Chapter 4

REPEATED GAMES WITH COMPLETE INFORMATION

SYLVAIN SORIN

Université Paris X and Ecole Normale Supérieure

Contents

0. Summary

The theory of repeated games is concerned with the analysis of behavior in long-term interactions as opposed to one-shot situations; in this framework new objects occur in the form of threats, cooperative plans, signals, etc. that are deeply related to "real life" phenomena like altruism, reputation or cooperation. More precisely, repeated games with complete information, also called supergames, describe situations where a play corresponds to a sequence of plays of the same stage game and where the payoffs are some long-run average of the stage payoffs. Note that unlike general repeated games [see, for example, Mertens, Sorin and Zamir (1992)] the stage game is the same (the state is constant; compare with stochastic games; see the chapter on 'stochastic games' in a forthcoming volume of this Handbook) and known to the players (the state is certain; compare with games of incomplete information, Chapters 5 and 6 in this Handbook).

1. Introduction and notation

A *repeated game* results when a given game is played a large number of times and, when deciding what to do at each stage, a player may take into account what happened at all previous stages (or more precisely what he knows about it). The payoff is an average of the stage payoffs.

More formally let $G = G_1$ be the following strategic form game: I is the finite set of players with generic element i (we also write I for its cardinality). Each player i has a finite non-empty set of moves (or actions) S^i and a payoff function g^i from $S = \Pi_{j \in I} S^j$ into $\mathbb{R}$. X^i will denote the set of randomized or mixed moves of i, i.e. probabilities on S^i. For x in $X = \Pi_i X^i$, $g(x)$ stands for the usual multilinear extension of g and is the expected vector payoff if each player i plays x^i.

To G is associated a *supergame* Γ, played in stages: at stage 1, all players choose a move simultaneously and independently, thus defining a move profile, that is an I-tuple $s_1 = \{s_1^i\}$ of moves in S. s_1 is then announced to all players and the game proceeds to stage 2. (Note that we are assuming full monitoring; all past behavior is observed by everyone. For a more general framework see Section 5.) Inductively at stage $n + 1$, knowing the previous sequence of move profiles $(s_1, s_2, \ldots, s_n)$, all players again choose their moves simultaneously and independently. This choice is then told to all and the game proceeds to the next stage.

A *history* (resp. a *play*) is a finite (resp. infinite) sequence of elements of S; and the set of such sequences will be denoted by H (resp. H_∞). H_n is the subset of n-stage histories. Histories are the basic ingredients of repeated games; they

allow the players to coordinate their behavior. Note that in the present framework histories are known by all players, but in more general models (see Section 5) they will lead to differentiated information.

A *pure strategy* for player i in Γ is, by the above description, a mapping from H to S^i, specifying after each history the action to select. A *mixed strategy* is a probability distribution on the set of pure strategies. Since Γ is a game with perfect recall, Kuhn's theorem implies that it is enough to work with behavioral strategies, a *behavioral strategy* σ^i of player i being a mapping from H to X^i. Alternatively, σ^i can be represented by a sequence $\{\sigma^i_n\}$, σ^i_n being a mapping from H_{n-1} to X^i that describes the "strategy of player i at stage n". Write Σ^i for the corresponding set and $\Sigma = \Pi \, \Sigma^i$.

Each pure strategy profile σ induces a play h_∞ in a natural way. Formally: $s_1 = \sigma(\emptyset)$, $s_{n+1} = \sigma(s_1, s_2, \ldots, s_n)$ and $h_\infty = (s_1, \ldots, s_n, \ldots)$. Accordingly, each σ in Σ (or in the set of mixed strategies) defines a probability, say P_σ, on $(H_\infty, \mathcal{H}_\infty)$, where $\mathcal{H}_\infty$ is the product σ-algebra on $H_\infty = S^\infty$ (and similarly $\mathcal{H}_n$ on H_n); we denote by E_σ the expectation operator corresponding to probability P_σ.

To complete the description of Γ it remains to define a payoff function φ from Σ to $\mathbb{R}^I$. The theory of repeated games deals with mappings that are some kind of average of the sequence of stage payoffs $(g_1 = g(s_1), \ldots, g_n = g(s_n), \ldots)$ associated with a play. This is (with the stationary structure of information) the main difference from multimove games where the payoff can be *any* function on plays. Three classes will be analyzed here.

(i) *The finite game* G_n. The payoff is the arithmetic average of the sum of the payoffs for the n first stages and is denoted by $\bar{\gamma}_n$; hence $\bar{\gamma}_n(\sigma) = E_\sigma(\bar{g}_n)$, where $\bar{g}_n = (1/n) \sum_{m=1}^n g(s_m)$, $n \in \mathbb{N}$. G_n is the usual n-stage game where we normalize the payoffs to allow for a comparative study as n varies.

(ii) *The discounted game* G_λ. Here φ is the geometric average of the infinite stream of payoffs; it is written $\bar{\gamma}_\lambda$ with $\bar{\gamma}_\lambda(\sigma) = E_\sigma(\sum_{m=1}^\infty \lambda(1-\lambda)^{m-1} \times g(s_m))$, $\lambda \in (0, 1]$. G_λ is thus the game with discount factor λ (where again the payoff is normalized).

In each of these two cases Γ is a well-defined game in strategic form, so that the usual concepts (like equilibrium) apply. The situation is a little more delicate in the final case.

(iii) *The infinite game* G_∞. The payoff is taken here as some limit of $\bar{g}_n$. Different definitions are possible, because the above limit may not exist and one may choose liminf or limsup or some Banach limit, and because one can take the expectation first or the limit first. Finally, especially if the infinite game is considered as an approximation of a long but finite game, some uniformity conditions may be required for equilibrium.

We will use mainly the following definitions: σ is a *lower* (resp. *upper*) equilibrium if $\bar{\gamma}_n(\sigma)$ converges to some $\bar{\gamma}(\sigma)$ as n goes to infinity, and for each

τ^i in Σ^i and each i one has: liminf (resp. limsup) $\bar{\gamma}_n^i(\tau^i, \sigma^{-i}) \leqslant \bar{\gamma}^i(\sigma)$, where as usual σ^{-i} stands for the $(I-1)$ tuple induced by σ on $I\backslash\{i\}$.

Similarly, σ is a *uniform* equilibrium if $\bar{\gamma}_n(\sigma)$ converges and, moreover, $\forall \varepsilon > 0$, $\exists N$, $n \geqslant N \Rightarrow \bar{\gamma}_n^i(\tau^i, \sigma^{-i}) \leqslant \bar{\gamma}_n^i(\sigma) + \varepsilon$, for each τ^i and each i. In words, for any positive ε, σ is an ε-equilibrium in any sufficiently long game G_n.

When the payoff function is unspecified, the result will be independent of its particular choice.

Remark. One can also work with the random variables $\bar{g}_n$ and say that a deviation is profitable if limsup $\bar{g}_n$ increases with probability one.

Recall, finally, that a subgame perfect equilibrium of Γ is a strategy profile σ such that for all h in H, $\sigma[h]$ is an equilibrium in Γ, where $\sigma[h]$ is defined on H by $\sigma[h](h') = \sigma(h, h')$ and (h, h') stands for the history h followed by h'.

The main aim of the theory is to study the behavior of long games. Hence, we will consider the asymptotic properties of G_n as n goes to infinity or G_λ as λ goes to 0, as well as the limit game G_∞.

(Note once and for all that the 0-sum case is trivial: each player can play his optimal strategy i.i.d. and the value is constant – compare with Chapter 5 and the chapter on 'stochastic games' in a forthcoming volume of this Handbook.)

Each of these approaches has its own advantages and drawbacks and to compare them is very instructive. G_n corresponds to the "real" finite game, but usually the actual length is unknown or not common knowledge (see Subsection 7.1.2). Here the existence of a last stage has a disturbing backwards effect. G_λ has some nice properties (compactness, stationary structure) but cannot be studied inductively and here the discount factor has to be known precisely. Note that G_λ can be viewed as some $G_{\tilde{n}}$, where $\tilde{n}$ is an integer-valued random variable, finite a.s., whose law (but not the actual value) is known by the players. On the other hand, the use of G_∞ is especially interesting if a uniform equilibrium exists.

A few more definitions are needed to state the results.

Given a normal form game $\Gamma = (\Sigma, \varphi)$, the set of achievable payoffs is $\Delta = \{d \in \mathbb{R}^I; \exists \sigma \in \Sigma, \varphi(\sigma) = d\} = \varphi(\Sigma)$; it is denoted by D_n, D_λ and D_∞ for G_n, G_λ and G_∞, respectively.

Similarly, the set of Nash equilibrium payoffs is $\mathscr{E} = \{d \in \mathbb{R}^I; \exists \sigma \in \Sigma$ that is an equilibrium in Γ with $\varphi(\sigma) = d\}$; it is denoted by E_n, E_λ or E_∞ in the respective cases. Finally, $\mathscr{E}'$ – and specifically, E_n', E_λ' and E_∞' – will denote the set of subgame perfect equilibrium payoffs.

D is the set of *feasible* payoffs with (public pure) correlated strategies in G_1, or equivalently, if Co denotes the convex hull: $D = Co\, D_1 = Co\, g(S)$. (This corresponds to the convex combination of payoffs in the original game G.) In

fact we shall see that repetition will allow us to mimic this public correlation in a verifiable way (because of the pure support).

The *minimax* level is defined by $v^i = \min_{X^{-i}} \max_{X^i} g^i(x^i, x^{-i})$ (recall that $X^{-i} = \Pi_{j \neq i} X^j$ is the set of vectors of mixed actions of the opponents of i). If $x^{-i}(i)$ realizes the above minimum, it will be referred to as a punishing strategy of players in $I\backslash\{i\}$ against i, and $x^i(i)$ will be the best reply of player i to it. V with components v^i is the *threat point*.

Finally, E is the set of *individually rational* (i.r. for short) and feasible payoffs: $E = \{d \in D; \forall i \in I, d^i \geqslant v^i\}$.

We will be interested in studying the asymptotic behavior of the sets $D_n, D_\lambda, E_n, \ldots$ (all convergence of sets will be with respect to the Hausdorff topology) and in describing D_∞, E_∞ and E_∞'. We shall see that the sets D and E will play a crucial role.

Before letting the parameters vary, we note that the games (Σ, φ) for the first two classes (i) and (ii) have compact pure strategy spaces and jointly continuous payoffs; hence the following properties hold.

Proposition 1.1. *D_n and D_λ are non-empty, path-connected, compact sets.*

Proposition 1.2 (Nash). *E_n, E_n', E_λ and E_λ' are non-empty, compact sets.*

Remarks. It is easy to see that neither D_n nor D_λ is necessarily convex, and neither E_n nor E_λ connected. On the other hand, both D and E are convex, compact, and non-empty, since E contains E_1.

The following easy result illustrates one aspect of repetition: the possibility of convexifying the joint payoffs.

Proposition 1.3. (i) *D_n converges to D as n goes to infinity.*
  (ii) *The same is true for D_λ as λ goes to* 0.
  (iii) *$D_\infty = D$.*

Proof. Note first that the random stage payoff takes its values in the closed convex set D and hence expectation, average and limits share the same property so that $\varphi(\sigma)$ belongs to D for all σ; thus $\Delta \subset D$ (but $Co(\Delta) = D$). Now for every $\varepsilon > 0$, there exists some integer p such that any point d in D can be ε-approximated by a barycentric rational combination of points in $g(S)$, say $d' = \Sigma_m (q_m/p)g(s_m)$. Thus the strategy profile σ defined as: play cycles of length p consisting of q_1 times s_1, q_2 times s_2, and so on, induces a payoff near d' in G_n for n large enough.

  (ii) follows from (i) since the above strategy satisfies $\bar{\gamma}_\lambda(\sigma) \to d'$ as $\lambda \to 0$.

(iii) is obtained by taking for σ a sequence of strategies σ_k, used during n_k stages, with $\|\bar{\gamma}_{nk}^\mu(\sigma_k) - d\| \le 1/k$. □

Note that D_n may differ from D for all n, but one can show that D_λ coincides with D as soon as $\lambda \le 1/I$ [Sorin (1986a)].

It is worth noting that the previous construction associates a play with a payoff, and hence it is possible for the players to observe any deviation. This point will be crucial in the future analysis.

The next three sections are devoted to the study of various equilibrium concepts in the framework of repeated games, using both the asymptotic approach and that of limit games. Section 2 deals with strategic or Nash equilibria, Section 3 with subgame perfection, and Section 4 with correlated and communication equilibria.

2. Nash equilibria

To get a rough feeling for some of the ideas involved in the construction of equilibrium strategies, consider an example with two players having two strategies each, Friendly and Aggressive. In a repeated framework, an equilibrium will be composed of a plan, like playing (F, F) at each stage, and of a threat, like: "play A forever as soon as the other does so once". Note that in this way one can also sustain a plan like playing (F, F) on odd days and (A, A) otherwise, or even playing (F, F) at stage n, for n prime (which is very inefficient), as well as other convex combinations of payoffs. On the one hand new good equilibria (in the sense of being Pareto superior) will appear, but the set of all equilibrium payoffs will be much greater than in the one-shot game.

In a discounted game two new aspects arise. One is related to the relative weight of the present versus the future (some punishment may be too weak to prevent deviations), but this failure disappears when looking at asymptotic properties. The second one is due to the stationary structure of the game: the strategy induced by an equilibrium, given a history consistent with it, is again an equilibrium in the initial game. For example, if a "deviation" is ignored at one stage, then there is an equilibrium in which similar "deviations" at all stages are ignored. We shall nevertheless see that this constraint will generically not decrease the set of equilibrium payoffs.

In finite games, there cannot be any threat on the last day; hence by induction some constraints arise that may prevent some of the previous plan/threat combinations. Nevertheless in a large class of games, the asymptotic results are roughly similar to those above.

Let us now present the formal analysis.

A first result states that all equilibrium payoffs are in E; obviously they need to be achievable and i.r. Formally:

Proposition 2.0. $\mathscr{E} \subset E$.

Proof. Obviously $\mathscr{E} \subset D$. Now let d be in E and σ be an associated equilibrium strategy profile. Then player i can, after any history h, use a best reply to $\sigma^{-i}(h)$. This gives him a (stage, and hence total) payoff greater than v^i in G_n or G_λ. As for G_∞ (if the payoff is not defined through limits of expectations), let g_m^i denote the random payoff of player i at stage m, then the random variables $z_m = g_m^i - E(g_m^i | \mathcal{H}_{m-1})$ are bounded, uncorrelated and with zero mean and hence by an extension of the strong law of large numbers converge a.s. in Cesaro mean to 0. Since $E(g_m^i | \mathcal{H}_{m-1}) \geq v^i$, this implies that player i can guarantee v^i and hence $d^i \geq v^i$ as well. $\square$

It follows that to prove the equality of the two sets, it will be sufficient to represent points in E as equilibrium payoffs.

We now consider the three models.

2.1. The infinitely repeated game G_∞

The following basic result is known as the Folk theorem and is the cornerstone of the theory of repeated games. It states that the set of Nash equilibrium payoffs in an infinitely repeated game coincides with the set of feasible and individually rational payoffs in the one-shot game so that the necessary condition for a payoff to be an equilibrium payoff obtained in Proposition 2.0 is also sufficient.

Most of the results in this field will correspond to similar statements but with other hypotheses regarding the kind of equilibria, the type of repeated game or the nature of the information for the players.

Theorem 2.1. $E_\infty = E$.

Proof. Let d be in E and h a play achieving it (Proposition 1.3). The equilibrium strategy is defined by two components: a *cooperative behavior* and *punishments* in the case of deviation. Explicitly, σ is: play according to h as long as h is followed; if the actual history differs from h for the first time at stage n, let player i be the first (in some order) among those whose move differs from the recommendation at that stage and switch to $x(i)$ i.i.d. from stage $n + 1$ on. Note that it is crucial for defining σ that h is a play (not a probability distribution on plays). The corresponding payoff is obviously d. Assume now that player i does not follow h at some stage and denote by $N(s^i)$ the set of subsequent stages where he plays s^i. The law of large numbers implies that $(1/\#N(s^i)) \sum_{n \in N(s^i)} g_n^i$ converges a.s. to $g(s^i, x^{-i}(i)) \leq v^i$ as $\#N(s^i)$ goes to ∞ and hence limsup $g_n^i \leq v^i$, a.s. Moreover, it is easy to see that σ

defines a uniform equilibrium, since the total gain by deviation is uniformly bounded. This proves that $E \subset E_\infty$ and hence the result by the previous proposition. $\square$

Note that since we are looking only for Nash equilibria, it may be better for one player not to punish. This point will be taken into account in the next section. For a nice interpretation of and comments on the Folk Theorem, see Kurz (1978). Conceptual problems arise when dealing with a continuum of players; see Kaneko (1982).

2.2. The discounted game G_λ

Note first that in this case the asymptotic set of equilibrium payoffs may differ from E, see Forges, Mertens and Neyman (1986). A simple example is the following three-person game, where player 3 is a dummy:

$$\begin{pmatrix} (1,0,0) & (0,1,0) \\ (0,1,0) & (1,0,1) \end{pmatrix}.$$

This being basically a constant-sum game between players 1 and 2, it is easy to see that for all values of the discount factor λ, the only equilibrium (optimal) strategies in G_λ are $(1/2, 1/2)$ i.i.d. for both, leading to the payoff $(1/2, 1/2, 1/4)$. Hence the point $(1/2, 1/2, 1/2)$ in E cannot be obtained. In particular this implies that Pareto payoffs cannot always be approached as equilibrium payoffs in repeated games even with low discount rates.

In fact this phenomenon does not occur in two-person games or when a generic condition is satisfied [Sorin (1986a)].

Theorem 2.2. *Assume $I = 2$ or that there exists a payoff vector d in E with $d^i > v^i$ for all i. Then E_λ converges to E.*

The idea, as in the Folk Theorem, is to define a play that the players should follow and to punish after a deviation. If $I \geq 3$, the play is cyclic and corresponds to a strictly i.r. payoff near the requested payoff. It follows that for λ small enough, the one-stage gain from deviating (coefficient λ) will be smaller than the loss (coefficient $1 - \lambda$) of getting at most the i.r. level in the future. If $I = 2$ and the additional condition is not satisfied, either $E = \{V\}$ or only one player can profitably deviate and the result follows. $\square$

2.3. The n-stage game G_n

It is well known that E_n may not converge to E, the classical example being the

Prisoner's Dilemma described by the following two-person game:

$$\begin{pmatrix} (3,3) & (0,4) \\ (4,0) & (1,1) \end{pmatrix},$$

where $E_n = \{(1,1)\}$ for all n. This property is not related to the existence of dominant strategies; a similar one holds with a mixed equilibrium in

$$\begin{pmatrix} (2,0) & (0,1) \\ (0,1) & (1,0) \end{pmatrix}.$$

In fact, these games are representative of the following class [Sorin (1986a)]:

Proposition 2.3.1. *If $E_1 = \{V\}$, then $E_n = \{V\}$ for all n.*

Proof. Let σ be an equilibrium in G_n and denote by $H(\sigma)$ the set of histories having positive probability under σ. Note first that on all histories of length $(n-1)$ in $H(\sigma)$, σ induces V, by uniqueness of the equilibrium in G_1. Now let m be the smallest integer such that after each history in $H(\sigma)$ with length strictly greater than m, σ leads to V. Assume $m \geqslant 0$ and take a history, say h, of length m in $H(\sigma)$ with $\sigma(h)$ not inducing V. It follows that one player has a profitable deviation at that stage and cannot be punished in the future. $\square$

The following result is typical of the field and shows that a good equilibrium payoff can play a dissuasive role and prevent backwards induction effects:

Theorem 2.3.2 [Benoit and Krishna (1987)]. *Assume that for all i there exists $e(i)$ in E_1 with $e^i(i) > v^i$. Then E_n converges to E.*

Proof. The idea is to split the stages into a cooperative phase at the beginning and a reward/punishment phase of fixed length at the end. During the first part the players are requested to follow a cyclic history leading to a strictly i.r. payoff approximating the required point in E. The second phase corresponds to playing a sequence of R cycles of length I, leading to $(e(1), \ldots, e(I))$. Note that this part consists of equilibria and hence no deviation is profitable. On the other hand, a deviation during the first period is observable and the players are then requested to switch to $x(i)$ for the remaining stages if i deviates. It follows that, by choosing R large enough, the one-shot gain is less than $R \times (e^i(i) - v^i)$ and hence the above strategy is an equilibrium. Letting n grow sufficiently large gives the result. $\square$

Note that the above proof also shows the following: if E contains a strictly i.r. payoff, a necessary and sufficient condition for E_n to converge to E is that for all i there exists n_i and $e^i(i)$ in E_{n_i} with $e^i(i) > v^i$.

In conclusion, repetition allows for coordination (and hence new payoffs) and threats (new equilibria). Moreover, for a large class of games, the set of equilibria increases drastically with repetition and one has continuity at ∞: lim $E_n = \lim E_\lambda = E_\infty = E$; every feasible i.r. payoff can be sustained by an equilibrium. On the other hand, this set seems too large (it includes the threat point V) and a first attempt to reduce it is to ask for subgame perfection.

3. Subgame perfect equilibria

The introduction of the requirement of perfection will basically not change the basic results concerning the limit game. Going back to the example at the beginning of Section 2, the length of the punishment (playing A) can be adapted to the deviation, but can remain finite and hence its impact on the payoff is zero.

On the other hand, the specific features of the discounted game (fixed point property) and of the finite game (backwards induction) will have a much larger impact, being applied on each history. For example, if A is a dominant move, playing A at each stage will be the only subgame perfect equilibrium strategy of the finite repeated game.

As in the previous section we will consider each type of game (and recall that $\mathscr{E}' \subset \mathscr{E}$).

3.1. G_∞

The first result is an analog of the Folk Theorem, showing that the equilibrium set is not reduced by requiring perfection. In fact, the possibly incredible threat of everlasting punishment can be adapted so that the same play will still be supported by a perfect equilibrium.

Theorem 3.1 [Aumann and Shapley (1976), Rubinstein (1976)]. $E'_\infty = E$.

Proof.   The cooperative aspect of the equilibrium is like in the Folk Theorem. The main difference is in the punishment phase; if the payoff is defined through some limiting average it is enough to punish a deviator during a finite number of stages and then to come back to the original cooperative play. It is not advantageous to deviate; it does not harm to punish. Explicitly, if a deviation happens at stage n, punish until the deviator's average payoff is within $1/n$ of the required payoff. Deviations during the punishment phase are ignored. (To get more in the spirit of subgame perfection, one might require inductively the punisher to be punished if he is not punishing. For this to be done, since a deviation may not be directly observable during the punishment phase, some statistical test has to be used.)   □

The interpretation of the "Perfect Folk Theorem" is that punishments can be enforced either because they do not hurt the punisher or because higher levels of punishment are available against a player who would not punish. This second idea will be used below.

Remarks. (1) Note that a priori the previous construction will not work in G_n or G_λ since there a profitable deviation during a finite set of stages counts, and on the other hand the hierarchy of punishment phases may lead to longer and longer phases.

(2) For similar results with different payoffs or concepts, see Rubinstein (1979a, 1980).

3.2. G_λ

A simple and useful result in this framework, which is due to Friedman (1985), states that any payoff that strictly dominates a one-shot equilibrium payoff is in E_λ' for λ small enough. (The idea is, as usual, to follow a play that generates the payoff and to switch to the equilibrium if a deviation occurs.)

In order to get the analog of Theorem 2.2, not only is an interior condition needed (recall the example in Subsection 2.2), but also a dimensional condition, as shown by the following example due to Fudenberg and Maskin (1986a). Player 1 chooses the row, player 2 the column and player 3 the matrix in the game with payoffs:

$$\begin{pmatrix} (1,1,1) & (0,0,0) \\ (0,0,0) & (0,0,0) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} (0,0,0) & (0,0,0) \\ (0,0,0) & (1,1,1) \end{pmatrix}.$$

Let w be the worst subgame perfect equilibrium payoff in G_λ. Then one has $w \geq \lambda g_1 + (1 - \lambda)w$, where g_1 is any payoff achievable at stage 1 when two of the players are using their equilibrium strategies. It is easily seen that for any triple of randomized moves there exists one player's best reply that achieves at least $1/4$, i.e. $g_1 \geq 1/4$; hence $w \geq 1/4$ so that $(0,0,0)$ cannot be approached in E_λ'.

A generic result is due to Fudenberg and Maskin (1986a):

Theorem 3.2. *If E has a non-empty interior, then E_λ' converges to E.*

Proof. This involves some nice new ideas and can be presented as follows. First define a play leading to the payoff, then a family of plans, indexed by I, consisting of some punishment phase [play $x(i)$] and some reward phase [play $h(i)$ inducing an i.r. payoff $f(i)$]. Now if at some stage of the game player i is the first (in some order) deviator, the plan i is played from then on until a new possible deviation.

To get the equilibrium condition, the length R of the punishment phase has to be adapted and the the rewards must provide an incentive for punishing, i.e. for all i, j one needs $f^i(j) > f^i(i)$ (here the dimensional condition is used). Finally, if the discount factor is small enough, the loss in punishing is compensated by the future bonus.

The proof itself is much more intricate. Care has to be taken in the choice of the play leading to a given payoff; it has to be smooth in the following sense: given any initial finite history the remaining play has to induce a neighboring payoff. Moreover, during the punishment phase some profitable and non-observable deviation may occur [recall that $x(i)$ consists of mixed actions] so that the actual play following this phase will have to be a random variable $h'(i)$ with the following property: for all players j, $j \neq i$, the payoff corresponding to R times $x(i)$, then $h(i)$ is equal to the one actually obtained during the punishment phase followed by $h'(i)$. At this point we use a stronger version of Proposition 1.3 which asserts that for all λ small enough, any payoff in D can be exactly achieved by a smooth play in G_λ. [Note that $h'(i)$ has also to satisfy the previous conditions on $h(i)$.] $\square$

Remarks. (1) The original proof deals with public correlation and hence the plays can be assumed "stationary". Extensions can be found in Fudenberg and Maskin (1991), Neyman (1988) (for the more general class of irreducible stochastic games) or Sorin (1990).

(2) Note that for two players the result holds under weaker conditions; see Fudenberg and Maskin (1986a).

3.3. G_n

More conditions are needed in G_n than in G_λ to get a Folk Theorem-like result. In fact, to increase the set of subgame perfect equilibria by repeating the game finitely many times, it is necessary to start with a game having multiple equilibrium payoffs.

Lemma 3.3.1.   *If $E_1' = E_1$ has exactly one point, then $E_n' = E_1'$ for all n.*

Proof. By the perfection requirement, the equilibrium strategy at the last stage leads to the same payoff, whatever the history, and hence backwards induction gives the result. $\square$

Moreover, a dimension condition is also needed, as the following example due to Benoit and Krishna (1985) shows. Player 1 chooses the row, player 2 the column and player 3 the matrix, with payoffs as follows:

$$\begin{pmatrix} (3,3,3) & (0,0,0) \\ (0,0,0) & (0,0,0) \\ (0,1,1) & (0,0,0) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} (1,1,1) & (2,2,2) \\ (0,1,1) & (0,1,1) \\ (0,1,1) & (0,0,0) \end{pmatrix}.$$

One has $V = (0,0,0)$; $(2,2,2)$ and $(3,3,3)$ are in E_1 but players 2 and 3 have the same payoffs. Let w_n be the worst subgame perfect equilibrium payoff for them in G_n. Then by induction $w_n \geq 1/2$ since for every strategy profile one of the two can, by deviating, get at least $1/2$. (If player 1 plays middle with probability less than $1/2$, player 2 plays left; otherwise, player 3 chooses right.) Hence E'_n remains far from E.

A general result concerning pure equilibria (with compact action spaces) is the following:

Theorem 3.3.2 [Benoit and Krishna (1985)]. *Assume that for each i there exists $e(i)$ and $f(i)$ in E_1 (or in some E_n) with $e^i(i) > f^i(i)$, and that E has a non-empty interior. Then E'_n converges to E.*

Proof. One proof can be constructed by mixing the ideas of the proofs in Subsections 2.3 and 3.2. Basically the set of stages is split into three phases; during the last phase, as in Subsection 2.3, cycles of $(e(1), \ldots, e(I))$ will be played. Hence no deviations will occur in phase 3 and one will be able to punish "late" deviations (i.e. in phase 2) of player i, say, by switching to $f(i)$ for the remaining stages. In order to take care of deviations that may occur before and to be able to decrease the payoff to V, a family of plans as in Subsection 3.2 is used. One first determines the length of the punishment phase, then the reward phase; this gives a bound on the duration of phase 2 and hence on the length of the last phase. Finally, one gets a lower bound on the number of stages to approximate the required payoff. $\square$

As in Subsection 3.2 more precise results hold for $I = 2$; see Benoit and Krishna (1985) or Krishna (1988).

An extension of this result to mixed strategies seems possible if public correlation is allowed. Otherwise the ideas of Theorem 3.2 may not apply, because the set of achievable payoffs in the finite game is not convex and hence future equalizing payoffs cannot be found.

3.4. The recursive structure

When studying subgame perfect equilibria (SPE for short) in G_λ, one can use the fact that after any history, the equilibrium conditions are similar to the initial ones, in order to get further results on E_λ while keeping λ fixed.

The first property arising from dynamic programming tools and using only the continuity in the payoffs due to the discount factor (and hence true in any multistage game with continuous payoffs) can be written as follows:

Proposition 3.4.1. *A strategy profile is a SPE in G_λ iff there is no one-stage profitable deviation.*

Proof. The condition is obviously necessary. Assume now that player i has a profitable deviation against the given strategy σ, say τ^i. Then there exists some integer N, such that θ^i defined as "play τ^i on histories of length less than N and σ^i otherwise", is still better than σ^i. Consider now the last stage of a history of length less than N, where the deviation from σ^i to θ^i increase i's payoff. It is then clear that to always play σ^i, except at that stage of this history where τ^i is played, is still a profitable deviation; hence the claim. $\square$

This criterion is useful to characterize all SPE payoffs.

We first need some notation. Given a bounded set F of $\mathbb{R}^I$, let $\Phi_\lambda(F)$ be the set of Nash equilibrium payoffs of all one-shot games with payoff $\lambda g + (1 - \lambda)f$, where f is any mapping from S to F.

Proposition 3.4.2. *E'_λ is the largest (in terms of set inclusion) bounded fixed point of Φ_λ.*

Proof. Assume first $F \subset \Phi_\lambda(F)$. Then, at each stage n, the future expected payoff given the history, say f_n in F, can be supported by an equilibrium leading to a present payoff according to g and some future payoff f_{n+1} in F. Let σ be the strategy defined by the above family of equilibria. It is clear that in G_λ σ yields the sequence f_n of payoffs, and hence by construction no one-stage deviation is profitable. Then, using the previous proposition, $\Phi_\lambda(F) \subset E'_\lambda$. On the other hand, the equilibrium condition for SPE implies $E'_\lambda \subset \Phi(E'_\lambda)$ and hence the result. $\square$

Along the same lines one has $E'_\lambda = \bigcap_n \Phi^n_\lambda(D')$ for any bounded set D' that contains D. These ideas can be extended to a much more general setup; see the following sections.

Note that when working with Nash equilibria the recursive structure is available only on the equilibrium path and that when dealing with G_n one loses the stationarity.

Restricting the analysis to pure strategies and using the compactness of the equilibrium set (strategies and payoffs) allows for nice representations of all pure SPE; see Abreu (1988). Tools similar to the following, introduced by Abreu, were in fact used in the previous section.

Given $(I + 1)$ plays $[h; h(i), i \in I]$, a *simple strategy profile* is defined by requiring the players to follow h and inductively to switch to $h(i)$ from stage $n + 1$ on, if the last deviation occurred at stage n and was due to player i.

Lemma 3.4.3. $[h(0); h(i), i \in I]$ *induces a SPE in G_λ iff for all $j = 0, \ldots, I$,* $[h(j); h(i), i \in I]$ *defines an equilibrium in G_λ.*

Proof. The condition is obviously necessary and sufficiency comes from Proposition 3.4.1. $\square$

Define $\sigma(i)$ as the pure SPE leading to the worst payoff for i in G_λ and denote by $h^*(i)$ the corresponding cooperative play.

Lemma 3.4.4. $[h^*(j); h^*(i), i \in I]$ *induces a SPE.*

Proof. Since $h^*(j)$ corresponds to a SPE, no deviation [leading, by $\sigma(j)$, to some other SPE] is profitable a fortiori if it is followed by the worst SPE payoff for the deviator. Hence the claim by the previous lemma. $\square$

We then obtain:

Theorem 3.4.5 [Abreu (1988)]. *Let σ be a pure SPE in G_λ and h be the corresponding play. Then $[h; h^*(i), i \in I]$ is a pure SPE leading to the same play.*

These results show that extremely simple strategies are sufficient to represent all pure SPE; only $(I + 1)$ plays are relevant and the punishments depend only on the deviator, not on his action or on the stage.

3.5. Final comments

In a sense it appears that to get robust results that do not depend on the exact specification of the length of the game (assumed finite or with finite mean), the approach using the limit game is more useful. Note nevertheless that the counterpart of an "equilibrium" in G_∞ is an ε-equilibrium in the finite or discounted game (see also Subsection 7.1.1). The same phenomena of "discontinuity" occur in stochastic games (see the chapter on 'stochastic games' in a forthcoming volume of this Handbook) and even in the zero-sum case for games with incomplete information (Chapter 5 in this Handbook).

4. Correlated and communication equilibria

We now consider the more general situation where the players can observe signals. In the framework of repeated games (or multimove games) several such extensions are possible depending on whether the signals are given once or at each stage, and whether their law is controlled by the players or not. These mechanisms increase the set of equilibrium payoffs, but under the hypothesis of full monitoring and complete information lead to the same results. (Compare with Chapter 6 in this Handbook.)

Recall that given a normal form game $\Gamma = (\Sigma, \varphi)$ and a correlation device $C = (\Omega, \mathcal{A}, P; \mathcal{A}^i)$, $i \in I$, consisting of a probability space and sub σ-algebras of $\mathcal{A}$, a correlated equilibrium is an equilibrium of the extended game Γ_C having as strategies, say μ^i for i, $\mathcal{A}^i$-measurable mappings from Ω to Σ^i, and as payoff $\Phi(\mu) = \int \varphi(\mu(\omega)) \, P(d\omega)$. In words, ω is chosen according to P and $\mathcal{A}^i$ is i's information structure. Similarly, in a multimove game the notion of an extensive form correlated equilibrium can be defined with the help of private filtrations, say $\mathcal{A}_n^i$ for player i – i.e. there is new information on ω at each stage – and by requiring μ_n^i to be $\mathcal{A}_n^i \otimes \mathcal{H}_n$ measurable on $\Omega \times H_n$. Finally, for communication equilibria [see Forges (1986)], the probability induced by P on $\mathcal{A}_{n+1}$ is $\mathcal{A}_n^i \otimes \mathcal{H}_n$ measurable, i.e. the law of the signal at each stage depends on the past history, including the moves of the players.

Let us consider repeated games with a correlation device (resp. extensive correlation device; communication device). We first remark that the set of feasible payoffs is the same in any extended game and hence the analog of Proposition 1.3 holds.

For any of these classes we consider the union of the sets of equilibrium payoffs when the device varies and we shall denote it by cE_∞, CE_∞ and KE_∞, respectively. It is clear that the main difference from the previous analysis (without information scheme) comes from the threat point, since now any player can have his payoff reduced to $w^i = \min_{Y^{-i}} \max_{X^i} g^i(x^i, y^{-i})$, where Y^{-i} stands for the probabilities on S^{-i} (correlated moves of the opponent to i) and this set is strictly larger than X^{-i} for more than two players. Hence the new threat point W will usually differ from V and the set to consider will be $CE = \{d \in D: \forall i \in I, d^i \geq w^i\}$. Then one shows easily that $cE_\infty = CE_\infty = KE_\infty = CE$.

There is a deep relationship between these concepts and repeated games (or multimove games) in the sense that given a strategy profile σ, $C_n = (H_n, \mathcal{H}_n, P_\sigma)$ is a correlation device at stage n (where in the framework of Sections 1–3, the private σ-algebra is $\mathcal{H}_n$ for all players). This was first explicitly used in games with incomplete information when constructing a *jointly controlled lottery* [see Aumann, Maschler and Stearns (1968)]. For extensions of these tools under partial monitoring, see the next section.

5. Partial monitoring

Only partial results are available when one drops the assumption of full monitoring, namely that after each stage all players are told (or can observe) the previous moves of their opponents. In fact the first models in this direction are due to Radner and Rubinstein and also incorporate some randomness in the payoffs (moral hazard problems). We shall first cover results along these lines. Basically one looks for sufficient conditions to get results similar to the Folk Theorem or for Pareto payoffs to be achievable. In a second part we will present recent results of Lehrer, where the structure of the game is basically as in Section 1 except for the signalling function, and one looks for a characterization of E_∞ in terms of the one-stage game data.

5.1. Partial monitoring and random payoffs (see also the chapter on 'principal-agent models' in a forthcoming volume of this Handbook)

5.1.1. One-sided moral hazard

The basic model arises from *principal–agent* situations and can be represented as follows. Two players play sequentially; the first player (principal) chooses a *reward function* and then with that knowledge the second player (agent) chooses a move. The outcome is random but becomes *common knowledge* and depends only on the choice of player 2, which player 1 does not know. Formally, let Ω be the set of outcomes. The actions of player 1 are measurable mappings from Ω to some set S. Denote by T the actions set of player 2 and by Q_t the corresponding probabilities on Ω. The payoff functions are real continuous bounded measurable mappings, f on $\Omega \times S$ for player 1 and g on $\Omega \times S \times T$ for player 2. Assume, moreover, some *revelation condition* (RC), namely that there exists some positive constant K such that, for all positive ε, if $E_{s,t}g \geq E_{s,t'}g + \varepsilon$, then $|\int \omega \, dQ_t - \int \omega \, dQ_{t'}| \geq K\varepsilon$. In words, this means that profitable deviations of player 2 generate a different distribution of outcomes.

It is easy to see that generically one-shot Nash equilibria are not efficient in such games. The interest of repetition is then made clear by the following result:

Theorem 5.1.1 [Radner (1981)]. *Assume that a feasible payoff d strictly dominates a one-shot Nash equilibrium payoff e. Then $d \in E_\infty$.*

Proof. The idea of the proof is to require both players to use the strategy combination leading to d, as in the Folk Theorem. A deviation from player 1 is observable and one then requires that both players switch to the equilibrium

payoff e. The main difficulty arises from the fact that the deviations of player 2 are typically non-observable (even if he is using a pure strategy the Q_t may have non-disjoint support). Both players have to use some statistical test, based for example on the law of large numbers, to check with probability one whether player 2 was playing a profitable deviation, using RC. In such a case they again both switch to e. □

By requiring the above punishment to last for a finite number of stages (adapted to the precision of the test), one may even obtain a form of "subgame perfection" (note that there are no subgames, but one may ask for an equilibrium condition given any common knowledge history); see again Radner (1981).

Similar results with alternative economic content have been obtained by Rubinstein (1979a, 1979b) and Rubinstein and Yaari (1983).

Going back to the previous model, it can also be shown [Radner (1985)] that the modified strategies described above lead to an equilibrium in G_λ if the discount factor is small enough, and that they approach the initial payoff d. A similar remark about perfection applies and hence formally the following holds:

Theorem 5.1.2.   *Let d be feasible, $e \in E_1$, and assume $d \geqslant e$. Then for all $\varepsilon > 0$ there exists λ^* such that for all $\lambda \leqslant \lambda^*$, d is ε-close to E'_λ.*

Other classes of strategies with related properties have been introduced and studied by Radner (1986c).

5.1.2. Two sided moral hazard

A model where both players have private information on the history has been introduced and studied by Radner (1986a) under the name *partnership game*. Here the players are simultaneously choosing moves in some sets S and T. The outcome is again random with some law Q_{st}. At each stage the information of each player consists of his move and of the outcome; moreover, his own stage payoff depends only on this information and the revelation condition is still required. Then the analogy of the previous Theorem 5.1.1 holds [Radner (1986a)]. Here also the construction of the strategies is based on some statistical test and uses review and punishment phases.

Nevertheless, if one studies G_λ the previous arguments are no longer valid. More precisely, since none of the moves is observable it may be worthwhile for one player to deviate from the prescribed strategy when the sequence of records of outcomes starts to differ significantly from the mean and to try to "correct" it in order to avoid the punishment phase. (Note that when the

payoff is not discounted, by the strong law of large numbers, there is no gain in doing so.)

In fact, an example of a partnership game due to Radner, Myerson and Maskin (1986) shows that E_λ may be uniformly (in λ) bounded away from the Pareto boundary. Schematically, the payoffs depend upon an outcome that may be good or bad and the game is symmetrical. If, at equilibrium, the future payoff is independent of the outcome one obtains only one-shot equilibrium payoffs. Thus, this future payoff has to be discriminating (higher for a good outcome than for a bad) and hence cannot be Pareto optimal in expectation. (See also the example in the next section.)

5.1.3. Public signals and recursive structure

We now turn to results that are not based on the use of statistical tests but rather on the recursive structure.

A first model due to Abreu, Pearce and Stachetti (1986, 1990) considers an oligopoly with compact pure strategy sets where the I firms are only told, after each stage, the price, which is a random function of the moves with a fixed support. One can see that in this case Nash and "subgame perfect" equilibria coincide and, moreover, the recursive properties still hold. This allows us to give a nice description of the set of equilibrium payoffs by using its extreme points.

Finally, in a recent work, Fudenberg, Levine and Maskin (1989) succeed in getting a theorem analogous to Theorem 3.2 in the following framework. Consider a game where after each stage each player gets some private information on a random signal depending on the moves of all players at that stage. We call public those strategies that depend only on events known to all players. Note first that an equilibrium in the discounted game restricted to public strategies is an equilibrium in the original game (given a best reply to public strategies, taking its conditional expectation on public events, is still a best reply) and that one can define "subgame perfect public equilibria" by introducing subgames related to public events. The tools of Subsection 3.4 are then applicable, and sufficient conditions are given, basically on the independence of the conditional laws of the signals as function of the moves of each player – the strategy of the others being fixed – to ensure that the corresponding set of payoffs converges to E as λ goes to 0. More precisely, it is shown that a smooth convex set F of payoffs included in E and at a small Hausdorff distance from it satisfies $F \subset \Phi_\lambda(F)$ for λ small enough. The main difficulty is to check the inclusion on extreme points. In fact, the above conditions allow us to compute explicitly the future payoffs by solving linear equations.

Note that here a dimension condition is needed, even in the two-player case. Let us consider the following game, due to Fudenberg and Maskin (1986b).

The payoff matrix is

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

the moves of player 2 are announced and a public signal with values (α, β) has the following distribution:

$$\begin{pmatrix} (3/4, 1/4) & (1/2, 1/2) \\ (1/2, 1/2) & (1/4, 3/4) \end{pmatrix}.$$

Then $(1, 1)$ is the only point in E'_λ. In fact, denote by w the worst SPE, by s and t the corresponding random moves of both players at stage 1, and by $w_{L\alpha} (\geq w)$ the expected payoff after Left and α, and so on. We note first that if $s = 1$ one has $w \geq \lambda + (1 - \lambda)(3/4 w_{L\alpha} + 1/4 w_{L\beta})$, and hence $w \geq 1$. Otherwise one has:

$$w = t((1 - \lambda)(1/2 w_{L\alpha} + 1/2 w_{L\beta}))$$
$$+ (1 - t)(-\lambda + (1 - \lambda)(1/4 w_{R\alpha} + 3/4 w_{R\beta}))$$

$$\geq t(\lambda + (1 - \lambda)(3/4 w_{L\alpha} + 1/4 w_{L\beta}))$$
$$+ (1 - t)((1 - \lambda)(1/2 w_{R\alpha} + 1/2 w_{R\beta})),$$

so that $t(1 - \lambda)w_{L\beta} + (1 - t)(1 - \lambda)w_{R\beta} \geq 4\lambda + (1 - \lambda)w$. Substituting this into the first equality yields $w \geq t + 1$.

Note that 0 is a subgame perfect public equilibrium payoff in G_∞ (even if 2's moves are not announced) by asking the players to use their dominated move at each stage where the empirical past frequency of α is greater than 1/2. [Compare with Sorin (1986b).]

On the other hand, 0 can be obtained as a perfect equilibrium in G_λ if 1's moves are observable by asking him to follow a history consisting of a sequence of 1 and -1 inducing a payoff increasing to 0 and playing again the same move in the case of a deviation from -1 to 0. In the previous framework player 1 could pretend to punish even if he did not and hence the punishment was not credible and player 2 would deviate.

It is important to remark that in these games the signals can be used as a correlation device or an extensive correlation device (recall Section 4 and see also Subsection 5.2). In particular, the set of equilibria can be larger than the set of public equilibria and can contain payoffs that are not i.r. (but in CE), if for example a subgroup of players get some common signal, unknown to the

others. (But if there are two players and one is more informed than the other one can always assume public strategies.)

Finally, similar results are used in the framework of games with long-run and short-run players [Fudenberg and Levine (1989b)].

5.2. Signalling functions

The results of this section are due mainly to Lehrer. We consider the infinitely repeated game G_∞ of Section 1, but after each stage n, each player i is only told $q_n^i = Q^i(s_n)$, s_n being the *I*-action at that stage and Q^i being i's signalling (deterministic) function, defined on S with values in some set Q. Each player's strategy is then required to be measurable with respect to his private information. Hence a pure strategy σ_n^i is a mapping from sequences $(q_1^i, \ldots, q_{n-1}^i)$ to S^i and perfect recall is assumed.

Let us first consider the case of two players and a general signalling function (we shall assume in this section non-trivial information, namely that each player may, by playing some move, get some information about his opponent's move, so that the players can communicate through their actions – the other case is much simpler to analyze).

It is easy to see that, since the signals are not common knowledge, equilibrium strategies do not induce, after finitely many stages, an equilibrium in the remaining game but rather a correlated equilibrium (see Section 4).

One is thus led to consider extensive form correlated equilibria and in fact these are much easier to characterize.

We first define two relations on actions by

$$s^i \sim t^i \Leftrightarrow Q^{-i}(t^i, s^{-i}) = Q^{-i}(s^i, s^{-i}) \quad \text{for all } s^{-i}$$

(in words, in a one-shot game player $-i$ has no way to distinguish whether player i is playing s^i or t^i); and

$$s^i > t^i \Leftrightarrow s^i \sim t^i \text{ and } Q^i(t^i, s^{-i}) \neq Q^i(t^i, t^{-i}) \text{ implies}$$

$$Q^i(s^i, s^{-i}) \neq Q^i(s^i, t^{-i}) \quad \text{for all } s^{-i}, t^{-i}$$

(player i gets more information on $-i$'s move by playing s^i than t^i).

The crucial point is that player i can mimic a pure strategy, say τ^i, by any other σ^i with $\sigma^i(h) > \tau^i(h)$, for all h, without being detected. [Inductively, at each stage n he uses an action $s_n^i > \tau^i(h)$, h being the history that would have occurred had he used $\{t_m^i\}$, $m < n$, up to now.]

Let P be the set of probabilities on S (correlated moves). The set of equilibrium payoffs will be characterized through the following sets (note that, as in the Folk Theorem, they depend only on the one-shot game):

$$A^i = \{p \in P: \sum_{s^{-i}} p(s^i, s^{-i})g^i(s^i, s^{-i}) \geq \sum_{s^{-i}} p(s^i, s^{-i})g^i(t^i, s^{-i}) \text{ for all } s^i$$

$$\text{and all } t^i \text{ with } t^i > s^i\} .$$

$$B^i = A^i \cap X = \{x \in X: g^i(s^i, x^{-i}) \geq g^i(t^i, x^{-i}) \text{ for all } s^i, t^i \text{ with } x^i(s^i) > 0$$

$$\text{and } t^i > s^i\} .$$

Write IR for the set of i.r. payoffs and E_∞ (resp. cE_∞, CE_∞, KE_∞) for the set of Nash (resp. correlated, extensive form correlated, communication) equilibrium payoffs in the sense of upper, $\mathcal{L}$ or uniform. lE_∞ and lCE_∞ will denote lower equilibrium payoffs [recall paragraph (iii) in Section 1].

Theorem 5.2.1 [Lehrer (1992a)].   (i) $cE_\infty = CE_\infty = KE_\infty = g(\bigcap_i A^i) \cap IR$. (ii) $lCE_\infty = \bigcap_i g(A^i) \cap IR$.

Proof. The proof of this result (and of the following) is quite involved and introduces new and promising ideas. Only a few hints will be presented here.

For (ii), the inclusion from left to right is due to the fact that given correlated strategies, each player can modify his behavior in a non-revealing way to force the correlated moves at each stage to belong to A^i.

Similarly, for the corresponding inclusion in (i) one obtains by convexity that if a payoff does not belong to the right-hand set, one player can profitably deviate on a set of stages with positive density. To prove the opposite inclusion in (i) consider p in $\bigcap_i A^i$. We define a probability on histories by a product $\otimes p_n$; each player is told his own sequence of moves and is requested to follow it. p_n is a perturbation of p, converging to p as $n \to \infty$, such that each I-move has a positive probability and independently each recommended move to one player is announced with a positive probability to his opponent. It follows that a profitable deviation, say from the recommended s^i to t^i, will eventually be detected if $t^i \not\sim s^i$. To control the other deviations ($t^i \sim s^i$ but $t^i \not> s^i$), note first that, since the players can communicate through their moves, one can define a code, i.e. a mapping from histories to messages. The correlated device can then be used to generate, at infinitely many fixed stages, say n_k, random times m_k in (n_{k-1}, n_k): at the stages following n_k the players use a finite code to report the signal they got at time m_k. In this case also a deviation, if used with

a positive density, will eventually occur at some stage m_k where moreover the opponent is playing a revealing move and hence will be detected. Obviously from then on the deviator is punished to his minimax. To obtain the same result for correlated equilibria, let the players use their moves as signals to generate themselves the random times m_k [see Sorin (1990)].

Finally, the last inclusion in (ii) follows from the next result. $\quad\square$

Theorem 5.2.2 [Lehrer (1989)]. $\quad IE_\infty = \bigcap_i Co\ g(B^i) \cap IR (= ICE_\infty)$.

Proof. It is easy to see that $Co\ g(B^i) = g(A^i)$ and hence a first inclusion by part (ii) of the previous theorem. To obtain the other direction let us approximate the reference payoff by playing on larger and larger blocks M_k, cycles consisting of extreme points in B^i [if $k \equiv i$ (mod 2)]. On each block, alternatively, one of the players is then playing a sequence of pure moves; thus a procedure like in the previous proof can be used. $\quad\square$

A simpler framework in which the results can be extended to more than two players is the following: each action set S^i is equipped with a partition $\underline{S}^i$ and each player is informed only about the elements of the partitions to which the other players' actions belong. Note that in this case the signal received by a player is independent of his identity and of his own move. The above sets B^i can now be written as

$$C^i = \{x \in X: g^i(x) \geq g(x^{-i}, y^i)\ \text{for all}\ y^i\ \text{with}\ \underline{y}^i = \underline{x}^i\}$$

where $\underline{x}^i$ is the probability induced by x^i on $\underline{S}^i$.

Theorem 5.2.3 [Lehrer (1990)]. (i) $E_\infty = Co\ g(\bigcap_i C^i) \cap IR$. (ii) $IE_\infty = \bigcap_i Co\ g(C^i) \cap IR$

Proof. It already follows in this case that the two sets may differ. On the other hand, they increase as the partitions get finer (the deviations are easier to detect) leading to the Folk Theorem for discrete partitions – full monitoring.

For (ii), given a strategy profile σ, note that at each stage n, conditional to $h_n = (x_1, \ldots, x_{n-1})$, the choices of the players are independent and hence each player i can force the payoff to be in $g(C^i)$; hence the inclusion of IE_∞ in the right-hand set. On the other hand, as in Theorem 5.2.2, by playing alternately in large blocks to reach extreme points in C^1, then $C^2, \ldots$, one can construct the required equilibrium.

As for E_∞, by convexity if a payoff does not belong to the right-hand set, there is for some i a set of stages with positive density where, with positive

probability, the expected move profiles, conditioned on h_n, are not in C^i. Since h_n is common knowledge, player i can profitably deviate.

To obtain an equilibrium one constructs a sequence of increasing blocks on each of which the players are requested to play alternately the right strategies in $\bigcap_i C^i$ to approach the convex hull of the payoffs. These strategies may induce random signals so that the players use some statistical test to punish during the following block if some deviation appears. $\square$

For the extension to correlated equilibria, see Naudé (1990).

Finally a complete characterization is available when the signals include the payoffs:

Theorem 5.2.4 [Lehrer (1992b)].   *If $g^i(s) \neq g^i(t)$ implies $Q^i(s) \neq Q^i(t)$ for all i, s, t, then $E_\infty = lE_\infty = Co \; g(\bigcap_i C^i) \cap IR$.*

Proof.   To obtain this result we first prove that the signalling structure implies $\bigcap_i Co \; g(B^i) \cap IR = Co \; g(\bigcap_i B^i) \cap IR$. Then one uses the structure of the extreme points of this set to construct equilibrium strategies. Basically, one player is required to play a pure strategy and can be monitored as in the proof of Theorem 5.2.1(i); the other player's behavior is controlled through some statistical test. $\square$

While it is clear that the above ideas will be useful in getting a general formula for E_∞, this one is still not available. For results in this direction, see Lehrer (1991, 1992b).

When dealing with more than two players new difficulties arise since a deviation, even when detected by one player, has first to be attributed to the actual deviator and then this fact has to become common knowledge among the non-deviators to induce a punishment.

For non-atomic games results have been obtained by Kaneko (1982), Dubey and Kaneko (1984) and Masso and Rosenthal (1989).

6. Approachability and strong equilibria

In this section we review the basic works that deal with other equilibrium concepts.

6.1. Blackwell's theorem

The following results, due to Blackwell (1956), are of fundamental importance in many fields of game theory, including repeated games and games with

incomplete information. [A simple version will be presented here; for extensions see Mertens, Sorin and Zamir (1992).]

Consider a two-person game G_1 with finite action sets S and T and a random payoff function g on $S \times T$ with values in $\mathbb{R}^k$, having a finite second-order moment (write f for its expectation). We are looking for an extension of the minimax theorem to this framework in G_∞ (assuming full monitoring) and hence for conditions for a player to be able to approach a (closed) set C in $\mathbb{R}^k$ – namely to have a strategy such that the average payoff will remain, in expectation and with probability one, close to C, after a finite number of stages. C is excludable if the complement of some neighborhood of it is approachable by the opponent.

To state the result we introduce, for each mixed action x of player 1, $P(x) = Co\{f(x, t): t \in T\}$ and similarly $Q(y) = Co\{f(s, y): s \in S\}$ for each mixed action y of player 2.

Theorem 6.1.1 *Assume that, for each point $d \not\subseteq C$ there exists x such that if c is a closest point to d in C, the hyperplane orthogonal to $[cd]$ through c separates d from $P(x)$. Then C is approachable by player 1.*

An optimal strategy is to use at each stage n a mixed action having the above property, with $d = \bar{g}_{n-1}$.

Proof. This is proved by showing by induction that, if d_n denotes the distance from $\bar{g}_n$, the average payoff at stage n, to C, then $E(d_n^2)$ is bounded by some K/n. Furthermore, one constructs a positive supermartingale converging to zero, which majorizes d_n^2. $\square$

If the set C is convex we get a minimax theorem, due to the following:

Theorem 6.1.2. *A convex set C is either approachable or excludable; in the second case there exists y with $Q(y) \cap C = \emptyset$.*

Proof. Note that the following sketch of the proof shows that the result is actually stronger: if $Q(y) \cap C = \emptyset$ for some y, C is clearly excludable (by playing y i.i.d.). Otherwise, by looking at the game with real payoff $\langle d - c, f \rangle$, the minimax theorem implies that the condition for approachability in the previous theorem holds. $\square$

Blackwell also showed that Theorem 6.1.2 is true for any set in $\mathbb{R}$, but that there exist sets in $\mathbb{R}^2$ that are neither approachable nor excludable, leading to the problem of "weak approachability", recently solved by Vieille (1989) which showed that every set is asymptotically approachable or excludable by a family of strategies that depend on the length of the game. This is related to

the definitions of $\lim v_n$ and v_∞ in zero-sum games (see Chapter 5 and the chapter on "stochastic games" in a forthcoming volume of this Handbook).

6.2. Strong equilibria

As seen previously, the Folk Theorem relates non-cooperative behavior (Nash equilibria) in G_∞ to cooperative concepts (feasible and i.r. payoffs) in the one-shot game. One may try to obtain a smaller cooperative set in G_1, such as the Core, and to investigate what its counterpart in G_∞ would be. This problem has been proposed and solved in Aumann (1959) using his notion of strong equilibrium, i.e., a strategy profile such that no coalition can profitably deviate.

Theorem 6.2.1. *The strong equilibrium payoffs in G_∞ coincide with the β-Core of G_1.*

Proof. First, if d is a payoff in the β-Core, there exists some (correlated) action achieving it that the players are requested to play in G_∞. Now for each subset $I\backslash J$ of potential deviators, there exists a correlated action σ^J of their opponent that prevents them from obtaining more than $d^{I\backslash J}$, and this will be used as a punishment in the case of deviation.

   On the other hand, if d does not belong to the β-Core there exists a coalition J that possesses, given each history and each corresponding correlated move $I\backslash J$ tuple of its complement, a reply giving a better payoff to its members. $\square$

Note the similarity with the Folk Theorem, with the β-characteristic function here playing the role of the minimax (as opposed to the α-one and the maximin).

   If one works with games with perfect information, one has the counterpart of the classical result regarding the sufficiency of pure strategies:

Theorem 6.2.2 [Aumann (1961)]. *If G_1 has perfect information the strong equilibria of G_∞ can be obtained with pure strategies.*

Proof. The result, based on the convexity of the β-characteristic function and on Zermelo's theorem, emphasizes again the relationship between repetition and convexity. $\square$

   Finally, Mertens (1980) uses Blackwell's theorem to obtain the convexity and superadditivity of the β-characteristic function of G_1 by proving that it

coincides with the α-characteristic function (and also the β-characteristic function) of G_∞.

7. Bounded rationality and repetition

As we have already pointed out, repetition alone, when finite, may not be enough to give rise to cooperation (i.e., Nash equilibria and a fortiori subgame perfect equilibria of the repeated game may not achieve the Pareto boundary). On the other hand, empirical data as well as experiments have shown that some cooperation may occur in this context [for a comprehensive analysis, see Axelrod (1984)].

We will review here some models that are consistent with this phenomenon. Most of the discussion below will focus on the Prisoner's Dilemma but can be easily extended to any finite game.

7.1. Approximate rationality

7.1.1. ε-equilibria

The intuitive idea behind this concept is that deviations that induce a small gain can be ignored. More precisely, σ will be an ε-equilibrium in the repeated game if, given any history (or any history consistent with σ), no deviation will be more than ε-profitable in the remaining game [see Radner (1980, 1986b)]. Consider the Prisoner's Dilemma (cf. Subsection 2.3):

Theorem 7.1. $\forall \varepsilon > 0, \forall \delta > 0, \exists N$ *such that for all* $n \geqslant N$ *there exists an* ε-*equilibrium in* G_n *inducing a payoff within* δ *of the Pareto point* $(3, 3)$.

Proof. Define σ as playing cooperatively until the last N_0 stages (with $N_0 \geqslant 1/\varepsilon$), where both players defect. Moreover, each player defects forever as soon as the other does so once. It is easy to see that any defection will induce an (average) gain less than ε, and hence the result for N large enough. $\square$

The above view implicitly contains some approximate rationality in the behavior of the players (they neglect small mistakes).

7.1.2. Lack of common knowledge

This approach deals with games where there is lack of common knowledge on some specific data (strategy or payoff), but common knowledge of this

uncertainty. Then even if all players know the true data, the outcome may differ from the usual framework by a contamination effect – each player considers the information that the others may have.

The following analysis of repeated games is due to Neyman (1989). Consider again the finitely repeated Prisoner's Dilemma and assume that the length of the game is a random variable whose law P is common knowledge among the players. (We consider here a closed model, including common knowledge of rationality.) If P is the point mass at n we obtain G_n and "$E_n = \{1, 1\}$" is common knowledge. On the other hand, for any λ there exists P_λ such that the corresponding game is G_λ if the players get no information on the actual length of the game. Consider now non-symmetric situations and hence a general information scheme, i.e. a correlation device with a mapping $\omega \mapsto n(\omega)$ corresponding to the length of the game at ω.

Recall that an event A is of mutual knowledge of order k [say $mk(k)$] at ω if $K^{i_0} \circ \cdots \circ K^{i_k}(\omega) \subset A$, for all sequences $i_0, \ldots, i_k$, where K^i is the knowledge operator of player i (for simplicity, assume Ω is countable and then $K^i(B) = \cap \{C : B \subset C, C \text{ is } \mathscr{A}^i\text{-measurable}\}$; hence K^i is independent of P). Thus $mk(0)$ is public knowledge and $mk(\infty)$ common knowledge.

It is easy to see that at any ω where "$n(\omega)$" is $mk(k)$, $(1, 1)$ will be played during the last $k + 1$ stages [and this fact is even $mk(0)$], but Neyman has constructed an example where even if $n(\omega) = n$ is $mk(k)$ at ω, cooperation can occur during $n - k - 1$ stages, so that even with large k, the payoff converges to Pareto as $n \mapsto \infty$.

The inductive hierarchy of K^i at ω will eventually reach games with length larger than $n(\omega)$, where the strategy of the opponent justifies the initial sequence of cooperative moves.

Thus, replacing a closed model with common knowledge by a local one with large mutual knowledge leads to a much richer and very promising framework.

7.2. Restricted strategies

Another approach, initiated by Aumann, Kurz and Cave [see Aumann (1981)], requests the players to use subclasses of "simple" strategies, as in the next two subsections.

7.2.1. Finite automata

In this model the players are required to use strategies that can be implemented by finite automata. The formal description is as follows: A finite automaton (say for player i) is defined by a finite set of states K^i and two mappings, α from $K^i \times S^{-i}$ to K^i and β from K^i to S^i. α models the way the

internal memory or state is updated as a function of the old memory and of the previous moves of the opponents. β defines the move of the player as a function of his internal state. Note that given the state and β, the action of i is known, so it is not necessary to define α as a function of S^i.

To represent the play induced by an automaton, we need in addition to specify the initial state k_0^i. Then the actions are constructed inductively by $\alpha(k_0^i)$, $\alpha(\beta(k_0^i, s_1^{-i})) = \alpha(k_1^i)$, ...

Games where both players are using automata have been introduced by Neyman (1985) and Rubinstein (1986).

Define the size of an automaton as the cardinality of its set of states and denote by $G(\kappa)$ the game where each player i is using as pure strategies automata of size less than κ^i.

Consider again the n-stage Prisoner's Dilemma. It is straightforward to check that given Tit for Tat (start with the the cooperative move and then at each following stage use the move used by the opponent at the previous stage) for both players, the only profitable deviation is to defect at the last stage. Now if $\kappa^i < n$, none of the players can "count" until the last stage, so if the opponent plays stationary, any move actually played at the last stage has to be played before then. It follows that for $2 \leq \kappa^i < n$, Tit for Tat is an equilibrium in G_n. Actually a much stronger result is available:

Theorem 7.2.1 [Neyman (1985)]. *For each integer* m, $\exists N$ *such that* $n \geq N$ *and* $n^{1/m} \leq \kappa^i \leq n^m$ *implies the existence of a Nash equilibrium in* $G_n(\kappa^1, \kappa^2)$ *with payoff greater than* $3 - 1/m$ *for each player.*

Proof. Especially in large games, even if the memory of the players is much larger than the length of the game (namely polynomial), Pareto optimality is almost achievable.

The idea of the proof relies on the observation that the cardinality of the set of histories is an exponential function of the length of the game. It is now possible to "fill" all the memory states by requiring both players to remember "small" histories, i.e. by answering in a prespecified way after such histories (otherwise the opponent defects for ever) and then by playing cooperatively during the remaining stages. Note that no internal state will be available to count the stages and that cooperative play arises during most of the game. $\square$

It is easy to see that in this framework an analog of Theorem 2.3 is available.

Similar results using Turing machines have been obtained by Megiddo and Widgerson (1986); see also Zemel (1989).

The model introduced in Rubinstein (1986) is different and we shall discuss the related version of Abreu and Rubinstein (1988). Both players are required to use finite automata (and no mixture is allowed) but there is no fixed bound

on the memory. The main change is in the preference function, which is strictly increasing in the payoff and strictly decreasing in the size [in Rubinstein (1986) some lexicographic order is used]. A complete structure of the corresponding set of equilibria is then obtained with the following striking aspect: $\kappa^1 = \kappa^2$; moreover, during the cycle induced by the automata each state is used only once; and finally both players change their moves simultaneously. In particular, this implies that in 2×2 two-person games the equilibrium payoffs have to lie on the "diagonals".

Considering now two-person, zero-sum games, an interesting question is to determine the worth of having a memory much larger than the memory of the other player: note that the payoff in $G_\infty(\kappa^1, \kappa^2)$ is well defined, hence also its value $V(\kappa', \kappa^2)$. Denote by V the value of the original G_1 and by $\bar{V}$ the minimax in pure strategies. This problem has been solved by Ben Porath (1986):

Theorem 7.2.2. *For any polynomial P, $\lim_{\kappa^2 \to \infty} V(P(\kappa^2), \kappa^2) = V$. There exists some exponential function Ψ such that $\lim_{\kappa^2 \to \infty} V(\Psi(\kappa^2), \kappa^2) = \bar{V}$.*

Proof. The second part is not difficult to prove, player 1 can identify player 2's automaton within $\Psi(\kappa^2)$ stages.

For the first part, player 2 uses an optimal strategy in the one-shot game to generate κ^2 random moves and then follows the corresponding distribution to choose an automaton generating these moves. The key point is, using large deviation tools, to show that the probability, with this procedure, of producing a sequence of κ^2 pairs of moves biased by more than ε is some exponential function, ψ, of $-\kappa^2 \cdot \varepsilon^2$. Since player 1 can have at most κ^1 different behaviors, the average payoff will be greater than $V + \varepsilon$ with a probability less than $P(\kappa^2)\psi(-\kappa^2 \cdot \varepsilon^2)$. □

7.2.2. Strategies with bounded recall

Another way to approach bounded rationality is to assume that players have bounded recall. Two classes of strategies can be introduced according to the following definitions: σ^i is of I- (resp. II)-bounded recall (BR) of size k if, for all histories h, $\sigma^i(h)$ depends only upon the last k components of h (resp. the last k moves of player $-i$).

It is easy to see that Tit for Tat can be implemented by a II-bounded recall strategy with $k = 1$; to punish forever after a deviation can be reached by a I-BR but not by a II-BR, and to punish forever after two deviations cannot be achieved with BR strategies (if the first deviation occurred a long time ago, the player will not remember it). Note nevertheless that with II-BR strategies the players can maintain the average frequency of deviations as low as required.

Using I-BR strategies Lehrer (1988) proves a result similar to Theorem 7.2.2

by using tools from information theory. (Note that in both cases player 1 does not need to know the moves of player 2.)

This area is currently very active and new results include the study of the complexity of a strategy and its relation with the size of an equivalent automaton [Kalai and Stanford (1988)], an analog of Theorems 3.1 and 3.2 in pure strategies for finite automata [Ben Porath and Peleg (1987)], and the works of Ben Porath, Gilboa, Kalai, Megiddo, Samet, Stearns and others on complexity. For a recent survey, see Kalai (1990).

To end these two subsections one should also mention the work of Smale (1980) on the Prisonner's Dilemma, in which the players are restricted to strategies where the actions at each stage depend continuously on some vector-valued parameter. The analysis is then performed in relation to dynamical systems.

7.3. Pareto optimality and perturbed games

The previous results, as well as sections 2 and 3, have shown that under quite general conditions a kind of Folk Theorem emerges; rationality and repetition enables cooperation. Note nevertheless that the previous procedures lead to a huge set of equilibrium payoffs (including all one-shot Nash equilibrium payoffs and even the threat point V). A natural and serious question was then to ask under which conditions would long-term interaction and utility maximizing behavior lead to cooperation; in other words, whether we would necessarily achieve Pareto points as equilibrium payoffs.

It is clear from the previous results that repetition is necessary and that complete rationality or bounded rationality alone would not be sufficient. In fact, one more ingredient – perturbation or uncertainty – is needed. Note that a similar approach was initiated by Selten (1975) in his work on perfect equilibria.

A first result in this direction was obtained in a very stimulating paper by Kreps, Milgrom, Roberts and Wilson (1982). Consider the finitely repeated Prisoner's Dilemma and assume that with some arbitrarily small but positive probability one of the players is a kind of automaton: he always uses Tit for Tat rather than maximizing. Then for sufficiently long games all the sequential equilibrium payoffs will be close to the cooperative outcome. The proof relies in particular on the following two facts: first, if the equilibrium strategies were non-cooperative, the perturbed player may play Tit for Tat thus pretending to be the automaton and thereby convincing his opponent that this is in fact the case; second, Tit for Tat induces payoffs that are close to the diagonal.

These suggestive and important ideas will be needed when trying to extend this result by dropping some of the conditions. The above result in fact

depends crucially on Tit for Tat (inducing itself almost the cooperative outcome as the best reply) being the only perturbation. More precisely a result of Fudenberg and Maskin (1986a) indicates that by choosing the perturbation in an adequate way the set of sequential equilibrium payoffs of a sufficiently long but finitely repeated game would approach any prespecified payoff. Now if all perturbations are allowed, each of the players may pretend to be a different automaton, advantageous from his own point of view.

One is thus lead to consider two-person games with common interest: one payoff strongly Pareto dominates all the others. Assume then that each player's strategy is ε-perturbed by some probability distribution having as support the set of II-BR strategies of some size k. Then the associated repeated game possesses equilibria in pure strategies and all the corresponding payoffs are close to the cooperative (Pareto) outcome $P(G)$. Formally, if pE_n^ε(resp. pE_λ^ε) denotes the set of pure equilibria payoffs in the n-stage (resp. λ-discounted) perturbed game, one has:

Theorem 7.3 [Aumann and Sorin (1989)]. $\lim_{\varepsilon\to 0} \lim_{n\to\infty} pE_n^\varepsilon = \lim_{\varepsilon\to 0} \lim_{\lambda\to 0} pE_\lambda^\varepsilon = P(G)$.

Proof.   To prove the existence of a pure equilibrium, one considers Pareto points in the payoff space generated by pure strategies in the perturbed game. One then shows that these are sustained by equilibrium strategies.

Now assuming the equilibrium to be not optimal, one player could deviate and mimic his best BR perturbation. Note that the corresponding history has positive probability under the initial strategies. Moreover, for n large enough (or λ small enough) a best reply on histories inconsistent with the "main" strategy is to identify the BR strategy used and then to maximize against it. For this to hold it is crucial to use II-BR perturbations: the moves used during this identification phase will eventually be forgiven and hence no punishment forever can arise. Finally, the game being with common interest a high payoff for one player implies the same for the other so that the above procedure would lead to a payoff close to the cooperative outcome; hence the contradiction.   □

The crucial properties of the set of perturbations used in the proof are: (1) identifiability (each player has a strategy such that, after finitely many stages he can predict the behavior of his opponent, if this opponent is in the perturbed mode); (2) the asymptotic payoff corresponding to a best reply to a perturbation is history independent. [For example, irreducible automata could be used; see Gilboa and Samet (1989).]

The extension to more than two players requires new tools since, even with bounded recall, two players can build everlasting events (e.g. punish during two stages if the other did so at the previous stage).

To avoid non-Pareto mixed equilibria one has to ask for some kind of perfection (or equivalently more perturbation) to avoid events of common knowledge of rationality (i.e. histories in which the probability of facing an opponent who is in the perturbed mode is 0 and common knowledge).

More recently, similar results, when a long-run player can build a reputation leading to Pareto payoffs against a sequence of short-run opponents, have been obtained by Fudenberg and Levine (1989a).

8. Concluding remarks

Before ending let us mention a connected field, multimove games, where similar features (especially the recursive structure) can be observed (and in fact were sometimes analyzed previously in specific examples). In this class of games the strategy sets have the same structure as in repeated games but the payoff is defined only on the set of plays and does not necessarily come from a stage payoff. A nice sampling can be founded in *Contributions to the Theory of Games, Vol. III* [Dresher, Tucker and Wolfe (1957)], and deals mainly with two-person games.

A game with two-move information lag was extensively studied by Dubins (1957), Karlin (1957), Ferguson (1967) and others, introducing new ideas and tools. The case with three-move information lag is still open. A general formulation and basic properties of games with information lag can be found in Scarf and Shapley (1957). A deep analysis of games of survival (or ruin) in the general case can be found in Milnor and Shapley (1957), using some related works of Everett (1957) on "recursive games". [For some results in the non-zero-sum case and ideas of the difficulties there, see Rosenthal and Rubinstein (1984).]

The properties of multimove games with perfect information are studied in Chapter 3 of this Handbook. The extension of those to general games seems very difficult [see, for example the very elegant proof of Blackwell (1969) for $\mathcal{G}_\delta$ games] and many problems are still open.

To conclude, we make two observations.

The first is that it is quite difficult to draw a well-defined frontier for the field of repeated games. Games with random payoffs are related to stochastic games; games with partial monitoring, as well as perturbed games, are related to games with incomplete information; sequential bargaining problems and games with multiple opponents are very close, ... To get a full overview of the field the reader should also consult Chapters 5, 6 and 7, and the chapter on 'stochastic games' in a forthcoming volume of this Handbook.

The second comment is that not only has the domain been very active in the last twenty years but that it is still extremely attractive. The numerous recent ideas and results allow us to unify the field and a global approach seems

conceivable [see the nice survey of Mertens (1987)]. Moreover, many concepts that are now of fundamental importance in other areas originate from repeated games problems (like selection of equilibria, plans, signals and threats, approachability, reputation, bounded complexity, and so on). In particular, the applications to economics (see, for example, Chapters 7, 8, 9, 10 and 11 in this Handbook) as well as to biology (see the chapter on 'biological games' in a forthcoming volume of this Handbook) have been very successful.

Bibliography

Abreu, D. (1986) 'Extremal equilibria of oligopolistic supergames', *Journal of Economic Theory*, 39: 191–225.

Abreu, D. (1988) 'On theory of infinitely repeated games with discounting', *Econometrica*, 56: 383–396.

Abreu, D. and A. Rubinstein (1988) 'The structure of Nash equilibria in repeated games with finite automata', *Econometrica*, 56: 1259–1282.

Abreu, D., D. Pearce and E. Stacchetti (1986) 'Optimal cartel equilibria with imperfect monitoring', *Journal of Economic Theory*, 39: 251–269.

Abreu, D., D. Pearce and E. Stacchetti (1990) 'Toward a theory of discounted repeated games with imperfect monitoring', *Econometrica*, 58: 1041–1063.

Aumann, R.J. (1959) 'Acceptable points in general cooperative *n*-person games', in: A.W. Tucker and R. Luce, eds., *Contributions to the theory of games*, Vol. IV, A.M.S. 40. Princeton: Princeton University Press, pp. 287–324.

Aumann, R.J. (1960) 'Acceptable points in games of perfect information', *Pacific Journal of Mathematics*, 10: 381–387.

Aumann, R.J. (1961) 'The core of a cooperative game without side payments', *Transactions of the American Mathematical Society* 98: 539–552.

Aumann, R.J. (1967) 'A survey of cooperative games without side payments', in: M. Shubik, ed., *Essays in mathematical economics in honor of Oskar Morgenstern*. Princeton: Princeton University Press, pp. 3–27.

Aumann, R.J. (1981) 'Survey of repeated games', *Essays in game theory and mathematical economics in honor of Oskar Morgenstern*. Mannheim: Bibliographisches Institüt, pp. 11–42.

Aumann, R.J. (1986) 'Repeated games', in: G.R. Feiwel, ed., *Issues in contemporary microeconomics and welfare*. London: Macmillan, pp. 209–242.

Aumann, R.J. and L.S. Shapley (1976) 'Long-term competition – A game theoretic analysis', preprint.

Aumann, R.J. and S. Sorin (1989) 'Cooperation and bounded recall', *Games and Economic Behavior*, 1: 5–39.

Aumann, R.J., M. Maschler and R. Stearns (1968) 'Repeated games of incomplete information: An approach to the non-zero sum case', *Report to the U.S. A.C.D.A. ST-143*, Chapter IV, pp. 117–216, prepared by Mathematica.

Axelrod, R. (1984) *The evolution of cooperation*. New York: Basic Books.

Benoit, J.P. and V. Krishna (1985) 'Finitely repeated games', *Econometrica*, 53: 905–922.

Benoit, J.P. and V. Krishna (1987) 'Nash equilibria of finitely repeated games', *International Journal of Game Theory*, 16: 197–204.

Ben-Porath, E. (1986) 'Repeated games with finite automata', preprint.

Ben-Porath, E. and B. Peleg (1987) 'On the folk theorem and finite automata', preprint.

Blackwell, D. (1956) 'An analog of the minimax theorem for vector payoff's', *Pacific Journal of Mathematics*, 6: 1–8.

Blackwell, D. (1969) 'Infinite G_δ games with imperfect information', *Applicationes Mathematicae* X: 99–101.

Cave, J. (1987) 'Equilibrium and perfection in discounted supergames', *International Journal of Game Theory*, 16: 15–41.

Dresher, M., A.W. Tucker and P. Wolfe, eds. (1957) *Contributions to the theory of games*, Vol. III, A.M.S. 39. Princeton: Princeton University Press.

Dubey, P. and M. Kaneko (1984) 'Information patterns and Nash equilibria in extensive games: I', *Mathematical Social Sciences*, 8: 11–139.

Dubins, L.E. (1957) 'A discrete evasion game', in: M. Dresher, A.W. Tucker and P. Wolfe, eds., *Contributions to the theory of games III*, A.M.S. 39. Princeton: Princeton University Press, pp. 231-255.

Evertett, H. (1957) 'Recursive games', in: M. Dresher, A.W. Tucker and P. Wolfe, eds., *Contributions to the theory of games III*, A.M.S. 39. Princeton: Princeton University Press, pp. 47–78.

Ferguson, T.S. (1967) 'On discrete evasion games with a two-move information lag', *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I: pp. 453–462. Berkeley U.P.

Forges, F., (1986) 'An approach to communication equilibria', *Econometrica*, 54: 1375-1385.

Forges, F., J.-F. Mertens and A. Neyman (1986) 'A counterexample to the folk theorem with discounting', *Economics Letters*, 20: 7.

Friedman, J. (1971) 'A noncooperative equilibrium for supergames', *Review of Economic Studies*, 38: 1–12.

Friedman, J. (1985) 'Cooperative equilibria in finite horizon noncooperative supergames', *Journal of Economic Theory*, 35: 390–398.

Fudenberg, D., D. Kreps and E. Maskin (1990) 'Repeated games with long-run and short-run players', *Review of Economic Studies*, 57: 555–573.

Fudenberg, D. and D. Levine (1989a) 'Reputation and equilibrium selection in games with a patient player', *Econometrica*, 57: 759–778.

Fudenberg, D. and D. Levine (1989b) 'Equilibrium payoffs with long-run and short-run players and imperfect public information', preprint.

Fudenberg, D. and D. Levine (1991) 'An approximate folk theorem with imperfect private information', *Journal of Economic Theory*, 54: 26–47.

Fudenberg, D. and E. Maskin (1986a) 'The folk theorem in repeated games with discounting and with incomplete information', *Econometrica*, 54: 533–554.

Fudenberg, D. and E. Maskin (1986b) 'Discounted repeated games with unobservable actions I: One-sided moral hazard', preprint.

Fudenberg, D. and E. Maskin (1991) 'On the dispensability of public randomizations in discounted repeated games', *Journal of Economic Theory*, 53: 428–438.

Fudenberg, D., D. Levine and E. Maskin (1989) 'The folk theorem with imperfect public information', preprint.

Gilboa, I. and D. Samet (1989) 'Bounded versus unbounded rationality: The tyranny of the weak', *Games and Economic Behavior*, 1: 213–221.

Hart, S. (1979) 'Lecture notes on special topics in game theory', IMSSS-Economics, Stanford University.

Kalai, E. (1990) 'Bounded rationality and strategic complexity in repeated games', in: T. Ichiishi, A. Neyman and Y. Tauman, eds., *Game theory and applications*. New York: Academic Press, pp. 131–157.

Kalai, E. and W. Stanford (1988) 'Finite rationality and interpersonal complexity in repeated games', *Econometrica* 56: 397–410.

Kaneko, M. (1982) 'Some remarks on the folk theorem in game theory', *Mathematical Social Sciences*, 3: 281–290; also Erratum in 5, 233 (1983).

Karlin, S. (1957) 'An infinite move game with a lag', in: M. Dresher, A.W. Tucker and P. Wolfe, eds., *Contributions to the theory of games III*, A.M.S. 39. Princeton: Princeton University Press, pp. 255–272.

Kreps, D., P. Milgrom, J. Roberts and R. Wilson (1982) 'Rational cooperation in the finitely repeated prisoner's dilemma', *Journal of Economic Theory*, 27: 245–252.

Krishna, V. (1988) 'The folk theorems for repeated games', Proceedings of the NATO-ASI conference: Models of incomplete information and bounded rationality, Anacapri, 1987, to appear.

Kurz, M. (1978) 'Altruism as an outcome of social interaction', *American Economic Review*, 68: 216–222.

Lehrer, E. (1988) 'Repeated games with stationary bounded recall strategies', *Journal of Economic Theory*, 46: 130–144.

Lehrer, E. (1989) 'Lower equilibrium payoffs in two-player repeated games with non-observable actions', *International Journal of Game Theory*, 18: 57–89.

Lehrer, E. (1990) 'Nash equilibria of n-player repeated games with semistandard information', *International Journal of Game Theory*, 19: 191–217.

Lehrer, E. (1991) 'Internal correlation in repeated games', *International Journal of Game Theory*, 19: 431–456.

Lehrer, E. (1992a) 'Correlated equilibria in two-player repeated games with non-observable actions', *Mathematics of Operations Research*, 17: 175–199.

Lehrer, E. (1992b) 'Two-player repeated games with non-observable actions and observable payoffs', *Mathematics of Operations Research*, 200–224.

Lehrer, E. (1992c) 'On the equilibrium payoffs set of two player repeated games with imperfect monitoring', *International Journal of Game Theory*, 20: 211–226.

Masso J. and R. Rosenthal (1989) 'More on the "Anti-Folk Theorem"', *Journal of Mathematical Economics*, 18: 281–290.

Megiddo, N. and A. Widgerson (1986) 'On plays by means of computing machines', in: J.Y. Halpern, ed., *Theoretical aspects of reasoning about knowledge*. Morgan Kaufman Publishers, pp. 259–274.

Mertens, J.-F. (1980) 'A note on the characteristic function of supergames', *International Journal of Game Theory*, 9: 189–190.

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

Mertens, J.-F., S. Sorin and S. Zamir (1992) *Repeated games*, book to appear.

Milnor, J. and L.S. Shapley (1957) 'On games of survival', in: M. Dresher, A.W. Tucker and P. Wolfe, eds., *Contributions to the theory of games III*, A.M.S. 39. Princeton: Princeton University Press, pp. 15–45.

Myerson, R. (1986) 'Multistage games with communication', *Econometrica*, 54: 323–358.

Naudé, D. (1990) 'Correlated equilibria with semi-standard information', preprint.

Neyman, A. (1985) 'Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma', *Economics Letters*, 19: 227–229.

Neyman, A. (1988) 'Stochastic games', preprint.

Neyman, A. (1989) 'Games without common knowledge', preprint.

Radner, R. (1980) 'Collusive behavior in non-cooperative epsilon-equilibria in oligopolies with long but finite lives', *Journal of Economic Theory*, 22: 136–154.

Radner, R. (1981) 'Monitoring cooperative agreements in a repeated principal-agent relationship', *Econometrica*, 49: 1127–1147.

Radner, R. (1985) 'Repeated principal–agent games with discounting', *Econometrica*, 53: 1173–1198.

Radner, R. (1986a) 'Repeated partnership games with imperfect monitoring and no discounting', *Review of Economic Studies*, 53: 43–57.

Radner, R. (1986b) 'Can bounded rationality resolve the prisoner's dilemma', in: A. Mas-Colell and W. Hildenbrand, eds., *Essays in honor of Gérard Debreu*. Amsterdam: North-Holland, pp. 387–399.

Radner, R. (1986c) 'Repeated moral hazard with low discount rates', in: W. Heller, R. Starr and D. Starrett, eds., *Essays in honor of Kenneth J. Arrow*. Cambridge: Cambridge University Press, pp. 25–63.

Radner, R., R.B. Myerson and E. Maskin (1986) 'An example of a repeated partnership game with discounting and with uniformly inefficient equilibria', *Review of Economic Studies*, 53: 59–69.

Rosenthal, R. and A. Rubinstein (1984) 'Repeated two-player games with ruin', *International Journal of Game Theory*, 13: 155–177.

Rubinstein, A. (1976) 'Equilibrium in supergames', preprint.

Rubinstein, A. (1979a) 'Equilibrium in supergames with the overtaking criterion', *Journal of Economic Theory*, 21: 1–9.

Rubinstein, A. (1979b) 'Offenses that may have been committed by accident – an optimal policy of redistribution', in: S.J. Brams, A. Schotter and G. Schwödiauer, eds., *Applied game theory*, Berlin: Physica-Verlag, pp. 406–413.

Rubinstein, A. (1980) 'Strong perfect equilibrium in supergames', *International Journal of Game Theory*, 9: 1–12.

Rubinstein, A. (1986) 'Finite automata play the repeated prisoner's dilemma', *Journal of Economic Theory*, 39: 83–96.

Rubinstein, A. and M. Yaari (1983) 'Repeated insurance contracts and moral hazard', *Journal of Economic Theory*, 30: 74–97.

Samuelson, L. (1987) 'A note on uncertainty and cooperation in a finitely repeated prisoner's dilemma', *International Journal of Game Theory*, 16: 187–195.

Scarf, H. and L.S. Shapley (1957) 'Games with partial information', in: M. Dresher, A.W. Tucker and P. Wolfe, eds., *Contributions to the theory of games III*, A.M.S. 39. Princeton: Princeton University Press, pp. 213–229.

Selten, R. (1975) 'Reexamination of the perfectness concept for equilibrium points in extensive games', *International Journal of Game Theory*, 4: 25–55.

Smale, S. (1980) 'The prisoner's dilemma and dynamical systems associated to non-cooperative games', *Econometrica*, 48: 1617–1634.

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

Sorin, S. (1986b) 'Asymptotic properties of a non-zero stochastic game', *International Journal of Game Theory*, 15: 101–107.

Sorin, S. (1988) 'Repeated games with bounded rationality', Proceedings of the NATO-ASI Conference: Models of incomplete information and bounded rationality, Anacapri, 1987, to appear.

Sorin, S. (1990) 'Supergames', in: T. Ichiishi, A. Neyman and Y. Tauman, eds., *Game theory and applications*. New York: Academic Press, pp. 46–63.

Vieille, N. (1989) 'Weak approachability', to appear in *Mathematics of Operations Research*.

Zemel, E. (1989) 'Small talk and cooperation: a note on bounded rationality', *Journal of Economic Theory*, 49: 1–9.