 0 chance chooses $k \in \{1, 2\}$ with probability $(1/2, 1/2)$. The result is told to player I (the maximizer) but not to player II (who knows only the probability $(1/2, 1/2)$).

- At stage m, $m=1,2,\ldots$ player I chooses $i_m \in I$ and player II chooses $j_m \in J$ and (i_m, j_m) is told to both players.

- After stage n, player II pays player I $\frac{1}{n}\sum_{m=1}^{n} a^k_{i_m j_m}$ where $A^1 = (a^1_{ij})$ and $A^2 = (a^2_{ij})$ are two $I \times J$ matrices known to both players.

Denote this game by $\Gamma_n(\frac{1}{2})$, and its (minmax) value by $v_n(\frac{1}{2})$.

<u>Example 5.1</u> $A^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$; $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

It is easily seen that $v_1(\frac{1}{2}) = 1/2$ and I's optimal strategy is to play his dominating strategy: $i = 1$ if $k = 1$ and $i = 2$ if $k = 2$. However, this strategy is *completely revealing* (CR), i.e. after the first stage II will deduce from the move of I the k chosen and from then on he can guarantee not to pay more than 0. Thus the CR strategy yields player I a pay-off $\frac{1}{2n}$ which tends to 0 as $n \to \infty$.

The other extreme behaviour of I is to play without using his information.

This will be a *non-revealing* (NR) strategy since, being independent of k, player I's move will give no information to II (about k). In this case each stage of $\Gamma_n(\frac{1}{2})$ is equivalent to the following one-stage game:

$$\Delta(\tfrac{1}{2}) = \tfrac{1}{2} A^1 + \tfrac{1}{2} A^2 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

This game has the value $\frac{1}{4}$ so in any $\Gamma_n(\frac{1}{2})$ with $n \geq 2$, player I does better by not using his information than by fully using it immediately. Later we shall show that $\lim_{n\to\infty} v_n(\frac{1}{2}) = \frac{1}{4}$, thus asymptotically player I cannot get more than what he gets by playing NR.

<u>Example 5.2</u> $A^1 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$; $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$.

Repeating the same discussion as in the previous example we have: By playing CR, player I can guarantee 0. By playing NR he can guarantee the value of $\Delta(\frac{1}{2}) = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$ which is $-\frac{1}{4}$. Since clearly $v_n(\frac{1}{2}) = 0 \ \forall \ n$ it follows that in this case the CR is the best strategy for I.

<u>Example 5.3</u> $A^1 = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}$, $A^2 = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}$.

- By playing CR player I guarantees 0 since the value of each matrix is 0.
- By playing NR he guarantees the value of $\Delta(\frac{1}{2}) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix}$ which is again 0.
- We claim that player I can do better than 0, he can guarantee 1 in any $\Gamma_n(\frac{1}{2})$. To do that player I 'prepares' two coins C^1 and C^2. Both have the outcomes $\{1, 2\}$ with probabilities $(\frac{3}{4}, \frac{1}{4})$ in C^1 and $(\frac{1}{4}, \frac{3}{4})$ in C^2. If the game is A^k he uses C^k to choose $i \in \{1, 2\}$ and then plays that move i in all stages.

The only information player II obtains is the outcome i of the coin without knowing which coin was used. However, the probability distribution of k is updated as follows:

$$P(k = 1 | i = 1) = \tfrac{3}{4} ; \quad P(k = 1 | i = 2) = \tfrac{1}{4} .$$

So, if $i = 1$, the expected (row) pay-off is:

$$\tfrac{3}{4}(4, 0, 2) + \tfrac{1}{4}(0, 4, -2) = (3, 1, 1) .$$

If $i = 2$ the expected pay-off is

$$\tfrac{1}{4}(4, 0, -2) + \tfrac{3}{4}(0, 4, 2) = (1, 3, 1) .$$

In any case, the expected pay-off is at least 1. We shall see later that $\lim_{n\to\infty} v_n(\frac{1}{2}) = 1$.

Limit of Value and Value of Limit

As a first step in the development of the theory it is important to clarify the notion of value for repeated games in general. As we mentioned in previous lectures one would like basically to model a *many times repeated game*. Two approaches suggest themselves: The first one which we used in discussing the examples may be called *limit of value*, and consists of considering the value of the n-stage game Γ_n (with pay-offs divided by n), letting $n \to \infty$. In the second approach *value of limit*, one defines the infinite stage game Γ_∞ and considers its value. The problem in defining Γ_∞ is the lack of an obvious candidate for a pay-off function, since the expectation of $\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} g_m$, where g_m is the pay-off at stage m, may fail to exist. As we mentioned in relation to stochastic games, to overcome this difficulty we either define some kind of limiting average or we define the value of Γ_∞ *directly* without defining the pay-offs. This is what we did in definition 3.14 and this will be our attitude whenever we treat the infinite game.

Unlike the situation in stochastic games where the two approaches yield the same value, in repeated incomplete information games, they may differ. To see how this can happen let us first observe:

Remark 5.4 If v is the value of Γ_∞ then $\lim v_n$ exists and equals to v.

To see this note first that a strategy σ (or τ) in Γ_∞ defines uniquely a strategy σ_n (or τ_n) in Γ_n for n=1,2,... . This may be called the *n-stage projection* of σ (or τ). Our remark then follows from definition 3.14 which implies that if a strategy ε-*guarantees* v in Γ_∞, its n-stage projection ε-guarantees v in Γ_n for n sufficiently large.

In view of remark 5.4 the only divergence which may occur is when limit of value exists while the value of limit does not. The first example of this kind was provided by Zamir (1973). Due to time constraints we do not analyze that example here but we shall see this phenomenon later on in our lectures.

Let us now reconsider our first model generalized in the obvious way.

- The states of nature are A^k, $k \in K = \{1,...,K\}$, which are I × J pay-off matrices of a zero-sum two-person game in which $I = \{1,...,I\}$ and $J = \{1,...,J\}$ are the pure strategy sets of player I and II respectively.

The state of nature is chosen according to a given probability vector $p \in P = \{p = (p^1,...,p^K) \mid p^k \geq 0,\ \forall\ k;\ \sum_k p^k = 1\}$.

We denote the repeated games by $\Gamma_n(p)$ and their values by $v_n(p)$.

Lemma 5.5 $v_n(p)$ is concave on P for all n=1,2,... .

Proof. Let p_1, p_2 be in P and α in [0, 1] such that $\alpha p_1 + (1-\alpha)p_2 = p$. Consider the two games $\Gamma'_n(\alpha, p_1, p_2)$ and $\Gamma''_n(\alpha, p_1, p_2)$ defined as follows:

- In Γ' chance chooses $r \in \{1, 2\}$ with probability $(\alpha, 1 - \alpha)$; both players are informed about the outcome. Then $\Gamma_n(p_r)$ is played.
- Γ''_n is defined in a similar way but only player I knows the r chosen.
- The above description is common knowledge.

Note that player I has the same strategy set in both games while player II's strategy set in Γ''_n is contained in that of Γ'_n . Thus, denoting by v'_n and v''_n the values of the games, it follows that $v'_n \leq v''_n$.

Now clearly $v'_n = \alpha v_n(p_1) + (1 - \alpha) v_n(p_2)$. On the other hand Γ''_n has the same value as $\Gamma_n(p)$ since for player I the knowledge of r is useless (he will know k), and for player II, k is chosen (in two steps) with probability $\alpha p_1 + (1 - \alpha)p_2 = p$. Hence $v''_n = v_n(p)$, and the result follows.

Q.E.D.

In considering the value of $\Gamma_n(p)$ we make use of the minmax theorem which says actually that an optimal strategy of player I guarantees the value even if player II knows that it is being used. Now given a strategy σ of player I in Γ_n , player II can compute before each stage m a *posterior probability* p_m *on* K , that is, the conditional probability distribution on K given σ and given the history up to that stage. The random variable p_m plays a very fundamental role in the theory; the role of *state variable* in the dynamic programming approach. The use of this approach is possible due to the following theorem which we mention here without proof (see Mertens and Zamir, 1971-72).

<u>Theorem 5.6</u> The game $\Gamma_n(p)$ has the same value as the game in which player I announces his strategy and at stage m a new game $\Gamma_{n-m+1}(p_m)$ is played.

The most important consequence of this theorem is the following *recursive formula* for v_n .

$$v_{n+1}(p) = \frac{1}{n+1} \max_s \{\min_j(\sum_k p^k s^k A^k)_j + n \sum_i \bar{s}_i v_n(p_i)\} \quad . \tag{5.1}$$

Here $s = (s^k)_{k\in K}$ is the first stage strategy of player I, i.e. $s^k = (s^k_i)_{i\in I}$ — a probability vector on I . $\bar{s} = \sum_k p^k s^k$ and p_i is the probability vector on K given by $p^k_i = p_k\, s^k_i/\bar{s}_i$.

<u>Lemma 5.7</u> For all $p \in P$ the sequence $v_n(p)$ is decreasing.

Proof. $v_{n+1}(p) \leq v_n(p)$ is proved inductively using (5.1) and the concavity of v_n (Lemma 5.5) which implies: $\sum_i \bar{s}_i\, v_n(p_i) \leq v_n(p)$ Q.E.D.

<u>Definition 5.8</u> The nonrevealing (NR) game is the one-stage game, denoted by $\Delta(p)$, in which the pay-off matrix is $\sum_k p^k A^k$. The value of the NR game is denoted by u(p) .

This is the game in which none of the players is informed about the choice of A^k .

Lemma 5.9 For all n, $v_n(p) \geq (\text{cav } u)(p)$ on P.

Here Cav u is the smallest concave function on P which is greater or equal to u.

Proof. By using an optimal strategy of $\Delta(p)$ in each stage of $\Gamma_n(p)$, player I guarantees $u(p)$ per stage thus $v_n(p) \geq u(p)$. Since v_n is concave, the result follows.

Q.E.D.

Lemma 5.10 For each n, $v_n(p)$ is Lipschitz.

Proof. It follows from the easily proved observation that if A and B are two pay-off functions of the same dimension then:

$$|\text{val}(A) - \text{val}(B)| \leq \max_{ij} |a_{ij} - b_{ij}|$$

Q.E.D.

Corollary 5.11 As $n \to \infty$, v_n uniformly converges on P to a concave function v which satisfies $v(p) \geq (\text{Cav } u)(p)$.

Proof. The proof follows from the monotonicity, the Lipschitz property, and the concavity of v_n combined with the compactness of P.

For notational simplicity only, let us assume from now on (unless otherwise specified) two states of nature $K = \{1, 2\}$. Then P can be identified with the unit interval $[0, 1]$, where $p \in [0, 1]$ is the probability of A^1.

To get a deeper understanding of the monotone convergence of v_n, let us recall the sequence $(p_n)_{n=1}^{n}$ of posterior probabilities (thus random variables in $[0, 1]$), and observe that this is a martingale bounded in P.

Lemma 5.12 For any strategy σ of player I in $\Gamma_n(p)$ we have:

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

Here E is the expectation with respect to the probability induced by σ and p.

Proof. Since p_m is a martingale with expectation p (which is p_1) we have:

$$E(\sum_{m=1}^{n} (p_{m+1} - p_m)^2) = E(\sum_{m=1}^{n} (p_{m+1} - p_m))^2 = E(p_{n+1} - p_1)^2 \leq p(1 - p) \quad .$$

The result now follows by using Cauchy-Schwartz inequality.

Q.E.D.

The expectation $E|p_{m+1} - p_m|$ is a measure for the amount of information revealed in stage m by player I. In particular, if, at that stage, he plays NR (i.e., independently of k) then $p_{m+1} = p_m$ and thus $E|p_{m+1} - p_m|$. The next lemma says that if player I does not play NR his extra gain is somehow proportional to the information he reveals. At any stage m let $s_m = (s_m^1, s_m^2)$ be the one-stage strategy played by player I (i.e., play the mixed strategy s_m^k if the state is A^k). Let t_m be the mixed strategy of player II and let $g_m(s_m, t_m)$ be the conditional expected pay-off (given p_m) at that stage, then:

<u>Lemma 5.13</u> For all s_m and t_m :

$$|g_m(s_m, t_m) - g_m(\bar{s}_m, t_m)| \leq cE_m(|p_{m+1} - p_m|) , \tag{5.2}$$

where $c = 2 \max_{i,j,k} |a^k_{ij}|$, $\bar{s}_m$ is the NR strategy $\bar{s}_m = p_m s^1_m + (1 - p_m)s^2_m$, and E_m is the conditional expectation given p_m .

We omit the proof which is a matter of straightforward verification (see lemma 2 in Zamir, 1971-72).

<u>Lemma 5.14</u> For all $p \in P$, $v_n(p) \leq (\text{Cav } u)(p) + O(1/\sqrt{n})$.

Proof. For any strategy σ of player I compute p_m and let player II play at stage m a mixed strategy t_m which is optimal in $\Delta(p_m)$. Denote this (response) strategy of player II by τ and by $\rho_n(\sigma, \tau)$ the expected average pay-off for σ and τ .

Since $\bar{s}_m$ is an NR strategy, $g_m(\bar{s}_m, t_m) \leq u(p_m) \leq (\text{Cav } u)(p_m)$. Using (5.2), averaging over m and using the Jensen's inequality for Cav u we obtain:

$$\rho_n(\sigma, \tau) \leq (\text{Cav } u)(p) + \frac{c}{n} \sum_{m=1}^{n} E(|p_{m+1} - p_m|) .$$

Combining this with lemma 5.12 we conclude that for each σ there exists τ such that

$$\rho_n(\sigma, \tau) \leq (\text{Cav } u)(p) + \frac{c}{\sqrt{n}} \sqrt{p(1 - p)} .$$

Q.E.D.

The following theorem, due to Aumann and Maschler (1967), is a corollary of what we have so far.

<u>Theorem 5.15</u> (i) $\lim_{n\to\infty} v_n(p) = (\text{Cav } u)(p) \ \forall \ p \in P$ and the convergence is uniform.

(ii) There exists $c > 0$ such that

$$0 \leq v_n(p) - (\text{Cav } u)(p) \leq \frac{c\sqrt{p(1 - p)}}{\sqrt{n}} \quad \text{for all } p \in P \text{ and all } n .$$

Zamir has shown (1971-72) that the bound $O(1/\sqrt{n})$ for the speed of convergence is the best uniform upper bound. This was done by the following.

<u>Example 5.16</u> Consider the game in which:

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

Here $u(p) = \text{val}\begin{pmatrix} p + 2 & p - 2 \\ -p - 2 & -p + 2 \end{pmatrix} = 0 \ \forall \ p \in [0, 1]$.

We state without proving that $v_n(p) \geq p(1-p)/\sqrt{n}$ for all p and all n. (Here $(\text{Cav } u)(p) = u(p) = 0 \;\; \forall \; p \in P$.)

Remark 5.17 The central fact to emphasize in theorem 5.15 is that $\Gamma_n(p)$ cannot be analyzed for a single $0 < p < 1$ unless we study the whole family of games $\Gamma_n(p)$; $p \in P$.

Examples Revisited

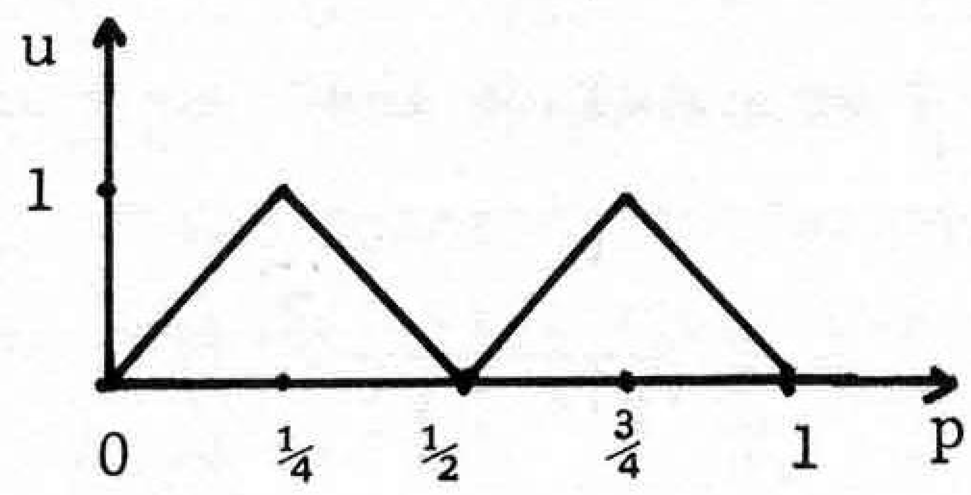
Example 5.1: $u(p) = \text{val}\begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix} = p(1-p)$.

Since $u(p)$ is concave $\lim v_n(p) = (\text{Cav } u)(p) = p(1-p)$. In particular $\lim v_n(\tfrac{1}{2}) = \tfrac{1}{4}$.

Example 5.2: $u(p) = \text{val}\begin{pmatrix} -p & 0 \\ 0 & -(1-p) \end{pmatrix} = -p(1-p)$, whose concavification is 0.

Therefore $\lim v_n(p) = (\text{Cav } u)(p) = 0 \;\; \forall \; p \in [0, 1]$.

Example 5.3: Here $u(p) = \text{val}\begin{pmatrix} 4p & 4(1-p) & 2(2p-1) \\ 4p & 4(1-p) & 2(1-2p) \end{pmatrix}$ is the following function:



Therefore $(\text{Cav } u)(\tfrac{1}{2}) = \tfrac{1}{2}u(\tfrac{1}{4}) + \tfrac{1}{2}u(\tfrac{3}{4}) = 1$.

The value of $\Gamma_\infty(p)$

By remark 5.4, if $\Gamma_\infty(p)$ has a value it must be $(\text{Cav } u)(p)$. To prove this (see Definition 3.14) one has to show that:

(a) for each $\varepsilon > 0$, Player I can guarantee $(\text{Cav } u)(p) - \varepsilon$;

(b) for each $\varepsilon > 0$, Player II can guarantee $(\text{Cav } u)(p) + \varepsilon$.

The proof of (a) is the easier part. It is even true that Player 1 *can guarantee* Cav u, i.e. he has a strategy σ such that $\rho_n(\sigma, \tau) \geq (\text{Cav } u)(p)$ for all n and all τ of player II. This strategy is quite transparent in Example 5.3. There player I constructs a type-dependent lottery in such a way that given the outcome of the lottery the (posterior) probability of A^1 is either $\tfrac{1}{4}$ or $\tfrac{3}{4}$ with equal probabilities. According to the outcome of the lottery he then plays optimally in $\Delta(\tfrac{1}{4})$ or $\Delta(\tfrac{3}{4})$ in all stages of the game. In such a strategy, the revelation part is only in the first step, which we may call the *splitting* part. That is the part in which the first stage posterior is distributed in the 'right way' so that $E(p_1) = p$ and $Eu(p_1) = (\text{Cav } u)(p)$. After the splitting part, player I plays an NR strategy which is an optimal strategy in $\Delta(p_1)$.

The fact that this splitting can always be done in the desired way yields the following (see Mertens and Zamir, 1971-72, Lemma 2 or Sorin, 1980, Lemma 2.17).

Lemma 5.18 If player I can guarantee $f(p)$ in $\Gamma_\infty(p)$, he can also guarantee $(\text{Cav } f)(p)$.

Corollary 5.19 Player I can guarantee $(\text{Cav } u)(p)$ in $\Gamma_\infty(p)$.

Proof. Player I can guarantee $u(p)$ per stage by playing at every stage, and independently of his type, an optimal strategy in $\Delta(p)$. The result now follows by Lemma 5.18.

To prove (b) let τ_n be an optimal strategy of player II in $\Gamma_n(p)$. Let N_2 N_3 ... be large integers (to be specified later) and consider the following strategy τ of player 2 in $\Gamma_\infty(p)$: At the first stage play τ_1. At the next $2N_2$ stages play N_2 times τ_2 and so on. After $1 + 2N_2 + \ldots + mN_m$ stages play N_{m+1} times τ_{m+1}. At the beginning of each 'block' player II ignores the history, as if the game newly started. With this τ player II's average pay-off for the first $(1 + 2N_2 + \ldots + mN_m)$ stages is at most:

$$(v_1 + 2N_2v_2 + \ldots + mN_mv_m)/(1 + 2N_2 + \ldots + mN_m) \quad . \qquad (5.3)$$

Now given $\varepsilon > 0$ we can choose $N_2, N_3, \ldots$ so that the expression in (5.3) will be at least $v_m - \frac{\varepsilon}{2}$ for sufficiently large m. Since $\lim v_m = \text{Cav } u$ this is at least $(\text{Cav } u)(p) - \varepsilon$ for sufficiently large m.

This concludes the second result of Aumann and Maschler.

Theorem 5.20 For all $p \in P$, $(\text{Cav } u)(p)$ is the value of $\Gamma_\infty(p)$.

Admittedly, the above-described strategy of player II to ε-guarantee Cav u is far from being appealing. Even for very moderate m, τ_m may be practically nonfeasible to compute even by the largest existing computer. In contrast, we shall now describe another very elegant, appealing and easily computable ε-optimal strategy for the uninformed player, player II. This strategy relies on a fundamental paper of Blackwell (1956).

Blackwell considered a two-person game with a "pay-off matrix" B whose elements $\{b_{ij} | i \in I, \ j \in J\}$ are vectors in the K-dimensional Euclidean space R^K. The game is infinitely repeated. After stage m, both players are told the vector pay-off $g_m \in R^K$ reached at that stage so that the total information up to this stage is the $m + 1$ "history" $h_{m+1} = (g_1, \ldots, g_m)$. A strategy of a player is a sequence of mappings from histories to probability distributions on his pure strategies (I or J).

Definition 5.21 A set $S \subset R^K$ is *approachable* for player II with τ_0 if for each $\varepsilon > 0$ there exists N_0 such that for all σ of player I, and all $n \geq N_0$ $E_{\sigma,\tau_0}(d(S,\bar{g}_n)) < \varepsilon$, where $d(\cdot,\cdot)$ is the distance in R^K, $\bar{g}_n = (1/n) \sum_{m=1}^{n} g_m$ and E_{σ,τ_0} is the expectation with respect to σ and τ_0.

S is *excludable* by player I with σ_0 if there exists $\delta > 0$ and N_0 such that for all τ and all $n \geq N_0$, $E(d(S, \bar{g}_n)) > \delta$.

Similar definitions are obtained by inversing the roles of the players. S is approachable for a player if he has a strategy with which it is approachable for him.

For each $t = (t_1,\ldots,t_J)$, a probability distribution on J, denote $R_{II}(t)$ convex hull of $\{ \sum_{j \in J} t_j b_{ij} \; ; \; i \in I\}$. Hence, if player II uses t his expected pay-off will be in $R_{II}(t)$. The following theorem is the only part of Blackwell's results needed here:

Theorem 5.22 (Blackwell 1956): Let S be a closed set in R^K. If for each $x \notin S$ there exists $t(x)$, a probability vector on J such that if y in S is the closest point to x, the hyperplane perpendicular to the line $x - y$ through y separates x from $R_{II}(t(x))$, then S is approachable for player II. An approaching strategy is given by:

- at stage 1 or if $\bar{g}_n \in S$ play anything;
- otherwise play $t_{m+1} = t(\bar{g}_m)$, $n \geq 1$.

With this theorem at hand we now construct a strategy of the uninformed player which ε-guarantees $(\text{Cav } u)(p)$.

Step 1. Let $H = \{x \in R^K | \alpha \cdot x = \alpha \cdot p\}$ be the supporting hyperplane to Cav u at the point p, i.e. $\alpha \in R^K$ satisfies:

$$(\text{Cav } u)(p) = \alpha \cdot p \quad \text{and} \quad u(q) \leq \alpha \cdot q \quad \text{for all } q \in P .$$

(As usual, $x \cdot y$ denotes the dot product in R^K.)

Step 2. Consider the set $S = \{y \in R^K | y^k \leq \alpha^k, \text{ for all } k \in K\}$, i.e. the 'corner set' in R^K defined by α. It is enough to show a strategy of player II with respect to which S would be approachable for him, since this would mean that the average expected pay-off up to state n will be at most $\alpha \cdot p + \varepsilon = (\text{Cav } u)(p) + \varepsilon$ for n large enough.

Step 3. Let $x_n \in R^K$ be the average vector pay-off at the end of stage n - 1, and let y_n be the point in S closest to x_n. The approaching strategy for player II is as follows. At stage n:

- If $y_n = x_n$ (i.e., $x_n \in S$) play anything.

- If $x_n \notin S$ let $p' \in P$ be a vector in the direction of $x_n - y_n$. Play t_n which is optimal in $\Delta(p')$.

Note that the hyperplane H' through y_n is perpendicular to p'.
$H' = \{y \in R^K | p' \cdot y = p' \cdot y_n\}$ separates x_n from X (since S is convex). Thus in view of Theorem 5.22 it remains to show that $R_{II}(t_n)$ is on the same side of H' as S. In fact, since t_n is optimal in $\Delta(p')$ we have
$\sum_k p'^k sA^k t_n \leq u(p') \leq \alpha \cdot p'$ for all mixed strategies s of player I.

Now remark that if $p'^k > 0$ then $y_n^k = \alpha^k$ so that $\alpha \cdot p' = y_n \cdot p' < x_n \cdot p'$, i.e., when t_n is used, the resulting expected vector pay-off for that stage is on the opposite side of H' from x_n, that is to say, on the same side as S.

Remark 5.23 Comparing Definition 3.14 and the notion of approachability in Definition 5.21 we actually prove a somewhat stronger result than needed. Not only that, for each $\varepsilon > 0$, player II has a strategy τ_ε which guarantees $\rho_n(\sigma, \tau_\varepsilon) < (\text{Cav } u)(p) + \varepsilon$ for large enough n for every σ, but he has *one strategy* τ which does this for all $\varepsilon > 0$.

Remark 5.24 When in the above-treated model the informed player is player II, the minimizer, then the Aumann-Maschler's result reads: $\lim v_n$ = value of Γ_∞ = Vex u, where Vex u is the largest convex function f satisfying $f(q) \leq u(q)$ for all $q \in P$.

Chapter 6

REPEATED GAMES WITH INCOMPLETE INFORMATION (II)

Incomplete Information on Both Sides

The first model of incomplete information for both players was given by Aumann and Maschler (1967) and was the natural generalization of their first asymmetric model treated in the previous chapter.

The Model. The states of nature are $I \times J$ matrices A^{ks} where $k \in K = \{1,\ldots,K\}$, $s \in S = \{1,\ldots,S\}$. $p \in P$ and $q \in Q$ are probability distributions on K and S , respectively.

At stage 0 , chance chooses the state of nature according to the product probability $p \times q$, i.e. $\Pr(A^{ks}) = p^k q^s \quad \forall\ k, s$. Player I is told the value of k and player II is told the value of s . (That is, K and S are the type sets of players I and II respectively.)

At stage m , $m=1,2,\ldots$ player I chooses $i_m \in I$ and player II chooses $j_m \in J$ and (i_m, j_m) is announced.

In the n-repeated game, denoted by $\Gamma_n(p, q)$, the pay-off is $\frac{1}{n} \sum_{m=1}^{n} a^{ks}_{i_m j_m}$, and the value is denoted by $v_n(p, q)$. In the infinitely repeated game $\Gamma_\infty(p, q)$ we again define the value $v_\infty(p, q)$ without defining a pay-off function (Definition 3.14).

Remark 6.1 Note that in our model the types of the players are chosen independently. We shall later refer to this as *the independent case* in contrast to *the dependent case* to be introduced later.

The *nonrevealing game* (NR), denoted by $\Delta(p, q)$ is the zero-sum two-person game with the matrix pay-off $\sum_{k,s} p^k q^s A^{ks}$. Its value is denoted by $u(p, q)$.

For any real function $f(p, q)$ defined on $P \times Q$ we denote by $\operatorname{Cav}_p f(\cdot, q)$ the concavification with respect to p , the value of q being fixed. $\operatorname{Vex}_q f(p,\cdot)$ is defined similarly. With minor abuse of notation we write $\operatorname{Cav}_p f(p,q)$ and $\operatorname{Vex}_q f(p,q)$ instead of $(\operatorname{Cav}_p f(\cdot, q))(p)$ and $(\operatorname{Vex}_q f(p, \cdot))(q)$, respectively.

The Infinitely Repeated Game $\Gamma_\infty(p, q)$

We recall without repeating the notion of strategies in $\Gamma_\infty(p, q)$. Note that for player I a strategy σ can be looked at as a K-tuple $\sigma = (\sigma^k)_{k\in K}$ where σ^k is a usual infinite game strategy (used by player I if he is of type k). A similar description is valid for the strategies of player II.

Definition 6.2 $f(p, q)$ is said to be *the minmax* of $\Gamma_\infty(p, q)$ if:

(i) For each strategy τ of player II, $\forall\ \varepsilon > 0$ there is σ of player I and N such that $\rho_n(\sigma, \tau) > f(p, q) - \varepsilon$ for all $n \geq N$.

(ii) $\forall\ \varepsilon > 0$, there is $N(\varepsilon)$ and a strategy τ_ε of player II such that $\rho_n(\sigma, \tau_\varepsilon) < f(p, q) + \varepsilon$ for all σ and all $n > N(\varepsilon)$.

The notion of *maxmin* is defined similarly.

Condition (ii) says that player II can guarantee $f + \varepsilon$ in terms of lim sup. Part (i) asserts that he cannot guarantee anything lower than f even in terms of lim inf.

Theorem 6.3 The minmax of $\Gamma_\infty(p, q)$ equals $\underset{q}{\text{Vex}}\ \underset{p}{\text{Cav}}\ u(p, q)$.

The maxmin of $\Gamma_\infty(p, q)$ equals $\underset{p}{\text{Cav}}\ \underset{q}{\text{Vex}}\ u(p, q)$.

Proof. We prove only the first part, the second follows then similarly.

Step 1. If player II ignores his private information (s) and plays NR, the game $\Gamma_\infty(p, q)$ reduces then to $\bar{\Gamma}_\infty(p)$ with lack of information on one side defined by the matrices $A^k = \sum_s q^s A^{ks}$ and the probability p on K. By Theorem 5.20, in this game player II can guarantee $(\text{Cav}\ \bar{u})(p)$ where $\bar{u}$ is the value of $\sum_k p^k A^k = \sum_{k,s} p^k q^s A^{ks}$ which is just $u(p, q)$. That is, player II can guarantee Cav $u(p, q)$ in the stronger sense of Remark 5.23: he has a strategy τ which guarantees $\rho_n(\sigma, \tau) < \underset{p}{\text{Cav}}\ u(p, q) + \varepsilon$ for all σ and for all $\varepsilon > 0$ for n large enough.

Step 2. By Lemma 5.18 used for the uninformed player II, he can also guarantee (in the same sense) $\underset{q}{\text{Vex}}\ \underset{p}{\text{Cav}}\ u(p, q)$.

This concludes the proof of (a somewhat stronger version than) (ii) in the definition of minmax. The proof of (i) is more technical, therefore we only outline the idea and main points in the proof.

- Given a strategy τ, player I can compute the posteriors q_m on S. Now $(q_m)_{m=1}^\infty$, being a martingale bounded in the simplex Q, converges with probability 1. In terms of information this means that far enough in the game, player II will reveal almost no information. Player I can therefore play NR during a large number of stages N in order "*to exhaust the maximal amount of information from* τ." Afterwards the situation is almost the same as if player II plays NR so player I can obtain $u(p, q_N)$, hence $\underset{p}{\text{Cav}}\ u(p, q_n)$. His expected pay-off is (up to an ε), $E\ \underset{p}{\text{Cav}}\ u(p, q_N)$ which is at least $\underset{q}{\text{Vex}}\ \underset{p}{\text{Cav}}\ u(p, q)$.

The technical steps which turn this idea into a formal proof are:

1. $E_{p,\sigma,\tau} \sum_s \sum_{m=1}^{\infty} (q_{m+1}^s - q_m^s) \leq \sum_s q^s(1 - q^s)$ for all σ. Thus let σ^* be the strategy of player I which ε-achieves the supremum of this quantity up to stage N.

2. Since q_m depends only on q, τ, h_m and j_m, the 'average' NR strategy σ_0 which will have the same distribution on j_m will produce the same q_m as σ^* and thus will do the same job as σ^*.

3. For any strategy σ which coincides with σ_0 up to the stage N we get for all $n \geq N$, $E(\sum_s |q_n^s - q_N^s|) \leq M\sqrt{\varepsilon}$ for some constant M.

4. Given τ and $\varepsilon > 0$, player I plays σ_0 up to stage N, then does the 'splitting of p' to $(p_i)_{i\in I}$ with probabilities $(\lambda_i)_{i\in I}$ such that $\sum_i \lambda_i p_i = p$ and $\sum \lambda_i u(p_i, q_N) = \text{Cav}\, u(p, q_N)$ and then play optimally in $\Delta(p_i, q_N)$. Q.E.D.

<u>Corollary 6.4</u> The infinite game has a value if and only if:

$$\underset{p}{\text{Cav}}\ \underset{q}{\text{Vex}}\ u(p, q) = \underset{q}{\text{Vex}}\ \underset{p}{\text{Cav}}\ u(p, q) \quad . \tag{6.1}$$

The following example provides a game in which (6.1) does not hold and hence $\Gamma_\infty(p, q)$ *does not have a value.*

<u>Example 6.5</u> (See Mertens and Zamir 1971-72) Let $K = \{1, 2\}$, $S = \{1, 2\}$ and:

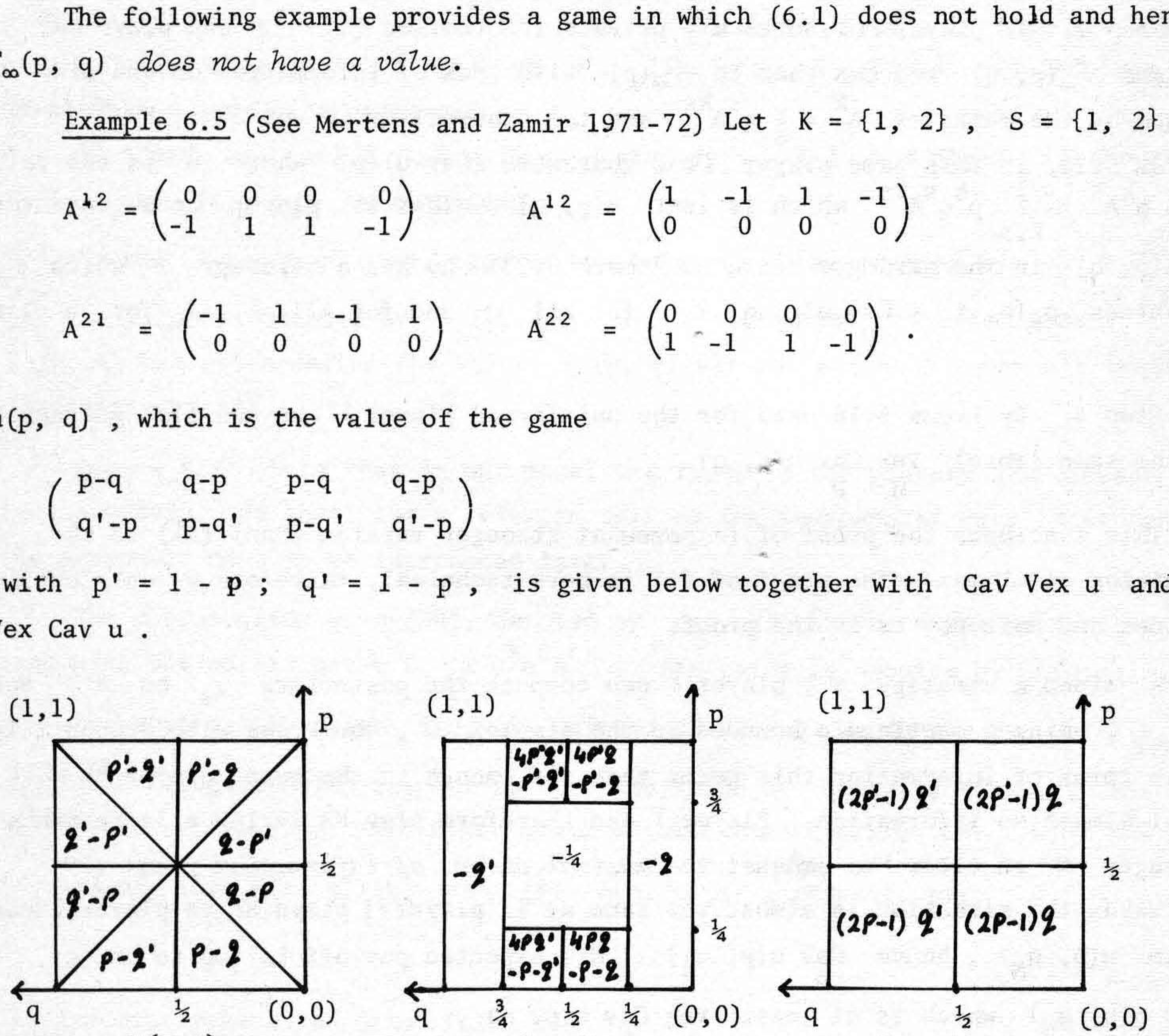
$$A^{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix} \qquad A^{12} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

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

$u(p, q)$, which is the value of the game

$$\begin{pmatrix} p-q & q-p & p-q & q-p \\ q'-p & p-q' & p-q' & q'-p \end{pmatrix}$$

with $p' = 1 - p$; $q' = 1 - p$, is given below together with Cav Vex u and Vex Cav u .



u(p,q) Cav Vex u Vex Cav u

The Finite Games $\Gamma_n(p,q)$

With the nonexistence of value for $\Gamma_\infty(p, q)$ there still remains the question of existence of the limit of $v_n(p, q)$, the value of the n-repeated games $\Gamma_n(p, q)$. This is answered by the following theorem which we bring without proof (see Mertens and Zamir, 1971-72).

Theorem 6.6 For all $(p, q) \in P \times Q$, $\lim_{n\to\infty} v_n(p, q)$ exists and is the only simultaneous solution of the following two functional equations.

$$v(p, q) = \operatorname{Vex}_q \max \{u(p, q), v(p, q)\} \qquad (6.2)$$

$$v(p, q) = \operatorname{Cav}_p \min \{u(p, q), v(p, q)\} . \qquad (6.3)$$

Remark 6.7 Let us show that when Cav Vex u = Vex Cav u this is also the unique solution of (6.2) and (6.3), as it should be.

- First observe that Vex Cav u *is a solution of (6.2)*. In fact, notice that Cav and Vex are monotone operators, therefore, on one hand:

$$\text{Vex Cav } u = \text{Vex max } \{u, \text{Cav } u\} \geq \text{Vex max } \{u, \text{Vex Cav } u\} .$$

On the other hand, Vex Cav u $\leq$ max {u, Vex Cav u} . Taking Vex from both sides yields the other inequality and so Vex Cav u = Vex max {u, Vex Cav u }.

- Similarly, Cav Vex u is a solution of (6.3). Therefore, if

Cave Vex u = Vex Cav u , this is a common solution of (6.2) and (6.3). That it is the *only* common solution follows from:

- Any v which satisfies (6.2) and (6.3) satisfies Cav Vex u $\leq$ v $\leq$ Vex Cav u. In fact, from (6.2), v $\geq$ Vex u . Since by (6.3) v is concave we have v $\geq$ Cav Vex u . The second inequality is obtained similarly.

Extensions of the Model

Of the variants of the above-described basic model let us mention two. One is the direction of allowing a more general mechanism for revealing information than just through the moves. The other is in allowing a more general structure of prior information and *dependence* between the types of the two players.

Signaling matrices. We modify our model by introducing two matrices H_I^{ks} and H_{II}^{ks} of dimensions $I \times J$ and with elements $h_I^{ks}(ij)$ and $h_{II}^{ks}(ij)$ in some finite set H . If the state of nature drawn at stage 0 is ks and if at stage m the players choose i_m and j_m , then player I is informed of $h_I^{ks}(i_m,j_m)$ and player II of $h_{II}^{ks}(i_m,j_m)$. When $h_I^{ks}(ij) = h_{II}^{ks}(ij) = (i, j)$ for all k and s this is the usual model which we shall therefore call the *standard signaling case*.

Aumann and Maschler (1968) proved their result for the incomplete information on one side. $v = \lim v_n = \text{Cav } u$ for general signaling matrices, of course after redefining appropriately the NR game and its value u. It turns out that the signaling matrices for the informed player are immaterial for this result (they may have an effect on v_n but not on its limit).

Definition 6.8 For $p \in P$ a one-stage strategy of player I is nonrevealing (NR) if for each $j \in J$, the distribution on the letters of H in the row $h^k_{II}(ij)$ is the same for all k. (That is, this is a strategy after which the posterior on K cannot change.)

Denote by $NR(p)$ the set of nonrevealing strategies of player I at p. (Note that $NR(p)$ may be empty, but is nonempty when p is an extreme point of P.) Define the NR game $\Delta(p)$ as the game in which player I is restricted to $NR(p)$ if it is not empty. Finally:

$$u(p) = \begin{cases} \text{value of } \Delta(p) & \text{if } NR(p) \neq \phi \\ -\infty & \text{if } NR(p) = \phi \end{cases} .$$

With this definition Aumann and Maschler proved:

Theorem 6.9 $\lim_{n\to\infty} v_n(p)$ and $v(p)$ exist and both equal to $(\text{Cav } u)(p)$.

The generalization of Blackwell's approachability strategy for the uninformed player was done by Kohlberg (1975).

For incomplete information on both sides Mertens and Zamir (1980) proved the above-stated results about minmax; maxmin and $\lim v_n$ for signaling matrices which are independent of the state of nature. The model they treated was more general also in another respect, namely, they treated the dependent case which shall be explained briefly now.

The Dependent Case. As we remarked before (Remark 6.1), the Aumann Maschler model for incomplete information on both sides assumed that the types of the two players are chosen independently. In such a model the probability distribution of a player on the types of his opponent is independent on his own type.

The Model. The set K is the set of states of world and $p \in P$ is a probability distribution on K. K^I and K^{II} are two partitions of K. (The elements of K^I and K^{II} are the *types* of players I and II, respectively.) The signaling matrices H_I and H_{II} are the same for all states of nature (and this is a very crucial assumption without which the results are not valid).

As mentioned above all results to the special case were extended to this general case where $u(p)$ is the value of the one-stage game in which both players are restricted to NR strategies, i.e., strategies which produce a probability distribution on the signals of the opponent which is independent on the state of the world k (no matter what the opponent does).

The main difficulty was the extension of the operators Cav and Vex since we no longer have the natural variables p for concavification and q for convexification. The key to the right generalization is the following observation. If the distribution on K is $p \in P$, since any one-stage strategy of player I is K^I-measurable, the resulting posterior distribution on K given his move at that stage will be in the set $\Pi_1(p) \subset P$ defined by:

$$\Pi_I(p) = \{(\alpha^1 p^1, \ldots, \alpha^K p^K) \in P \mid (\alpha^k)_{k \in K} \text{ is } K^I\text{-measurable}\} .$$

Similarly,

$$\Pi_{II}(p) = \{(\beta^1 p^1, \ldots, \beta^K p^K) \in P \mid (\beta^k)_{k \in K} \text{ is } K^{II}\text{-measurable}\} .$$

Clearly for any $p \in P$ both $\Pi_I(p)$ and $\Pi_{II}(p)$ are nonempty convex and compact subsets of P. A real function f defined on P will be called I-concave if for any $p_0 \in P$, f(p) restricted to $\Pi_I(p_0)$ is concave. The notion of II-convex is defined similarly. Then we define $\underset{I}{\mathrm{Cav}}\, f$ and $\underset{II}{\mathrm{Vex}}\, f$ in the natural way and we have:

Theorem 6.10 (a) The minmax of $\Gamma_\infty(p)$ is $\underset{II}{\mathrm{Vex}}\, \underset{I}{\mathrm{Cav}}\, u(p)$.

(b) The maxmin of $\Gamma_\infty(p)$ is $\underset{II}{\mathrm{Cav}}\, \underset{I}{\mathrm{Vex}}\, u(p)$.

(c) For each $p \in P$, $\lim v_n(p)$ exists and is the only solution v of the following two equations:

(i) $v(p) = \underset{II}{\mathrm{Vex}} \max \{u(p), v(p)\}$

(ii) $v(p) = \underset{I}{\mathrm{Cav}} \min \{u(p), v(p)\}$.

The existence of a unique solution to (i) and (ii) is an interesting duality theorem that can be proved without any mention of game theory (see Mertens and Zamir 1977b, and Sorin 1986, forthcoming).

Speed of Convergence and the Normal Distribution

We have seen that in the case of incomplete information on one side and standard signaling, the speed of convergence of $v_n(p)$ is bounded by $O(1/\sqrt{n})$ and this is the best bound. This turns out to be the case also for incomplete information on both sides with standard signaling. When signaling is by H_I and H_{II} independent of the state of nature we have a higher bound of $O(1/n^{\frac{1}{3}})$ and this is the best bound (Zamir, 1973a).

Let us recall example 5.16 with $A^1 = \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix}$, $A^2 = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$. For this game $(\mathrm{Cav}\, u)(p) = u(p) = 0 \;\; \forall \; p \in P$ and

$$p(1 - p)/\sqrt{n} \leq v_n(p) \leq \sqrt{p(1 - p)}/\sqrt{n} \quad . \tag{6.4}$$

The order $O(1/\sqrt{n})$ may be explained by the following argument. Since $\text{Cav}\, u(p) = u(p)$, the informed player must 'essentially' ignore his information and play the same mixed strategy $(\frac{1}{2}, \frac{1}{2})$ at each stage independently of his type. The average pay-off will be a random variable with variation (namely standard deviation) of the order of $(1/\sqrt{n})$. In the example under consideration the informed player can take advantage of this natural variation by 'pretending' to play $(\frac{1}{2}, \frac{1}{2})$, but actually deviating slightly from it. This deviation is exactly of the order that the uninformed player might expect as random but is actually used to the advantage of the informed player.

The following, quite surprising, result (Mertens and Zamir, 1976) shows a much closer connection to the Central Limit Theorem than outlined above: the normal distribution appears explicitly.

Theorem 6.11 For the game in Example 5.16: $\lim_{n\to\infty} \sqrt{n}\, v_n(p) = \phi(p)$, where $\phi(p)$ is the standard normal density evaluated at its p-quantile, i.e.

$$\phi(p) = (1/\sqrt{2\pi}) \exp(-\tfrac{1}{2}x_p^2) \quad \text{and} \quad (1/\sqrt{2\pi}) \int_{-\infty}^{x_p} \exp(-\tfrac{1}{2}x^2)dx = p \quad . \tag{6.5}$$

In a recent unpublished result, Mertens and Zamir showed that this is generally true at least for the case of two states of nature. Whenever the error term $v_n(p) - \lim v_n(p)$ is of the order of $(1/\sqrt{n})$ the coefficient is the normal density function. The proof has nothing to do with the above intuitive argument. The normal distribution comes out as a solution of a certain differential equation. At this point, this result is quite mysterious and the "something behind" is still to be discovered.

The result of Theorem 6.11 is intimately related, actually equivalent, to the following optimization problem, which makes no mention of game theory (Mertens and Zamir, 1977). Let $X(p) = p_0, p_1, \ldots p_n$ be a martingale with values in $[0, 1]$ and $p_0 \equiv p$. Let $M(p)$ be the set of all such martingales.

Theorem 6.12
$$\lim_{n\to\infty} \Big(\sup_{X(p)\in M(p)} (1/\sqrt{n}) \sum_{m=0}^{n} E|p_{m+1} - p_m| \Big) = \phi(p) ,$$
where $\phi(p)$ is given by (6.5).

The connection between the two problems should be clear by now. The $(p_m)_{m=1}^n$ are the posterior probabilities of A^1 (given that player I is using a certain strategy σ). $E|p_{m+1} - p_m|$ is a measure of the information revealed by player I at the m-th stage. This is also his pay-off at stage m (compared to $u(p_m) = 0$ that he would get if he would play NR). Therefore, the objective function to be maximized is $\sum_{m=0}^{n} E|p_{m+1} - p_m|$. But since $(p_m)_{m=1}^n$ is a martingale bounded in

[0, 1] , this expression is bounded by $O(1/\sqrt{n})$ and is in fact of that order.

Relations to Stochastic Games

By now it should be clear that repeated games of incomplete information are fundamentally different from stochastic games. In repeated games of incomplete information the state of nature is *fixed* but may be *unknown* by some players. It is the 'state of mind' of the players that changes along the play. In stochastic games, on the other hand, the state of nature *changes* randomly but it is *known* to all players. This difference is well manifested in the results: infinite undiscounted stochastic games have a value while incomplete information infinitely repeated games have no value in general (except for some special cases). Nevertheless there is a close relation between the two models which consists mainly of the fact that some incomplete information games can be transformed to equivalent stochastic games. To see that, consider the following examples.

Example 6.13 In a two-person zero-sum game Γ_{12} there are two states of nature {1, 2} chosen with equal probabilities. The pay-off matrices are:

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} ; \quad A_2 = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} .$$

No player gets any prior information and both get the same signals according to the signaling matrices

$$H_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} ; \quad H_2 = \begin{pmatrix} a & e \\ c & d \end{pmatrix} .$$

Since the values of both A_1 and A_2 are 0 it follows that:

- As soon as signal b or e is announced a pay-off -1 is made and the games moves to an absorbing state with value 0 (we denote this by -1, →0*).
- As long as neither b nor e was announced, any signal a yields a pay-off 0 for that stage and the same game is repeated again, similarly for c and d .

We summarize this as follows:

$$\begin{pmatrix} 0, \to\text{same} & -1, \to 0^* \\ 1, \to\text{same} & 2, \to\text{same} \end{pmatrix} .$$

This means that our game is equivalent to the infinitely repeated game with one absorbing state: $\begin{pmatrix} 0 & 0^* \\ 1 & 2 \end{pmatrix}$. (We omitted the -1 since it does not affect the evaluation of the pay-off sequence which is 0 from a finite stage on.)

Now this game has clearly the value 1 , therefore this is also the value of the infinitely repeated game in our example.

Example 6.14 Consider the game Γ_{23} which is of the same type as Γ_{12} of the previous example, but with states of nature $\{2, 3\}$ and:

$$A_2 = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \; ; \qquad A_3 = \begin{pmatrix} 0 & 0 \\ -4 & -2 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} a & e \\ c & d \end{pmatrix} \; ; \qquad H_3 = \begin{pmatrix} f & e \\ c & d \end{pmatrix} .$$

This is equivalent to the infinitely repeated game $\begin{pmatrix} 0^* & -1 \\ -1 & -1 \end{pmatrix}$ with value -1 .

Example 6.15 (Mertens, 1982) Consider now the same type of game with three states of nature $\{1, 2, 3\}$ chosen with probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The pay-off and the signaling matrices are as before:

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} \; ; \qquad A_2 = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \; ; \qquad A_3 = \begin{pmatrix} 0 & 0 \\ -4 & -2 \end{pmatrix}$$

$$H_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \; ; \qquad H_2 = \begin{pmatrix} a & e \\ c & d \end{pmatrix} \; ; \qquad H_3 = \begin{pmatrix} f & e \\ c & d \end{pmatrix} .$$

Using the same reasoning as before we reduce our game to:

$$\begin{pmatrix} 0 \begin{array}{l} \xrightarrow{2/3} \Gamma_{12} \\ \xrightarrow[1/3]{} A_3 \end{array} & -\frac{2}{3} \begin{array}{l} \xrightarrow{2/3} \Gamma_{23} \\ \xrightarrow[1/3]{} A_1 \end{array} \\ \\ -\frac{2}{3}, \rightarrow \text{same} & -\frac{2}{3}, \rightarrow \text{same} \end{pmatrix}$$

By our previous examples $v(\Gamma_{12}) = 1$, $v(\Gamma_{23}) = -1$. Also $v(A_1) = v(A_3) = 0$. Therefore our game is equivalent to the repeated game $\frac{2}{3}\begin{pmatrix} 1^* & -1^* \\ -1 & 1 \end{pmatrix}$.

This is a special stochastic game called "The Big Match" (Blackwell and Ferguson 1968). The foregoing reduction of some games with incomplete information to games with absorbing states is due to Kohlberg and Zamir (1974). By induction on the number of states of nature they proved this type of reduction to the family of repeated two-person zero-sum games with incomplete information in which:

(i) The players have no prior information on the state of nature, i.e., $K^I = K^{II} = K$.

(ii) The signals are the same for both players.

(iii) The signals tell the players at least each other's pure strategy choices.

As a matter of fact, it was this work of Kohlberg and Zamir that motivated the generalization of Blackwell and Ferguson's results about the Big Match to general stochastic games with absorbing states. This generalization, which was accomplished by Kohlberg (1974), accelerated the research on stochastic games which was concluded in a very satisfactory way by the works of Bewley-Kohlberg (1976a, 1976b, 1978) and Mertens-Neyman (1981).

Aumann, R.J. and M. Maschler (1966). Game-theoretic aspects of gradual disarmament. Mathematica ST-80, Ch.V, 1-55.

Aumann, R.J. and M. Maschler (1967). Repeated games with incomplete information: a survey of recent results. Mathematica ST-116, Ch.III, 287-403.

Aumann, R.J. and M. Maschler (1967). Repeated games with incomplete information: the zero-sum extensive case. Mathematica ST-143, Ch.III, 37-116.

Bewley, T. and E. Kohlberg (1976a). The asymptotic theory of stochastic games. Math. Oper. Res. 1, 197-208.

Bewley, T. and E. Kohlberg (1976b). The asymptotic solution of a recursion equation occuring in stochastic games. Math. Oper. Res. 1, 321-336.

Bewley, T. and E. Kohlberg (1978). On stochastic games with stationary optimal strategies. Math. Oper. Res. 3, 104-125.

Blackwell, D. (1956). An analog of the minmax theorem for vector pay-offs. Pacific J. Math. 6, 1-8.

Blackwell, D. and T.S. Ferguson (1968). The big match. Ann. Math. Statist. 39, 159-163.

Kohlberg, E. (1974). Repeated games with absorbing states. Ann. Statist. 2, 724-738.

Kohlberg, E. (1975). Optimal strategies in repeated games of incomplete information. Internat. J. Game Theory 4, 7-24.

Kohlberg, E. and S. Zamir (1974). Repeated games of incomplete information: the symmetric case. Ann. Statist. 2, 1040-1041.

Mertens, J-F. (1971-72). Repeated games: an overview of the zero-sum case. Advances in Economic Theory, W. Hildenbrand (ed.). Cambridge University Press: Cambridge, 175-182.

Mertens, J-F. and A. Neyman (1982). Stochastic games. Internat. J. Game Theory 10, 53-66.

Mertens, J-F. and S. Zamir (1971-1972). The value of two-person zero-sum repeated games with lack of information on both sides. Internat. J. Game Theory 1, 39-64.

Mertens, J-F. and S. Zamir (1977a). The maximal variation of a bounded martingale. Israel J. of Mathe. 27, 252-276.

Mertens, J-F. and S. Zamir (1977b). A duality theory on a pair of simultaneous functional equations. J. Mathe. Anal. App. 60, 550-558.

Mertens, J-F. and S. Zamir (1980). Minmax and maxmin of repeated games with incomplete information. Internat. J. Game Theory 9, 201-215.

Sorin, S. (1980). An introduction to two-person zero-sum repeated games with incomplete information. IMSSS-Economics, Stanford University, TR 312.

Zamir, S. (1971-72). On the relation between finitely and infinitely repeated games with incomplete information. Internat. J. Game Theory 1, 179-198.

Zamir, S. (1973a). On repeated games with general information function. Internat. J. Game Theory 2, 215-229.

Zamir, S. (1973b). On the notion of value for games with infinitely many stages. Ann. Statist. 1, 791-796.