New Approaches and Recent Advances in Two-Person Zero-Sum Repeated Games

Sylvain Sorin*
Laboratoire d'Econométrie
Ecole Polytechnique
1 rue Descartes
75005 Paris, France

and

Equipe Combinatoire et Optimisation
UFR 921, Université Pierre et Marie Curie - Paris 6
4 place Jussieu 75230 Paris, France

1 Preliminaries

In repeated games where the payoff is accumulated along the play, the players face a problem since they have to take into account the impact of their choices both on the current payoff and on the future of the game.

When considering long games this leads to two alternative cases. Whenever the previous problem can be solved in a "robust" way the game possesses a **uniform value**. In the other situation optimal strategies are very sensitive to the exact specification of the duration of the process. The **asymptotic approach** consists in studying the values of games with finite expected length along a sequence with length going to infinity and the questions are then the existence of a limit and its dependence w.r.t. the sequence.

A typical example is the famous Big Match (Blackwell and Ferguson, 1968) described by the following matrix:

	α	β
a	1^*	0^*
b	0	1

This corresponds to a stochastic game where, as soon as Player 1 plays a, the game reaches an absorbing state with a constant payoff corresponding to the entry played at that stage. Both the n-stage value v_n and the λ-discounted value v_λ are equal to 1/2 and are also independent of the additional information transmitted along the play to the players. Moreover, under standard signaling (meaning that the past play is public knowledge), or with only known past payoffs, the uniform value exists.

*Prepared for a plenary lecture at the International Symposium on Dynamic Games and Applications, Adelaide, Australia, December 18-21, 2000.

This is no longer the case with general signals: for example, when Player 1 has no information on Player 2's moves the max min is 0 (Kohlberg, 1974). It follows that the existence of a uniform value for stochastic games depends on the signalling structure on moves (Mertens and Neyman, 1981; Coulomb, 1992, 1999, 2001). On the other hand, the asymptotic behavior does not (Shapley, 1953).

We now describe more precisely the model of two-person zero-sum repeated game Γ that we consider. We are given a parameter space M and a function g from $I \times J \times M$ to $\mathbb{R}$: for each $m \in M$ this defines a two-person zero-sum game with action spaces I and J for players 1 and 2 respectively and payoff function g. To simplify the presentation we will first consider the case where all sets are finite, avoiding measurability issues. The initial parameter m_1 is chosen at random and the players receive some initial information about it, say a_1 (resp. b_1) for Player 1 (resp. Player 2). This choice is performed according to some probability π on $M \times A \times B$, where A and B are the signal sets of each player. In addition, after each stage the players obtain some further information about the previous choice of actions and both the previous and the current values of the parameter. This is represented by a map Q from $M \times I \times J$ to probabilities on $M \times A \times B$. At stage t given the state m_t and the moves (i_t, j_t), a triple $(m_{t+1}, a_{t+1}, b_{t+1})$ is chosen at random according to the distribution $Q(m_t, i_t, j_t)$. The new parameter is m_{t+1}, and the signal a_{t+1} (resp. b_{t+1}) is transmitted to Player 1 (resp. Player 2). A play of the game is thus a sequence $m_1, a_1, b_1, i_1, j_1, m_2, a_2, b_2, i_2, j_2, \ldots$ while the information of Player 1 before his play at stage t is a private history of the form $(a_1, i_1, a_2, i_2, \ldots, a_t)$ and similarly for Player 2. The corresponding sequence of payoffs is $g_1, g_2, \ldots$ with $g_t = g(i_t, j_t, m_t)$. (Note that it is not known to the players except if included in the signals.)

A strategy σ for Player 1 is a map from private histories to $\Delta(I)$, the space of probabilities on the set I of actions and τ is defined similarly for Player 2. Such a couple (σ, τ) induces, together with the components of the game, π and Q, a distribution on plays, hence on the sequence of payoffs.

There are basically two ways of handling the game repeated a large number of times that are described as follows (see Aumann and Maschler, 1995, Mertens, Sorin and Zamir, 1994):

1) The first one corresponds to the "compact case". One considers a sequence of evaluations of the stream of payoffs converging to the "uniform distribution on the set of integers, $\mathbb{N}$". For each specific evaluation, under natural assumptions on the action spaces and on the reward and transition mappings, the strategy spaces will be compact for a topology for which the payoff function will be continuous, hence the value will exist.

Two typical examples correspond to:

1.1) the **finite n-stage** game Γ_n with payoff given by the average of the first n rewards:

$$\gamma_n(\sigma, \tau) = E_{\sigma,\tau}\left(\frac{1}{n}\sum\nolimits_{t=1}^{n} g_t\right).$$

In the finite case (all sets considered being finite), this reduces to a game with finitely many pure strategies.

1.2) the λ-**discounted game** Γ_λ with payoff equal to the discounted sum of the rewards:

$$\gamma_\lambda(\sigma, \tau) = E_{\sigma,\tau}\left(\sum\nolimits_{t=1}^{\infty} \lambda(1-\lambda)^{t-1} g_t\right).$$

The values of these games are denoted by v_n and v_λ respectively. The study of their asymptotic behavior, as n goes to ∞ or λ goes to 0 is the study of the **asymptotic game**.

Extensions consider games with random duration or random duration process (Neyman, 2003, Neyman and Sorin, 2001).

2) An alternative analysis considers the whole family of "long games". It does not specify payoffs in some infinite game like $\liminf \frac{1}{n}\sum_{t=1}^{n} g_t$ or a measurable function defined on plays (see Maitra and Sudderth, 1998), but requires uniformity properties of the strategies.

Explicitly, $\underline{v}$ is the **maxmin** if the two following conditions are satisfied:

- Player 1 can **guarantee** $\underline{v}$: for any $\varepsilon > 0$, there exists a strategy σ of Player 1 and an integer N such that for any $n \geq N$ and any strategy τ of Player 2:

  $$\gamma_n(\sigma, \tau) \geq \underline{v} - \varepsilon.$$

  (It follows from the uniformity in τ that if Player 1 can guarantee f both $\liminf_{n\to\infty} v_n$ and $\liminf_{\lambda\to 0} v_\lambda$ will be greater than f.)
- Player 2 can **defend** $\underline{v}$: for any $\varepsilon > 0$ and any strategy σ of Player 1, there exist an integer N and a strategy τ of Player 2 such that for all $n \geq N$:

  $$\gamma_n(\sigma, \tau) \leq \underline{v} + \varepsilon.$$

  (Note that to satisfy this requirement is stronger than to contradict the previous condition; hence the existence of $\underline{v}$ is an issue.)

  A dual definition holds for the **minmax** $\overline{v}$. Whenever $\underline{v} = \overline{v}$, the game has a **uniform value**, denoted by v_∞. Remark that the existence of v_∞ implies:

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

We will be concerned here mainly with the asymptotic approach that relies more on the recursive structure and the related value operator and less on the construction of strategies. We will describe several recent ideas and results: the extension and the study of the Shapley operator, the variational approach to the asymptotic game, the use of the dual game, the limit game and the relation with differential games of fixed duration.

2 The Operator Approach

In this section the games are analyzed through the recursive relations that basically extend the dynamic programming principle.

2.1 Operators and Games

2.1.1 Stochastic Games and Shapley Operator

Consider a stochastic game with parameter space Ω, action spaces I and J and real bounded payoff function g from $\Omega \times I \times J$. This corresponds to the model of Section 1 with $M = \Omega$ and where the initial probability π is the law of a parameter ω_1 announced to both. In addition at each stage $t+1$, the transition $Q(\cdot|\omega_t, i_t, j_t)$ determines the new parameter ω_{t+1} and the signal for each player a_{t+1} or b_{t+1} contains at least the information ω_{t+1}. It follows that Ω will be the natural state space on which v_n and v_λ are defined.

Explicitly, let $X = \Delta(I)$ and $Y = \Delta(J)$ and extend by bilinearity g and Q to $X \times Y$. The Shapley (1953) operator Ψ acts on the set $\mathcal{F}$ of real bounded measurable functions f on Ω as follows:

$$\Psi(f)(\omega) = \mathtt{val}_{X\times Y} \left\{ g(\omega, x, y) + \int_\Omega f(\omega') Q(d\omega'|\omega, x, y) \right\} \tag{1}$$

where $\mathtt{val}_{X\times Y}$ stands for the value operator:

$$\mathtt{val}_{X\times Y} = \max_X \min_Y = \min_Y \max_X .$$

A basic property is that Ψ is non-expansive on $\mathcal{F}$ endowed with the uniform norm:

$$\|\Psi(f) - \Psi(g)\| \leq \|f - g\| = \sup_{\omega\in\Omega} |f(\omega) - g(\omega)|.$$

Ψ determines the family of values through:

$$n v_n = \Psi^n(0), \qquad \frac{v_\lambda}{\lambda} = \Psi\left((1-\lambda)\frac{v_\lambda}{\lambda}\right). \tag{2}$$

The same relations hold for general state and action spaces when dealing with a complete subspace of $\mathcal{F}$ on which Ψ is well defined and which is stable under Ψ.

2.1.2 Non-expansive Mappings

The asymptotic approach of the game is thus related to the following problems: given T a non-expansive mapping on a linear normed space Z, study the iterates $T^n(0)/n = v_n$ as n goes to ∞ and the behavior of $\lambda z_\lambda = v_\lambda$, where z_λ is the fixed point of the mapping $z \mapsto T((1-\lambda)z)$, as λ goes to 0.

Kohlberg and Neyman (1981) proved the existence of a linear functional, f, of norm 1 on Z such that:

$$\lim_{n\to\infty} f(v_n) = \lim_{n\to\infty} \|v_n\| = \lim_{\lambda\to 0} f(v_\lambda) = \lim_{\lambda\to 0} \|v_\lambda\| = \inf_{z\in Z} \|T(z) - z\| \qquad (3)$$

Then they deduce that if Z is reflexive and strictly convex, there is weak convergence to one point, and if the dual space Z^* has a Frechet differentiable norm, the convergence is strong.

In our framework the norm is the uniform norm on a space of real bounded functions and is not strictly convex, see however section 5.3.

Neyman (2003) proved that if v_λ is of bounded variation in the sense that for any sequence λ_i decreasing to 0,

$$\sum_i \|v_{\lambda_{i+1}} - v_{\lambda_i}\| < \infty, \qquad (4)$$

then $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.

2.1.3 ε-Weighted and Projective Operators

Back to the framework of Section 2.1.1, it is also natural to introduce the ε-weighted operator:

$$\Phi(\varepsilon, f)(\omega) = \mathtt{val}_{X\times Y} \left\{\varepsilon g(\omega, x, y) + (1-\varepsilon)\int_\Omega f(\omega')Q(d\omega'|\omega, x, y)\right\}, \qquad (5)$$

related to the initial Shapley operator by:

$$\Phi(\varepsilon, f) = \varepsilon\Psi\left(\frac{(1-\varepsilon)f}{\varepsilon}\right). \qquad (6)$$

Then one has:

$$v_n = \Phi\left((1/n), v_{n-1}\right), \qquad v_\lambda = \Phi(\lambda, v_\lambda) \qquad (7)$$

which are the basic recursive equations for the values. The asymptotic study relies thus on the behavior of $\Phi(\varepsilon, \cdot)$, as ε goes to 0. Obviously, if v_n or v_λ converges uniformly, the limit w will satisfy:

$$w = \Phi(0, w) \qquad (8)$$

hence $\Phi(\varepsilon, \cdot)$ also appears as a perturbation of the "projective" operator $\mathcal{P}$ which gives the evaluation today of the payoff f tomorrow:

$$\mathcal{P}(f)(\omega) = \Phi(0, f)(\omega) = \mathtt{val}_{X\times Y}\int_\Omega f(\omega')Q(d\omega'|\omega, x, y). \qquad (9)$$

Explicitly, one has

$$\Phi(\varepsilon, f)(\omega) = \mathtt{val}_{X\times Y}\left\{\int_\Omega f(\omega')Q(d\omega'|\omega, x, y) + \varepsilon h(\omega, x, y)\right\}, \tag{10}$$

with $h(\omega, x, y) = g(\omega, x, y) - \int_\Omega f(\omega')Q(d\omega'|\omega, x, y)$.

One can also consider $\Phi(0, f)$ as a function of Ψ defined by:

$$\Phi(0, f)(\omega) = \lim_{\varepsilon\to 0} \varepsilon\Psi\left(\frac{(1-\varepsilon)}{\varepsilon} f\right)(\omega) = \lim_{\varepsilon\to 0} \varepsilon\Psi\left(\frac{f}{\varepsilon}\right)(\omega) \tag{11}$$

thus $\Phi(0, .)$ is the recession operator associated to Ψ. Note that this operator is independent of g and relates to the stochastic game only through the transition Q.

2.1.4 Games with Incomplete Information

A similar representation is available in the framework of repeated games with incomplete information, Aumann and Maschler (1995).

We will describe here the simple case of independent information and standard signaling. In the setup of Section 1, the parameter space M is a product $K\times L$ endowed with a product probability $\pi = p\otimes q \in\Delta(K)\times\Delta(L)$ and the initial signals are $a_1 = k_1$, $b_1 = \ell_1$. Hence the players have partial private information on the parameter (k_1, ℓ_1). This one is fixed for the duration of the play $((k_t, \ell_t) = (k_1, \ell_1))$ and the signals to the players reveal the previous moves $a_{t+1} = b_{t+1} = (i_t, j_t)$. A one-stage strategy of Player 1 is an element x in $\mathbf{X} = \Delta(I)^K$ (resp. y in $\mathbf{Y} = \Delta(J)^L$ for Player 2).

We represent now this game as a stochastic game. The basic state space is $\chi = \Delta(K)\times\Delta(L)$ and corresponds to the beliefs of the players on the parameter along the play. The transition is given by a map Π from $\chi\times\mathbf{X}\times\mathbf{Y}$ to probabilities on χ with $\Pi((p(i), q(j))|(p, q), x, y) = x(i)y(j)$, where $p(i)$ is the conditional probability on K given the move i and $x(i)$ the probability of this move (and similarly for the other variable). Explicitly: $x(i) = \sum_k p^k x_i^k$ and $p^k(i) = (p^k x_i^k)/(x(i))$. Ψ is now an operator on the set of real bounded saddle (concave/convex) functions on χ, Rosenberg and Sorin (2001):

$$\Psi(f)(p, q) = \mathtt{val}_{\mathbf{X}\times\mathbf{Y}}\left\{g(p, q, x, y) + \int_\chi f(p', q')\Pi(d(p', q')|(p, q), x, y)\right\} \tag{12}$$

with $g(p, q, x, y) = \sum_{k,\ell} p^k q^\ell g\big(k, \ell, x^k, y^\ell\big)$. Then one establishes recursive formula for v_n and v_λ, Mertens, Sorin and Zamir (1994), similar to the ones described in section 2.1.1.

Note that by the definition of Π, the state variable is updated as a function of the one-stage strategies of the players, which are not public information during the play. The argument is thus first to prove the existence of a value (v_n or v_λ) and then using the minmax theorem to construct an equivalent game, in the sense of having the same sequence of values, in which one-stage strategies are announced. This last game is now reducible to a stochastic game.

2.1.5 General Recursive Structure

More generally a recursive structure holds for games described in Section 1 and we follow the construction in Mertens, Sorin and Zamir (1994), Sections III.1, III.2, IV.3.

Consider for example a game with lack of information on one side (described as in Section 2.1.4 with L of cardinal 1) and with signals so that the conditional probabilities of Player 2 on the parameter space are unknown to Player 1, but Player 1 has probabilities on them. In addition Player 2 has probabilities on those beliefs of Player 1 and so on.

The recursive structure thus relies on the construction of the universal belief space, Mertens and Zamir (1985), that represents this infinite hierarchy of beliefs: $\Xi = M \times \Theta^1 \times \Theta^2$, where Θ^i, homeomorphic to $\Delta(M \times \Theta^{-i})$, is the type set of Player i, The signaling structure in the game, just before the moves at stage t, describes an information scheme that induces a consistent distribution on Ξ. This is referred to as the entrance law $\mathcal{P}_t \in \Delta(\Xi)$. The entrance law $\mathcal{P}_t$ and the (behavioral) strategies at stage t (say α_t and β_t) from type set to mixed move set determine the current payoff and the new entrance law $\mathcal{P}_{t+1} = H(\mathcal{P}_t, \alpha_t, \beta_t)$. This updating rule is the basis of the recursive structure. The stationary aspect of the repeated game is expressed by the fact that H does not depend on t. The Shapley operator is defined on the set of real bounded functions on $\Delta(\Xi)$ by:

$$\Psi(f)(\mathcal{P}) = \sup_{\alpha} \inf_{\beta} \{g(\mathcal{P}, \alpha, \beta) + f(H(\mathcal{P}, \alpha, \beta))\},$$

(there is no indication at this level that sup inf commutes for all f) and the usual relations hold, see Mertens, Sorin and Zamir (1994) Section IV.3:

$$(n+1)v_{n+1}(\mathcal{P}) = \mathtt{val}_{\alpha \times \beta} \{g(\mathcal{P}, \alpha, \beta) + n v_n(H(\mathcal{P}, \alpha, \beta))\},$$
$$v_\lambda(\mathcal{P}) = \mathtt{val}_{\alpha \times \beta} \{\lambda g(\mathcal{P}, \alpha, \beta) + (1-\lambda) v_\lambda(H(\mathcal{P}, \alpha, \beta))\},$$

where $\mathtt{val}_{\alpha \times \beta} = \sup_\alpha \inf_\beta = \inf_\beta \sup_\alpha$ is the value operator for the "one stage game on $\mathcal{P}$".

We have here a "deterministic" stochastic game: in the framework of a regular stochastic game, it would correspond to working at the level of distributions on the state space, $\Delta(\Omega)$.

2.2 Variational Inequalities

We use here the previous formulations to obtain properties on the asymptotic behavior of the values, following Rosenberg and Sorin (2001).

2.2.1 A Basic Inequality

We first introduce sets of functions that will correspond to upper and lower bounds on the sequences of values. This allows us, for certain classes of games, to identify the asymptotic value through variational inequalities. The starting point is the next inequality.

Given $\delta > 0$, assume that the function f from Ω to $\mathbb{R}$ satisfies, for all R large enough

$$\Psi(Rf) \leq (R+1)f + \delta.$$

This gives $\Psi(R(f+\delta)) \leq (R+1)(f+\delta)$ and implies:

$$\limsup_{n\to\infty} v_n \leq f + \delta,$$

as well as

$$\limsup_{\lambda\to 0} v_\lambda \leq f + \delta.$$

In particular if f belongs to the set $\mathcal{C}^+$ of functions satisfying the stronger condition: for all $\delta > 0$ there exists R_δ such that $R \geq R_\delta$ implies

$$\Psi(Rf) \leq (R+1)f + \delta \tag{13}$$

and one obtains that both $\limsup_{n\to\infty} v_n$ and $\limsup_{\lambda\to 0} v_\lambda$ are less than f.

2.2.2 Finite State Space

We first apply the above results to **absorbing games**: these are stochastic games where all states except one are absorbing, hence the state can change at most once. It follows that the study on Ω can be reduced to that at one point. In this case, one has easily:

i) $\Psi(f) \leq \|g\|_\infty + f$,

ii) $\Psi(Rf) - (R+1)f$ is strictly decreasing in f:
In fact, let $g - f = d > 0$, then

$$\Psi(Rg) - \Psi(Rf) \leq \Psi(R(f+d)) - \Psi(Rf) \leq Rd = (R+1)(g-f) - d,$$

so that

$$\Psi(Rf) - (R+1)f - (\Psi(Rg) - (R+1)g) \geq d,$$

iii) $\Psi(Rf) - (R+1)f$ is decreasing in R, for $f \geq 0$:

$$\Psi((R+R')f) - (R+R'+1)f \leq \Psi(Rf) - (R+1)f.$$

Define $\mathcal{C}^-$ in a way symmetric to (13). From $i), ii)$ and $iii)$, there exists an element $f \in \mathcal{C}^+ \cap \mathcal{C}^-$ and it is thus equal to both $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$. This extends the initial proof of Kohlberg (1974).

In the framework of **recursive games** where the payoff in all non absorbing states (say Ω_0) is 0, the Shapley operator is defined on real functions f on Ω_0 (with an obvious extension $\overline{f}$ to Ω) by:

$$\Psi(f)(\omega) = \mathtt{val}_{X\times Y} \int_\Omega \overline{f}(\omega')Q(d\omega'|\omega, x, y).$$

It follows that condition (13) reduces to

$$\Psi(f) \le f \quad \text{and moreover} \quad f(\omega) < 0 \quad \text{implies} \quad \Psi(f)(\omega) < f(\omega), \qquad (14)$$

which defines a set $\mathcal{E}^+$. Everett (1957) has shown that the closure of the set $\mathcal{E}^+$ and of its symmetric $\mathcal{E}^-$ have a non-empty intersection from which one deduces that $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$ exists and is the only element of this intersection.

A recent result of Rosenberg and Vieille (2000) drastically extends this property to recursive games with lack of information on both sides. The proof relies on an explicit construction of strategies. Let w an accumulation point of the family v_λ as λ goes to 0. Player 1 will play optimally in the discounted game with a small discount factor if w is larger than $\varepsilon > 0$ at the current value of the parameter and optimally in the projective game $\Psi(0, w)$ otherwise. The sub-martingale property of the value function and a bound on the upcrossings of $[0, \varepsilon]$ are used to prove that $\liminf_{n\to\infty} v_n \ge w$, hence the result.

2.2.3 Simple Convergence

More generally, when Ω is not finite, one can introduce the larger class of functions $\mathcal{S}^+$ where in condition (13) only simple convergence is required:
for all $\delta > 0$ and all ω, there exists $R_{\delta,\omega}$ such that $R \ge R_{\delta,\omega}$ implies

$$\Psi(Rf)(\omega) \le (R+1)f(\omega) + \delta \qquad (15)$$

or

$$\theta^+(f)(\omega) = \limsup_{R\to\infty}\{\Psi(Rf)(\omega) - (R+1)f(\omega)\} \le 0. \qquad (16)$$

In the case of continuous functions on a compact set Ω, an argument similar to point $ii)$ above implies that $f^+ \ge f^-$ for any functions $f^+ \in \mathcal{S}^+$ and $f^- \in \mathcal{S}^-$ (defined similarly with $\theta^-(f^-) \le 0$). Hence the intersection of the closures of $\mathcal{S}^+$ and $\mathcal{S}^-$ contains at most one point.

This argument suffices for the class of games with incomplete information on both sides: any accumulation point w of the family v_λ as $\lambda \to 0$ belongs to the closure of $\mathcal{S}^+$, hence by symmetry the existence of a limit follows. A similar argument holds for $\limsup_{n\to\infty} v_n$.

In the framework of (finite) absorbing games with incomplete information on one side, where the parameter is both changing and unknown, Rosenberg (2000) used similar tools in a very sophisticated way to obtain the first general results of existence of an asymptotic value concerning this class of games. First she shows that any w as above belongs to the closure of $\mathcal{S}^+$. Then that at any point (p, q), $\limsup_{\lambda\to 0} v_\lambda(p, q) \leq w(p, q)$ which again implies convergence. A similar analysis is done for $\lim_{n\to\infty} v_n$.

Remarks

1) Many of the results above only used the following two properties of the operator Ψ, Sorin (2004):
   Ψ is monotonic
   Ψ reduces the constants: for all $\delta > 0$, $\Psi(f + \delta) \leq \Psi(f) + \delta$.
2) The initial and basic proof of convergence of $\lim_{\lambda\to 0} v_\lambda$ for stochastic games relies on the finiteness of the sets involved (Ω, I and J). Bewley and Kohlberg (1976a) used an algebraic approach and proved that v_λ is an algebraic function of λ, from which existence of $\lim_{\lambda\to 0} v_\lambda$ and equality with $\lim_{n\to\infty} v_n$ follows.
3) The results sketched above correspond to three levels of proofs:
   a) The non emptiness of the intersection of the closure of $\mathcal{C}^+$ and $\mathcal{C}^-$. This set contains then one point, namely $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.
   b) For continuous functions on a compact set Ω: any accumulation point of the family of values (as the length goes to ∞) belongs to the intersection of the closures of $\mathcal{S}^+$ and $\mathcal{S}^-$, which contains at most one element.
   c) A property of some accumulation point (related to $\mathcal{S}^+$ or $\mathcal{S}^-$) and a contradiction if two accumulation points differ.

2.3 The Derived Game

We follow here Rosenberg and Sorin (2001). Still dealing with the Shapley operator, condition (15) can be written in a simpler form. This relies, using the expression (10), on the existence of a limit:

$$\varphi(f)(\omega) = \lim_{\varepsilon\to 0^+} \frac{\Phi(\varepsilon, f)(\omega) - \Phi(0, f)(\omega)}{\varepsilon}$$

extending a result of Mills (1956), see also Mertens, Sorin and Zamir (1994), Section 1.1, Ex. 6. More precisely $\varphi(f)(\omega)$ is the value of the "derived game" with payoff $h(\omega, x, y)$, see (10), played on the product of the subsets of optimal strategies in $\Phi(0, f)$. The relation with (16) is given by:

$$\theta^+(f) = \theta^-(f) = \begin{cases} \varphi(f) & \text{if } \Phi(0, f) = f \\ +\infty & \text{if } \Phi(0, f) > f \\ -\infty & \text{if } \Phi(0, f) < f \end{cases}$$

In the setup of games with incomplete information, the family $v_\lambda(p, q)$ is uniformly Lipschitz and any accumulation point as $\lambda \to 0$ is a saddle function $w(p, q)$ satisfying: $\Phi(0, w) = w$. Thus one wants to identify one point in the set of solutions of equation (8) which contains in fact all saddle functions.

For this purpose, one considers the set $\mathcal{A}^+$ of continuous saddle functions f on $\Delta(K) \times \Delta(L)$ such that for any positive strictly concave perturbation η on $\Delta(K)$: $\varphi(f + \eta) \leq 0$. The proof that w belongs to $\mathcal{A}^+$, which is included in the closure of $\mathcal{S}^+$, shows then the convergence of the family v_λ. A similar argument holds for v_n, which in addition implies equality of the limits. Note that the proof relies on the explicit description of $\varphi(f)$ as the value of the derived game.

In addition one obtains the following geometric property. Given f on $\Delta(K)$, say that p is an extreme point of f, $p \in \mathcal{E} f$, if $(p, f(p))$ cannot be expressed as a convex combination of a finite family $\{(p_i, f(p_i))\}$. Then one shows that for any $f \in \mathcal{A}^-$, $f(p, q) \leq u(p, q)$ holds at any extreme point p of $f(\cdot, q)$, where u is the value of the non-revealing game or equivalently:

$$u(p, q) = \mathtt{val}_{\Delta(I) \times \Delta(J)} \sum\nolimits_{k,\ell} p^k q^\ell g(k, \ell, x, y).$$

Hence $v = \lim_{n \to \infty} v_n = \lim_{\lambda \to 0} v_\lambda$ is a saddle continuous function satisfying both inequalities:

$$\begin{aligned} p \in \mathcal{E} v(\cdot, q) &\Rightarrow v(p, q) \leq u(p, q) \\ q \in \mathcal{E} v(p, \cdot) &\Rightarrow v(p, q) \geq u(p, q) \end{aligned} \tag{17}$$

and it is easy to see that it is the only one, Laraki (2001a).

Given a function f on a compact convex set C, let us denote by $\mathtt{Cav}_C f$ the smallest function concave and greater than f on C. By noticing that, for a function f continuous and concave on a compact convex set C, the property $p \in \mathcal{E} f \Rightarrow f(p) \leq g(p)$ is equivalent to $f = \mathtt{Cav}_C \min(g, f)$ one recovers the famous characterization of v due to Mertens and Zamir (1971):

$$v = \mathtt{Cav}_{\Delta(K)} \min(u, v) = \mathtt{Vex}_{\Delta(L)} \max(u, v) \tag{18}$$

where $\mathtt{Cav}_{\Delta(K)} f(p, q)$ stands for $\mathtt{Cav}_{\Delta(K)} f(., q)(p)$.

2.4 The Splitting Game

This section deals again with games with incomplete information on both sides as defined in Section 2.1.4 and follows Laraki (2001a). The operator approach is then extended to more general games. Recall that the recursive formula for the discounted value is:

$$v_\lambda(p, q) = \Phi(\lambda, v_\lambda)(p, q) = \mathtt{val}_{\mathbf{X} \times \mathbf{Y}} \{\lambda g(p, q, x, y) + (1 - \lambda) E(v_\lambda(p', q'))\}$$

where (p', q') is the new posterior distribution and E stands for the expectation induced by p, q, x, y. Denoting by $\mathbf{X}_\lambda(p, q)$ the set of optimal strategies of Player 1 in $\Phi(\lambda, v_\lambda)(p, q)$ and using the concavity of v_λ one deduces:

$$\max_{\mathbf{X}_\lambda(p,q)} \min_{Y} \{g(p, q, x, y) - E(v_\lambda(p', q))\} \geq 0.$$

Let w be an accumulation point of the family $\{v_\lambda\}$ as λ goes to 0 and let $p \in \mathcal{E} w(., q)$. Then the set of optimal strategies for Player 1 in $\Phi(0, w)(p, q)$ is included in the set $NR^1(p)$ of non-revealing strategies (namely with $x(i)\|p(i) - p\| = 0$, recall (8)), hence by uppersemicontinuity one gets from above:

$$\max_{NR^1(p)} \min_{Y} \{g(p, q, x, y) - E(w(p', q))\}$$

$$= \max_{X} \min_{Y} \{g(p, q, x, y) - w(p, q)\} \geq 0$$

hence $u(p, q) \geq w(p, q)$ which is condition (17).

We now generalize this approach.

Since the payoff $g(p, q, x, y) = \sum p^k q^\ell x_i^k y_j^\ell g(k, \ell, i, j)$ is linear it can be written as $\sum_{i,j} x(i) y(j) g(p(i), q(j), i, j)$ so that the Shapley operator is:

$$\Psi(f)(p, q) = \mathtt{val}_{X \times Y} \left\{ \sum\nolimits_{i,j} [g(p(i), q(j), i, j) + f(p(i), q(j))] x(i) y(j) \right\}$$

and one can consider that Player 1's strategy set is the family of random variables from I to $\Delta(K)$ with expectation p. In short, rather than deducing the state variables from the strategies, the state variables are now taken as strategies. The second step is to change the payoffs (introducing a perturbation of the order of $n^{-1/2}$ in v_n) and to replace $g(p(i), q(j), i, j)$ by the value of the "local game" at this state $u(p(i), q(j))$. There is no reason now to keep the range of the martingale finite so that the operator becomes:

$$\mathcal{S}(f)(p, q) = \mathtt{val}_{\Delta_p^2(K) \times \Delta_q^2(L)} E\{u(\widetilde{p}, \widetilde{q}) + f(\widetilde{p}, \widetilde{q})\},$$

where $\Delta_p^2(K)$ is the set of random variables $\widetilde{p}$ with values in $\Delta(K)$ and expectation p. The corresponding game is called the "splitting game". The recursive formula is now different:

$$\frac{w_\lambda}{\lambda} = \mathcal{S}\left(\frac{(1-\lambda)}{\lambda} w_\lambda\right)$$

but with the same proof as above, it leads to the existence of $w = \lim_{\lambda \to 0} w_\lambda$ satisfying the same functional equations:

$$w = \mathtt{Cav}_{\Delta(K)} \min(u, w) = \mathtt{Vex}_{\Delta(L)} \max(u, w).$$

These tools then allow us to extend the operator $u \mapsto M(u) = w$ (existence and uniqueness of a solution) to more general products of compact convex sets $P \times Q$ and functions u, see Laraki (2001b).

3 Duality Properties

3.1 Incomplete Information, Convexity and Duality

Consider a two-person zero-sum game with incomplete information on one side defined by sets of actions S and T, a finite parameter space K, a probability p on K and for each k a real payoff function G^k on $S \times T$. Assume S and T convex and for each k, G^k bounded and bilinear on $S \times T$. The game is played as follows: $k \in K$ is selected according to p and told to Player 1 (the maximizer) while Player 2 only knows p.

In normal form, Player 1 chooses $\mathbf{s} = \{s^k\}$ in S^K, Player 2 chooses t in T and the payoff is

$$G^p(\mathbf{s}, t) = \sum_k p^k G^k(s^k, t).$$

Assume that this game has a value

$$v(p) = \sup_{S^K} \inf_T G^p(\mathbf{s}, t) = \inf_T \sup_{S^K} G^p(\mathbf{s}, t),$$

then v is concave and continuous on the set $\Delta(K)$ of probabilities on K.

Following De Meyer (1996a) one introduces, given $z \in \mathbb{R}^k$, the "dual game" $G^*(z)$ where Player 1 chooses k, then Player 1 plays s in S (resp. Player 2 plays t in T) and the payoff is

$$h[z](k, s; t) = G^k(s, t) - z^k.$$

Translating in normal form, Player 1 chooses $(p, \mathbf{s})$ in $\Delta(K) \times S^K$, Player 2 chooses t in T and the payoff is $\sum_k p^k h[z](k, s^k; t) = G^p(\mathbf{s}, t) - \langle p, z \rangle$.

Then the game $G^*(z)$ has a value $w(z)$, which is convex and continuous on $\mathbb{R}^K$ and the following duality relations holds:

$$w(z) = \max_{p \in \Delta(K)} \{v(p) - \langle p, z \rangle\} = \Lambda_s(v)(z), \tag{19}$$

$$v(p) = \inf_{z \in \mathbb{R}^K} \{w(z) + \langle p, z \rangle\} = \Lambda_i(w)(p), \tag{20}$$

Two consequences are:

Property 3.1. Given z, let p achieve the maximum in (19) and $\mathbf{s}$ be ε-optimal in G^p: then $(p, \mathbf{s})$ is ε-optimal in $G^*(z)$.

Given p, let z achieve the infimum up to ε in (20) and t be ε-optimal in $G^*(z)$: then t is also 2ε-optimal in G^p.

Property 3.2. Let G' be another game on $K \times S \times T$ with corresponding primal and dual values v' and w'. Since Fenchel's transform is an isometry one has

$$\|v - v'\|_{\Delta(K)} = \|w - w'\|_{\mathbb{R}^K}.$$

3.2 The Dual of a Repeated Game with Incomplete Information

We consider now repeated games with incomplete information on one side as introduced in 2.1.4. (with L reduced to one point), and study their duals, following De Meyer (1996b). Obviously the previous analysis applies when working with mixed strategies in the normalized form.

3.2.1 Dual Recursive Formula

The use of the dual game will be of interest for two purposes: construction of optimal strategies for the uninformed player and asymptotic analysis. In both cases the starting point is the recursive formula in the original game.

$$\begin{aligned} F(p) &= \Phi(\varepsilon, f)(p) \\ &= \mathtt{val}_{x\in\mathbf{X}, y\in\mathbf{Y}} \left\{ \varepsilon \sum_k p^k x^k G^k y + (1-\varepsilon) \sum_i x(i) f(p(i)) \right\}, \end{aligned} \tag{21}$$

where we write G^k_{ij} for $g(k, i, j)$. Then one obtains:

$$\begin{aligned} F^*(z) &= \max_{p\in\Delta(K)} \{F(p) - \langle p, z\rangle\} \\ &= \max_{p,x} \min_y \left\{ \varepsilon \sum_k p^k x^k G^k y + (1-\varepsilon) \sum_i x(i) f(p(i)) - \langle p, z\rangle \right\}. \end{aligned}$$

We represent now the couple (p, x) in $\Delta(K)\times\Delta(I)^K$ as an element π in $\Delta(K\times I)$: p is the marginal on K and x^k the conditional probability on I given k:

$$F^*(z) = \max_\pi \min_y \left\{ \varepsilon \sum_{i,k} \pi(i,k) G^k_i y + (1-\varepsilon) \sum_i \pi(i) f(\pi(.|i)) - \langle p, z\rangle \right\}.$$

The concavity of f w.r.t. p implies the concavity of $\sum_i \pi(i) f(\pi(.|i))$ w.r.t. π. This allows us to use the minmax theorem leading to:

$$F^* = \min_y \max_\pi \left\{ \varepsilon \sum_{i,k} \pi(i,k) G^k_i y + (1-\varepsilon) \sum_i \pi(i) f(\pi(.|i)) - \langle p, z\rangle \right\},$$

hence, since $p^k = \sum_i \pi(i)\pi(k|i)$:

$$\begin{aligned} F^*(z) = \min_y \max_{\pi(i)} \Bigg\{ \sum_i \pi(i)(1-\varepsilon) \max_{\pi(.|i)} \Bigg[f(\pi(.|i)) \\ - \left\langle \pi(.|i), \frac{1}{1-\varepsilon} z - \frac{\varepsilon}{1-\varepsilon} G_i y \right\rangle \Bigg] \Bigg\} \end{aligned}$$

where $G_i y$ is the vector $\{G_i^k y\}$. This finally leads to the dual recursive formula:

$$F^*(z) = \min_y \max_i (1-\varepsilon) f^* \left(\frac{1}{1-\varepsilon} z - \frac{\varepsilon}{1-\varepsilon} G_i y \right). \tag{22}$$

The main advantage of dealing with (22) rather than with (21) is that the state variable is known by Player 2 (who controls y and observes i) and evolves smoothly from z to $z + (\varepsilon/(1-\varepsilon))(z - G_i y)$.

3.2.2 Properties of Optimal Strategies

Rosenberg (1998) extended the previous duality to games having at the same time incomplete information and stochastic transition on the parameters. There are then two duality operators (D^1 and D^2) corresponding to the private information of each player. D^1 associates to each function on $\Omega \times \Delta(K) \times \Delta(L)$ a function on $\Omega \times \Delta(K) \times \mathbb{R}^L$. The duality is taken with respect to the unknown parameter of Player 1 replacing q by a vector in $\mathbb{R}^L$. The extension of formula (22) to each dual game allows us to deduce properties of optimal strategies in this dual game for each player. In the discounted case, Player 1 has stationary optimal strategies on a private state space of the form $\Omega \times \Delta(K) \times \mathbb{R}^L$. The component on Ω is the publicly known stochastic parameter; the second component p is the posterior distribution on $\Delta(K)$ that is induced by the use of x: it corresponds to the transmission of information to Player 2; the last one is a vector payoff indexed by the unknown parameter $\ell \in L$ that summarizes the past sequence of payoffs. Similarly, in the finitely repeated game, Player 1 has an optimal strategy which is Markovian on $\Omega \times \Delta(K) \times \mathbb{R}^L$. Obviously dual properties hold for Player 2.

Recall that as soon as lack of information on both sides is present the recursive formula does not allow us to construct inductively optimal strategies (except in specific classes, like games with almost perfect information where a construction similar to the one above could be done, Ponssard and Sorin (1982)). It simply expresses a property satisfied by an alternative game having the same sequence of values, but not the same signals along the play, hence not the same strategy sets. However the use of the dual game allows us, through Property 3.1 (Section 3.1), to deduce optimal strategies in the primal game from optimal strategies in the dual game, and hence to recover an inductive procedure for constructing optimal strategies.

Further properties of the duality operators have been obtained in Laraki (2000). First one can apply the (partial 2) duality operator D^2 to the (partial 1) dual game $D^1(\Gamma)$, then the duality transformations commute and other representations of the global dual game $D^1 \circ D^2(\Gamma) = D^2 \circ D^1(\Gamma)$ are established.

3.2.3 Asymptotic Analysis and Approximate Fixed Points

This section follows De Meyer and Rosenberg (1999). Going back to the class of games with incomplete information on one side, Aumann and Maschler's theorem on the convergence of the families v_n or v_λ to $\mathtt{Cav}_{\Delta(K)}u$ will appear as a consequence of the convergence of the conjugate functions w_n (value of the dual game Γ_n^*) or $w_\lambda(z)$ (for Γ_λ^*) to the Fenchel conjugate of u.

Explicitly let

$$\Lambda_s(u)(z) = \max_{\Delta(K)}\{u(p) - \langle p, z\rangle\},$$

then the bi-conjugate

$$\Lambda_i \circ \Lambda_s(u)(z) = \min_{z\in\mathbb{R}^K}\{\Lambda_s(u)(z) - \langle p, z\rangle\}$$

equals $\mathtt{Cav}_{\Delta(K)}u$. Using property 3.2 in Section 3.1 it is enough to prove the convergence of w_n or w_λ to $\Lambda_s(u)$. Heuristically one deduces from (22) that the limit w should satisfy:

$$w(z) = (1-\varepsilon)\min_Y \max_X \left\{ w(z) + \frac{\varepsilon}{(1-\varepsilon)}\langle \nabla w(z), z - xGy\rangle \right\},$$

which leads to the partial differential equation:

$$-w(z) + \langle \nabla w(z), z\rangle + u(-\nabla w(z)\} = 0, \tag{23}$$

where we recall that $u(q) = \min_Y \max_X \left\{\sum_k q^k x G^k y\right\}$.

Fenchel duality gives:

$$\Lambda_s u(z) - u(-q) = \langle q, z\rangle$$

for $-q\in\partial\Lambda_s u(z)$, which shows that $\Lambda_s u$ is a solution (in a weak sense) of (23). The actual proof uses a general property of approximate operators and fixed points that we described now. Consider a family of operators Ψ_n on a Banach space Z with the following contracting property:

$$\|\Psi_{n+1}(f) - \Psi_{n+1}(g)\| \le \left(\frac{n}{n+1}\right)^a \|f-g\|,$$

for some positive constant a, and n large enough. For example, Ψ is non-expansive and $\Psi_{n+1}(.) = \Phi\left(\frac{1}{n+1}, .\right)$. Define a sequence in Z by $f_0 = 0$ and $f_{n+1} = \Psi_{n+1}(f_n)$. Then if a sequence g_n satisfies an approximate induction in the sense that, for some positive b and n large enough:

$$\|\Psi_{n+1}(g_n) - g_{n+1}\| \le \frac{1}{(n+1)^{1+b}},$$

and g_n converges to g, then f_n converges to g also.

The result on the convergence of w_n follows by choosing g_n as a smooth perturbation of $\Lambda_s u$, like $g_n(z) = E\left[\Lambda_s u\left(z + (X/\sqrt{n})\right)\right]$, X being a cantered reduced normal random variable.

A similar property holds for the sequence w_λ.

3.2.4 Speed of Convergence

The recursive formula and its dual also play a crucial role in the recent deep and astonishing result of Mertens (1998). Given a game with incomplete information on one side with finite state and action spaces but allowing for measurable signal spaces the speed of convergence of v_n to its limit is bounded by $C\big((\ln n)/n\big)^{1/3}$ and this is the best bound. (Recall that the corresponding order of magnitude is $n^{-1/2}$ for standard signaling and $n^{-1/3}$ for state independent signals – even allowing for lack of information on both sides.)

3.3 The Differential Dual Game

This section follows Laraki (2002) and starts again from equation (22). The recursive formula for the value w_n of the dual of the n stage game can be written, since $w_n(z)$ is convex, as:

$$w_n(z) = \min_y \max_x \left(1 - \frac{1}{n}\right) w_{n-1}\left(\frac{1}{(1-(1/n))}\left(z - \frac{1}{n}xGy\right)\right). \qquad (24)$$

This leads us to consider w_n as the n^{th} discretization of the upper value of a differential game.

Explicitly consider the differential game (of fixed duration) on $[0, 1]$ with dynamic $\zeta(t)\in\mathbb{R}^K$ given by:

$$d\zeta/dt = x_t G y_t, \quad \zeta(0) = -z$$

$x_t\in X$, $y_t\in Y$ and terminal payoff $\max_k \zeta^k(1)$.

Given a partition $\Pi = \{t_0 = 0, ..., t_k, ...\}$ with $\theta_k = t_k - t_{k-1}$ and $\sum_{k=1}^{\infty}\theta_k = 1$ we consider the discretization of the game adapted to Π. Let $W_\Pi^+(t_k, \zeta)$ denote the upper value (correspondingly to the case where Player 2 plays first) of the game starting at time t_k from state ζ. It satisfies:

$$W_\Pi^+(t_k, \zeta) = \min_y \max_x W_\Pi^+(t_{k+1}, \zeta + \theta_{k+1}xGy).$$

In particular if Π_n is the uniform discretization with mesh $(1/n)$ one obtains:

$$W_{\Pi_n}^+(0, \zeta) = \min_y \max_x W_{\Pi_n}^+\left(\frac{1}{n}, \zeta + \frac{1}{n}xGy\right)$$

and by time homogeneity, $W_{\Pi_n}^+\left(\frac{1}{n}, \zeta\right) = \left(1 - \frac{1}{n}\right) W_{\Pi_{n-1}}^+\left(0, \frac{\zeta}{1-(1/n)}\right)$, so that:

$$W_{\Pi_n}^+(0, \zeta) = \min_y \max_x \left(1 - \frac{1}{n}\right) W_{\Pi_{n-1}}^+\left(0, \frac{\zeta + (1/n)xGy}{1-(1/n)}\right). \qquad (25)$$

Hence (24) and (25) prove that $w_n(z)$ and $W^+_{\Pi_n}(0, -z)$ satisfy the same recursive equation. They have the same initial value for $n = 1$ hence they coincide.

Basic results of the theory of differential games (see e.g. Souganidis (1999)) show that the game starting at time t from state ζ has a value $\varphi(t, \zeta)$, which is the only viscosity solution, uniformly continuous in ζ uniformly in t, of the following partial differential equation with boundary condition:

$$\frac{\partial \varphi}{\partial t} + u(\nabla \varphi) = 0, \quad \varphi(1, \zeta) = \max_k \zeta^k. \tag{26}$$

One thus obtains $\varphi(0, -z) = \lim_{n\to\infty} W^+_{\Pi_n}(0, -z) = \lim_{n\to\infty} w_n(z) = w(z)$. The time homogeneity property gives $\varphi(t, \zeta) = (1-t)\varphi\left(0, \zeta/(1-t)\right)$, so that w is a solution of

$$f(x) - \langle x, \nabla f(x)\rangle - u(-\nabla f(x)) = 0, \qquad \lim_{\alpha\to 0} \alpha f\,(x/\alpha) = \max_k \left\{ -x^k \right\}$$

which is the previous equation (23) but with a limit (recession) condition.

One can identify the solution of (26), written with $\psi(t, \zeta) = \varphi(1-t, \zeta)$ as satisfying:

$$\frac{\partial \psi}{\partial t} + L(\nabla \psi) = 0 \quad \psi(0, \zeta) = b(\zeta)$$

with L continuous, b uniformly Lipschitz and convex. Hence, using Hopf's representation formula, one obtains:

$$\psi(t, \zeta) = \sup_{p\in\mathbb{R}^K} \inf_{q\in\mathbb{R}^K} \{b(q) + \langle p, \zeta - q\rangle - tL(p)\}$$

which gives here:

$$\psi(t, \zeta) = \sup_{p\in\mathbb{R}^K} \inf_{q\in\mathbb{R}^K} \left\{ \max_k q^k + \langle p, \zeta - q\rangle + tu(p) \right\}$$

and finally $w(z) = \psi(1, -z) = \sup_{p\in\Delta(K)}\{u(p) - \langle p, z\rangle\} = \Lambda_s u(z)$, as in section 3.2.3.

In addition the results in Souganidis (1985) concerning the approximation schemes give a speed of convergence of $(\delta(\Pi))^{1/2}$ of W_Π to φ (where $\delta(\Pi)$ is the mesh of the subdivision Π), hence by duality one obtains Aumann and Maschler's (1995) bound:

$$\|v_n - \mathtt{Cav}_{\Delta(K)} u\| \le \frac{C}{\sqrt{n}}, \qquad \|v_\lambda - \mathtt{Cav}_{\Delta(K)} u\| \le C\sqrt{\lambda}$$

for some constant C.

A last result is a direct identification of the limit. Since w is the conjugate of a concave continuous function v on $\Delta(K)$ and $\varphi(t,\zeta) = (1-t)w\,(-\zeta/(1-t))$ the conditions on φ can be translated as conditions on v. More precisely the first order conditions in terms of local sub- and super-differentials imply that φ is a viscosity subsolution (resp. supersolution) of (26) if and only if v satisfies the first (resp. second) inequality in the variational system (17). In our framework this gives

$$p \in \mathcal{E}v \Rightarrow v(p) \leq u(p) \text{ and } v(p) \geq u(p),\ \forall p,$$

so that $v = \mathtt{Cav}_{\Delta(K)}u$.

4 The Game in Continuous Time

4.1 Repeated Games and Discretization

The main idea here is to consider a repeated game (in the compact case, i.e. with finite expected length) as a game played between time 0 and 1, the length of stage n being simply its relative weight in the evaluation. Non-negative numbers θ_n with $\sum_{n=1}^{\infty}\theta_n = 1$ define a partition Π of [0, 1] with $t_0 = 0$ and $t_n = \sum_{m\leq n}\theta_m$. The repeated game with payoff $\sum_n g_n\theta_n$ corresponds to the game in continuous time where changes in the moves can occur only at times t_m. The finite n-stage game is represented by the uniform partition Π_n with mesh $(1/n)$ while the λ-discounted game is associated to the partition Π_λ with $t_m = 1-(1-\lambda)^m$. In the framework of section 2.1.5. one obtains a recursive formula for the value $W_\Pi(t,\mathcal{P})$ of the game starting at time t with state variable $\mathcal{P}$:

$$W_\Pi(t_n,\mathcal{P}) = \mathtt{val}_{\alpha\times\beta}(\theta_{n+1}g(\mathcal{P},\alpha,\beta) + W_\Pi(t_{n+1}, H(\mathcal{P},\alpha,\beta))).$$

The fact that the payoff is time-independent is expressed by the relation:

$$W_\Pi(t_n,.) = (1-t_n)W_{\Pi[t_n]}(0,.)$$

where $\Pi[t_m]$ is the renormalization to the whole interval [0, 1] of Π restricted to $[t_m, 1]$. By enlarging the state space and incorporating the payoff as new parameter, say ζ, we obtain new functions $L_\Pi(t,\zeta,\mathcal{P})$ with

$$L_\Pi(t_n,\zeta,\mathcal{P}) = \mathtt{val}_{\alpha\times\beta}L_\Pi(t_{n+1},\zeta+\theta_{n+1}g(\mathcal{P},\alpha,\beta), H(\mathcal{P},\alpha,\beta))$$

and

$$L_\Pi(t_n,\zeta,\mathcal{P}) = (1-t_n)L_{\Pi[t_n]}\left(0, \frac{\zeta}{1-t_n}, \mathcal{P}\right).$$

This time normalization explains why the PDE obtained as a limit is homogeneous.

A first heuristic approach in this spirit is in Mertens and Zamir (1976a) where they study, for a specific example of repeated game with lack of information on one side, the limit of the "normalized error term" $\eta_n(p) = \sqrt{n}\,(v_n(p) - v_\infty(p))$, on $[0, 1]$. From the recursive formula for v_n, they deduce another one for the sequence η_n and obtain the following equation for the limit: $\varphi\varphi'' + 1 = 0$. It follows then that $\varphi(p)$ is the normal density evaluated at its p-quantile.

Consider now a simple variation of the Big Match game where Player 1 knows the true game while Player 2 does not and the payoffs are as follows:

	α	β
a	1^*	0^*
b	0	0

Game 1: Probability p

	α	β
a	0^*	0^*
b	0	1

Game 2: Probability $1 - p$

Sorin (1984) derives from the recursive formula the following equation for the limit of v_n: $(2 - p)\varphi(p) = (1 - p) - (1 - p)^2\varphi'(p)$ which leads to $\varphi(p) = (1 - p)\,\{1 - \exp(-p/(1 - p))\}$. Note that this function is not algebraic, which could not be the case for stochastic games nor for games with incomplete information on one side. (Moreover it is also equal to the max min $\underline{v}$.)

4.2 The Limit Game

The recursive formula may also, by exhibiting properties of optimal strategies, allow us to define an auxiliary game in continuous time, considered as a representation of the "limit game" on $[0, 1]$. Two examples are as follows.

A first class, Sorin (1984), corresponds to specific absorbing games with incomplete information on one side of the form:

a	a_1^{k*}	$\dots^*$	a_J^{k*}
b	b_1^k	$\dots$	b_J^k

Game k: Probability p^k

From the recursive formula one deduces that both players can be restricted to strategies independent of the past. One constructs then a game on $[0, 1]$ where Player 1's strategies are stopping times ρ^k corresponding to the first occurrence of a in game k, while Player 2's strategies are measurable functions f from $[0, 1]$ to $\Delta(J)$. The payoff is the integral from 0 to 1 of the instantaneous payoff at time t, $\sum_k p^k g_t^k(\rho^k, f)$ with

$$g_t^k(\rho, f) = \int_0^t a^k(f(s))\rho(ds) + (1 - \rho([0, t]))b^k(f(t)),$$

where $a^k(f) = \sum_j a_j^k f_j$ and similarly for b^k. This game has a value ν and it is easy to show that $\nu = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$. In fact discretizations of ε-optimal strategies in the limit game define strategies in G_n or G_λ and the payoff is continuous.

A much more elaborate construction is in De Meyer (1999). The starting point is again the asymptotic expansion of v_n for games with incomplete information on one side. More precisely the dual recursive formula for $\eta_n = \sqrt{n}\,(v_n - v_\infty)$ leads on one hand to an heuristic second-order PDE (E) and on the other to properties of optimal strategies for both players. One shows that any regular solution of (E) would be the limit of η_n. De Meyer constructs a family of games $\chi(z,t)$ on [0, 1], endowed with a Brownian filtration where strategies for each player are adapted stochastic processes and the payoff is defined through a stochastic integral on [0, 1]. The existence of a value $W(z,t)$ and optimal strategies in $\chi(z,t)$ are then established. One deduces that, under optimal strategies of the players, the state variable $Z_s(z,t)$ in $\chi(z,t)$, $t \le s \le 1$, follows a stochastic differential equation. The value being constant on such trajectories one obtains that $W(z,0)$ is a solution to (E) – where the regularity remains to be proved.

Note that this approach is somehow a dual of the one used in differential games where the initial model is in continuous time and is analyzed through discretization. Here the game on [0, 1] is an idealization of discrete time model with a large number of stages.

4.3 Repeated Games and Differential Games

The first example of resolution of a repeated game trough a differential game is due to Vieille (1992). Consider a repeated game with vector payoffs described by a function g from $I\times J$ to $\mathbb{R}^K$. Given a compact set C in $\mathbb{R}^K$ let $f(z) = -d(z,C)$ where d is the euclidean distance and defines the n stage repeated game with standard information G_n. The sequence of payoffs is $g_1 = g(i_1 j_1), \cdots, g_n$ with average $\overline{g}_n$ and the reward is $f(\overline{g}_n)$.

The game was introduced by Blackwell (1956) who proved the existence of a uniform value (in the sense of Section 1) when C is convex or $K = 1$. He gave also an example of a game in $\mathbb{R}^2$ with no uniform value.

We consider here the asymptotic approach. The value of the G_n is $v_n = V_{\Pi_n}(0,0)$ where V_{Π_n} satisfies $V_{\Pi_n}(1,z) = f(z)$ and:

$$V_{\Pi_n}(t_k, z) = \mathtt{val}_{X\times Y} E_{x,y}\{V_{\Pi_n}(t_{k+1}, z + \theta_{k+1} G_{ij})\}$$

with $X = \Delta(I)$, $Y = \Delta(J)$. The idea is to replace the above equation by the two equations:

$$W^-_{\Pi_n}(t_k, z) = \max_X \min_Y W^-_{\Pi_n}(t_{k+1}, z + \theta_{k+1} G_{ij}),$$

$$W^+_{\Pi_n}(t_k, z) = \min_Y \max_X W^+_{\Pi_n}(t_{k+1}, z + \theta_{k+1} G_{ij}),$$

hence to approximate v_n by the lower and upper values of the discretization of a differential game Γ played on $X \times Y$ between time 0 and 1, with terminal payoff $f\left(\int_0^1 g_u du\right)$ and deterministic differential dynamic given by:

$$\frac{dz}{dt} = x_t G y_t.$$

The main results used are, see e.g. Souganidis (1999):

1) $W^-_{\Pi_n}$ and $W^+_{\Pi_n}$ converge to some functions W^- and W^+ as n goes to ∞,

2) W^- is a viscosity solution on [0, 1] of the equation:

$$\frac{\partial U}{\partial t} + \max_X \min_Y \langle \nabla U, xGy \rangle = 0, \qquad U(1, z) = g(z)$$

which is condition (26) with a new limit condition,

3) this solution is unique.

A similar result for W^+ and the property: $\max_X \min_Y xGy = \min_Y \max_X xGy$ finally imply: $W^- = W^+$ and we denote this value by W.

Hence if $W(0, 0) = 0$, for any $\varepsilon > 0$ there exists N such that if $n \geq N$ Player 1 can force an outcome within ε of C in the lower n^{th} discretization Γ^-_n . The fact that Player 1 can do the same in the original game where the payoff is random relies on a uniform law of large numbers. For L large enough, playing i.i.d. the mixed move x in the m^{th} block between stages mL (included) and $(m+1)L$ (excluded) will generate in G_{nL} an average path near the one generated by x at stage m of Γ^-_n.

Otherwise, $W(0, 0) \leq 2\delta < 0$, in this case Player 2 can avoid a δ-neighborhood of C and a symmetric argument applies.

Altogether the above construction shows that any set is either weakly approachable ($\overline{g}_n$ will be near C with high probability) or weakly excludable ($\overline{g}_n$ will be near the complement of a neighborhood C with high probability)

A second example, Laraki (2002), was described in the earlier section 3.3.

Note that in both cases the random aspect due to the use of mixed moves was eliminated, either by taking expectation or by working with the dual game.

5 Alternative Methods and Further Results

5.1 Dynamic Programming Setup

In the framework of dynamic programming (one person stochastic game), Lehrer and Sorin (1992) gave an example where $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ both exist and differ.

They also proved that uniform convergence (on Ω) of v_n is equivalent to uniform convergence of v_λ and then the limits are the same.

However this condition alone does not imply existence of the uniform value, v_∞, see Lehrer and Monderer (1994), Monderer and Sorin (1993).

5.2 A Limit Game with Double Scale

Another example of a game where the play in Γ_n between stages $t_1 n$ and $t_2 n$ is approximated by the play in the limit game between time t_1 and t_2 is in Sorin (1989). The framework is simple since there are no signals. However one cannot work directly in continuous time because of the presence of two properties: some moves are exceptional in the sense that they induce some change in the state and the number of times they occur has to be taken into account; as for the other moves only the frequency matters. The analysis is done through a "semi normalization" of Γ_n by a game G_L. Each stage ℓ in L corresponds to a large block of stages in Γ_n and the strategies used in G_L at stage ℓ are the summary of the ones used on the block ℓ in Γ_n according to the above classification. One then shows that both $\liminf_{n\to\infty} v_n$ and $\liminf_{\lambda\to 0} v_\lambda$ are greater than $\limsup_{L\to\infty} \mathtt{val} G_L$ and the result follows.

One should add that these sets of reduced strategies were introduced by Mertens and Zamir (1976b) for the uniform approach: they proved the existence of the min max $\overline{v}$ and of the max min $\underline{v}$ and showed that they may differ. See also Waternaux (1983).

5.3 Non-expansive Mappings and Convexity

A proof of the convergence of v_n in the framework of one-sided incomplete information repeated games; using Kohlberg and Neyman's Theorem (result 2.1.2), was achieved by Mertens; see Mertens, Sorin and Zamir (1994), Chapter V, Exercise 5. Convergence of the sequence of norms $\|v_n\|$ implies convergence of the dual values hence of the primal values via Fenchel 's transform. Let v be the limit. Then the linear functional f appearing in Kohlberg and Neyman's result is identified at each extreme point of v and leads to $v = \mathtt{Cav} u$.

5.4 Asymptotic and Uniform Approaches

There are several deep connections between the two approaches (recall Section 1), in addition to the fact that the existence of a uniform value implies convergence of the limiting values under very general conditions (even with private information upon the duration) (Neyman (2003), Neyman and Sorin (2001)).

a) Under standard signaling ($a_t = b_t = (i_t, j_t)$) a bounded variation condition on the discounted values, see (4), is a sufficient condition for the existence of a uniform value in stochastic games, Mertens and Neyman (1981). In addition an optimal strategy is constructed stage after stage by computing at stage t a discounted factor λ_t as a function of the past history of payoffs and then playing once optimally in Γ_{λ_t}.
b) A general conjecture states that in (finitely generated) games where Player 1's information includes Player 2's information the equality: max min $=$

$\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$ holds, Sorin (1984, 1985), Mertens (1987). Somehow Player 1 could, using an optimal strategy of Player 2 in the limit game, define a map from histories to [0, 1]. Given the behavior of Player 2 at stage n, this map induces a time t and Player 1 plays an optimal strategy in the limit game at this time.

Finally recent results along the uniform approach include:

- proof of existence and characterization of max min and min max in absorbing games with signals, Coulomb (1999, 2001),
- proof of existence of max min and equality with $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ in recursive games with lack of information on one side, Rosenberg and Vieille (2000), see point b) above.

REFERENCES

[1] Aumann R.J. and Maschler M. (1995), *Repeated Games with Incomplete Information*, M.I.T. Press (with the collaboration of R. Stearns).

[2] Bewley T. and Kohlberg E. (1976a), The asymptotic theory of stochastic games, *Mathematics of Operations Research*, **1**, 197–208.

[3] Bewley T. and Kohlberg E. (1976b), The asymptotic solution of a recursion equation occurring in stochastic games, *Mathematics of Operations Research*, **1**, 321–336.

[4] Blackwell D. (1956), An analog of the minmax theorem for vector payoffs, *Pacific Journal of Mathematics*, **6**, 1–8.

[5] Blackwell D. and Ferguson T. (1968), The Big Match, *Annals of Mathematical Statistics*, **39**, 159–163.

[6] Coulomb J.-M. (1992), Repeated games with absorbing states and no signals, *International Journal of Game Theory*, **21**, 161–174.

[7] Coulomb J.-M. (1996), A note on 'Big Match', *ESAIM: Probability and Statistics*, **1**, 89–93, http://www.edpsciences.com/ps/.

[8] Coulomb, J.-M. (1999), Generalized Big Match, *Mathematics of Operations Research*, **24**, 795–816.

[9] Coulomb, J.-M. (2001), Repeated games with absorbing states and signaling structure, *Mathematics of Operations Research*, **26**, 286–303.

[10] De Meyer B. (1996a), Repeated games and partial differential equations, *Mathematics of Operations Research*, **21**, 209–236.

[11] De Meyer B. (1996b), Repeated games, duality and the Central Limit theorem, *Mathematics of Operations Research*, **21**, 237–251.

[12] De Meyer B. (1999), From repeated games to Brownian games, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **35**, 1–48.

[13] De Meyer B. and Rosenberg D. (1999), "Cav u" and the dual game, *Mathematics of Operations Research*, **24**, 619–626.

[14] Everett H. (1957), Recursive games, in *Contributions to the Theory of Games, III*, M. Dresher, A.W. Tucker and P. Wolfe (eds.), Annals of Mathematical Studies, 39, Princeton University Press, 47–78.

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

[16] Kohlberg E. and Neyman A. (1981), Asymptotic behavior of non expansive mappings in normed linear spaces, *Israel Journal of Mathematics*, **38**, 269–275.

[17] Laraki R. (2000), Duality and games with incomplete information, preprint.

[18] Laraki R. (2001a), Variational inequalities, systems of functional equations and incomplete information repeated games, *SIAM Journal of Control and Optimization*, **40**, 516–524.

[19] Laraki R. (2001b), The splitting game and applications, *International Journal of Game Theory*, **30**, 359–376.

[20] Laraki R. (2002), Repeated games with lack of information on one side: the dual differential approach, *Mathematics of Operations Research*, **27**, 419–440.

[21] Lehrer E. and Monderer D. (1994), Discounting versus averaging in dynamic programming, *Games and Economic Behavior*, **6**, 97–113.

[22] Lehrer E. and Sorin S. (1992), A uniform Tauberian theorem in dynamic programming, *Mathematics of Operations Research*, **17**, 303–307.

[23] Maitra A. and Sudderth W. (1998), Finitely additive stochastic games with Borel measurable payoffs, *International Journal of Game Theory*, **27**, 257–267.

[24] Mertens J.-F. (1972), The value of two-person zero-sum repeated games: the extensive case, *International Journal of Game Theory*, **1**, 217–227.

[25] Mertens J.-F. (1987), Repeated games, in *Proceedings of the International Congress of Mathematicians, Berkeley, 1986*, A. M. Gleason (ed.), American Mathematical Society, 1528–1577.

[26] Mertens J.-F. (1998), The speed of convergence in repeated games with incomplete information on one side, *International Journal of Game Theory*, **27**, 343–359.

[27] Mertens J.-F. (2002), Stochastic games, in *Handbook of Game Theory, 3*, R. J. Aumann and S. Hart (eds.), North-Holland, 1809–1832.

[28] Mertens J.-F. and Neyman A. (1981), Stochastic games, *International Journal of Game Theory*, **10**, 53–66.

[29] Mertens J.-F., S. Sorin and Zamir S. (1994), *Repeated Games*, CORE D.P. 9420-21–22.

[30] Mertens J.-F. and Zamir S. (1971), The value of two-person zero-sum repeated games with lack of information on both sides, *International Journal of Game Theory*, **1**, 39–64.

[31] Mertens J.-F. and Zamir S. (1976a), The normal distribution and repeated games, *International Journal of Game Theory*, **5**, 187–197.

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

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

[34] Mills H. D. (1956), Marginal values of matrix games and linear programs, in *Linear Inequalities and Related Systems*, H. W. Kuhn and A. W. Tucker (eds.), Annals of Mathematical Studies, 38, Princeton University Press, 183–193.

[35] Monderer D. and Sorin S. (1993), Asymptotic properties in dynamic programming, *International Journal of Game Theory*, **22**, 1–11.

[36] Neyman A. (2003), Stochastic games and non-expansive maps, Chapter 26 in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), NATO Science Series C 570, Kluwer Academic Publishers.

[37] Neyman A. and Sorin S. (2001), Zero-sum two-person games with public uncertain duration process, Cahier du Laboratoire d'Econometrie, Ecole Polytechnique, 2001–013.

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

[39] Rosenberg D. (1998), Duality and Markovian strategies, *International Journal of Game Theory*, **27**, 577–597.

[40] Rosenberg D. (2000) Zero-sum absorbing games with incomplete information on one side: asymptotic analysis, *SIAM Journal on Control and Optimization*, **39**, 208–225.

[41] Rosenberg D. and Sorin S. (2001), An operator approach to zero-sum repeated games, *Israel Journal of Mathematics*, **121**, 221–246.

[42] Rosenberg D. and Vieille N. (2000), The maxmin of recursive games with lack of information on one side, *Mathematics of Operations Research*, **25**, 23–35.

[43] Shapley L. S. (1953), Stochastic games, *Proceedings of the National Academy of Sciences of the U.S.A*, **39**, 1095–1100.

[44] Sorin S. (1984), Big Match with lack of information on one side (Part I), *International Journal of Game Theory*, **13**, 201–255.

[45] Sorin S. (1985), Big Match with lack of information on one side (Part II), *International Journal of Game Theory*, **14**, 173–204.

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

[47] Sorin S. (2002), *A First Course on Zero-Sum Repeated Games*, Springer.

[48] Sorin S. (2003), The operator approach to zero-sum stochastic games, Chapter 27 in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), NATO Science Series C 570, Kluwer Academic Publishers.

[49] Sorin S. (2004), Asymptotic properties of monotonic non-expansive mappings, *Discrete Events Dynamic Systems*, **14**, 109–122.

[50] Souganidis P.E. (1985), Approximation schemes for viscosity solutions of Hamilton–Jacobi equations, *Journal of Differential Equations*, **17**, 781–791.

[51] Souganidis P.E. (1999), Two player zero sum differential games and viscosity solutions, in *Stochastic and Differential Games*, M. Bardi, T.E.S. Raghavan and T. Parthasarathy (eds.), Birkhauser, 70–104.

[52] Vieille N. (1992), Weak approachability, *Mathematics of Operations Research*, **17**, 781–791.

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

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