CHAPTER 2

Advances in Zero-Sum Dynamic Games

Rida Laraki[*,†], Sylvain Sorin[‡]

[*]CNRS in LAMSADE (Université Paris-Dauphine), France
[†]Economics Department at Ecole Polytechnique, France
[‡]Mathematics, CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, Univ Paris Diderot, Sorbonne Paris Cité, Paris, France

Contents

Handbook of game theory
http://dx.doi.org/10.1016/B978-0-444-53766-9.00002-1

Abstract

The survey presents recent results in the theory of two-person zero-sum repeated games and their connections with differential and continuous-time games. The emphasis is made on the following points:

(1) A general model allows to deal simultaneously with stochastic and informational aspects.
(2) All evaluations of the stage payoffs can be covered in the same framework (and not only the usual Cesàro and Abel means).
(3) The model in discrete time can be seen and analyzed as a discretization of a continuous time game. Moreover, tools and ideas from repeated games are very fruitful for continuous time games and vice versa.
(4) Numerous important conjectures have been answered (some in the negative).
(5) New tools and original models have been proposed. As a consequence, the field (discrete versus continuous time, stochastic versus incomplete information models) has a much more unified structure, and research is extremely active.

Keywords: repeated, stochastic and differential games, discrete and continuous time, Shapley operator, incomplete information, imperfect monitoring, asymptotic and uniform value, dual game, weak and strong approachability.

JEL Codes: C73. C61, C62

AMS Codes: 91A5, 91A23, 91A25

2.1. INTRODUCTION

The theory of repeated games focuses on situations involving multistage interactions, where, at each period, the Players are facing a stage game in which their actions have two effects: they induce a stage payoff, and they affect the future of the game (note the difference with other multimove games like pursuit or stopping games where there is no stage payoff). If the stage game is a fixed zero-sum game G, repetition adds nothing: the value and optimal strategies are the same as in G. The situation, however, is drastically different for nonzero-sum games leading to a family of so-called Folk theorems: the use of plans and threats generates new equilibria (Sorin's (1992) chapter 4 in Handbook of Game Theory (HGT1)).

In this survey, we will concentrate on the zero-sum case and consider the framework where the stage game belongs to a family G^m, $m \in M$, of two-person zero-sum games played on action sets $I \times J$. Two basic classes of repeated games that have been studied and analyzed extensively in previous volumes of HGT are *stochastic games* (the subject of Mertens's (2002) chapter 47 and Vieille's (2002) chapter 48 in HGT3) and *incomplete information games* (the subject of Zamir's (1992) chapter 5 in HGT1). The reader is referred to these chapters for a general introduction to the topic and a presentation of the fundamental results.

In stochastic games, the parameter m, which determines which game is being played, is a publicly known variable, controlled by the Players. It evolves over time and its value m_{n+1} at stage $n+1$ (called the state) is a random stationary function of the triple (i_n, j_n, m_n) which are the moves, respectively the state, at stage n. At each period, both Players share the same information and, in particular, know the current state. On the other hand, the state is changing and the issue for the Players at stage n is to control both the current payoff g_n (induced by (i_n, j_n, m_n)) and the next state m_{n+1}.

In incomplete information games, the parameter m is chosen once and for all by nature and kept fixed during play, but at least one Player does not have full information about it. In this situation, the issue is the trade-off between using the information (which increases the set of strategies in the stage game) and revealing it (this decreases the potential advantage for the future).

We will see that these two apparently very different models—evolving known state versus unknown fixed state—are particular incarnations of a more general model and share many common properties.

2.1.1 General model of repeated games (RG)

The general presentation of this section follows Mertens et al. (1994). To make it more accessible, we will assume that all sets (of actions, states, and signals) are finite; in the general case, measurable and/or topological hypotheses are in order, but we will not treat such issues here. Some theorems will be stated with compact action spaces. In that case, payoff and transition functions are assumed to be continuous.

Let M be a parameter space and g a function from $I \times J \times M$ to $\mathbb{R}$. For every $m \in M$, this defines a two Player zero-sum game with action sets I and J for Player 1 (the maximizer) and Player 2, respectively, and with a payoff function $g(\cdot, m)$. 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 of nature is performed according to some initial probability distribution π on $A \times B \times M$, where A and B are the signal sets of each Player. The game is then played in discrete time.

At each stage $n = 1, 2, \ldots,$ Player 1 (resp. Player 2) chooses an action $i_n \in I$ (resp. $j_n \in J$). This determines a stage payoff $g_n = g(i_n, j_n, m_n)$, where m_n is the current value of the state parameter. Then, a new value m_{n+1} of the parameter is selected and the Players get some information about it. This is generated by a map Q from $I \times J \times M$ to the set of probability distributions on $A \times B \times M$. More precisely, at stage $n+1$, a triple $(a_{n+1}, b_{n+1}, m_{n+1})$ is chosen according to the distribution $Q(i_n, j_n, m_n)$ and a_{n+1} (resp. b_{n+1}) is transmitted to Player 1 (resp. Player 2).

Note that each signal may reveal some information about the previous choice of actions (i_n, j_n) and/or past and current values (m_n and m_{n+1}) of the parameter:

Stochastic games (with standard signaling: perfect monitoring) (Mertens, 2002) correspond to *public signals* including the parameter: $a_{n+1} = b_{n+1} = \{i_n, j_n, m_{n+1}\}$.

Incomplete information repeated games (with standard signaling) (Zamir, 1992) correspond to an absorbing transition on the parameter ($m_n = m_1$ for every n) and no further information (after the initial one) on the parameter, but previous actions are observed: $a_{n+1} = b_{n+1} = \{i_n, j_n\}$.

A *play* of the game induces 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 move at stage n is a private history of him of the form $(a_1, i_1, a_2, i_2, \ldots, a_n)$ and similarly for Player 2. The corresponding sequence of payoffs is $g_1, g_2, \ldots$ and it is not known to the Players (unless it can be deduced from the signals).

A (behavioral) strategy σ for Player 1 is a map from Player 1 private histories to $\Delta(I)$, the space of probability distributions on the set of actions I: in this way, σ defines the probability distribution of the current stage action as a function of the past events known to Player 1: a behavioral strategy τ for Player 2 is defined similarly. Together with the components of the game, π and Q, a pair (σ, τ) of behavioral strategies induces a probability distribution on plays, and hence on the sequence of payoffs. $\mathsf{E}_{(\sigma,\tau)}$ stands for the corresponding expectation.

2.1.2 Compact evaluations

Once the description of the repeated game is specified, strategy sets are well defined as well as the play (or the distribution on plays) that they induce. In turn, a play determines a flow of stage payoffs $g = \{g_n; n \geq 1\}$, indexed by the positive integers in $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$. Several procedures have been introduced to evaluate this sequence.

Compact evaluations associate to every probability distribution $\mu = \{\mu_n; n \geq 1\}$ on $\mathbb{N}^*$ a game G_μ with evaluation function $\langle \mu, g \rangle = \sum_n g_n \mu_n$. μ_n is interpreted as the (normalized) weight of stage n. Under standard assumptions on the data of the game, the strategy sets are convex-compact and the payoff function is bilinear and continuous in the product topology. Consequently, Sion's minmax theorem implies that the game has a value, denoted by v_μ.

2.1.3 Asymptotic analysis

The *asymptotic analysis* focuses on the problems of existence and the characterization of the asymptotic value (or limit value) $v = \lim v_{\mu^r}$ along a sequence of distributions μ^r with maximum weight (interpreted as the mesh in the continuous time discretization)

$\|\mu^r\| = \sup_n \mu_n^r \to 0$, and the dependence of this limit on a particular chosen sequence (see also Section 9.1.1). The connection with continuous time games is as follows. The RG is considered as the discretization in time of some continuous-time game (to be defined) played between time 0 and 1 and such that the duration of stage n is μ_n.

Two standard and well studied RG evaluations are:

(i) The *finitely repeated n-stage game* G_n, $n \geq 1$, with the Cesàro average of the stream of payoffs $\bar{g}_n = \frac{1}{n}\sum_{r=1}^{n} g_r$ and value v_n,

(ii) The λ*-discounted repeated game* G_λ, $\lambda \in]0,1]$, with the Abel average of the stream of payoffs $\bar{g}_\lambda = \sum_{r=1}^{\infty} \lambda(1-\lambda)^{r-1} g_r$ and value v_λ.

More generally, instead of deterministic weights, we can also consider stochastic evaluations. This has been introduced by Neyman and Sorin (2010) under the name "random duration process." In this framework, μ_n is a random variable, the law of which depends upon the previous path of the process (see Section 2.3).

2.1.4 Uniform analysis

A drawback of the previous approach is that even if the asymptotic value exists, the optimal behavior of a Player in the RG may depend heavily upon the exact evaluation (n for finitely repeated games, λ for discounted games) (Zamir, 1973). In other words, the value of the game with many stages is well defined, but one may need to know the exact duration to play well. The uniform approach considers this issue by looking for strategies that are almost optimal in any sufficiently long RG (n large enough, λ small enough). More precisely:

Definition 2.1. *We will say that $\underline{v}$ is the uniform maxmin if the following two conditions are satisfied:*

(i) *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 all $n \geq N$ and every strategy τ of Player 2:*

$$\mathsf{E}_{(\sigma,\tau)}(\bar{g}_n) \geq \underline{v} - \varepsilon,$$

(ii) *Player 2 can* defend *$\underline{v}$: for all $\varepsilon > 0$ and any strategy σ of Player 1, there exist an integer N and a strategy τ of Player 2 such that for all $n \geq N$:*

$$\mathsf{E}_{(\sigma,\tau)}(\bar{g}_n) \leq \underline{v} + \varepsilon.$$

Note the strong requirement of uniformity with respect to both n and τ in (i) and with respect to n in (ii). In particular existence has to be proved. A dual definition holds for the *uniform minmax* $\bar{v}$.

Whenever $\underline{v} = \bar{v}$, the game is said to have a *uniform value*, denoted by v_∞.

The existence of the uniform value does not always hold in general RG (for example, for games with incomplete information on both sides or stochastic games with signals on the moves (imperfect monitoring)). However, its existence implies the existence of the asymptotic value for decreasing evaluation processes (μ_n^r decreasing in n). In particular: $v_\infty = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.

The next sections will focus on the following points:

- in the compact case, corresponding to the asymptotic approach, a unified analysis is provided through the extended Shapley operator that allows us to treat general repeated games and arbitrary evaluation functions;
- the links between the asymptotic approach and the uniform approach, initiated by the tools in the Theorem of Mertens and Neyman (1981);
- the connection with differential games: asymptotic approach versus games of fixed duration, uniform approach versus qualitative differential games.

2.2. RECURSIVE STRUCTURE

2.2.1 Discounted stochastic games

The first and simplest recursive formula for repeated games was established by Shapley (1953), who characterizes the λ-discounted value of a finite stochastic game with state space Ω as the only solution of the equation (recall that g is the payoff and Q the transition):

$$v_\lambda(\omega) = \mathtt{val}_{X\times Y}\Big\{\lambda g(x,y,\omega) + (1-\lambda)\sum_{\omega'} Q(x,y,\omega)[\omega']v_\lambda(\omega')\Big\}. \qquad [2.1]$$

where $X = \Delta(I)$ and $Y = \Delta(J)$ are the spaces of mixed moves, $\mathtt{val}_{X\times Y} = \sup_{x\in X}\inf_{y\in Y} = \inf_{y\in Y}\sup_{x\in X}$ is the value operator (whenever it exists, which will be the case in almost all the chapter where moreover, max and min are achieved). Also, for a function $h : I\times J \to \mathbb{R}$, $h(x,y)$ denotes $\mathsf{E}_{x,y}h = \sum_{i,j} x(i)y(j)h(i,j)$ which is the bilinear extension to $X\times Y$.

This formula expresses the value of the game as a function of the current payoff and the value from the next stage and onwards. Since the Players know the initial state ω and learn at each stage the new state ω', they can perform the analysis for each state separately and can use the recursive formula to compute an optimal strategy for each ω. In particular, they have an optimal stationary strategy and the "state" ω of the stochastic game is the natural "state variable" to compute the value and optimal strategies.

The *Shapley operator* $\mathbf{T}$ associates to a function f from Ω to $\mathbb{R}$ the function $\mathbf{T}(f)$ defined as:

$$\mathbf{T}(f)(\omega) = \mathtt{val}_{X\times Y}\{g(x,y,\omega) + \sum_{\omega'} Q(x,y,\omega)[\omega']f(\omega')\} \qquad [2.2]$$

Thus, v_λ is the only fixed point of the mapping $f \mapsto \lambda\mathbf{T}\left(\frac{(1-\lambda)f}{\lambda}\right)$.

2.2.2 General discounted repeated games

2.2.2.1 Recursive structure

The result of the previous section can be extended to any general repeated game, following Mertens (1987) and Mertens et al. (1994). Let us give a brief description.

The recursive structure relies on the construction of the universal belief space, Mertens and Zamir (1985), which represents the infinite hierarchy of beliefs of the Players: $\Theta = M \times \Theta^1 \times \Theta^2$, where Θ^i, homeomorphic to $\Delta(M \times \Theta^{-i})$, is the universal space of types of Player i (where $i = 1, 2, -i = 3 - i$).

A consistent probability ρ on Θ is such that the conditional probability induced by ρ at θ^i coincides with θ^i itself, both as elements of $\Delta(M \times \Theta^{-i})$. The set of consistent probabilities is denoted by $\mathbb{P} \subset \Delta(\Theta)$. The signaling structure in the game, just before the actions at stage n, describes an information scheme (basically a probability on $M \times \hat{A} \times \hat{B}$ where $\hat{A}$ is a general signal space to Player 1 and $\hat{B}$ for Player 2) that induces a consistent probability $\mathcal{P}_n \in \mathbb{P}$ (see Mertens et al., 1994, Sections III.1, III.2, IV.3). This is referred to as the "entrance law." Taking into account the existence of a value for any finite repeated game with compact evaluation, one can assume that the strategies used by the Players are announced to both. The entrance law $\mathcal{P}_n$ and the (behavioral) strategies at stage n (say α_n and β_n), which can be represented as measurable maps from type sets to mixed actions sets, determine the current payoff and the new entrance law $\mathcal{P}_{n+1} = H(\mathcal{P}_n, \alpha_n, \beta_n)$. Thus the initial game is value-equivalent to a game where at each stage n, before the moves of the Players, a new triple of parameter and signals to the Players is generated according to $\mathcal{P}_n$, and $\mathcal{P}_{n+1}$ is determined given the stage behavioral strategies. This updating rule is the basis of the *recursive structure* for which $\mathbb{P}$ is the "state space." The stationary aspect of the repeated game is expressed by the fact that H does not depend on the stage n.

The (generalized) Shapley operator $\mathbf{T}$ is defined on the set of real-bounded functions on $\mathbb{P}$ as:

$$\mathbf{T}(f)(\mathcal{P}) = \sup_{\alpha} \inf_{\beta} \left\{ g(\mathcal{P}, \alpha, \beta) + f(H(\mathcal{P}, \alpha, \beta)) \right\}. \tag{2.3}$$

Then the usual recursive equations hold (see Mertens et al. (1994), Section IV.3). For the discounted game, one has:

$$v_\lambda(\mathcal{P}) = \mathtt{val}_{\alpha \times \beta} \left\{ \lambda g(\mathcal{P}, \alpha, \beta) + (1 - \lambda) v_\lambda(H(\mathcal{P}, \alpha, \beta)) \right\} \tag{2.4}$$

where $\mathtt{val}_{\alpha \times \beta} = \sup_\alpha \inf_\beta = \inf_\beta \sup_\alpha$ is the value operator for the "one stage game at $\mathcal{P}$." This representation corresponds to a "deterministic" stochastic game on the state space $\mathbb{P} \subset \Delta(\Theta)$.

Hence to each compact repeated game G, one can associate an *auxiliary game* Γ having the same compact values on $\mathbb{P}$. The discounted values satisfy the recursive equation [2.4]. However, the play and strategies in the two games differ, since, in the auxiliary game, an additional signal corresponding to the stage behavioral strategies is given to the payers.

2.2.2.2 Specific classes of repeated games

In the framework of a standard stochastic game with state space Ω, the universal belief space representation of the previous section would correspond to the level of probabilities on the state space, thus $\mathbb{P} = \Delta(\Omega)$. One recovers the initial Shapley formula [2.1] by letting $\mathcal{P}$ be the Dirac mass at ω, in which case (α, β) reduce to (x, y) (i.e., only the ω-component of $(\Delta(I) \times \Delta(J))^{\Omega}$ is relevant), $H(\mathcal{P}, \alpha, \beta)$ corresponds to $Q(\omega, x, y)$, and finally $v_\lambda(H(\mathcal{P}, \alpha, \beta)) = \mathsf{E}_{Q(\omega,x,y)} v_\lambda(\cdot)$.

Let us describe explicitly the recursive structure in the framework of repeated games with incomplete information (independent case with standard signaling). M is a product space $K \times L$, π is a product probability $\pi(k, l) = p^k \times q^l$ with $p \in \Delta(K)$, $q \in \Delta(L)$, and the first signals to the Players are given by: $a_1 = k$ and $b_1 = \ell$. Given the parameter $m = (k, \ell)$, each Player knows his own component and holds a prior on the other Player's component. From stage 1 on, the parameter m is fixed and the information of the Players after stage n is $a_{n+1} = b_{n+1} = \{i_n, j_n\}$.

The auxiliary stochastic game Γ corresponding to the recursive structure is as follows: the "state space" Ω is $\Delta(K) \times \Delta(L)$ and is interpreted as the space of beliefs on the realized value of the parameter. $\mathbf{X} = \Delta(I)^K$ and $\mathbf{Y} = \Delta(J)^L$ are the type-dependent mixed action sets of the Players; g is extended on $\Omega \times \mathbf{X} \times \mathbf{Y}$ as $g(p, q, x, y) = \sum_{k,\ell} p^k q^\ell g(k, \ell, x^k, y^\ell)$, where $g(k, \ell, x^k, y^\ell)$ denotes the expected payoff at $m = (k, \ell)$ where Player 1 (resp. 2) plays according to $x^k \in \Delta(I)$ (resp. $y^\ell \in \Delta(J)$). Given (p, q, x, y), let $\overline{x}(i) = \sum_k x^k(i) p^k$ be the total probability of action i by Player 1 and $p(i)$ be the conditional probability on K given the action i, explicitly $p^k(i) = \frac{p^k x^k(i)}{\overline{x}(i)}$ (and similarly for y and q). Since actions are announced in the original game, and stage strategies are known in the auxiliary game, these posterior probabilities are known by the Players, so one can work with $\Theta = \Omega$ and take as $\mathbb{P} = \Delta(\Omega)$. Finally, the transition Q (from Ω to $\Delta(\Omega)$) is defined by the following formula:

$$Q(p, q, x, y)(p', q') = \sum_{i,j;(p(i),q(j))=(p',q')} \overline{x}(i)\overline{y}(j).$$

The probability to move from (p, q) to (p', q') under (x, y) is the probability of playing the actions that will generate these posteriors. The resulting form of the Shapley operator is:

$$\mathbf{T}(f)(p, q) = \sup_{x \in \mathbf{X}} \inf_{y \in \mathbf{Y}} \left\{ \sum_{k,\ell} p^k q^\ell g(k, \ell, x^k, y^\ell) + \sum_{i,j} \overline{x}(i)\overline{y}(j) f(p(i), q(j)) \right\} \qquad [2.5]$$

where with the previous notations:

$$\sum_{i,j} \overline{x}(i)\overline{y}(j)f(p(i),q(j)) = \mathsf{E}_{Q(p,q,x,y)}\left[f(p',q')\right] = f(H(p,q,x,y))$$

and again:

$$v_\lambda = \lambda \mathbf{T}\left[\frac{(1-\lambda)}{\lambda} v_\lambda\right].$$

The corresponding equations for v_n and v_λ are due to Aumann and Maschler (1966–67) and are reproduced in Aumann and Maschler (1995), and Mertens and Zamir (1971).

Recall that the auxiliary game Γ is "equivalent" to the original one in terms of values but uses different strategy spaces. In the true game, the strategy of the opponent is unknown, hence the computation of the posterior distribution is not feasible.

Most of the results extend to the *dependent case*, introduced by Mertens and Zamir (1971). In addition to the space M endowed with the probability π, there are two signaling maps from M to A for Player 1 (resp. B for Player 2) that correspond to the initial (correlated) information of the Players on the unknown parameter. $\pi(.|a)$ then denotes the belief of Player 1 on M given his signal a (and similarly for Player 2).

2.2.3 Compact evaluations and continuous time extension

The recursive formula expressing the discounted value through the Shapley operator can be extended for values of games with the same plays but alternative evaluations of the stream of payoffs. Introduce, for $\varepsilon \in [0,1]$ the operator Φ given by:

$$\Phi(\varepsilon, f) = \varepsilon \mathbf{T}\left(\frac{(1-\varepsilon)f}{\varepsilon}\right).$$

Then v_λ is a fixed point of $\Phi(\lambda, .)$:

$$v_\lambda = \Phi(\lambda, v_\lambda) \tag{2.6}$$

and v_n (the value of the n-stage game) satisfies the recursive formula:

$$v_n = \Phi\left(\frac{1}{n}, v_{n-1}\right), \tag{2.7}$$

with $v_0 = 0$. Note that [2.7] is equivalent to:

$$nv_n = \mathbf{T}((n-1)v_{n-1}) = \mathbf{T}^n(0), \tag{2.8}$$

More generally, any probability μ on the integers induces a partition $\Pi = \{t_n; n \geq 0\}$ of $[0,1]$ with $t_0 = 0, t_n = \sum_{m=1}^{n} \mu_m$. Consequently, the repeated game is naturally represented as a discretization of a continuous time game played between times 0 and 1, where the actions are constant on each subinterval (t_{n-1}, t_n) with length μ_n, which is the weight of stage n in the original game. Let v_Π (or equivalently v_μ) denote its value. The recursive equation can then be written as:

$$v_\Pi = \Phi(t_1, v_{\Pi_{t_1}}) \quad [2.9]$$

where Π_{t_1} is the renormalization on $[0,1]$ of the restriction of Π to the interval $[t_1, 1]$.

The difficulty with the two recursive formulas [2.7] and [2.9], expressing v_n and v_Π, is the lack of stationarity compared to [2.6]. One way to deal with this issue is to add the time variable to the state space and to define $V_\Pi(t_n)$ as the value of the RG starting at time t_n, i.e., with evaluation μ_{n+m} for the payoff g_m at stage m. The total weight (length) of this game is no longer 1 but $1 - t_n$. With this new time variable, one obtains the equivalent recursive formula:

$$V_\Pi(t_n) = (t_{n+1} - t_n)\mathbf{T}\left(\frac{V_\Pi(t_{n+1})}{t_{n+1} - t_n}\right) \quad [2.10]$$

which is a functional equation for V_Π. Observe that the stationarity properties of the game induce time homogeneity:

$$V_\Pi(t_n) = (1 - t_n)V_{\Pi_{t_n}}(0). \quad [2.11]$$

By taking the linear extension of $V_\Pi(t_n)$, one can now define, for every partition Π, a function $V_\Pi(t)$ on $[0,1]$. A key lemma is the following (see Cardaliaguet et al. 2012):

Lemma 2.1. *Assume that the sequence μ_n is decreasing. Then V_Π is $2C$-Lipschitz in t, where* $C = \sup_{i,j,m} |g(i,j,m)|$.

Similarly, one can show that RG with random duration satisfies a recursive formula such as the one described above (Neyman and Sorin, 2010). Explicitly, an *uncertain duration process* $\Theta = \langle(A, \mathcal{B}, \mu), (s_n)_{n\geq 0}, \theta\rangle$ is a triple where θ is an integer-valued random variable defined on a probability space $(A, \mathcal{B}, \mu)$ with finite expectation $E(\theta)$, and each signal s_n (sent to the Players after their moves at stage n) is a measurable function defined on the probability space $(A, \mathcal{B}, \mu)$ with finite range S. An equivalent representation is through a random tree with finite expected length where the nodes at distance n correspond to the information sets at stage n. Given such a node ζ_n, known to the Players, its successor at stage $n+1$ is chosen at random according to the subtree defined by Θ at ζ_n. One can define the random iterate $\mathbf{T}^\Theta$ of the nonexpansive map $\mathbf{T}$ (see Neyman, 2003). Then, a recursive formula analogous to [2.8] holds for the value v_Θ of the game with uncertain duration Θ (see Theorem 3 in Neyman and Sorin, 2010):

$$E(\theta)v_\Theta = \mathbf{T}^\Theta(0). \quad [2.12]$$

Note that the extension to continuous time has not yet been done for RG with random duration.

2.3. ASYMPTOTIC ANALYSIS

The asymptotic analysis aims at finding conditions for: (1) the existence of the asymptotic values $\lim_{\lambda\to 0} v_\lambda$, $\lim_{n\to\infty} v_n$, and more generally $\lim_{r\to\infty} v_{\mu^r}$, (2) equality of those limits, (3) a characterization of the asymptotic value by a formula or variational inequalities expressed from the basic data of the game. Most of the results presented in this section belong to the class of finite stochastic games with incomplete information which will be our benchmark model. While the existence of the asymptotic value is still an open problem in this framework, it has been solved in some particular important subclasses.

2.3.1 Benchmark model

A zero-sum stochastic game with incomplete information is played in discrete time as follows. Let I, J, K, L, and Ω be finite sets. At stage $n = 0$, nature chooses independently $k \in K$ and $l \in L$ according to some probability distributions $p \in \Delta(K)$ and $q \in \Delta(L)$. Player 1 privately learns his type k, and Player 2 learns l. An initial state $\omega_1 = \omega \in \Omega$ is given and known to the Players. Inductively, at stage $n = 1, 2, \ldots$ knowing the past history of moves and states $h_n = (\omega_1, i_1, j_1, \ldots., i_{n-1}, j_{n-1}, \omega_n)$, Player 1 chooses $i_n \in I$ and Player 2 chooses $j_n \in J$. The payoff at stage n is $g_n = g(k, l, i_n, j_n, \omega_n)$. The new state $\omega_{n+1} \in \Omega$ is drawn according to the probability distribution $Q(i_n, j_n, \omega_n)(\cdot)$ and (i_n, j_n, ω_{n+1}) is publicly announced. Note that this model encompasses both stochastic games and repeated games with incomplete information (with standard signaling).

Let $\mathcal{F}$ denote the set of real-valued functions f on $\Delta(K) \times \Delta(L) \times \Omega$ bounded by C, concave in p, convex in q, and $2C$-Lipschitz in (p, q) for the L_1 norm, so that for every $(p_1, q_1, p_2, q_2, \omega)$ one has:

$$|f(p_1, q_1, \omega) - f(p_2, q_2, \omega)| \leq 2C(\|p_1 - p_2\|_1 + \|q_1 - q_2\|_1),$$

where $\|p_1 - p_2\|_1 = \sum_{k\in K} |p_2^k - p_2^k|$ and similarly fo q.

In this framework, *the Shapley operator* $\mathbf{T}$ associates to a function f in $\mathcal{F}$ the function:

$$\mathbf{T}(f)(p, q, \omega) = \mathtt{val}_{x\in\Delta(I)^K \times\ y\in\Delta(J)^L} \left[\begin{array}{c} g(p, q, x, y, \omega) \\ \sum_{i,j,\tilde{\omega}} \overline{x}(i)\overline{y}(j) Q(i, j, \omega)(\tilde{\omega}) f(p(i), q(j), \tilde{\omega}) \end{array} \right] \quad [2.13]$$

where $g(p, q, x, y, \omega) = \sum_{i,j,k,l} p^k q^l x^k(i) y^l(j) g(k, l, i, j, \omega)$ is the expected stage payoff.

$\mathbf{T}$ maps $\mathcal{F}$ to itself (see Laraki, 2001a, 2004). The associated *projective operator* (corresponding to the game where only the future matters) is:

$$\mathbf{R}(f)(p, q, \omega) = \mathtt{val}_{x\in\Delta(I)^K \times\ y\in\Delta(J)^L} \left[\sum_{i,j,\tilde{\omega}} \overline{x}(i)\overline{y}(j) Q(i, j, \omega)(\tilde{\omega}) f(p(i), q(j), \tilde{\omega}) \right]. \quad [2.14]$$

Any accumulation point (for the uniform norm on $\mathcal{F}$) of the equi-Lipschitz family $\{v_\lambda\}$, as λ goes to zero or of $\{v_n\}$, as n goes to infinity, is a fixed point of the projective operator. Observe, however, that any function in $\mathcal{F}$ which is independent of ω is a fixed point of $\mathbf{R}$.

2.3.2 Basic results

2.3.2.1 Incomplete information

When Ω is a singleton, the game is a repeated game with incomplete information (Aumann and Maschler, 1995). Moreover, when information is incomplete on one side (L is a singleton), Aumann and Maschler proved the existence of the asymptotic value:

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

and provided the following famous explicit formula:

$$v(p) = \texttt{Cav}_{\Delta(K)}(u)(p)$$

where:

$$u(p) = \texttt{val}_{\Delta(I)\times\Delta(J)} \sum_k p^k g(k, x, y)$$

is the value of the *nonrevealing game* and $\texttt{Cav}_C$ is the concavification operator: given ϕ, a real bounded function defined on a convex set C, $\texttt{Cav}_C(\phi)$ is the smallest concave function on C greater than ϕ.

The Aumann-Maschler proof works as follow. A splitting lemma (Zamir (1992), Proposition 3.2.) shows that if Player 1 can guarantee a function $f(p)$, he can guarantee $\texttt{Cav}_{\Delta(K)}(f)(p)$. Since Player 1 can always guarantee the value of the nonrevealing game $u(p)$ (by ignoring his information), he can also guarantee $\texttt{Cav}_{\Delta(K)}(u)(p)$. As for Player 2, by playing a best response in the nonrevealing game stage by stage given his updated belief, he can prevent Player 1 to obtain more than $\texttt{Cav}_{\Delta(K)}(u)(p)$ up to some error term which is at most, using martingale arguments, vanishing on average.

For RG of incomplete information on both sides, Mertens and Zamir (1971) proved the existence of $v = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$. They identified v as the unique solution of the system of functional equations with unknown real function ϕ on $\Delta(I) \times \Delta(J)$:

$$\phi(p, q) = \texttt{Cav}_{p\in\Delta(K)} \min\{\phi, u\}(p, q), \quad \phi(p, q) = \texttt{Vex}_{q\in\Delta(L)} \max\{\phi, u\}(p, q). \quad [2.15]$$

u is again the value of the nonrevealing game with $u(p, q) = \texttt{val}_{\Delta(I)\times\Delta(J)} \sum_{k,\ell} p^k q^\ell$ $g(k, \ell, x, y)$. The operator $u \mapsto \phi$ given by [2.15] will be called the Mertens-Zamir system and denoted by $\mathbf{MZ}$. It associates to any continuous function w on $\Delta(K) \times \Delta(L)$, a unique concave-convex continuous function $\mathbf{MZ}(w)$ (see Laraki, 2001b).

One of Mertens and Zamir's proofs is as follows. Using sophisticated reply strategies, one shows that $h = \liminf v_n$ satisfies:

$$h(p,q) \geq \mathtt{Cav}_{p\in\Delta(K)}\mathtt{Vex}_{q\in\Delta(L)} \max\{h,u\}(p,q). \qquad [2.16]$$

Define inductively dual sequences of functions $\{c_n\}$ and $\{d_n\}$ on $\Delta(K) \times \Delta(L)$ by $c_0 \equiv -\infty$ and

$$c_{n+1}(p,q) = \mathtt{Cav}_{p\in\Delta(K)}\mathtt{Vex}_{q\in\Delta(L)} \max\{c_n,u\}(p,q)$$

and similarly for $\{d_n\}$. Then they converge respectively to c and d satisfying:

$$\begin{aligned} c(p,q) &= \mathtt{Cav}_{p\in\Delta(K)}\mathtt{Vex}_{q\in\Delta(L)} \max\{c,u\}(p,q), \\ d(p,q) &= \mathtt{Vex}_{q\in\Delta(L)}\mathtt{Cav}_{p\in\Delta(K)} \min\{d,u\}(p,q) \end{aligned} \qquad [2.17]$$

A comparison principle is then used to deduce that $c \geq d$. In fact, consider an extreme point (p_0, q_0) of the convex hull of the set where $d - c$ is maximal. Then one shows that the $\mathtt{Vex}$ and $\mathtt{Cav}$ operators in the above formula [2.15] at (p_0, q_0) are trivial (there is no use of information) which implies $c(p_0,q_0) \geq u(p_0,q_0) \geq d(p_0,q_0)$. Finally $h \geq c$ implies by symmetry that $\lim_{n\to\infty} v_n$ exists.

2.3.2.2 Stochastic games

For stochastic games (K and L are reduced to a singleton), the existence of $\lim_{\lambda\to 0} v_\lambda$ in the finite case (Ω, I, J finite) was first proved by Bewley and Kohlberg (1976a) using algebraic arguments: the optimality equations for strategies and values in the Shapley operator can be written as a finite set of polynomial equalities and inequalities and thus define a semi-algebraic set in some euclidean space $\mathbb{R}^N$. By projection, v_λ has an expansion in Puiseux series hence has a limit as λ goes to 0.

An alternative, more elementary approach has been recently obtained by Oliu-Barton (2013).

The existence of $\lim_{n\to\infty} v_n$ may be deduced from $\lim_{\lambda\to 0} v_\lambda$ by comparison arguments, see Bewley and Kohlberg (1976b) or, more generally, Theorem 2.1.

2.3.3 Operator approach

The operator approach corresponds to the study of the asymptotic value trough the Shapley operator. It was first introduced by Kearns et al. (2001) in the analysis of finite absorbing games (stochastic games with a single non-absorbing state). The author uses the additional information obtained from the derivative of the Shapley operator (which is defined from $\mathbf{R}$ to itself in this case) at $\lambda = 0$ to deduce the existence of $v = \lim_{\lambda\to 0} v_\lambda = \lim_{n\to\infty} v_n$ and a characterization of v through variational inequalities.

Rosenberg and Sorin (2001) extended this approach to general RG. This tool provides sufficient conditions under which v exists and exhibits a variational characterization.

2.3.3.1 Nonexpansive monotone maps

We introduce here general properties of operators that will be applied to repeated games through the Shapley operator $\mathbf{T}$. We study iterates of an operator T mapping $\mathcal{G}$ to itself, where $\mathcal{G}$ is a subspace of the space of real bounded functions on some set E.

Assume:

1) $\mathcal{G}$ is a convex cone, containing the constants and closed for the uniform norm, denoted by $\|\cdot\|$.

2) T is monotonic: $f \geq g$ implies $Tf \geq Tg$, and translates the constants $T(f+c) = T(f)+c$.

The second assumption implies in particular that T is non-expansive.

In our benchmark model, $E = \Delta(K) \times \Delta(L) \times \Omega$ and $\mathcal{G}{=}\mathcal{F}$. In general RG, one may think of E as the set of consistent probabilities on the universal belief space.

Define the iterates and fixed points of T as:

$$V_n = T^n[0], \quad V_\lambda = T[(1-\lambda)V_\lambda]$$

and by normalizing $v_n = V_n$, $v_\lambda = \lambda V_\lambda$. Introducing the family of operators:

$$\Phi(\varepsilon, f) = \varepsilon T\left[\frac{1-\varepsilon}{\varepsilon}f\right], \tag{2.18}$$

one obtains:

$$v_n = \Phi\left(\frac{1}{n}, v_{n-1}\right), \qquad v_\lambda = \Phi(\lambda, v_\lambda) \tag{2.19}$$

and we consider the asymptotic behavior of these families of functions (analogous to the n stage and discounted values) which thus relies on the properties of $\Phi(\varepsilon,\cdot)$, as ε goes to 0.

A basic result giving a sufficient condition for the existence and equality of the limits is due to Neyman (2003):

Theorem 2.1. *If v_λ is of bounded variation in the sense that:*

$$\sum_i \|v_{\lambda_{i+1}} - v_{\lambda_i}\| < \infty, \tag{2.20}$$

for any sequence λ_i decreasing to 0. Then $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.

The result extends to random duration processes ($\lim_{\mathsf{E}(\theta)\to+\infty} v_\Theta = \lim_{\lambda\to 0} v_\lambda$) when the expected duration of the current stage decreases along the play (Neyman and Sorin, 2010).

Following Rosenberg and Sorin (2001), Sorin (2004) defines spaces of functions that will correspond to upper and lower bounds on the families of values.

(i) *Uniform domination.* Let $\mathcal{L}^+$ be the space of functions $f \in \mathcal{G}$ that satisfy: there exists $M_0 \geq 0$ such that $M \geq M_0$ implies $T(Mf) \leq (M+1)f$. ($\mathcal{L}^-$ is defined similarly.)

(ii) *Pointwise domination.* When the set E is not finite, one can introduce the larger class $\mathcal{S}^+$ of functions satisfying $\limsup_{M\to\infty}\{T(Mf)(e) - (M+1)f(e)\} \leq 0$, $\forall e \in E$. ($\mathcal{S}^-$ is defined in a dual way).

The comparison criteria for uniform domination is expressed by the following result:

Theorem 2.2. *Rosenberg and Sorin (2001)*
If $f \in \mathcal{L}^+$, $\limsup_{n\to\infty} v_n$ and $\limsup_{\lambda\to 0} v_\lambda$ are less than f.
Consequently, if the intersection of the closure of $\mathcal{L}^+$ and $\mathcal{L}^-$ is not empty, then both $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ exist and coincide.

And for pointwise domination by the next property:

Theorem 2.3. *Rosenberg and Sorin (2001)*
Assume E compact. Let $\mathcal{S}_0^+$ (resp. $\mathcal{S}_0^-$) be the space of continuous functions in $\mathcal{S}^+$ (resp. $\mathcal{S}^-$). Then, for any $f^+ \in \mathcal{S}_0^+$ and $f^- \in \mathcal{S}_0^-$, $f^+ \geq f^-$.
Consequently, the intersection of the closures of $\mathcal{S}_0^+$ and $\mathcal{S}_0^-$ contains at most one point.

These two results provide sufficient conditions for the uniqueness of a solution satisfying the properties. The next one gives a sufficient condition for the existence of a solution.

T has the *derivative property* if for every $f \in \mathcal{G}$ and $e \in E$:

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

exists in $\overline{\mathbb{R}}$. If such a derivative exists, $\mathcal{S}^+$ is the set of functions $f \in \mathcal{G}$ that satisfy $\varphi^*(f)(e) \leq 0$, for all $e \in E$ and similarly for $\mathcal{S}^-$, $\varphi^*(f)(e) \geq 0$.

Theorem 2.4. *Rosenberg and Sorin (2001)*
Assume that T has the derivative property and E is compact. Let f be such that φ^* "changes sign" at f, meaning that there exist two sequences $\{f_n^-\}$ and $\{f_n^+\}$ of continuous functions converging to f such that $\varphi^*(f_n^-) \leq 0 \leq \varphi^*(f_n^+)$. Then, f belongs to the closures of $\mathcal{S}_0^+$ and $\mathcal{S}_0^-$.

Definition 2.2. *T has the recession property if $\lim_{\varepsilon\to 0} \Phi(\varepsilon, f)(\theta) = \lim_{\varepsilon\to 0} \varepsilon T(\frac{f}{\varepsilon})(\theta)$, written $R(f)(\theta)$, exists.*

Theorem 2.5. *Vigeral (2010b)*

Assume that T has the recession property and is convex. Then v_n (resp. v_λ) has at most one accumulation point.

The proof uses the inequality: $R(f+g) \leq T(f) + R(g)$ and relies on properties of the family of operators T_m defined by:

$$T_m(f) = \frac{1}{m} T^m(mf). \qquad [2.21]$$

2.3.3.2 Applications to RG

The Shapley operator $\mathbf{T}$ satisfies the derivative and recession properties, so the results of the previous section can be applied.

Absorbing games are stochastic games where the state can change at most once on a play Kearns et al. (2001).

Theorem 2.6. *Rosenberg and Sorin (2001)*

$\lim_{n\to\infty} v_n$ *and* $\lim_{\lambda\to 0} v_\lambda$ *exist and are equal in absorbing games with compact action spaces.*

Recursive games are stochastic games where the payoff is 0 until the state becomes absorbing (Everett, 1957).

Theorem 2.7. *Sorin (2003), Vigeral (2010c)*

$\lim_{n\to\infty} v_n$ *and* $\lim_{\lambda\to 0} v_\lambda$ *exist and are equal in recursive games with finite state space and compact action spaces.*

Notice that the algebraic approach (for stochastic games) cannot be used when action or state spaces are not finite. However, one cannot expect to get a proof for general stochastic games with finite state space, see Vigeral's counterexample in Section 9.

Pointwise domination is used to prove existence and equality of $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ through the derived game φ^* and the recession operator $\mathbf{R}$ in the following cases.

Theorem 2.8. *Rosenberg and Sorin (2001)*

$\lim_{n\to\infty} v_n$ *and* $\lim_{\lambda\to 0} v_\lambda$ *exist and are equal in repeated games with incomplete information.*

This provides an alternative proof of the result of Mertens and Zamir (1971). One shows that any accumulation point w of the family $\{v_\lambda\}$ (resp. $\{v_n\}$) as $\lambda \to 0$ (resp. $n \to \infty$) belongs to the closure of $\mathcal{S}_0^+$, hence, by symmetry, the existence of a limit follows using Theorem 2.3. More precisely, Theorem 2.4 gives the following characterization of the asymptotic value: given a real function f on a linear set X, denote by $\mathcal{E}f$ the projection on X of the extreme points of its epigraph. Then: $v = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$ exists and is the unique saddle continuous function satisfying both inequalities:

$$\mathbf{Q1} : p \in \mathcal{E}v(\cdot, q) \Rightarrow v(p,q) \le u(p,q), \mathbf{Q2} : q \in \mathcal{E}v(p,\cdot) \Rightarrow v(p,q) \ge u(p,q), \qquad [2.22]$$

where u is the value of the nonrevealing game. One then shows that this is equivalent to the characterization of Mertens and Zamir [2.15] (see also Laraki (2001a) and Section 3.4).

All the results above also hold for (decreasing) random duration processes (Neyman and Sorin, 2010).

Theorem 2.9. *Rosenberg (2000)*

$\lim_{n\to\infty} v_n$ *and* $\lim_{\lambda\to 0} v_\lambda$ *exist and are equal in finite absorbing games with incomplete information on one side.*

This is the first general subclass where both stochastic and information aspects are present and in which the asymptotic value, $\lim_{\lambda\to 0} v_\lambda = \lim_{n\to\infty} v_n$ exists.

Theorem 2.10. *Vigeral (2010b)*

$\lim_{n\to\infty} v_n$ *(resp.* $\lim_{\lambda\to 0} v_\lambda$*) exists in repeated games where one Player controls the transition and the family* $\{v_n\}$ *(resp.* $\{v_\lambda\}$*) is relatively compact.*

This follows from the convexity of $\mathbf{T}$ in that case. It applies in particular to dynamic programming or games with incomplete information on one side (see also Sections 5.3 and 5.4).

Vigeral (2010b) provides a simple stochastic game in which the sets $\mathcal{L}^+$ and $\mathcal{L}^-$ associated to $\mathbf{T}$ do not intersect, but the sets $\mathcal{L}^+_m$ and $\mathcal{L}^-_m$ associated to the operator $\mathbf{T}_m$ do intersect for some m large enough. This suggests that, for games in the benchmark model, the operator approach should be extended to iterations of the Shapley operator.

2.3.4 Variational approach

Inspired by the tools used to prove existence of the value in differential games (see Section 6), Laraki (2001a,b, 2010) introduced the variational approach to obtain existence of $\lim v_\lambda$ in a RG and to provide a characterization of the limit via variational inequalities. Rather than going to the limit between time t and $t+h$ in the dynamic programming equation, one uses the fact that v_λ is a fixed point of the Shapley operator [2.6] (as in the operator approach). Given an optimal stationary strategy of the maximizer (resp. minimizer) in the λ-discounted game, one deduces an inequality (involving both value and strategy) that must be satisfied as λ goes to zero. Finally, a comparison principle is used to deduce that only one function satisfies the two inequalities (which have to be sharp enough to specify a single element).

The extension of the variational approach to $\lim v_n$ or $\lim v_\mu$ was solved recently by Cardaliaguet et al. (2012) by increasing the state space to $\Omega \times [0,1]$, that is, by introducing time as a new variable, and viewing each evaluation as a particular discretization of the time interval $[0,1]$ (Section 2.3). From [2.10], one shows that accumulation points are viscosity solutions of a related differential equation and finally, comparison tools give uniqueness.

To understand the approach (and the exact role of the time variable), we first describe it for discounted games and then for general evaluations in three classes: RG with incomplete information, absorbing games, and splitting games.

2.3.4.1 Discounted values and variational inequalities

(A) RG with incomplete information. We follow Laraki (2001a) (see also Cardaliaguet et al., 2012). To prove (uniform) convergence, it is enough to show that $\mathcal{V}_0$, the set of accumulation points of the family $\{v_\lambda\}$, is a singleton. Let $\mathcal{V}$ be the set of fixed points of the projective operator $\mathbf{R}$ [2.14] and observe that $\mathcal{V}_0 \subset \mathcal{V}$.

Given $w \in \mathcal{V}_0$, denote by $\mathbf{X}(p,q,w) \subseteq \mathbf{X} = \Delta(I)^K$ the set of optimal strategies for Player 1 (resp. $\mathbf{Y}(p,q,w) \subseteq \mathbf{Y} = \Delta(J)^L$ for Player 2) in the projective operator $\mathbf{R}$ for w at (p,q). A strategy $x \in \mathbf{X}$ of Player 1 is called nonrevealing at p, $x \in NR_{\mathbf{X}}(p)$ if $p(i) = p$ for all $i \in I$ with $x(i) > 0$ and similarly for $y \in \mathbf{Y}$. The value of the nonrevealing game is given by:

$$u(p,q) = \mathtt{val}_{NR_{\mathbf{X}}(p)\times NR_{\mathbf{Y}}(q)} g(p,q,x,y)\ . \qquad [2.23]$$

Lemma 2.2. *Any $w \in \mathcal{V}_0$ satisfies:*

$$\textbf{\textit{P1}}: \quad \text{If } \mathbf{X}(p,q,w) \subset NR_{\mathbf{X}}(p) \text{ then } w(p,q) \le u(p,q) \qquad [2.24]$$

$$\textbf{\textit{P2}}: \quad \text{If } \mathbf{Y}(p,q,w) \subset NR_{\mathbf{Y}}(q) \text{ then } w(p,q) \ge u(p,q). \qquad [2.25]$$

The interpretation is straightforward. If by playing optimally a Player should not use asymptotically his information, he cannot get more that the value of the nonrevealing game, because the other Player has always the option of playing nonrevealing. What is remarkable is that the two above properties, plus geometry (concavity in p, convexity in q) and smoothness (continuity) are enough to characterize the asymptotic value.

The proof is simple. Let $v_{\lambda_n} \to w$ and x_n optimal for the maximizer in the Shapley equation $v_{\lambda_n}(p,q) = \Phi(\lambda_n, v_{\lambda_n})(p,q)$. Then, for every nonrevealing pure strategy j of Player 2, one has:

$$v_{\lambda_n}(p,q) \leq \lambda_n g(x_n, j, p, q) + (1-\lambda_n)\sum_i \bar{x}_n(i) v_{\lambda_n}(p_n(i), q).$$

Jensen's inequality implies $\sum_i \bar{x}_n(i) v_{\lambda_n}(p_n(i), q) \leq v_{\lambda_n}(p,q)$ hence $v_{\lambda_n}(p,q) \leq g(x_n, j, p, q)$. Thus, if $\bar{x} \in \mathbf{X}(p,q,w)$ is an accumulation point of $\{x_n\}$, one obtains $w(p,q) \leq g(\bar{x}, j, p, q)$, $\forall j \in J$. Since by assumption, $\mathbf{X}(p,q,w) \subset NR_{\mathbf{X}}(p)$, one gets $w(p,q) \leq u(p,q)$.

For uniqueness, the following comparison principle is established:

Lemma 2.3. *Let w_1 and w_2 be in $\mathcal{V}$ satisfying P_1 and P_2 respectively, then $w_1 \leq w_2$.*

The proof follows an idea by Mertens and Zamir (see Section 3.2.1.). Let (p_0, q_0) be an extreme point of the compact set where the difference $(w_1 - w_2)(p,q)$ is maximal. Then, one has that $\mathbf{X}(p_0, q_0, w_1) \subset NR_{\mathbf{X}}(p_0)$, and $\mathbf{Y}(p_0, q_0, w_2) \subset NR_{\mathbf{Y}}(q_0)$. In fact, both functions being concave, w_1 has to be strictly concave at p_0 (if it is an interior point) and the result follows (see the relation with [2.22] **Q1**, **Q2**). Thus, $w_1(p_0, q_0) \leq u(p_0, q_0) \leq w_2(p_0, q_0)$.

Consequently one obtains:

Theorem 2.11. $\lim_{\lambda\to 0} v_\lambda$ *exists and is the unique function in $\mathcal{V}$ that satisfies **P1** and **P2**.*

This characterization is equivalent to the Mertens-Zamir system (Laraki, 2001a; Rosenberg and Sorin, 2001) and to the two properties **Q1** and **Q2** established above [2.22].

(B) Absorbing games. We follow Laraki (2010). Recall that only one state is nonabsorbing, hence in the other states one can assume without loss of generality that the payoff is constant (and equal to the value). The game is thus defined by the following elements: two finite sets I and J, two (payoff) functions f, g from $I \times J$ to $\mathbb{R}$ and a function π from $I \times J$ to $[0,1]$: $f(i,j)$ is the current payoff, $\pi(i,j)$ is the probability of nonabsorption, and $g(i,j)$ is the absorbing payoff. Define $\pi^*(i,j) = 1 - \pi(i,j)$, $f^*(i,j) = \pi^*(i,j) \times g(i,j)$, and extend bilinearly any $\varphi : I \times J \to \mathbf{R}$ to $\mathbf{R}^I \times \mathbf{R}^J$ as usual: $\varphi(\alpha, \beta) = \sum_{i\in I, j\in J} \alpha^i \beta^j \varphi(i,j)$.

The variational approach proves the following new characterization of $\lim_{\lambda\to 0} v_\lambda$ as the value of a strategic one-shot game that can be interpreted as a "limit" game (see Section 8).

Theorem 2.12. *As $\lambda \to 0$, v_λ converges to v given by:*

$$v = \mathtt{val}_{((x,\alpha),(y,\beta))\in(\Delta(I)\times\mathbf{R}^I_+)\times(\Delta(J)\times\mathbf{R}^J_+)} \frac{f(x,y) + f^*(\alpha, y) + f^*(x,\beta)}{1 + \pi^*(\alpha, y) + \pi^*(x, \beta)} . \qquad [2.26]$$

Actually, if x_λ is an optimal stationary strategy for Player 1 in the λ-discounted game, then: $v_\lambda \leq \lambda f(x_\lambda, j) + (1-\lambda)(\pi(x_\lambda, j)v_\lambda + f^*(x_\lambda, j), \forall j \in J$. Accordingly:

$$v_\lambda \leq \frac{f(x_\lambda, j) + f^*(\frac{(1-\lambda)x_\lambda}{\lambda}, j)}{1 + \pi^*(\frac{(1-\lambda)x_\lambda}{\lambda}, j)}, \quad \forall j \in J. \qquad [2.27]$$

Consequently, given an accumulation point w of $\{v_\lambda\}$, there exists $\overline{x} \in \Delta(I)$ an accumulation point of $\{x_\lambda\}$ such that for all $\varepsilon > 0$, there exists $\overline{\alpha} \in \mathbf{R}_+^I$ $\left(\text{of the form } \frac{(1-\overline{\lambda})x_{\overline{\lambda}}}{\overline{\lambda}}\right)$ satisfying:

$$w \leq \frac{f(\overline{x}, j) + f^*(\overline{\alpha}, j)}{1 + \pi^*(\overline{\alpha}, j)} + \varepsilon, \quad \forall j \in J. \qquad [2.28]$$

By linearity, the last inequality extends to any $y \in \Delta(J)$. On the other hand, w is a fixed point of the projective operator and $\overline{x}$ is optimal there, hence:

$$w \leq \pi(\overline{x}, y)\, w + f^*(\overline{x}, y), \qquad \forall y \in \Delta(J). \qquad [2.29]$$

Inequality [2.29] is linear thus extends to $\beta \in \mathbf{R}_+^J$ and combining with [2.28] one obtains that w satisfies the inequality:

$$w \leq \sup_{(x,\alpha) \in \Delta(I) \times \mathbf{R}_+^I} \ \inf_{(y,\beta) \in \Delta(J) \times \mathbf{R}_+^J} \frac{f(x, y) + f^*(\alpha, y) + f^*(x, \beta)}{1 + \pi^*(\alpha, y) + \pi^*(x, \beta)}. \qquad [2.30]$$

Following the optimal strategy of Player 2 yields the reverse inequality. Uniqueness and characterization of $\lim v_\lambda$ is deduced from the fact that $\sup\inf \leq \inf\sup$.

2.3.4.2 General RG and viscosity tools

We follow Cardaliaguet et al. (2012).

(A) RG with incomplete information. Consider an arbitrarily evaluation probability μ on $\mathbb{N}^*$ inducing a partition Π. Let $V_\Pi(t_k, p, q)$ be the value of the game starting at time t_k. One has $V_\Pi(1, p, q) := 0$ and:

$$V_\Pi(t_n, p, q) = \mathtt{val}[\mu_{n+1} g(x, y, p, q) + \sum_{i,j} x(i)y(j)\, V_\Pi(t_{n+1}, p(i), q(j))]. \qquad [2.31]$$

Moreover V_Π belongs to $\mathcal{F}$. Given a sequence $\{\mu^m\}$ of decreasing evaluations ($\mu_n^m \geq \mu_{n+1}^m$), Lemma 2.1 implies that the family of $V_{\Pi(m)}$ associated to partitions $\Pi(m)$ is equi-Lipschitz.

Let $\mathcal{T}_0$ be the nonempty set of accumulation points (for the uniform convergence) as $\mu_1^m \to 0$. Let $\mathcal{T}$ be the set of real continuous functions W on $[0,1] \times \Delta(K) \times \Delta(L)$ such that for all $t \in [0,1]$, $W(t, ., .) \in \mathcal{V}$. Recall that $\mathcal{T}_0 \subset \mathcal{T}$. $\mathbf{X}(t, p, q, W)$ is the set of optimal strategies for Player 1 in $\mathbf{R}(W(t, ., .))$ and $\mathbf{Y}(t, p, q, W)$ is defined accordingly. Define two properties for a function $W \in \mathcal{T}$ and a $\mathcal{C}^1$ test function $\phi : [0,1] \to \mathbb{R}$ as follows:

P1$'$: If $t \in [0,1)$ is such that $\mathbf{X}(t,p,q,W)$ is nonrevealing and $W(\cdot,p,q) - \phi(\cdot)$ has a global maximum at t, then $u(p,q) + \phi'(t) \geq 0$.

P2$'$: If $t \in [0,1)$ is such that $\mathbf{Y}(t,p,q,W)$ is nonrevealing and $W(\cdot,p,q) - \phi(\cdot)$ has a global minimum at t then $u(p,q) + \phi'(t) \leq 0$.

The variational counterpart of Lemma 2.2 is obtained by using [2.31].

Lemma 2.4. *Any $W \in \mathcal{T}_0$ satisfies* ***P1′*** *and* ***P2′****.*

The proof is as follows. Let (t,p,q) as is **P1′**. Adding the function $s \mapsto (s-t)^2$ to ϕ if necessary, we can assume that this global maximum is strict. Let $W_{\mu^k} \to W$. Let Π^k be the partition associated to μ^k and $t^k_{n(k)}$ be a global maximum of $W_{\mu^k}(\cdot,p,q) - \phi(\cdot)$ on Π^k. Since t is a strict maximum, one has $t^k_{n(k)} \to t$, as $k \to \infty$.

Proceeding as in Section 2.3.4.1. one obtains with $x_k \in \mathbf{X}$ being optimal for W_{μ^k}:

$$0 \leq g(x_k,j,p,q)\left(t^k_{n(k)+1} - t^k_{n(k)}\right) + \left[W_{\mu^k}\left(t^k_{n(k)+1},p,q\right) - W_{\mu^k}\left(t^k_{n(k)},p,q\right)\right].$$

Since $t^k_{n(k)}$ is a global maximum of $W_m(\cdot,p,q) - \phi(\cdot)$ on Π^k one deduces:

$$0 \leq g(x_k,j,p,q)\left(t^k_{n(k)+1} - t^k_{n(k)}\right) + \left[\phi\left(t^k_{n(k)+1},p,q\right) - \phi\left(t^k_{n(k)},p,q\right)\right].$$

so $0 \leq g(\bar{x},j,p,q) + \phi'(t)$ for any accumulation point $\bar{x}$ of $\{x_k\}$. The result follows because $\bar{x}$ is nonrevealing by **P1′**. Uniqueness is proved by the following comparison principle:

Lemma 2.5. *Let W_1 and W_2 in $\mathcal{T}$ satisfy* **P1′**, **P2′**, *and also the terminal condition:*

$$W_1(1,p,q) \leq W_2(1,p,q), \quad \forall (p,q) \in \Delta(K) \times \Delta(L).$$

Then $W_1 \leq W_2$ on $[0,1] \times \Delta(K) \times \Delta(L)$.

Consequently, we obtain:

Theorem 2.13. *V_π converges uniformly to the unique point $W \in \mathcal{T}$ that satisfies the variational inequalities* **P1′** *and* **P2′** *and the terminal condition $W(1,p,q) = 0$. In particular, $v_\mu(p,q)$ converges uniformly to $v(p,q) = W(0,p,q)$ and $W(t,p,q) = (1-t)v(p,q)$, where $v = \mathbf{MZ}(u)$.*

To summarize the idea of the proof, (A) one uses viscosity solutions (smooth majorant or minorant) to obtain first-order conditions on the accumulation points of the sequence of values and on the corresponding optimal strategies. (B) One then considers two accumulation points (functions) and establishes property of an extreme point of the set of states where their difference is maximal. At this point, the optimal strategies have specific aspects (C) that imply that (A) gives uniqueness.

On the other hand, the variational approach proof is much stronger than needed. Continuous time is used as a tool to show eventually, that the asymptotic value $W(t) = (1-t)v$ is linear in t. However, one must first keep in mind that linearity is a valid conclusion only if the existence of the limit is known (which is the statement that needs to be shown). Second, if g in equation [2.31] is time-dependent in the game on $[0,1]$, the same proof and characterization still work (see Section 8.3.3. for continuous-time games with incomplete information).

(B) Absorbing games. Consider a decreasing evaluation $\mu = \{\mu_n\}$. Denote by v_μ the value of the associated absorbing game. Let $W_\mu(t_m)$ be the value of the game starting at time t_m defined recursively by $W_\mu(1) = 0$ and:

$$W_\mu(t_m) = \mathtt{val}_{(x,y)\in\Delta(I)\times\Delta(J)}\left[\mu_{m+1} f(x,y) + \pi(x,y) W_\mu(t_{m+1}) + (1-t_{m+1}) f^*(x,y)\right]. \tag{2.32}$$

Under monotonicity of μ, the linear interpolation of W_μ is a $2C$-Lipschitz continuous in $[0,1]$. Set for any $(t,a,b,x,\alpha,y,\beta) \in [0,1]\times\mathbb{R}\times\mathbb{R}\times\Delta(I)\times\mathbb{R}^I_+\times\Delta(J)\times\mathbb{R}^J_+$,

$$h(t,a,b,x,\alpha,y,\beta) = \frac{f(x,y) + (1-t)[f^*(\alpha,y) + f^*(x,\beta)] - \left[\pi^*(\alpha,y) + \pi^*(x,\beta)\right] a + b}{1 + \pi^*(\alpha,y) + \pi^*(x,\beta)}.$$

Define the Hamiltonian of the game as:

$$H(t,a,b) = \mathtt{val}_{((x,\alpha),(y,\beta))\in(\Delta(I)\times\mathbf{R}^I_+)\times(\Delta(J)\times\mathbf{R}^J_+)} h(t,a,b,x,\alpha,y,\beta).$$

Theorem 2.12 implies that H is well defined (the value exists).

Define two variational inequalities for a continuous function U on [0,1] as follows: for all $t \in [0,1)$ and any $\mathcal{C}^1$ function $\phi : [0,1] \to \mathbb{R}$:

R1: If $U(\cdot) - \phi(\cdot)$ admits a global maximum at $t \in [0,1)$ then $H^-(t, U(t), \phi'(t)) \geq 0$.

R2: If $U(\cdot) - \phi(\cdot)$ admits a global minimum at $t \in [0,1)$ then $H^+(t, U(t), \phi'(t)) \leq 0$.

Lemma 2.6. *Any accumulation point $U(\cdot)$ of $W_\mu(\cdot)$ satisfies **R1** and **R2**.*

That is, U is a viscosity solution of the HJB equation $H(t, U(t), \nabla U(t)) = 0$. The comparison principle is as one expects:

Lemma 2.7. *Let U_1 and U_2 be two continuous functions satisfying **R1-R2** and $U_1(1) \leq U_2(1)$. Then $U_1 \leq U_2$ on $[0,1]$.*

Consequently, we get:

Theorem 2.14. *W_μ converges to the unique C Lipschitz solution to* ***R1–R2*** *with $U(1) = 0$. Consequently, $U(t) = (1 - t)v$ and v_μ converges to v given by [2.26].*

(C) Splitting games. In RG with incomplete information on the one side, the use of information has an impact on the current payoff and on the state variable. However, the difference between the payoff of the mean strategy and the expected payoff is proportional to the L^1 variation of the martingale of posteriors at that stage hence on the average vanishes (Zamir, 1992, Propositions 3.8, 3.13, and 3.14). One can assume thus that the informed Player is playing optimally in the non-revealing game and the remaining aspect is the strategic control of the state variable: the martingale of posteriors. This representation, introduced in Sorin (2002) p. 50, leads to a continuous time martingale maximization problem $\max_{p_t} \int_0^1 u(p_t)\mathrm{d}t$ (where the max is over all càdlàg martingales $\{p_t\}$ on $\Delta(K)$ with $p_0 = p$). The value is clearly $Cav_{\Delta(K)}(u)(p)$. Moreover, the optimal strategy in the RG that does the splitting at the beginning is maximal in the continuous time problem and inversely, any maximal martingale is ϵ-optimal in the RG for a sufficiently fine time partition.

Since then, this representation has been extended to nonautonomous continuous time games with incomplete information on the one side (see Section 9.4).

The search for an analog representation for the asymptotic value of RG with incomplete information on *both* sides naturally leads to the splitting game (Sorin, 2002, p. 78). It is a stochastic game where each Player controls a martingale. A 1-Lipschitz continuous stage payoff function U from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$ is given and the RG is played as follows. At stage m, knowing the state variable (p_m, q_m) in $\Delta(K) \times \Delta(L)$, Player 1 chooses p_{m+1} according to some $\theta \in M_{p_m}^K$ and Player 2 chooses q_{m+1} according to some $\nu \in M_{q_m}^L$ where M_p^K stands for the set of probabilities on $\Delta(K)$ with expectation p (and similarly for M_q^L). The current payoff is $U(p_{m+1}, q_{m+1})$ and the new state (p_{m+1}, q_{m+1}) is announced to both Players.

Existence of $\lim v_\lambda$ and its identification as $\mathbf{MZ}(U)$ was proved in that case by Laraki (2001a). One shows that the associated Shapley operator (called the *splitting operator*):

$$\mathbf{T}(f)(p,q) = \mathtt{val}_{\theta \in M_p^K \times \nu \in M_q^L} \int_{\Delta(K)\times\Delta(L)} [U(p',q') + f(p',q')]\theta(\mathrm{d}p')\nu(\mathrm{d}q') \qquad [2.33]$$

preserves 1-Lipschitz continuity (for the L^1 norm on $\Delta(K) \times \Delta(L)$): note the difficulty is due to the fact that the strategy sets depend on the state. This result is a consequence of a splitting lemma proved in Laraki (2004), and guarantees that the discounted values form a family of equi-Lipschitz functions, so one can mimic the variational approach proof as in RG with incomplete information on both sides. For the extension of preservation of equi-Lipschitz continuity in MDP see Laraki and Sudderth (2004).

The extension to $\lim v_\mu$ for general evaluation μ appears in Cardaliaguet et al. (2012) using the same time extension trick as in the previous paragraph (A) starting from the Shapley equation:

$$V_\Pi(t_n, p, q) = \mathtt{val}_{\theta \in M_p^K \times \nu \in M_q^L} \int_{\Delta(K)\times\Delta(L)} [\mu_{n+1} U(p', q') + \ V_\Pi(t_{n+1}, p', q')]\theta(\mathrm{d}p')\nu(\mathrm{d}q'). \quad [2.34]$$

Laraki (2001b) extends the definition of the splitting game from $\Delta(K)$ and $\Delta(L)$ to any convex compact subsets C and D of $\mathbf{R}^n$ and uses it to show existence and uniqueness of a concave-convex continuous function ϕ on $C \times D$ that satisfies the Mertens-Zamir system: $\phi = \mathtt{Cav}_C \min(\phi, U) = \mathtt{Vex}_D \max(\phi, U)$ under regularity assumptions on C and D. Namely, they need to be face-closed (FC): the limiting set of any converging sequence of faces of C is also a face of C.

The FC condition on a set C is necessary and sufficient to guarantee that $\mathtt{Cav}_C\psi$ is continuous for every continuous function ψ (Laraki, 2004), hence for the splitting operator to preserve continuity (Laraki, 2001b). The proof of convergence of $\lim v_\lambda$ is more technical than on $\Delta(K) \times \Delta(L)$ and uses epi-convergence technics because the family $\{v_\lambda\}$ is not equi-Lipschitz.

The problem of existence and uniqueness of a solution of the Mertens-Zamir system without the FC condition or continuity of U remains open.

The above analysis extends to the non-autonomous case where U depends on t (Cardaliaguet et al., 2012). Any partition Π of $[0, 1]$ induces a value function $V_\Pi(t, p, q)$ that converges uniformly to the unique function satisfying:

P1$''$: If $t \in [0, 1)$ is such that $\mathbf{X}(t, p, q, W)$ is nonrevealing and $W(\cdot, p, q) - \phi(\cdot)$ has a global maximum at t, then $U(t, p, q) + \phi'(t) \geq 0$.

P2$''$: If $t \in [0, 1)$ is such that $\mathbf{Y}(t, p, q, W)$ is nonrevealing and $W(\cdot, p, q) - \phi(\cdot)$ has a global minimum at t, then $U(t, p, q) + \phi'(t) \leq 0$.

2.3.4.3 Compact discounted games and comparison criteria

An approach related to the existence of the asymptotic value for discounted games has been proposed in Sorin and Vigeral (2012). The main tools used in the proofs are:

- the fact that the discounted value v_λ satisfies the Shapley equation [2.26],
- properties of accumulation points of the discounted values, and of the corresponding optimal strategies,
- comparison of two accumulation points leading to uniqueness and characterization.

In particular, this allows to cover the case of absorbing and recursive games with compact action spaces and provides an alternative formula for the asymptotic value of absorbing games, namely:

$$v = \mathtt{val}_{X\times Y} W(x,y)$$

with:

$$W(x,y) = med\{f(x,y), \sup_{x', \pi^*(x',y)>0} \bar{f}^*(x',y), \inf_{y',\pi^*(x,y')>0} \bar{f}^*(x,y')\}$$

where $\bar{f}^*(x,y)$ is the expected absorbing payoff: $\pi^*(x,y)\bar{f}^*(x,y) = f^*(x,y)$.

2.4. THE DUAL GAME

In this section, we focus on repeated games with incomplete information. The dual game has been introduced by De Meyer (1996a,b) and leads to many applications.

2.4.1 Definition and basic results

Consider a two-person zero-sum game with incomplete information on one side defined by two sets of actions S and T, a finite parameter space K, a probability distribution $p \in \Delta(K)$, and for each $k \in K$ a real-valued 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 revealed to Player 1 (the maximizer) while Player 2 only knows p. In the normal form, Player 1 chooses $\mathbf{s} = \{s^k\}$ in S^K, Player 2 chooses t in T and the payoff is $G^p(\mathbf{s},t) = \sum_k p^k G^k(s^k,t)$. Let $\underline{v}(p) = \sup_{S^K} \inf_T G^p(\mathbf{s},t)$ and $\overline{v}(p) = \inf_T \sup_{S^K} G^p(\mathbf{s},t)$. Then both value functions are concave in p, the first thanks to the splitting procedure (see, e.g., Zamir, 1992, p. 118) and the second as an infimum of linear functions.

Following De Meyer (1996a,b), one defines for each $z \in \mathbb{R}^k$, the "dual game" $G^*(z)$, where Player 1 chooses $k \in K$ and $s \in S$, while Player 2 chooses $t \in T$ and the payoff is:

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

Define by $\underline{w}(z)$ and $\overline{w}(z)$ the corresponding maxmin and minmax.

Theorem 2.15. *De Meyer (1996a,b), Sorin (2002)*

The following duality relations hold:

$$\underline{w}(z) = \max_{p\in\Delta(K)} \{\underline{v}(p) - \langle p,z\rangle\}, \qquad \underline{v}(p) = \inf_{z\in\mathbb{R}^K} \{\underline{w}(z) + \langle p,z\rangle\}. \qquad [2.35]$$

$$\overline{w}(z) = \max_{p\in\Delta(K)} \{\overline{v}(p) - \langle p,z\rangle\}, \qquad \overline{v}(p) = \inf_{z\in\mathbb{R}^K} \{\overline{w}(z) + \langle p,z\rangle\}. \qquad [2.36]$$

In terms of strategies, one obtains the following correspondences:

Corollary 2.1.

1) *Given z (and $\varepsilon \geq 0$), let p attain the maximum in (2.35) and let* **s** *be ε-optimal in G^p, then $(p, \mathbf{s})$ is ε-optimal in $G^*(z)$.*

2) *Given p (and $\varepsilon \geq 0$), let z attain the infimum up to ε in (2.36) and let t be ε-optimal in $G^*(z)$, then t is also 2ε-optimal in G^p.*

To see the link with approachability (Section 7 and Sorin, 2002 Chapter 2) define

$$B = \{z \in \mathbb{R}^k : \ \overline{v}(p) \leq \langle p, z\rangle, \forall p \in \Delta(K)\}.$$

Then, B is the set of "reachable" vectors for Player 2 in the sense that for any $z \in B$, and any $\varepsilon > 0$, there exists a $t \in T$ such that $\sup_s G^k(s, t) \leq z^k + \varepsilon$, $\forall k \in K$. In particular $\overline{w}(z) \leq 0$ if and only if $z \in B$, which means that, from his point of view, the uninformed Player 2 is playing in a game with "vector payoffs."

2.4.2 Recursive structure and optimal strategies of the noninformed player

Consider now a RG with incomplete information on one side and recall the basic recursive formula for G_n:

$$(n+1)\ v_{n+1}(p) = \max_{x \in X^K} \min_{y \in Y} \left\{ \sum_k p^k g(k, x^k, y) + n \sum_i \bar{x}(i)\ v_n(p(i)) \right\} \qquad [2.37]$$

Let us consider the dual game G_n^* and its value w_n which satisfies:

$$w_n(z) = \max_{p \in \Delta(K)} \{v_n(p) - \langle p, z\rangle\}.$$

This leads to the dual recursive equation (De Meyer, 1996b):

$$(n+1) w_{n+1}(z) = \min_{y \in Y} \max_{i \in I} n w_n \left(\frac{n+1}{n} z - \frac{1}{n} g(i, y) \right). \qquad [2.38]$$

In particular, Player 2 has an optimal strategy in $G_{n+1}^*(z)$ that depends only on z (and on the length of the game). At stage 1, he plays y optimal in the dual recursive equation and from stage 2 on, given the move i_1 of Player 1 at stage 1, plays optimally in $G_n^*(\frac{n+1}{n} z - \frac{1}{n} g(i_1, y))$. A similar result holds for any evaluation μ of the stream of payoffs. Thus z is the natural state variable for Player 2.

Recall that the recursive formula for the primal game (2.37) allows the informed Player to construct inductively an optimal strategy since he knows $p(i)$. This is not the case for Player 2, who cannot compute $p(i)$, hence the first interest of the dual game is to obtain an explicit algorithm for optimal strategies of the uninformed Player via Corollary 2.1.

These properties extend to RG with incomplete information on both sides. In such a game, Player 2 must consider all possible realizations of $k \in K$ and so plays in a game with vector payoffs in $\mathbb{R}^K$. On the other hand, he reveals information and so generates a martingale $\tilde{q}$ on $\Delta(L)$.

There are thus two dual games: the Fenchel conjugate with respect to p (resp. q) allows to compute an optimal strategy for Player 2 (resp. 1). In the first dual, from $w_n(z, q) = \max_{p \in \Delta(K)} \{v_n(p, q) - \langle p, z \rangle\}$, De Meyer and Marino (2005) deduce that

$$(n+1) w_{n+1}(z, q) = \min_{y, \{z_{i,j}\}} \max_{i \in I} \sum_j \overline{y}(j) n w_n(z_{i,j}, q(j)) \qquad [2.39]$$

where the minimum is taken over all $y \in \Delta(J)^L$ and $z_{i,j} \in \mathbb{R}^K$ such that $\sum_j \overline{y}(j) z_{i,j} = \frac{n+1}{n} z - \frac{1}{n} g(q, i, y)$. Hence, Player 2 has an optimal strategy, which is Markovian with respect to (z, q).

A similar conclusion holds for stochastic games with incomplete information on both sides where Player 2 has a optimal strategy Markovian with respect to (z, q, ω) (Rosenberg, 1998).

To summarize, to a game with incomplete information on both sides are associated three games having a recursive structure:

- The usual auxiliary game related to the Shapley operator [2.6] with state parameter (p, q).
- For Player 2, a "dual game" where the value satisfies [2.39] and the state variable, known by Player 2, is (z, q).
- Similarly for Player 1.

2.4.3 The dual differential game

Consider a RG with incomplete information on one side. The advantage of dealing with the dual recursive formula [2.38] rather than with [2.37] is that the state variable evolves smoothly from z to $z + \frac{1}{n}(z - g(i, y))$ while the martingale $p(i)$ may have jumps. De Meyer and Rosenberg (1999) use the dual formula to provide a new proof of Aumann and Maschler's result via the study of approximate fixed points and derive a heuristic partial differential equation for the limit. This leads them to anticipate a link with differential game theory. This is made precise in Laraki (2002) where it is proved that w_n satisfying [2.38] is the value of the time discretization with mesh $\frac{1}{n}$ of a differential game on $[0, 1]$ with dynamics $\zeta(t) \in \mathbb{R}^K$ given by:

$$\frac{d\zeta}{dt} = g(x_t, y_t), \quad \zeta(0) = -z$$

$x_t \in X$, $y_t \in Y$, and terminal payoff $\max_k \zeta^k(1)$. Basic results of differential games of fixed duration (see Section 6) (Souganidis, 1985) show that the game starting at time t from state ζ has a value $\varphi(t, \zeta)$, which is the only viscosity solution of the following Hamilton-Jacobi equation on $[0, 1]$ with terminal condition:

$$\frac{\partial \varphi}{\partial t} + u(D\varphi) = 0, \quad \varphi(1, \zeta) = \max_k \zeta^k. \qquad [2.40]$$

Hence $\varphi(0, -z) = \lim_{n\to\infty} w_n(z) = w(z)$. Using Hopf's representation formula, one obtains:

$$\varphi(1-t, \zeta) = \sup_{a\in\mathbb{R}^K} \inf_{b\in\mathbb{R}^K} \{\max_k b^k + \langle a, \zeta - b\rangle + tu(a)\}$$

and finally $w(z) = \sup_{p\in\Delta(K)}\{u(p) - \langle p, z\rangle\}$. Hence $\lim_{\lambda\to 0} v_\lambda = \lim_{n\to\infty} v_n = \mathtt{Cav}_{\Delta(K)} u$, by taking the Fenchel conjugate. Moreover, this is true for any compact evaluation of payoffs.

An alternative identification of the limit can be obtained through variational inequalities by translating in the primal game, the viscosity properties in the dual expressed in terms of local sub/super-differentials. This leads exactly to the properties **P1** and **P2** in the variational approach (Section 3.4). The approach has been extended recently by Gensbittel (2012) to games with infinite action spaces.

Interestingly, the dynamics of this differential game is exactly the one introduced by Vieille (1992), to show, in the context of Blackwell approachability, that any set is either weakly approachable or weakly excludable (see Section 7).

2.4.4 Error term, control of martingales, and applications to price dynamics

The initial objective of De Meyer (1996a,b) when he introduced the dual game was to study the error term in the Aumann and Maschler's RG model. The proof of $Cav(u)$ theorem shows that $e_n(p) := v_n(p) - \lim_{n\to\infty} v_n(p) = O\left(n^{-\frac{1}{2}}\right)$. The precise asymptotic analysis of $e_n(p)$ was first studied by Mertens and Zamir (1976b, 1995) when $I = J = K$ have cardinality 2 (see also Heuer, 1992b). They show that the speed of convergence can be improved to $O\left(n^{-\frac{2}{3}}\right)$ except in a particular class of "fair games" where $u(p) = 0$ for every p (without information, no Player has an advantage). In this class, the limit $\Psi(p)$ of $\sqrt{n}e_n(p)$ is shown to be related to the normal density function using a differential equation obtained by passing to the limit in the recursive primal formula. Moreover, $\Psi(p)$ appears as the limit of the maximal normalized L^1 variation of a martingale. In fact, an optimal strategy of the informed Player in G_n induces a martingale $\{p^n_m; 1 \le m \le n\}$ on $\Delta(K)$ starting at p. This martingale has a n-stage L^1 variation $D^1_n(p^n) := \mathbb{E}[\sum_{m=1}^n \|p^n_m - p^n_{m-1}\|_1]$ and asymptotically $D^1_n(p^n)/\sqrt{n}$ is maximal:

$$\lim_n \frac{1}{\sqrt{n}} |\max_{q^n} D^1_n(q^n) - D^1_n(p^n)| \to 0,$$

where the maximum is taken over all martingales on $\Delta(K)$ with length n starting at p.

For games where $I = J$ are finite, the error term is analyzed in depth by De Meyer in a series of papers (De Meyer, 1996a,b, 1999) where the dual game and the central limit theorem play a crucial role. For this purpose, De Meyer introduces a heuristic limit game with incomplete information and its dual, which is a stochastic differential game played on the time interval $[0, 1]$. He proves that it has a value V which satisfies a dynamic programming principle that leads to a second-order PDE. De Meyer proves then that if V is smooth, then $\sqrt{n}e_n$ converges uniformly to the Fenchel conjugate of V.

The main application of this work is achieved in De Meyer and Moussa-Saley (2002). They show that when two asymmetrically informed risk neutral agents repeatedly exchange a risky asset for a numéraire, they are playing a "fair RG" with incomplete information. The model may be seen as a particular RG à la Aumann-Maschler where action sets are infinite. Their main result is a characterization of the limit of the martingales of beliefs induced by an optimal strategy of the informed Player. Those discrete-time martingales are mapped to the time interval $[0, 1]$ and are considered as piecewise constant stochastic processes. The limit process for the weak topology is shown to be a Brownian motion.

De Meyer (2010) extends the result to general spaces I, J, and $K = \mathbf{R}$ and shows that the limit diffusion process does not depend on the specific "natural" trading mechanism, but only on the initial belief $p \in \Delta(K)$: it belongs to the CMMV class (continuous time martingales of maximal variation). It contains the dynamics of Black and Scholes and Bachelier as special cases.

The main step of De Meyer's results is the introduction of a discrete-time stochastic control problem whose value is equal to v_n (the value of the n-stage game) and whose maximizers coincide with a posteriori martingales at equilibrium. This generalizes the above maximization of the L^1-variation of a martingale. The first idea is to measure the variation at stage m by the value of the one-stage game where the transmission of information by Player 1 corresponds to the step from p_m to p_{m+1}. The next extension is, starting from W a real valued function defined on the set of probabilities over $\mathbf{R}^d$, to define the W-variation of a martingale $p = \{p_0, p_1, \dots, p_n\}$ by:

$$V_n^W(p^n) := \mathbb{E}\left[\sum_{m=1}^{n} W(p_m - p_{m-1}|p_0, \dots, p_{m-1})\right].$$

De Meyer (2010) solved the problem for $d = 1$, corresponding to a financial exchange model with one risky asset, and Gensbittel (2013) extended the result to higher dimensions (corresponding to a portfolio of $d \geq 1$ risky assets). Under quite general conditions on W, he obtains:

Theorem 2.16. *De Meyer (2010) for $d = 1$, Gensbittel (2013) for $d \geq 1$.*

$$\lim_n \sqrt{n} V_n(t,p) = \max_{\{p_s\}} \mathbb{E}[\int_t^1 \phi\left(\frac{\mathrm{d} < p_s >}{\mathrm{d}s}\right) \mathrm{d}s],$$

where the maximum is over càdlàg martingales p_t on $\Delta(K)$ that start at p and

$$\phi(A) = sup_{\mu \in \Delta(\mathbf{R}^d): cov(\mu) = A} W(\mu).$$

Here, $< p_s >$ denotes the quadratic variation and $cov(\mu)$ the covariance. The maximizing process above is interpreted by De Meyer as the dynamics of the equilibrium price of the risky assets. When $d = 1$ or under some conditions on the correlation between the risky assets, this process belongs to the CMMV class.

2.5. UNIFORM ANALYSIS

We turn now to the uniform approach and first recall basic results described in the previous chapters of HGT by Zamir (1992), Mertens (2002), and Vieille (2002).

2.5.1 Basic results

2.5.1.1 Incomplete information

Concerning games with lack of information on one side, Aumann and Maschler (1966) show the existence of a uniform value, (see Aumann and Maschler, 1995) and the famous formula:

$$v_\infty(p) = \mathtt{Cav}_{p \in \Delta(K)} u(p), \quad [2.41]$$

first for games with *standard signaling* (or perfect monitoring, i.e., where the moves are announced), then for general signals on the moves. u is as usual the value of the nonrevealing game and the construction of an optimal strategy for the uninformed Player is due to Kohlberg (1975).

For games with lack of information on both sides and standard signaling, Aumann and Maschler (1967) show that the maxmin and minmax exist (see Aumann and Maschler, 1995). Moreover, they give explicit formulas:

$$\underline{v}(p,q) = \mathtt{Cav}_{p \in \Delta(K)} \mathtt{Vex}_{q \in \Delta(L)} u(p,q), \ \overline{v}(p,q) = \mathtt{Vex}_{q \in \Delta(L)} \mathtt{Cav}_{p \in \Delta(K)} u(p,q). \quad [2.42]$$

They also construct games without a value. For several extensions to the dependent case and state independent signaling structure, mainly due to Mertens and Zamir (see Mertens et al., 1994).

2.5.1.2 Stochastic games

In the framework of stochastic games with standard signaling, the first proof of existence of a uniform value was obtained for the "Big Match" by Blackwell and Ferguson (1968), and then for absorbing games by Kearns et al. (2001). The main result is due to Mertens and Neyman (1981):

$$v_\infty \textit{ exists for finite stochastic games.} \tag{2.43}$$

The proof uses two ingredients:

(i) properties of the family $\{v_\lambda\}$ obtained by Bewley and Kohlberg (1976a) through their algebraic characterization,

(ii) the knowledge of the realized payoff at each stage n, to build an ε-optimal strategy as follows. One constructs a map $\bar{\lambda}$ and a sufficient statistics L_n of the past history at stage n such that σ is, at that stage, an optimal strategy in the game with discount parameter $\bar{\lambda}(L_n)$.

2.5.1.3 Symmetric case

A first connection between incomplete information games and stochastic games is obtained in the so-called symmetric case. This corresponds to games where the state in M is constant and may not be known by the Players, but their information during the play is symmetric (hence includes their actions). The natural state space is the set of probabilities on M and the analysis reduces to a stochastic game on $\Delta(M)$, which is no longer finite but on which the state process is regular (martingale), see Kohlberg and Zamir (1974), Forges (1982), and for alternative tools that extend to the non-zero-sum case, Neyman and Sorin (1998).

2.5.2 From asymptotic value to uniform value

Recall that the existence of v_∞ implies that it is also the limit of any sequence v_μ (with μ decreasing) or more generally v_θ (random duration with $\mathsf{E}(\theta) \longrightarrow \infty$).

On the other hand, the proof in Mertens and Neyman (1981) shows that in a stochastic game with standard signaling the following holds:

Theorem 2.17. *Assume that $w :]0,1] \to \mathbb{R}^\Omega$ satisfies:*

1) $\|w(\lambda) - w(\lambda')\| \le \int_\lambda^{\lambda'} f(x)\mathrm{d}x$, *for* $0 < \lambda < \lambda' < 1$, *with* $f \in L^1(]0,1])$,

2) $\Phi(\lambda, w(\lambda)) \ge w(\lambda)$, *for every* $\lambda > 0$ *small enough.*

Then Player 1 can guarantee $w(\lambda)$.

In the initial framework of finite stochastic games, one can take $w(\lambda) = v_\lambda$, hence (2) follows and one deduces property (1) from the fact that v_λ is semi-algebraic in λ.

More generally, this approach allows to prove the existence of v_∞ for continuous games with compact action spaces that are either absorbing (Mertens et al., 2009), recursive (Vigeral, 2010c) using the operator approach of Rosenberg and Sorin (2001), or definable (Bolte et al., 2013) (see Section 9.3.4).

2.5.3 Dynamic programming and MDP

Stronger results are available in the framework of general dynamic programming: this corresponds to a one person stochastic game with a state space Ω, a correspondence C from Ω to itself (with non empty values), and a real-valued, bounded payoff g on Ω. A play is a sequence $\{\omega_n\}$ satisfying $\omega_{n+1} \in C(\omega_n)$.

Littman and Stone (2005) give an example where $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ both exist and differ. They also prove that uniform convergence (on Ω) of v_n is equivalent to uniform convergence of v_λ and then the limits are the same. For a recent extension to the continuous time framework, see Oliu-Barton and Vigeral (2013).

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

Recent advances have been obtained by Renault (2011) introducing new notions like the values $v_{m,n}$ (resp. $\nu_{m,n}$) of the game where the payoff is the average between stage $m+1$ and $m+n$ (resp. the minimum of all averages between stage $m+1$ and $m+\ell$ for $\ell \leq n$).

Theorem 2.18. *Assume that the state space Ω is a compact metric space.*

(1) *If the family of functions v_n is uniformly equicontinuous, then $\lim_{n\to\infty} v_n = v$ exists, the convergence is uniform and:*

$$v(\omega) = \inf_{n\geq 1} \sup_{m\geq 0} v_{m,n}(\omega) = \sup_{m\geq 0} \inf_{n\geq 1} v_{m,n}(\omega).$$

(2) *If the family of functions $\nu_{m,n}$ is uniformly equicontinuous, then the uniform value v_∞ exits and*

$$v_\infty(\omega) = \inf_{n\geq 1} \sup_{m\geq 0} \nu_{m,n}(\omega) = \sup_{m\geq 0} \inf_{n\geq 1} \nu_{m,n}(\omega) = v(\omega)$$

For part (1), no assumption on Ω is needed and $\{v_n\}$ totally bounded suffices.

For (2), the construction of an ε-optimal strategy is by concatenation of strategies defined on large blocks, giving good payoffs while keeping the "level" of the state. Condition (2) plays the role of (i) in Mertens and Neyman's proof. It holds, for example, if g is continuous and C is non-expansive.

In particular for Markov Decision Process (finite state space K, move space I, and transition probability from $K \times I$ to K), the natural state space is $X = \Delta(K)$. In the case of partial observation (signal space A and transition probability from $K \times I$ to $K \times A$), the natural state space is $\Delta_f(X)$ on which C is non-expansive and the previous result implies:

Theorem 2.19. *Renault (2011)*

MDP processes with finite state space and partial observation have a uniform value.

This extends previous tools and results by Rosenberg et al. (2002).

Further developments to the continuous time setup lead to the study of asymptotic and uniform value in control problems defined as follows: the differential equation $\dot{x}_s = f(x_s, u_s)$ describes the control by $\mathbf{u} \in \mathbf{U}$ (measurable functions from $[0, +\infty)$ to U) of the state $x \in \mathbb{R}^n$ and one defines the value function $V_{m,t}$, $m, t \in \mathbb{R}^+$ by:

$$V_{m,t}(x) = \sup_{\mathbf{u} \in \mathbf{U}} \frac{1}{t} \int_m^{m+t} g(x_s, u_s) \mathrm{d}s, \quad x_0 = x.$$

Quincampoix and Renault (2011) prove that if g is continuous, the (feasible) state space X is bounded and the nonexpansiveness condition

$$\forall x, y \in X, \qquad \sup_{u \in U} \inf_{v \in U} \langle x - y, f(x, u) - f(y, v) \rangle \leq 0$$

holds, then the uniform value exists, the convergence $V_t (= V_{0,t}) \to V_\infty$ is uniform and:

$$V_\infty(x) = \inf_{t \geq 1} \sup_{m \geq 0} V_{m,t}(x) = \sup_{m \geq 0} \inf_{t \geq 1} V_{m,t}(x).$$

For similar results in the framework of differential games, see Bardi (2009).

2.5.4 Games with transition controlled by one player

Consider now a game where Player 1 controls the transition on the parameter: basic examples are stochastic games where the transition is independent of Player's 2 moves, or games with incomplete information on one side (with no signals); but this class also covers the case where the parameter is random, its evolution independent of Player 2's moves, and Player 1 knows more than Player 2.

Basically, the state space will be the beliefs of Player 2 on the parameter, which are variables controlled and known by Player 1. The analysis in Renault (2012) first constructs an auxiliary stochastic game on this space, and then reduces the analysis of the game to a dynamic programming problem by looking at stage by stage best reply of Player 2 (whose moves do not affect the future of the process). The finiteness assumption on the basic data implies that one can apply Theorem 2.18 part 2 to obtain:

Theorem 2.20. *In the finite case, games with transition controlled by one Player have a uniform value.*

The result extends previous work of Rosenberg et al. (2004) and also the model of Markov games with lack of information on one side introduced by Renault (2006) (see also Krausz and Rieder, 1997): here the parameter follows a Markov chain and is known at each stage by Player 1 while Player 2 knows only the initial law. The moves are observed. Neyman (2008) extends the analysis to the case with signals and constructs an optimal strategy for the informed Player.

This class is very interesting but an explicit formula for the value is not yet available.

Here is an example: there are two states $k = 1, 2$. At each stage, the state changes with probability ρ and the initial distribution is $(1/2, 1/2)$. The payoff is given by:

$k=1$	L	R
T	1	0
B	0	0

$k=2$	L	R
T	0	0
B	0	1

If $\rho = 0$ or 1, the game reduces to a standard game with incomplete information on one side with value $1/4$. By symmetry, it is enough to consider the interval $[1/2, 1]$; for $\rho \in [1/2, 2/3]$ the value is $\rho/(4\rho - 1)$, and still unknown otherwise (Marino, 2005; Hörner et al., 2010).

2.5.5 Stochastic games with signals on actions

Consider a stochastic game and assume that the signal to each Player reveals the current state but not necessarily the previous action of the opponent. By the recursive formula for v_λ and v_n, or more generally v_Θ, these quantities are the same as in the standard signaling case since the state variable is not affected by the change in the information structure. However, for example, in the Big Match, when Player 1 has no information on Player 2's action the max min is 0 (Kearns et al., 2001) and the uniform value does not exist anymore.

It follows that the existence of a uniform value for stochastic games depends on the signaling structure on actions. However, one has the following property:

Theorem 2.21. *Maxmin and minmax exist in finite stochastic games with signals on actions.*

This result, due to Coulomb (2003), and Rosenberg et al. (2003a) is extremely involved and relies on the construction of two auxiliary games, one for each Player.

Consider the maxmin and some discount factor λ. Introduce an equivalence relation among the mixed actions y and y' of Player 2 facing the mixed action x of Player 1 by $y \sim y'$ if they induce the same transition on the signals of Player 1 for each action i having significant weight ($\geq L\lambda$) under x. Define now the maxmin value of a discounted game where the payoff is the minimum with respect to an equivalence class of Player 2. This quantity will satisfy a fixed point equation defined by a semialgebraic set and will play the role of $w(\lambda)$ in Theorem 2.17. It remains to show, first for Player 1, that this auxiliary payoff indeed can be achieved in the real game. Then for Player 2, he will first follow a strategy realizing a best reply to σ of Player 1 up to a stage where the equivalence relation will allow for an indistinguishable switch in action. He will then change his strategy to obtain a good payoff from then on, without being detected. Obviously, a dual game is defined for the minmax (involving the structure of signals for Player 2).

An illuminating example, due to Coulomb (2003), is as follows:

	α	β	γ
a	1^*	0^*	L
b	0	1	L

Payoffs ($L \geq 1$)

	α	β	γ
a	?	?	?
b	A	B	A

Signals to Player 1

Given a strategy σ of Player 1, Player 2 will start by playing $(0, \varepsilon, 1-\varepsilon)$ and switch to $(1-\varepsilon, \varepsilon, 0)$ when the probability under σ of playing a in the future, given the distribution $(1-\varepsilon, \varepsilon)$ on the signals (A, B), is small enough. Hence the maxmin is 0.

For a nice overview, see Rosenberg et al. (2003b), Coulomb (1992, 1996, 1999, 2001, 2003), and Neyman and Sorin (2003).

2.5.6 Further results

Different examples include Sorin (1984), Sorin (1985a), and Chapter 6 in Sorin (2002) where several stochastic games with incomplete information are analyzed. Among the new tools are approachability strategies for games with vector payoffs and absorbing states and the use of a time change induced by an optimal strategy in the asymptotic game, to play well in the uniform game.

Rosenberg and Vieille (2000) consider a recursive game with lack of information on one side. The initial state is chosen in a finite set K according to some $p \in \Delta(K)$. After each stage, the moves are announced and Player 1 knows the state. If one leaves K, the payoff is absorbing and denoted by a. Denote by $\pi(p, x, y)$ the probability to stay in K and by $\tilde{p}$ the random conditional probability on K. The Shapley operator is:

$$\mathbf{T}(f)(p) = \mathtt{val}_{X^K \times Y}\left\{\pi(p, x, y)\mathsf{E}f(\tilde{p}) + (1 - \pi(p, x, y))\mathsf{E}(a)\right\}. \qquad [2.44]$$

Consider $w(p)$ an accumulation point of $v_\lambda(p)$. To prove that Player 1 can guarantee w, one alternates optimal strategies in the projective game if the current state satisfies $w(p_n) \leq \varepsilon$ and in G_λ (with $\|v_\lambda - w\| \leq \varepsilon^2$) otherwise. See Solan and Vieille (2002) for another extension.

In all these games with standard signaling, whenever Player 1 is fully informed, one has: $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda = \max\min$. A counter example in the general case is given in Section 9.3.2.

2.6. DIFFERENTIAL GAMES

2.6.1 A short presentation of differential games (DG)

Differential games (DG) are played in continuous time. A state space Z and control sets U for Player 1 and V for Player 2 are given. At each time t, the game is in some state z_t and each Player chooses a control ($u_t \in U, v_t \in V$). This induces a current payoff $\gamma_t = \gamma(z_t, t, u_t, v_t)$ and defines the dynamics $\dot{z}_t = f(z_t, t, u_t, v_t)$ followed by the state (see Friedman (1994) Chapter 22 in HGT2). Notice that in the autonomous case, if the Players use piece-wise constant controls on intervals of size δ, the induced process is like a RG.

There are many ways of defining strategies in differential games. For simplicity of the presentation, only nonanticipative strategies with delay are presented here. The main reasons are (1) they allow to put the game in normal form and (2) they are the most natural since they suppose that a Player always needs a delay (that may be chosen strategically) before reacting to a change in the behavior of the other Player.

Let $\mathcal{U}$ (resp. $\mathcal{V}$) denote the set of measurable control maps from $\mathbb{R}^+$ to U (resp. V). $\alpha \in \mathcal{A}$ (resp. $\beta \in \mathcal{B}$) is a nonanticipative strategy with delay (NAD) if α maps $\mathbf{v} \in \mathcal{V}$ to $\alpha(\mathbf{v}) = \mathbf{u} \in \mathcal{U}$ and there is $\delta > 0$ such that if $\mathbf{v}_s = \mathbf{v}'_s$ on $[0, t]$ then $\alpha(\mathbf{v}) = \alpha(\mathbf{v}')$ on $[0, t + \delta]$, for all $t \in \mathbb{R}^+$. A pair (α, β) defines a unique pair $(\mathbf{u}, \mathbf{v}) \in \mathcal{U} \times \mathcal{V}$ with $\alpha(\mathbf{v}) = \mathbf{u}$ and $\beta(\mathbf{u}) = \mathbf{v}$, thus the solution $\mathbf{z}$ is well defined. The map $t \in [0, +\infty[\mapsto (\mathbf{z}_t, \mathbf{u}_t, \mathbf{v}_t)$ specifies the trajectory $(\mathbf{z}, \mathbf{u}, \mathbf{v})(\alpha, \beta)$ (see, e.g., Cardaliaguet, 2007; Cardaliaguet and Quincampoix, 2008).

As in RG, there are many ways to evaluate payoffs.

Compact evaluations, or quantitative DG. This concerns the class of DG with total evaluation of the form:

$$\Gamma(\alpha, \beta)(z_0) = \int_0^T \gamma_t \, \mathrm{d}t + \overline{\gamma}(z_T) \tag{2.45}$$

where $\overline{\gamma}$ is some terminal payoff function, or:

$$\Gamma(\alpha, \beta)(z_0) = \int_0^\infty \gamma_t \mu(\mathrm{d}t)$$

where μ is a probability on $[0, +\infty]$ like $\frac{1}{T}\mathbf{1}_{[0,T]}\mathrm{d}t$ or $\lambda \exp(-\lambda t)\mathrm{d}t$.

The game is now well defined in normal form and the issues are the existence of a value, its characterization and properties of optimal strategies.

Uniform criteria, or qualitative DG. The aim is to control the asymptotic properties of the trajectories like: the state $\mathbf{z}_t$ should stay in some set C for all $t \in \mathbb{R}^+$ or from some time $T \geq 0$ on. Basic references include Krasovskii and Subbotin (1988), Cardaliaguet (1996), and Cardaliaguet et al. (2007).

2.6.2 Quantitative differential games

We describe very briefly the main tools in the proof of existence of a value, due to Evans and Souganidis (1984), but using NAD strategies, compare with Friedman (1994).

Consider the case defined by (2.45) under the following assumptions:

1) U and V are compact sets in $\mathbb{R}^K$,

2) $Z = \mathbb{R}^N$,

3) All functions f (dynamics), γ (running payoff), and $\bar{\gamma}$ (terminal payoff) are bounded, jointly continuous and uniformly Lipschitz in z,

4) Define the Hamiltonians $H^+(p, z, t) = \inf_v \sup_u \{\langle f(z, t, u, v), p\rangle + \gamma(z, t, u, v)\}$ and $H^-(p, z, t) = \sup_u \inf_v \{\langle f(z, t, u, v), p\rangle + \gamma(z, t, u, v)\}$ and assume that Isaacs's condition holds: $H^+(p, z, t) = H^-(p, z, t) = H(p, z, t)$, for all $(p, z, t) \in \mathbb{R}^N \times \mathbb{R}^N \times [0, T]$.

For $T \geq t \geq 0$ and $z \in Z$, consider the game on $[t, T]$ starting at time t from state z and let $\overline{v}[z, t]$ and $\underline{v}[z, t]$ denote the corresponding minmax and maxmin. Explicitly:

$$\overline{v}[z, t] = \inf_\beta \sup_\alpha [\int_t^T \gamma_s \mathrm{d}s + \bar{\gamma}(Z_T)]$$

where $\gamma_s = \gamma(\mathbf{z}_s, s, \mathbf{u}_s, \mathbf{v}_s)$ is the payoff at time s and $(\mathbf{z}, \mathbf{u}, \mathbf{v})$ is the trajectory induced by (α, β) and f on $[t, T]$ with $\mathbf{z}_t = z$. Hence $\mathbf{u}_s = \mathbf{u}_s(\alpha, \beta), \mathbf{v}_s = \mathbf{v}_s(\alpha, \beta)$, and $\mathbf{z}_s = \mathbf{z}_s(\alpha, \beta, z, t)$.

The first property is the following dynamic programming inequality:

Theorem 2.22. *For $0 \leq t \leq t + \delta \leq T$, $\overline{v}$ satisfies:*

$$\overline{v}[z, t] \leq \inf_\beta \sup_\alpha \left\{ \int_t^{t+\delta} \gamma(\mathbf{z}_s(\alpha, \beta, z, t), s, \mathbf{u}_s(\alpha, \beta), \mathbf{v}_s(\alpha, \beta))\mathrm{d}s + \overline{v}[\mathbf{z}_{t+\delta}(\alpha, \beta, z, t), t + \delta] \right\}. \tag{2.46}$$

In addition $\overline{v}$ is uniformly Lipschitz in z and t.

Property [2.46] implies in particular that for any $\mathcal{C}^1$ function Φ on $[0, T] \times Z$ with $\Phi[t, z] = \overline{v}[t, z]$ and $\Phi \geq \overline{v}$ in a neighborhood of (t, z) one has, for all $\delta > 0$ small enough:

$$\inf_\beta \sup_\alpha \left\{ \frac{1}{\delta} \int_t^{t+\delta} \gamma(\mathbf{z}_s(\alpha, \beta, z, t), s, \mathbf{u}_s(\alpha, \beta), \mathbf{v}_s(\alpha, \beta))\mathrm{d}s + \frac{\Phi[\mathbf{z}_{t+\delta}(\alpha, \beta, z, t), t + \delta] - \Phi[z, t]}{\delta} \right\} \geq 0. \tag{2.47}$$

Letting δ going to 0 implies that Φ satisfies the following property:

$$\inf_v \sup_u \left\{ \gamma(z, t, u, v) + \partial_t \Phi[z, t] + \langle D\Phi[z, t], f((z, t, u, v)\rangle \right\} \geq 0$$

which gives the differential inequality:

$$\partial_t \Phi[z,t] + H^+(D\Phi[z,t],z,t) \geq 0. \qquad [2.48]$$

The fact that any smooth local majorant of $\overline{v}$ satisfies [2.48] can be expressed as:

Proposition 2.1. *$\overline{v}$ is a viscosity subsolution of the equation $\partial_t W[z,t] + H^+$ $(DW[z,t],$ $z,t) = 0$.*

Obviously a dual property holds. One use then Assumption (3) and the next comparison principle:

Theorem 2.23. *Let W_1 be a viscosity subsolution and W_2 be a viscosity supersolution of*

$$\partial_t W[z,t] + H(DW[z,t],z,t) = 0$$

then $W_1[T,.] \leq W_2[T,.]$ implies $W_1[t,z] \leq W_2[z,t], \forall z \in Z, \forall t \in [0,T]$.

One obtains finally:

Theorem 2.24. *The differential game has a value:*

$$\overline{v}[z,t] = \underline{v}[z,t].$$

In fact, the previous Theorem 2.23 implies $\overline{v}[z,t] \leq \underline{v}[z,t]$.

Note that the comparison Theorem 2.23 is much more general and applies to W_1 u.s.c., W_2 l.s.c., H uniformly Lipschitz in p and satisfying: $|H(p,z_1,t_1) - H(p,z_2,t_2)| \leq C(1 + \|p\|)\|(z_1,t_1) - (z_2,t_2)\|$. Also $\overline{v}$ is in fact, even without Isaacs's condition, a viscosity solution of $\partial_t W[z,t] + H^+(DW[z,t],z,t) = 0$.

For complements, see, e.g., Souganidis (1999), Bardi and Capuzzo Dolcetta (1996) and for viscosity solutions Crandall et al. (1992).

The analysis has been extended by Cardaliaguet and Quincampoix (2008) to the symmetric case where the initial value of the state $z \in \mathbb{R}^N$ is random and only its law μ is known. Along the play, the Players observe the controls but not the state. Assuming $\mu \in M$, the set of measures with finite second moment, the analysis is done on M endowed with the L^2-Wasserstein distance by extending the previous tools and results to this infinite-dimensional setting.

Another extension involving mixed strategies when Isaacs' condition is not assumed is developed in Buckdahn et al. (2013).

2.6.3 Quantitative differential games with incomplete information

An approach similar to the one for RG has been introduced by Cardaliaguet (2007) and developed by Cardaliaguet and Rainer (2009a) to study differential games of fixed duration with incomplete information. Stochastic differential games with incomplete information have been analyzed by Cardaliaguet and Rainer (2009a) (see also Buckdahn et al., 2010).

The model works as follows. Let K and L be two finite sets. For each (k, ℓ), a differential game $\Gamma^{k\ell}$ on $[0, T]$ with control sets U and V is given. The initial position of the system is $z_0 = \{z_0^{k\ell}\} \in Z^{K\times L}$, the dynamics is $f^{k\ell}(z^{k\ell}, t, u, v)$, the running payoff is $\gamma^{k\ell}(z^{k\ell}, t, u, v)$, and the terminal payoff is $\bar{\gamma}^{k\ell}(z^{k\ell})$. $k \in K$ is chosen according to a probability distribution $p \in \Delta(K)$, similarly $\ell \in L$ is chosen according to $q \in \Delta(L)$. Both Players know p and q and in addition Player 1 learns k and Player 2 learns l. Then, the game $\Gamma^{k\ell}$ is played starting from $z_0^{k\ell}$. The corresponding game is $\Gamma(p, q)[z_0, 0]$. The game $\Gamma(p, q)[z, t]$ starting from $z = \{z^{k\ell}\}$ at time t is defined similarly. One main difference with the previous section is that even if Isaacs' condition holds, the Players have to use randomization to choose their controls in order to hide their private information. $\alpha \in \bar{\mathcal{A}}$ is the choice at random of an element in $\mathcal{A}$. Hence a strategy for Player 1 is described by a profile $\hat{\alpha} = \{\alpha^k\} \in \bar{\mathcal{A}}^K$ (α^k is used if the signal is k). The payoff induced by a pair of profiles $(\hat{\alpha}, \hat{\beta})$ in $\Gamma(p, q)[z, t]$ is $G^{p,q}[z, t](\hat{\alpha}, \hat{\beta}) = \sum_{k,\ell} p^k q^\ell G^{k\ell}[z, t](\alpha^k, \beta^\ell)$ where $G^{k\ell}[z, t](\alpha^k, \beta^\ell)$ is the payoff in the game $\Gamma^{k\ell}$ induced by the (random) strategies (α^k, β^ℓ).

Notice that $\Gamma(p, q)[z, t]$ can be considered as a game with incomplete information on one side where Player 1 knows which of the games $\Gamma(k, q)[z, t]$ will be played, where k has distribution p and Player 2 is uninformed. Consider the minmax in $\Gamma(p, q)[z, t]$:

$$\overline{V}(p, q)[z, t] = \inf_{\hat{\beta}} \sup_{\hat{\alpha}} G^{p,q}[z, t](\hat{\alpha}, \hat{\beta}) = \inf_{\{\beta^\ell\}} \sup_{\{\alpha^k\}} \sum_k p^k \left\{ \sum_\ell q^\ell G^{k\ell}[z, t](\alpha^k, \beta^\ell) \right\}$$

The dual game with respect to p and with parameter $\theta \in \mathbb{R}^K$ has a minmax given by:

$$\overline{W}(\theta, q)[z, t] = \inf_{\hat{\beta}} \sup_{\alpha \in \bar{\mathcal{A}}} \max_k \left\{ \sum_\ell q^\ell G^{k\ell}[z, t](\alpha, \beta^\ell) - \theta^k \right\}$$

and using (2.36):

$$\overline{W}(\theta, q)[z, t] = \max_{p \in \Delta(K)} \{\overline{V}(p, q)[z, t] - \langle p, \theta \rangle\}.$$

Note that $\overline{V}(p, q)[z, t]$ does not obey a dynamic programming equation: the Players observe the controls not the strategy profiles, and the current state is unknown but $\overline{W}(\theta, q)[z, t]$ will satisfy a subdynamical programming equation. First the max can be taken on $\mathcal{A}$, then if Player 2 ignores his information, one obtains:

Proposition 2.2.

$$\overline{W}(\theta, q)[z, t] \leq \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \overline{W}(\theta(t+\delta), q)[\mathbf{z}_{t+\delta}, t+\delta] \qquad [2.49]$$

where $\mathbf{z}_{t+\delta} = \mathbf{z}_{t+\delta}(\alpha, \beta, z, t)$ *and* $\theta^k(t+\delta) = \theta^k - \sum_\ell q^\ell \int_t^{t+\delta} \gamma^{k\ell}(\mathbf{z}_s^{k\ell}, s, \mathbf{u}_s, \mathbf{v}_s)\mathrm{d}s.$

Assume that the following Hamiltonian H satisfies Isaacs's condition:

$$\begin{aligned} H(z, t, \xi, p, q) &= \inf_v \sup_u \{\langle f(z, t, u, v), \xi\rangle + \sum_{k,\ell} p^k q^\ell \gamma^{k\ell}(z^{k\ell}, t, u, v)\} \\ &= \sup_u \inf_v \{\langle f(z, t, u, v), \xi\rangle + \sum_{k,\ell} p^k q^\ell \gamma^{k\ell}(z^{k\ell}, t, u, v)\}. \end{aligned}$$

Here $f(z, ., ., .)$ stands for $\{f^{k\ell}(z^{k\ell}, ., ., .)\}$ and $\xi = \{\xi^{k\ell}\}$.

Given $\Phi \in \mathcal{C}^2(Z \times [0, T] \times \mathbb{R}^K)$, let $\bar{L}\Phi(z, t, \bar{p}) = \max\{\langle D^2_{pp}\Phi(z, t, \bar{p})\rho, \rho\rangle; \rho \in T_{\bar{p}}\Delta(K)\}$ where $T_{\bar{p}}\Delta(K)$ is the tangent cone to $\Delta(K)$ at $\bar{p}$.

The crucial idea is to use (2.36) to deduce from (2.49) the following property on $\overline{V}$:

Proposition 2.3. $\overline{V}$ *is a viscosity subsolution for* H *in the sense that: for any given* $\bar{q} \in \Delta(L)$ *and any test function* $\Phi \in \mathcal{C}^2(Z \times [0, T] \times \mathbb{R}^K)$ *such that the map* $(z, t, p) \mapsto \overline{V}(z, t, p, \bar{q}) - \Phi(z, t, p)$ *has a local maximum on* $Z \times [0, T] \times \Delta(K))$ *at* $(\bar{z}, \bar{t}, \bar{p})$ *then*

$$\max\left\{\bar{L}\Phi(\bar{z}, \bar{t}, \bar{p}); \partial_t\Phi(\bar{z}, \bar{t}, \bar{p}) + H(\bar{z}, \bar{t}, D_z\Phi(\bar{z}, \bar{t}, \bar{p}), \bar{p}, \bar{q})\right\} \geq 0. \qquad [2.50]$$

A similar dual definition, with $\underline{L}$, holds for a viscosity supersolution.

Finally a comparison principle extending Theorem 2.23 proves the existence of a value $v = \overline{V} = \underline{V}$.

Theorem 2.25. *Let* F_1 *and* F_2: $Z \times [0, T] \times \Delta(K) \times \Delta(L) \mapsto \mathbb{R}$ *be Lipschitz and saddle (concave in p and convex in q). Assume that* F_1 *is a subsolution and* F_2 *a supersolution with* $F_1(., T, ., .) \leq F_2(., T, ., .)$, *then* $F_1 \leq F_2$ *on* $Z \times [0, T] \times \Delta\ (K)\ \times\ \Delta\ (L)$.

Basically the idea of the proof is to mimick the complete information case. However, the value operator is on the pair of profiles $(\hat{\alpha}, \hat{\beta})$. Hence, the infinitesimal version involves vectors in $U^K \times V^L$ while only the realized controls in U or V are observed; consequently, the dynamic programming property does not apply in the same space. The use of the dual games allows, for example, for Player 2 to work with a variable that depends only on the realized trajectory of his opponent. The geometrical properties (convexity) of the minmax imply that it is enough to characterize the extreme points and then Player 2 can play non-revealing. As a consequence, the dynamic programming inequality on the dual of the minmax involving a pair (α, β) induces an inequality with the infinitesimal value operator on $U \times V$ for the test function. The situation being symmetrical for Player 1, a comparison theorem can be obtained.

Using the characterization above, Souquière (2010) shows that in the case where f and γ are independent of z and the terminal payoff is linear, $v = \mathbf{MZ}(u)$, where u is the value of the corresponding non-revealing game, and thus recovers Mertens-Zamir's result through differential games. This formula does not hold in general (see examples in Cardaliaguet, 2008). However, one has the following approximation procedure. Given a finite partition Π of $[0, 1]$ define inductively V_Π by:

$$V_\Pi(z, t_m, p, q) = \mathbf{MZ}\left[\sup_u \inf_v \{V_\Pi(z + \delta_{m+1} f(z, t_m, u, v), t_{m+1}, p, q) + \delta_{m+1} \sum_{k\ell} p^k q^\ell \gamma^{k\ell}(z^{k\ell}, t_m, u, v)\}\right]$$

where $\delta_{m+1} = t_{m+1} - t_m$. Then using results of Laraki (2001b, 2004), Souquière (2010) proves that V_Π converges uniformly to v, as the mesh of Π goes to 0. This extends a similar construction for games with lack of information on one side in Cardaliaguet (2009), where moreover an algorithm for constructing ε-optimal strategies is provided. Hence, the $\mathbf{MZ}$ operator (which is constant in the framework of repeated games: this is the time homogeneity property) appears as the true infinitesimal operator in a nonautonomous framework.

Cardaliaguet and Souquière (2012) study the case where the initial state is random (with law μ) and known by Player 1 who also observes the control of Player 2. Player 2 knows μ but is blind: he has no further information during the play.

2.7. APPROACHABILITY

This section describes the exciting and productive interaction between RG and DG in a specific area: approachability theory, introduced and studied by Blackwell (1956).

2.7.1 Definition

Given an $I \times J$ matrix A with coefficients in $\mathbb{R}^k$, a two-person infinitely repeated game form G is defined as follows. At each stage $n = 1, 2, \ldots$, each Player chooses an element in his set of actions: $i_n \in I$ for Player 1 (resp. $j_n \in J$ for Player 2), the corresponding vector outcome is $g_n = A_{i_n j_n} \in \mathbb{R}^k$ and the couple of actions (i_n, j_n) is announced to both Players. $\bar{g}_n = \frac{1}{n} \sum_{m=1}^n g_m$ is the average vector outcome up to stage n. The aim of Player 1 is that $\bar{g}_n$ approaches a target set $C \subset \mathbb{R}^k$. Approachability theory is thus a generalization of max-min level in a (one shot) game with real payoff where C is of the form $[v, +\infty)$.

The asymptotic approach corresponds to the following notion:

Definition 2.3. *A nonempty closed set C in $\mathbb{R}^k$ is* **weakly approachable** *by Player 1 in G if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for any $n \geq N$ there is a strategy $\sigma = \sigma(n, \varepsilon)$ of Player 1 such that, for any strategy τ of Player 2:*

$$\mathsf{E}_{\sigma,\tau}(d_C(\bar{g}_n)) \leq \varepsilon.$$

where d_C stands for the distance to C.

If v_n is the value of the n-stage game with payoff $-\mathsf{E}(d_C(\bar{g}_n))$, weak-approachability means $v_n \to 0$. The uniform approach is expressed by the next definition:

Definition 2.4. *A nonempty closed set C in $\mathbb{R}^k$ is* **approachable** *by Player 1 in G if, for every $\varepsilon > 0$, there exists a strategy $\sigma = \sigma(\varepsilon)$ of Player 1 and $N \in \mathbb{N}$ such that, for any strategy τ of Player 2 and any $n \geq N$:*

$$\mathsf{E}_{\sigma,\tau}(d_C(\bar{g}_n)) \leq \varepsilon.$$

In this case, asymptotically the average outcome remains close in expectation to the target C, uniformly with respect to the opponent's behavior. The dual concept is excludability.

The "expected deterministic" repeated game form $G^\star$ is an alternative two-person infinitely repeated game associated, as the previous one, to the matrix A but where at each stage $n = 1, 2, \ldots$, Player 1 (resp. Player 2) chooses $u_n \in U = \Delta(I)$ (resp. $v_n \in V = \Delta(J)$), the outcome is $g_n^\star = u_n A v_n$ and (u_n, v_n) is announced. Hence $G^\star$ is the game played in "mixed actions" or in expectation. Weak $\star$approachability, $v_n^\star$, and $\star$approachability are defined similarly.

2.7.2 Weak approachability and quantitative differential games

The next result is due to Vieille (1992). Recall that the aim is to obtain an average outcome at stage n close to C.

First consider the game $G^\star$. Use as state variable the accumulated payoff $z_t = \int_0^t \gamma_s \mathrm{d}s$, $\gamma_s = u_s A v_s$ being the payoff at time s and consider the differential game Λ of fixed duration played on $[0, 1]$ starting from $z_0 = 0 \in \mathbb{R}^k$ with dynamics:

$$\dot{z}_t = u_t A v_t = f(u_t, v_t) \qquad [2.51]$$

and terminal payoff $-d_C(\mathbf{z}(1))$. Note that Isaacs's condition holds: $\max_u \min_v \langle f(z, u, v), \xi\rangle = \min_v \max_u \langle f(z, u, v), \xi\rangle = \mathtt{val}_{U\times V}\langle f(z, u, v), \xi\rangle$ for all $\xi \in \mathbb{R}^k$. The n stage game $G_n^\star$ appears then as a discrete time approximation of Λ and $v_n^\star = V_n(0, 0)$ where V_n satisfies, for $k = 0, \ldots, n-1$ and $z \in \mathbb{R}^k$:

$$V_n\left(\frac{k}{n}, z\right) = \mathtt{val}_{U\times V} V_n\left(\frac{k+1}{n}, z + \frac{1}{n}uAv\right) \qquad [2.52]$$

with terminal condition $V(1, z) = -d_C(z)$. Let $\Phi(t, z)$ be the value of the game played on $[t, 1]$ starting from z (i.e., with total outcome $z + \int_t^1 \gamma_s \mathrm{d}s$). Then basic results from DG imply (see Section 6):

Theorem 2.26.
(1) *$\Phi(z, t)$ is the unique viscosity solution on $[0, 1] \times \mathbb{R}^k$ of:*

$$\partial_t \Phi(z, t) + \mathtt{val}_{U\times V}\langle D\Phi(z, t), uAv\rangle = 0$$

with $\Phi(z, 1) = -d_C(z)$.
(2)

$$\lim_{n\to\infty} v_n^\star = \Phi(0, 0).$$

The last step is to relate the values in $G_n^\star$ and in G_n.

Theorem 2.27.

$$\lim v_n^\star = \lim v_n$$

The idea of the proof is to play by blocks in G_{Ln} and to mimic an optimal behavior in $G_n^\star$. Inductively at the m^{th} block of L stages in G_{Ln} Player 1 will play i.i.d. a mixed action optimal at stage m in $G_n^\star$ (given the past history) and $\gamma_m^\star$ is defined as the empirical distribution of actions of Player 2 during this block. Then the (average) outcome in G_{Ln} will be close to the one in $G_n^\star$ for large L, hence the result.

Corollary 2.2.
Every set is weakly approachable or weakly excludable.

2.7.3 Approachability and B-sets

The main notion was introduced by Blackwell (1956). $\pi_C(a)$ denotes the set of closest points to a in C.

Definition 2.5. *A closed set C in $\mathbb{R}^k$ is a* **B***-set for Player 1 (for a given matrix A), if for any $a \notin C$, there exists $b \in \pi_C(a)$ and a mixed action $u = \hat{u}(a)$ in $U = \Delta(I)$ such that the hyperplane through b orthogonal to the segment $[ab]$ separates a from uAV:*

$$\langle uAv - b, a - b\rangle \leq 0, \qquad \forall v \in V$$

The basic result of Blackwell (1956) is:

Theorem 2.28. *Let C be a* **B**-*set for Player 1. Then it is approachable in G and ⋆approachable in $G^\star$ by that Player. An approachability strategy is given by $\sigma(h_n) = \hat{u}(\bar{g}_n)$ (resp. $\sigma^\star(h_n^\star) = \hat{u}(\bar{g}_n^\star)$), where h_n (resp. $h_n^\star$) denotes the history at stage n.*

An important consequence of Theorem 2.28 is the next result due to Blackwell (1956):

Theorem 2.29. *A convex set C is either approachable or excludable.*

A further result due to Spinat (2002) characterizes minimal approachable sets:

Theorem 2.30. *A set C is approachable iff it contains a* **B**-*set.*

2.7.4 Approachability and qualitative differential games

To study ⋆approachability, we consider an alternative qualitative differential game Γ where both the dynamics and the objective differ from the previous quantitative differential game Λ. We follow here Assoulamani et al. (2009). The aim is to control asymptotically the average payoff that will be the state variable and the discrete dynamics is of the form:

$$\bar{g}_{n+1} - \bar{g}_n = \frac{1}{n+1}(g_{n+1} - \bar{g}_n).$$

The continuous counterpart is $\bar{\gamma}_t = \frac{1}{t}\int_0^t u_s A v_s \mathrm{d}s$. A change of variable $z_t = \bar{\gamma}_{e^t}$ leads to:

$$\dot{z}_t = u_t A v_t - z_t. \qquad [2.53]$$

which is the dynamics of an autonomous differential game Γ with $f(z,u,v) = uAv - z$, that still satisfies Isaacs' condition. In addition, the aim of Player 1 in Γ is to stay in a certain set C.

Definition 2.6. *A nonempty closed set C in $\mathbb{R}^k$ is a* **discriminating domain** *for Player 1, given f if:*

$$\forall a \in C, \quad \forall p \in NP_C(a), \qquad \sup_{v\in V}\inf_{u\in U} \langle f(a,u,v), p\rangle \le 0, \qquad [2.54]$$

where $NP_C(a) = \{p \in \mathbb{R}^K; d_C(a+p) = \|p\|\}$ is the set of proximal normals to C at a.

The interpretation is that, at any boundary point $x \in C$, Player 1 can react to any control of Player 2 in order to keep the trajectory in the half space facing a proximal normal p.

The following theorem, due to Cardaliaguet (1996), states that Player 1 can ensure remaining in a discriminating domain.

Theorem 2.31. *Assume that f satisfies Isaacs' condition, that $f(x, U, v)$ is convex for all x, v, and that C is a closed subset of $\mathbb{R}^k$. Then C is a discriminating domain if and only if for every z belonging to C, there exists a nonanticipative strategy $\alpha \in \mathcal{A}'$, such that for any $\mathbf{v} \in \mathbf{V}$, the trajectory $\mathbf{z}[\alpha(\mathbf{v}), \mathbf{v}, z](t)$ remains in C for every $t \geq 0$.*

The link with approachability is through the following result:

Theorem 2.32. *Let $f(z, u, v) = uAv - z$. A closed set $C \subset \mathbb{R}^k$ is a discriminating domain for Player 1, if and only if C is a $\mathbf{B}$-set for Player 1.*

The main result is then:

Theorem 2.33. *A closed set C is $\star$approachable in $G^\star$ if and only if it contains a $\mathbf{B}$-set.*

The direct part follows from Blackwell's proof. For the converse, first one defines a map Ψ from strategies of Player 1 in $G^\star$ to nonanticipative strategies in Γ. Next, given $\varepsilon > 0$ and a strategy σ_ε that ε-approaches C in $G^\star$, one shows that the trajectories in the differential game Γ that are compatible with $\alpha_\varepsilon = \Psi(\sigma_\varepsilon)$ approach C up to ε. Finally, one proves that the ω-limit set of any trajectory compatible with some α is a discriminating domain.

In particular, approachability and $\star$approachability coincide.

In a similar way, one can explicitly construct an approachability strategy in the repeated game G starting from a preserving strategy in Γ. The proof is inspired by the "extremal aiming" method of Krasovskii and Subbotin (1988) which is in the spirit of proximal normals and approachability.

2.7.5 Remarks and extensions

1. In both cases, the main ideas to represent a RG as a DG is first to take as state variable either the total payoff or the average payoff, but in both cases the corresponding dynamics is (asymptotically) smooth; the second aspect is to work with expectation so that the trajectory is deterministic.

2. For recent extensions of approachability condition for games on more general spaces, see Lehrer (2002) and Milman (2006). For games with signals on the outcomes, see Lehrer and Solan (2007) and Perchet (2009, 2011a,b) which provides a characterization for convex sets. Perchet and Quincampoix (2012) present a general perspective by working on the space of distribution on signals that can be generated during the play. Approachability is then analyzed in the space of probability distributions on $\mathbb{R}^n$ with the

L^2-Wasserstein distance using tools from As Soulaimani (2008) and Cardaliaguet and Quincampoix (2008).

2.8. ALTERNATIVE TOOLS AND TOPICS

2.8.1 Alternative approaches

2.8.1.1 A different use of the recursive structure

The use of the operator approach does not allow to deal easily with games with signals: the natural state space on which the value is defined is large. In their original approach, Mertens and Zamir (1971) and Mertens (1972) introduce thus majorant and minorant games having both simple recursive structure, i.e., small level in the hierarchy of beliefs in the auxiliary game.

Similarly, a sophisticated use of the recursive structure allows to obtain exact speed of convergence for games with incomplete information on one side (Mertens, 1998).

2.8.1.2 No signals

For a class of games with no signals, Mertens and Zamir (1976a) introduced a kind of normal form representation of the infinite game for the maxmin (and for the minmax): this is a collection of strategies (and corresponding payoffs) that they prove to be an exhaustive representation of optimal plays. In any large game, Player 1 can guarantee the value of this auxiliary game and Player 2 defend it.

A similar approach is used in Sorin (1989) for the asymptotic value. One uses a two-level scale: blocks of stages are used to identify the regular moves and sequence of blocks are needed for the exceptional moves.

2.8.1.3 State dependent signals

For RG with incomplete information, the analysis in Sorin (1985b) shows that the introduction of state dependent signals generates absorbing states in the space of beliefs, hence the natural study of absorbing games with incomplete information. In Sorin (1984,1985a), two classes are studied and the minmax, maxmin, and asymptotic values are identified.

2.8.1.4 Incomplete information on the duration

These are games where the duration is a random variable on which the Players have private information. For RG with incomplete information, these games have maxmin and minmax that may differ (Neyman, 2012a).

2.8.1.5 Games with information lag

There is a large literature on games with information lag starting with Scarf and Shapley (1957). A recent study in the framework of stochastic games is due to Levy (2002).

2.8.2 The "Limit Game"

2.8.2.1 Presentation

In addition to the convergence of the values $\{v_\mu\}$, one looks for a game $\mathcal{G}$ on $[0,1]$ with strategy sets $\mathbf{U}$ and $\mathbf{V}$ and value w such that:

(1) the play at time t in $\mathcal{G}$ would be similar to the play at stage $m = [tn]$ in G_n (or at the fraction t of the total weight of the game for general evaluation μ),

(2) ε-optimal strategies in $\mathcal{G}$ would induce 2ε-optimal strategies in G_n, for large n. More precisely, the history for Player 1 up to stage m in G_n defines a state variable that is used to define with a strategy U in $\mathcal{G}$ at time $t = m/n$ a move in G_n.

Obviously then, the asymptotic value exists and is w.

2.8.2.2 Examples

One example was explicitly described (strategies and payoff) for the Big Match with incomplete information on one side in Sorin (1984). $\mathbf{V}$ is the set of measurable maps f from $[0,1]$ to $\Delta(J)$. Hence Player 2 plays $f(t)$ at time t and the associated strategy in G_n is a piecewiese constant approximation. $\mathbf{U}$ is the set of vectors of stopping times $\{\rho^k\}$, $k \in K$, i.e., increasing maps from $[0,1]$ to $[0,1]$ and $\rho^k(t)$ is the probability to stop the game before time t if the private information is k.

The auxiliary differential game introduced by Vieille (1992) to study weak approachability, Section 7.2 is also an example of a limit game.

A recent example deals with absorbing games (Laraki, 2010). Recall the auxiliary game Γ corresponding to (2.26). Then one shows that given a strategy (x, α) ε-optimal in Γ, its image $x + \lambda\alpha$ (normalized) is 2ε-optimal in G_λ.

For games with incomplete information on one side, the asymptotic value $v(p)$ is the value of the splitting game (Section 3.4.2. C) with payoff $\int_0^1 u(p_t)\mathrm{d}t$, where u is the value of the nonrevealing game and p_t is a martingale in $\Delta(K)$ starting from p at $t = 0$.

For the uninformed Player, an asymptotically optimal strategy has been defined by Heuer (1992a) and extends to the case of lack of information on both sides. His approach has the additional advantage to show that, assuming that the Mertens-Zamir system has a solution v, then the asymptotic value exists and is v.

2.8.2.3 Specific Properties

This representation allows also to look for further properties, like stationarity of the expected payoff at time t along the plays induced optimal strategies, see Sorin et al. (2010) or robustness of optimal strategies: to play at stage n optimally in the discounted game with $\lambda_n = \frac{\mu_n}{\sum_{m \geq n} \mu_m}$ should be asymptotically optimal for the evaluation μ.

2.8.3 Repeated games and differential equations

2.8.3.1 RG and PDE

As already remarked by De Meyer (1996a, 1999), the fact that the asymptotic value of game satisfies a limit dynamic principle hence a PDE can be used in the other direction. To prove that a certain PDE has a solution, one constructs a family of simple finite RG (corresponding to time and space discretizations) and one then shows that the limit of the values exists.

For similar results in this direction see, e.g., Kohn and Serfaty, (2006), and Peres et al. (2009).

2.8.3.2 RG and evolution equations

We follow Vigeral (2010a). Consider again a nonexpansive mapping $\mathbf{T}$ from a Banach space X to itself. The nonnormalized n stage values satisfy $V_n = \mathbf{T}^n(0)$ hence:

$$V_n - V_{n-1} = -(Id - \mathbf{T})(V_{n-1})$$

which can be considered as a discretization of the differential equation:

$$\dot{x} = -Ax \qquad [2.55]$$

where the maximal monotone operator A is $Id - \mathbf{T}$.

The comparison between the iterates of $\mathbf{T}$ and the solution U of (2.55) is given by Chernoff's formula (see Brézis, 1973, p. 16):

$$\|U(t) - \mathbf{T}^n(U(0))\| \leq \|U'(0)\|\sqrt{t + (n-t)^2}.$$

In particular with $U(0) = 0$ and $t = n$, one obtains:

$$\|\frac{U(n)}{n} - v_n\| \leq \frac{\|\mathbf{T}(0)\|}{\sqrt{n}}.$$

It is thus natural to consider $u(t) = \frac{U(t)}{t}$ which satisfies an equation of the form:

$$\dot{x}(t) = \Phi(\varepsilon(t), x(t)) - x(t) \qquad [2.56]$$

where as usual $\Phi(\varepsilon, x) = \varepsilon\mathbf{T}\left(\frac{1-\varepsilon}{\varepsilon}x\right)$ and notice that (2.56) is no longer autonomous.

Define the condition (L) by:

$$\|\Phi(\lambda, x) - \Phi(\mu, x)\| \leq |\lambda - \mu|(C + \|x\|).$$

Theorem 2.34. *Let $u(t)$ be the solution of [2.56], associated to $\varepsilon(t)$.*

(a) *If $\varepsilon(t) = \lambda$, then $\|u(t) - v_\lambda\| \to 0$*

(b) *If $\varepsilon(t) \sim \frac{1}{t}$, then $\|u(n) - v_n\| \to 0$*

Assume condition (L).

(c) *If $\frac{\varepsilon'(t)}{\varepsilon^2(t)} \to 0$ then $\|u(t) - v_{\varepsilon(t)}\| \to 0$*

Hence $\lim v_n$ and $\lim v_\lambda$ mimic solutions of similar perturbed evolution equations and in addition one has the following robustness result:

Theorem 2.35. *Let u (resp. $\bar{u}$) be a solution of [2.56] associated to ε (resp. $\bar{\varepsilon}$).*

Then $\|u(t) - \bar{u}(t)\| \to 0$ as soon as

(i) $\varepsilon(t) \sim \bar{\varepsilon}(t)$ *as* $t \to \infty$ *or*

(ii) $|\varepsilon - \bar{\varepsilon}| \in L^1$.

2.8.4 Multimove games

We describe here very briefly some other areas that are connected to RG.

2.8.4.1 Alternative evaluations

There are games similar to those of Section 1 where the sequence of payoffs $\{g_n\}$ is evaluated through a single functional like limsup, or liminf.

Some of the results and of the tools are quite similar to those of RG, like "recursive structure," "operator approach," or construction by iteration of optimal strategies. A basic reference is Maitra and Sudderth (1996).

2.8.4.2 Evaluation on plays

More generally, the evaluation is defined here directly on the set of plays (equipped with the product topology) extending the games of Gale and Stewart (1953) and Blackwell (1969). The basic results, first in the perfect information case then in the general framework, are due to Martin (1975, 1998) and can be expressed as:

Theorem 2.36. *Borel games are determined.*

For the extension to stochastic games and related topics, see Maitra and Sudderth (1992, 1993, 1998, 2003). For a recent result in the framework of games with delayed information, see Shmaya (2011).

2.8.4.3 Stopping games

Stopping games (with symmetric information) have been introduced by Dynkin (1969). They are played on a probability space $(\Omega, \mathcal{A}, P)$ endowed with a filtration $\mathcal{F} = (\mathcal{F}_t)$ describing the common information of the Players as time go on. Each Player i chooses a measurable time θ_i when he stops and the game ends at $\theta = \min\{\theta_1, \theta_2\}$. The payoff is $g_i(\theta, \omega)$ if Player $i = 1, 2$ stops first the game and is $f(\theta, \omega)$ if they stop it simultaneously. Payoff functions are supposed to be uniformly integrable.

Neveu (1975) in discrete time and Lepeltier and Maingueneau (1984) in continuous time proved the existence of the value in *pure strategies* under the "standard" condition $g_1 \leq f \leq g_2$ (at each moment, each Player prefers the other to stop rather than himself).

Without this condition, mixed strategies are necessary to have a value. Rosenberg et al. (2001) proved the existence of the value in discrete time. As in Mertens and Neyman (1981), they let the Players play an optimal discounted strategy, where the discount factor may change from time to time, depending on their information. Shmaya and Solan (2004) provided a very elegant proof of the result based on a stochastic variation of Ramsey's theorem. Finally, Laraki and Solan (2005) proved the existence of the value in continuous time.

For stopping games with incomplete information on one side, the value may not exist even under the standard assumption (Laraki, 2000). This is due to the fact that Player 2 prefers to wait until Player 1 uses his information, while Player 1 prefers to use his information only if he knows that Player 2 will never stop.

For finite horizon stopping games with incomplete information on one side, the value exists in discrete time using Sion's minmax theorem. In continuous time, the value exists under the standard assumption and may be explicitly characterized by using viscosity solutions combined with BSDE technics (Grün, 2012). Without the standard condition, the value may not exist (Bich and Laraki, 2013). One example is as follows. One type of Player 1 has a dominant strategy: to stop at time zero. The other type prefers to stop just after 0, but before Player 2. However, Player 2 also prefers to stop just after zero but before type 2 of Player 1.

For a survey on the topic, see Solan and Vieille (2004) and for the related class of duels see Radzik and Raghavan (1994).

2.9. RECENT ADVANCES

We cover here very recent and important advances and in particular counter examples to conjectures concerning the asymptotic value and new approaches to multistage interactions.

2.9.1 Dynamic programming and games with an informed controller

2.9.1.1 General evaluation and total variation

Recall that an evaluation μ is a probability on $\mathbb{N}^*$ which corresponds to a discrete time process and the associated length is related to its expectation. However, some regularity has to be satisfied to express that the duration goes to ∞: $\|\mu\| \to 0$ is not sufficient (take for payoff an alternating sequence of 0 and 1 and for evaluation the sequences $\mu^n = (0, 1/n, 0, 1/n, \ldots)$ and $\nu^n = (1/n, 0, 1/n, 0, \ldots)$).

One measure of regularity is the *total variation*, $TV(\mu) = \sum_n |\mu_n - \mu_{n+1}|$, and a related ergodic theorem holds for a Markov chain P: there exists Q such that $\sum_n \mu_n P^n$ converges to Q as $TV(\mu) \to 0$ (Sorin, 2002, p. 105). Obviously for decreasing evaluations, one has $\mu_1 = \|\mu\| = TV(\mu)$.

One can define stronger notions of convergence associated to this criteria, following Renault (2013), Renault and Venel (2012):

v is a TV-asymptotic value if for each $\varepsilon > 0$, there exists $\delta > 0$ such that $TV(\mu) \leq \delta$ implies $\|v_\mu - v\| \leq \varepsilon$. (In particular then $v = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.)

Similarly v_∞ is a TV-uniform value if for each $\varepsilon > 0$, there exists $\delta > 0$ and σ^* strategy of Player 1 such that $TV(\mu) \leq \delta$ implies $E_{\sigma^*,\tau}\langle \mu, g\rangle \geq v_\infty - \varepsilon$, for all τ (and a dual statement for Player 2).

Obviously a TV-uniform value is a TV-asymptotic value.

2.9.1.2 Dynamic programming and TV-asymptotic value

In the framework of dynamic programming (see Section 5.3), Renault (2013) proves that given a sequence of evaluations $\{\mu^k\}$ with $TV(\mu^k) \to 0$, the corresponding sequences of values v_{μ^k} converge (uniformly) iff the family $\{v_{\mu^k}\}$ is totally bounded (pre-compact). In this case the limit is v^* with:

$$v^*(\omega) = \inf_{\mu\in\mathcal{M}} \sup_m v_{m\circ\mu}(\omega) = \inf_k \sup_m v_{m\circ\mu^k}(\omega)$$

where $\mathcal{M}$ is the set of evaluations and $m \circ \mu$ is the m-translation of the evaluation μ ($m \circ \mu_n = \mu_{m+n}$). In particular a TV-asymptotic value exists if Ω is a totally bounded metric space and the family $\{v_\mu, \mu \in \mathcal{M}\}$ is uniformly equicontinuous.

2.9.1.3 Dynamic programming and TV-uniform value

Renault and Venel (2012) provide condition for an MDP to have a TV-uniform value and give a characterization of it, in the spirit of gambling, namely: excessive property and invariance. They also introduce a new metric on $\Delta_f(X)$ where $X = \Delta(K)$ with K finite, that allows them to prove the existence of a TV-uniform value in MPD with finite state space and signals.

2.9.1.4 Games with a more informed controller

Renault and Venel (2012) prove that game with an informed controller has a TV-uniform value extending the result in Renault (2012).

An extension in another direction is due to Gensbittel et al. (2013). Define a game with a more informed controller as a situation where:

(i) the belief $\zeta_n^1(1) \in M^1(1) = \Delta(M)$ of Player 1 on the state m_n at stage n (i.e., the conditional probability given his information h_n^1) is more precise than the belief

$\zeta_n^2(1) \in M^2(1) = \Delta(M)$ of Player 2 (adding the information h_n^2 to h_n^1 would not affect $\zeta_n^1(1)$).

(ii) Player 1 knows the belief $\zeta_n^2(2) \in M^2(2) = \Delta(M^1(1))$ of Player 2 on his own beliefs.

(iii) Player 1 control the process $\{\zeta_n^2(2)\}$.

Then the game has a uniform value.

The analysis is done trough an auxiliary game with state space $M^2(2)$ and where the result of Renault (2012) applies. The difficulty is to prove that properties of optimal strategies in the auxiliary game can be preserved in the original one.

2.9.1.5 Comments

Note that the analysis in Cardaliaguet et al. (2012) shows in fact that a TV-asymptotic value exists in incomplete information, absorbing and splitting games. Even more, v_μ is close to v as soon as $\|\mu\|$ is small enough. However, these results are very specific to these classes.

Recently, Ziliotto (2014) has given examples of a finite stochastic game with no TV-asymptotic value and of an absorbing game with no TV-uniform value (actually the Big Match).

In particular, the above results in Section 9.1 do not extend from the one-Player to the two-Player case.

2.9.2 Markov games with incomplete information on both sides

Gensbittel and Renault (2012) proved recently the existence of $\lim_{n\to\infty} v_n$ in repeated games with incomplete information in which the types of Players 1 and 2 evolve according to two independent Markov chains $\mathbf{K}$ and $\mathbf{L}$ on two finite sets, respectively K and L. The parameter $\{k_m\}$ follows $\mathbf{K}$, starting from some distribution p at stage one and k_m is observed by Player 1 at stage m. The distribution of the state k_m is thus given by $p\mathbf{K}^m$. A similar situation, with $\mathbf{L}$ and q, holds for Player 2.

Moves are revealed along the play and the stage m payoff is given by $g_m = g(k_m, l_m, i_m, j_m)$.

The authors show that $v = \lim_{n\to\infty} v_n$ exists and is the unique continuous solution of a "Mertens-Zamir"-like system of equations with respect to a function $u^*(p,q)$. In addition to be concave/convex, the function v has to be invariant through the Markov chain: $v(p,q) = v(p\mathbf{K}, q\mathbf{L})$. As for u^*, it is the uniform limit $u_n^*(p,q)$ of the sequence of values of a n-stage repeated "non-revealing game" $\Gamma_n^*(p,q)$. This game is similar to the one introduced by Renault (2006) to solve the one side incomplete information case (see Section 5.4). In $\Gamma_n^*(p,q)$, Player 1 plays as in the original game but is restricted to use strategies that keep the induced beliefs of Player 2 constant on the partition

$\tilde{K}$ of K into recurrence classes defined by the Markov chain $\mathbf{K}$ (and similarly for Player 2).

The proof of uniform convergences of u_n^* is difficult and uses, in particular, a splitting lemma in Laraki (2004) to show equi-continuity of the family of functions $\{u_n^*\}$.

2.9.3 Counter examples for the asymptotic approach

2.9.3.1 Counter example for finite state stochastic games with compact action spaces

Vigeral (2013) gives an example of a continuous stochastic game on $[0, 1]^2$ with four states (two of them being absorbing) where the asymptotic value does not exists: the family $\{v_\lambda\}$ oscillates as $\lambda \to 0$. The payoff is independent of the moves and the transition are continuous. However, under optimal strategies, the probability to stay in each state goes to zero and the induced occupation measure on the two states oscillate as $\lambda \to 0$. An analog property holds for v_n.

This example answers negatively a long standing conjecture for stochastic games with finite state space. In fact, the operator and variational approaches were initially developed to prove the existence of $\lim_{\lambda\to 0} v_\lambda$ and $\lim_{n\to\infty} v_n$. They work in the class of irreversible games (in which, once one leaves a state, it cannot be reached again, such as absorbing, recursive, and incomplete information games). Outside this class, the asymptotic value may not exist.

2.9.3.2 Counter examples for games with finite parameter sets

A phenomena similar to the previous one is obtained by Ziliotto (2013) for a game, say G_4 played with countable action set and finite state space.

More interestingly, this game is obtained by a series of transformations preserving the value and starting from a game G_1 which is a finite stochastic game where the Players have symmetric information, hence know the moves, but do not observe the state. In that case, the game is equivalent to a stochastic game with standard signaling and a countable state space $\Omega' \subset \Delta(\Omega)$.

Explicitly, Player 1 has three moves: Stay1, Stay2, and Quit, and Player 2 has two moves: Left and Right. The payoff is -1 and the transition are as follows:

A	***Left***	***Right***
Stay1	A	$(\frac{1}{2}A + \frac{1}{2}B)$
Stay2	$(\frac{1}{2}A + \frac{1}{2}B)$	A
Quit	A^*	A^*

B	***Left***	***Right***
Stay1	A	B
Stay2	B	A
Quit	B^+	B^+

The initial state is A and A^* is an absorbing state with payoff -1. During the play, the Players will generate common beliefs on the states $\{A, B\}$. Once B^+ is reached, a dual game with a similar structure and payoff 1 is played. Now, Player 2 has three moves: Player 1 two moves: the transition is similar but different from the one above.

In an equivalent game G_2, only one Player plays at each stage and he receives a random signal on the transition (like Player 1 facing $(1/2, 1/2)$ in the above game).

Finally, G_3 is a stochastic game with known state and a countable state space corresponding to the beliefs of the Player on $\{A, B\}$, which are of the form $\{1/2^n, 1 - 1/2^n\}$.

In all cases, the families $\{v_\lambda\}$ and $\{v_n\}$ oscillate.

The main point is that the discrete time in the repeated interaction generates a discrete countable state space while the evolution of the parameter λ is continuous.

This spectacular result answers negatively two famous conjectures of Mertens (1987): existence of the asymptotic value in games with finite parameter spaces and its equality with the maxmin whenever Player 1 is more informed than Player 2. In fact, the example shows more, the asymptotic value may not exist even in the case of symmetric information!

2.9.3.3 Oscillations

To understand the nature and links between the previous examples, Sorin and Vigeral (2013) construct a family of configurations which are zero-sum repeated games in discrete time where the purpose is to control the law of a stopping time of exit. For a given discount factor $\lambda \in]0, 1]$, optimal stationary strategies define an inertia rate Q_λ ($1 - Q_\lambda$ is the normalized probability of exit during the game).

When two such configurations (1 and 2) are coupled, this induces a stochastic game where the state will move from one to the other in a way depending on the previous rates $Q_\lambda^i, i = 1, 2$ and the discounted value will satisfy:

$$v_\lambda^i = a^i Q_\lambda^i + (1 - Q_\lambda^i) v_\lambda^{-i}, \quad i = 1, 2$$

where a^i is the payoff in configuration i. The important observation is that the discounted value is a function of the ratio $\frac{Q_\lambda^1}{Q_\lambda^2}$. It can oscillate as $\lambda \to 0$, when both inertia rates converge to 0.

The above analysis shows that oscillations in the inertia rate and reversibility allow for nonconvergence of the discounted values.

2.9.3.4 Regularity and o-minimal structures

An alternative way to keep regularity in the asymptotic approach is to avoid oscillations. This is the case when the state space is finite (due to the algebraic property of v_λ). Recently, Bolte et al. (2013) extend the algebraic approach to a larger class of stochastic games with compact action sets.

The concept of o-minimal structure was introduced recently as an extension of semi-algebraic geometry through an axiomatization of its most important properties (see van den Dries, 1998). It consists in a collection of subsets of $\mathbf{R}^n$, for each $n \in \mathbb{N}$, called *definable* sets. Among other natural requirements, the collection need to be stable by linear projection, its "one-dimensional" sets must consist on the set of finite unions of intervals, and must at least contain all semi-algebraic sets. It shares most of the nice properties of semi-algebraic sets, in particular, finiteness of the number of connected components.

Theorem 2.37. *If (the graph of) the Shapley operator is definable, v_λ is of bounded variations hence the uniform value exists.*

This raises the question: under which condition is the Shapley operator definable?

A function f on $\Omega \times I \times J$ is separable if there are finite sets S and T, and functions a_s, b_t, and $c_{s,t}$ such that $f(i,j,\omega) = \sum_{s\in S}\sum_{t\in T} c_{s,t}(\omega)a_s(i,\omega)b_t(j,\omega)$. This is the case for polynomial games or games where one of the payers has finitely many actions. Then one proves that if transition and payoff functions are separable and definable in some o-minimal structure, then the Shapley operator is definable in the *same* structure.

The Stone-Weierstrass theorem is then used to drop the separability assumption on payoffs, leading to the following extension of fundamental result of Mertens and Nemann.

Theorem 2.38. *Any stochastic games with finitely many states, compact action sets, continuous payoff functions, and definable separable transition functions has a uniform value.*

An example shows that with semi-algebraic data, the Shapley operator may not be semi-algebraic, but belongs to a higher o-minimal structure. Hence, the open question: do a definable stochastic game has a definable Shapley operator in a larger o-minimal structure?

2.9.4 Control problem, martingales, and PDE

Cardaliaguet and Rainer (2009b) consider a continuous time game $\Gamma(p)$ on $[0,T]$ with incomplete information on one side and payoff function $\gamma^k(t,u,v)$, $k \in K$. Player 1 knows k and both Players know its distribution $p \in \Delta(K)$. The value of the nonrevealing local game is $U(t,p) = \mathtt{val}_{U\times V}\sum_k p^k\gamma^k(t,u,v)$. This is a particular differential game with incomplete information where the state variable is not controlled. Hence, all previous primal and dual PDE characterizations of the value $V(t,p)$ style hold (Section 6.3.). Moreover, there is a martingale maximization formulation (compare with the splitting game, Section 3.4.2 C). Explicitly:

Theorem 2.39. $V(\cdot\,,\cdot)$ *is:*

(a) *the smallest function, concave in p and continuous viscosity solution of $\frac{\partial f}{\partial t}(t,p) + U(t,p) \leq 0$, with boundary condition $f(1,p) = 0$.*

(b) *the value of the control problem $\max_{p_t} \mathbb{E}[\int_0^1 U(s,p_t)\mathrm{d}t]$ where the maximum is over all continuous time càdlàg martingales $\{p_t\}$ on $\Delta(K)$ that starts at p.*

Any maximizer in the above problem induces an optimal strategy of the informed Player in the game. Cardaliaguet and Rainer (2009b) provide more explicit computations and show in particular that, unlike in Aumann-Maschler's model where only one splitting is made at the beginning and no further information is revealed, in the nonautonomous case information may be revealed gradually or in the middle of the game, depending on how U varies with t. Such phenomena cannot be observed in RG: they are time-independent.

Grün (2012) extends these characterizations (PDE in the primal and dual games, martingale maximization) to stochastic continuous time games with incomplete information on one side.

Cardaliaguet and Rainer (2012) generalized the result to the case where $K = \mathbb{R}^d$ so that the PDE formulation is on $\Delta(\mathbb{R}^d)$, like in Cardaliaguet and Quincampoix (2008).

The martingale maximization problem appears also when the state evolves along the play according to a Markov chain (Cardaliaguet et al., 2013) (see Section 9.5.5.).

Very recently, Gensbittel (2013) generalizes the characterizations to continuous time games where Players are gradually informed, but Player 1 is more informed. Formally, information is modeled by a stochastic process $Z_t = (X_t, Y_t)$, where X_t is a private information for Player 1 and Y_t is a public information. This is a first step toward understanding De Meyer's financial model when Players get information gradually and not only at once.

2.9.5 New links between discrete and continuous time games

2.9.5.1 Multistage approach

The game is played in discrete time, stage after stage. However, stage n represents more the n^{th} interaction between the Players rather than a specific time event. In addition stage n has no intrinsic length: given an evaluation μ, its duration in the normalized game on $[0,1]$ is μ_n. Finally, the variation of the state variable at each stage is independent of the evaluation.

2.9.5.2 Discretization of a continuous time game

An alternative framework for the asymptotic approach is to consider an increasing frequency of the interactions between the Players. The underlying model is a game in continuous time played on $[0, +\infty]$ with a state variable Z_t which law depends upon the actions on the Players.

The usual repeated game G^1 corresponds to the version where the Players play at integer times $1, 2, \ldots$ (or that their moves are constant on the interval of time $[n, n+1]$). Then the law of Z_{n+1} is a function of Z_n, i_n, j_n.

In the game G^δ, the timing of moves is still discrete and the play is like above; however, stage n corresponds to the time interval $[n\delta, (n+1)\delta)$ and the transition from Z_n to Z_{n+1} will be a function of Z_n, i_n, j_n and of the length δ (see Neyman, 2012b).

In particular, as $\delta \to 0$ the play should be like in continuous time but with "smooth" transitions.

In addition, some evaluation criteria has to be given to integrate the process of payoffs on $[0, +\infty]$.

2.9.5.3 Stochastic games with short stage duration

A typical such model is a stochastic game with finite state space Ω where the state variable follows a continuous Markov chain controlled by the Players trough a generator $A(i,j), i \in I, j \in J$ on $\Omega \times \Omega$ (see the initial version of Zachrisson, 1964).

Consider the discounted case where the evaluation on $[0, +\infty]$ is given by re^{-rt}. We follow Neyman (2013). Given a duration $h \geq 0$ let $P_h = \exp(hA)$ be the transition kernel associated to A and stage duration h.

Then the value of the corresponding $G^{\delta,r}$ game satisfies:

$$v^{\delta,r}(\omega) = \mathtt{val}_{X\times Y}\Big\{(1 - e^{-r\delta})g(x, y, \omega) + e^{-r\delta} \sum_{\omega'} P_\delta(x, y, \omega)[\omega']v^{\delta,r}(\omega')\Big\} \qquad [2.57]$$

and converges, as $\delta \to 0$, to v^r solution of:

$$v^r(\omega) = \mathtt{val}_{X\times Y}\Big\{g(x, y, \omega) + \sum_{\omega'} A(x, y, \omega)[\omega']v^r(\omega')/r\Big\} \qquad [2.58]$$

which is also the value of the $r\delta$-discounted repeated game with transition $Id + \delta A/(1 - r\,\delta)$, for any δ small enough.

Optimal strategies associated to (2.58) induce stationary strategies that are approximatively optimal in the game $G^{\delta,r}$ for δ small enough. The convergence of the values to v^r holds for all games with short duration stages, not necessarily uniform (compare with the limit game 8.3).

Replacing the discounted evaluation by a decreasing evaluation θ on $[0, +\infty]$ gives similar results with Markov optimal strategies for the limit game G^θ of the family $G^{\delta,\theta}$.

In the finite case (I and J finite), the above system defining v^r is semi algebraic and $\lim_{r\to 0} v^r = v^0$ exists. Using an argument of Solan (2003), one shows that the convergence of $v^{\delta,r}$ to v^r is uniform in r.

In addition, Neyman (2013), considers a family of repeated games with vanishing stage duration and adapted generators and gives condition for the family of values to converge.

Finally, the existence of a uniform value (with respect to the duration t for all δ small enough) is studied, using Theorem 2.17.

2.9.5.4 Stochastic games in continuous time

Neyman (2012b) also introduced directly a game Γ played in continuous time. Basically, one requires the strategy of each Player to be independent of his previous short past behavior. In addition, Players use and observe mixed strategies. This implies that facing "simple strategies" a strategy is inducing a single history (on short time interval) and the play is well defined.

Optimal strategies associated to the game [2.58] induce optimal stationary strategies in the discounted game Γ^r, which value is v^r (see also Guo and Hernandez-Lerma, 2003 and 2005).

Consider the operator $\mathbf{N}$ on $\mathbb{R}^\Omega$ defined by :

$$\mathbf{N}[v](\omega) = \mathtt{val}_{X\times Y}\Big\{g(x,y,\omega) + \sum_{\omega'} A(x,y,\omega)[\omega']v(\omega')\Big\}. \qquad [2.59]$$

$\mathbf{N}$ defines a semigroup of operators with $S_0 = Id$ and $\lim_{h\to 0} \frac{S_h - Id}{h} = \mathbf{N}$. Then the value of the continuous game of length t, Γ_t is $S_t(0)/t$ (compare with [2.8] and [2.10]). Optimal strategies are obtained by playing optimally at time s in the game associated to $\mathbf{N}(S_{t-s}(0))(\omega_s)$.

Finally, the game has a uniform value.

2.9.5.5 Incomplete information games with short stage duration

Cardaliaguet et al. (2013) consider a model similar to 9.4.2 with a given discount factor r. Only Player 1 is informed upon the state which follows a continuous time Markov chain with generator A. The initial distribution is p and known by both Players. This is the "continuous time" analog of the model of Renault (2006); however, there are two main differences: the asymptotics is not the same (large number of stages versus short stage duration) and in particular the total variation of the sate variable is unbounded in the first model (take a sequence of i.i.d random variables on K) while bounded in the second. Note that in the initial Aumann-Maschler framework where the state is fixed during the play, the two asymptotics are the same.

Cardaliaguet et al. (2013) considers the more general case where the Markov chain is controlled by both Players (thus A is a family $A(i,j)$) and proves that the limit (as $\delta \to 0$) of the values $v^{\delta,r}$ of $G^{\delta,r}$, the r-discounted game with stage duration δ, exists and is the unique viscosity solution of an Hamilton-Jacobi equation with a barrier (compare with [2.50]):

$$\min \left\{ rv(p) - \mathtt{val}_{X\times Y}[rg(x,y,p) + \langle \nabla v(p), pA(x,y)\rangle], -\bar{L}v(p) \right\} = 0. \qquad [2.60]$$

A similar result holds for the case with lack of information on both sides where each Player controls his privately known state variable that evolves through a continuous time Markov chain.

Note that the underlying nonrevealing game corresponds to a game where the state follows a continuous time Markov chain, but the Players are not informed, which is in the spirit of Cardaliaguet and Quincampoix (2008). Its value satisfies:

$$u^{\delta,r}(p) = \mathtt{val}_{X\times Y}\left\{(1-e^{r\delta})g(x,y,p) + e^{r\delta}\mathsf{E}_{x,y}[v^{\delta,r}(pP_\delta(i,j))]\right\}. \qquad [2.61]$$

hence at the limit

$$u^r(p) = \mathtt{val}_{X\times Y}\left\{g(x,y,p) + \langle \nabla u^r(p)/r, pA(x,y)\rangle\right\}. \qquad [2.62]$$

2.9.6 Final comments

Recent developments in the field of two-Player zero-sum repeated games are promising and challenging. On the one hand, one sees the emergence of many new models, deep techniques, and a unifying theory of dynamic games, including both discrete- and continuous-time considerations and dealing with incomplete information, stochastic, and signaling aspects at the same time. On the other hand, we know now that v_λ and v_n may not converge even in finite RG and the characterization of the class of regular games (where both $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ exist and are equal) is an important challenge.

In any case, a reflexion is necessary on the modeling aspect, and especially on the relation between discrete and continuous time formulations. The natural way to take the limit in a long repeated interaction is by considering the *relative* length of a stage compared to the total length, obtaining thus a continuous-time representation on the compact interval $[0,1]$. However, in the presence of an exogenous duration process on $\mathbb{R}^+$ (like the law of a state variable), this normalization on $[0,1]$ is no longer possible, and one would like the actual length of a stage to go to zero, leading to a continuous time game on $\mathbb{R}^+$. Also, some continuous-time games do not have faithful discretization (such as stopping games), but this is not only due to the variable "time," it also occurs with discrete versus continuum strategy spaces.

Finally, we should stress here that all these recent advances are expected to have a fundamental impact on the study of *nonzero-sum* games as well, such as nonexistence of limiting equilibrium payoffs in discounted games, folk theorems in continuous-time stochastic games, modeling imperfect monitoring in continuous-time games, nonautonomous transition and payoff functions, etc.

ACKNOWLEDGMENTS

The authors thank N. Vieille for precise and useful comments.

This research was supported by grant ANR-10-BLAN 0112 (Laraki) and by the European Union under the 7th Framework Program "FP7-PEOPLE-2010-ITN", grant agreement number 264735-SADCO (Sorin).

REFERENCES

As Soulaimani, S. 2008. Viability with probabilistic knowledge of initial condition, application to optimal control, Set-Valued Anal. 16, 1037–1060.

Assoulamani, S., Quincampoix, M., Sorin, S., 2009. Repeated games and qualitative differential games: approachability and comparison of strategies, SIAM J. Control Optim. 48, 2461–2479.

Aumann, R.J., Maschler, M., 1995. Repeated Games with Incomplete Information. M.I.T. Press (with the collaboration of R. Stearns), Cambridge MA, USA.

Bardi, M., 2009. On differential games with long-time-average cost. In: Advances in Dynamic Games and their Applications, Bernhard P., Gaitsgory, V., Pourtalier, O., (Eds.), Annals of ISDG, vol. 10, Birkhauser, pp. 3–18.

Bardi, M., Capuzzo Dolcetta, I., 1996. Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhauser.

Bewley, T., Kohlberg, E., 1976a. The asymptotic theory of stochastic games. Math Operat. Res. 1, 197–208.

Bewley, T., Kohlberg, E., 1976b. The asymptotic solution of a recursion equation occurring in stochastic games. Math. Operat. Res. 1, 321–336.

Bich, P., Laraki, R., 2013. On the existence of approximated and sharing rule equilibria in discontinuous games. Preprint hal-00846143.

Blackwell, D., 1956. An analog of the minmax theorem for vector payoffs. Paci. J. Math. 6, 1–8.

Blackwell, D., 1969. Infinite G_δ games with imperfect information, Applicationes Mathematicae X, 99–101.

Blackwell, D., Ferguson, T., 1968. The Big Match. Ann. Math. Stat. 39, 159–163.

Bolte, J., Gaubert, S., Vigeral, G., 2013. Definable zero-sum stochastic games. Preprint. arXiv:1301.1967v1.

Brézis, H., 1973. Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert, North Holland, Amsterdam.

Buckdahn, R., Li, J., Quincampoix, M., 2013. Value function of differential games without Isaacs conditions: An approach with nonanticipative mixed strategies. Int. J. Game Theory 42, 989–1020.

Buckdahn, R., Cardaliaguet, P., Quincampoix, M., 2010. Some recent aspects of differential game theory. Dyn. Games Appl. 1, 74–114. ellman-Isaacs equations. SIAM J. Control Optim. 47, 444–475.

Cardaliaguet, P., 1996. A differential game with two Players and one target. SIAM J. Control Optim. 34, 1441–1460.

Cardaliaguet, P., 2007. Differential games with asymmetric information. SIAM J. Control Optim. 46, 816–838.

Cardaliaguet, P., 2008. Representation formulas for differential games with asymmetric information. J. Opimi. Theory Appl. 138, 1–16.

Cardaliaguet, P., 2009. Numerical approximation and optimal strategies for dfferential games with lack of information on one side. Advances in Dynamic Games and their Applications. In: Bernhard P., Gaitsgory, V., Pourtalier, O., (Eds.), Annals of ISDG, vol. 10. Birkhauser, pp. 159–176.

Cardaliaguet, P., Quincampoix, M., 2008. Deterministic differential games under probability knowledge of initial condition. Int. Game Theory Rev. 10, 1–16.

Cardaliaguet, P., Rainer, C., 2009a. Stochastic differential games with asymmetric information, Appl. Math. Optim., 59, 1–36.
Cardaliaguet, P., Rainer, C., 2009b. On a continuous time game with incomplete information. Math. Operat. Res. 34, 769–794.
Cardaliaguet, P., Rainer, C., 2012. Games with incomplete information in continuous time and for continuous types. Dyn. Games Appl. 2, 206–227.
Cardaliaguet, P., Souquière, A., 2012. A differential game with a blind Player. SIAM J. Control and Optimization 50, 2090–2116.
Cardaliaguet, P., Laraki, R., Sorin, S., 2012. A continuous time approach for the asymptotic value in two-person zero-sum repeated games. SIAM J. Control Optim. 50, 1573–1596.
Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P., 2007. Differential games through viability theory: Old and recent results. Advances in Dynamic Game Theory. In: Jorgensen S., Quincampoix, M., Vincent, T., (Eds.), Annals of ISDG, vol. 9. Birkhauser, pp. 3–36.
Cardaliaguet, P., Rainer, C., Rosenberg, D., Vieille, N., 2013. Markov games with frequent actions and incomplete information. Preprint hal 00843475.
Coulomb, J.-M., 1992. Repeated games with absorbing states and no signals. Int. J. Game Theory 21, 161–174.
Coulomb, J.-M., 1996. A note on 'Big Match'. ESAIM.: Probab. Stat. 1, 89–93, http://www.edpsciences.com/ps/.
Coulomb, J.-M., 1999. Generalized Big Match. Math. Operat. Res. 24, 795–816.
Coulomb, J.-M., 2001. Repeated games with absorbing states and signalling structure. Math. Operat. Res. 26, 286–303.
Coulomb, J.M., 2003a. Stochastic games without perfect monitoring. Int. J. Game Theory 32, 73–96.
Coulomb, J.M., 2003b. Games with a recursive structure. Stochastic Games and Applications. In: Neyman A. and Sorin, S., (eds), vol. C 570. NATO Science Series, Kluwer Academic Publishers, pp. 427–442.
Crandall, M.G., Ishii, H., Lions, P.-L., 1992. User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Soc. 27, 1–67.
De Meyer, B., 1996a. Repeated games and partial differential equations. Math. Operat. Res. 21, 209–236.
De Meyer, B., 1996b. Repeated games, duality and the Central Limit theorem. Math. Operat. Res. 21, 237–251.
De Meyer, B., 1999. From repeated games to Brownian games. Ann. Inst. Henri Poinc. Prob. Statist. 35, 1–48.
De Meyer, B., 2010. Price dynamics on a stock market with asymmetric information. Games Econ. Behav. 69, 42–71.
De Meyer, B., Marino, A., 2005. Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides. Cahier de la MSE 2005–27.
De Meyer, B., Moussa-Saley, H., 2002. On the strategic origin of Brownian motion in finance. Int. J. Game Theory 31, 285–319.
De Meyer, B., Rosenberg, D., 1999. "Cav u" and the dual game. Math. Operat. Res. 24, 619–626.
Dynkin, E.B., 1969. Game variant of a problem on optimal stopping. Soviet Math. Dokl. 10, 270–274.
Evans, L.C., Souganidis, P.E., 1984. Differential games and representation formulas for solutions of Hamilton-Jacobi equations. Indiana Univ. Math. J. 33, 773–797.
Everett, H., 1957. Recursive games. Dresher M., Tucker, A.W., Wolfe, P., (Eds.), Contributions to the Theory of Games, III, Annals of Mathematical Studies. vol. 39. Princeton University Press, pp. 47–78.
Forges, F., 1982. Infinitely repeated games of incomplete information: Symmetric case with random signals. Int. J. Game Theory 11, 203–213.
Friedman, A., 1994. Differential games. Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory. vol. 2. Elsevier, pp. 781–799.
Gale, D., Stewart, F.M., 1953. Infinite games with perfect information. Kuhn H., Tucker, A.W., (Eds.), Contributions to the Theory of Games, II, Annals of Mathematical Studies. vol. 28. Princeton University Press, pp. 245–266.

Gensbittel, F., 2012. Extensions of the Cav(u) theorem for repeated games with one sided information. Math. Operat. Res. Preprint, hal-00745575.
Gensbittel, F., 2013. Covariance control problems of martingales arising from game theory. SIAM J. Control Optim. 51(2), 1152–1185.
Gensbittel, F., 2013. Continuous-time limit of dynamic games with incomplete information and a more informed Player. Preprint hal 00910970.
Gensbittel, F., Renault, R., 2012. The value of Markov chain games with incomplete information on both sides. Preprint arXiv:1210.7221.
Gensbittel, F., Oliu-Barton, M., Venel, X., 2013. Existence of the uniform value in repeated games with a more informed controller. Preprint hal-00772043.
Grün, C., 2012. A BSDE approach to stochastic differential games with incomplete information. Stoch. Proc. Appl. 122(4), 1917–1946.
Grün, C., 2013. On Dynkin games with incomplete information. Preprint arXiv:1207.2320v1.
Guo, X., Hernandez-Lerma, O., 2003. Zero-sum games for continuous-time Markov chains with unbounded transition and average payoff rates. J. Appl. Prob. 40, 327–345.
Guo, X., Hernandez-Lerma, O., 2005. Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates. Bernoulli 11, 1009–1029.
Heuer, M., 1992a. Asymptotically optimal strategies in repeated games with incomplete information. Int. J. Game Theory 20, 377–392.
Heuer, M., 1992b. Optimal strategies for the uninformed Player. Int. J. Game Theory 20, 33–51.
Hörner, J., Rosenberg, D., Solan, E., Vieille, N., 2010. On a Markov game with one-sided incomplete information. Operat. Res. 58, 1107–1115.
Kohlberg, E., 1974. Repeated games with absorbing states. Ann Stat. 2, 724–738.
Kohlberg, E., 1975. Optimal strategies in repeated games with incomplete information. Int. J. Game Theory 4, 7–24.
Kohlberg, E., Zamir, S., 1974. Repeated games of incomplete information: the symmetric case. Ann. Stat. 2, 1040.
Kohn, R.V., Serfaty, S., 2006. A deterministic-control-based approach to motion by curvature. Communi. Pure Appl. Math. 59, 344–407.
Krasovskii, N.N., Subbotin, A.I., 1988. Game-Theoretical Control Problems. Springer.
Krausz, A., Rieder, U., 1997. Markov games with incomplete information. Math. Methods Operat. Res. 46, 263–279.
Laraki, R., 2000. Repeated games with incomplete information: a variational approach. PhD thesis, UPMC, Paris, France.
Laraki, R., 2001a. Variational inequalities, systems of functional equations and incomplete information repeated games. SIAM J. Control Optim. 40, 516–524.
Laraki, R., 2001b. The splitting game and applications. Int. J. Game Theory 30, 359–376.
Laraki, R., 2002. Repeated games with lack of information on one side: the dual differential approach. Math. Operat. Res. 27, 419–440.
Laraki, R., 2004. On the regularity of the convexification operator on a compact set, J. Convex Anal. 11, 209–234.
Laraki, R., 2010. Explicit formulas for repeated games with absorbing states. Int. J. Game Theory 39, 53–69.
Laraki, R., Solan, E., 2005. The value of zero-sum stopping games in continuous time. SIAM J. Control Optimi. 43, 1913–1922.
Laraki, R., Sudderth, W.D., 2004. The preservation of continuity and Lipschitz continuity by optimal reward operators. Mathe. Operat. Res. 29, 672–685.
Lehrer, E., 2002. Approachability in infinite dimensional spaces. Int. J. Game Theory 31, 253–268.
Lehrer, E., Monderer, D., 1994. Discounting versus averaging in dynamic programming. Games and Econ. Behav. 6, 97–113.
Lehrer, E., Solan, E., 2007. Learning to play partially specified equilibrium. Preprint.
Lehrer, E., Sorin, S., 1992. A uniform Tauberian theorem in dynamic programming. Math. Operat. Res. 17, 303–307.
Lepeltier, J.P., Mainguenau, M.A., 1984. Le jeu de Dynkin en théorie générale sans l' hypothèse de Mokobodski. Stochastics 13, 25–44.

Levy, Y., 2002. Stochastic games with information lag. Games Econ. Behav. 74, 243–256.
Maitra, A.P., Sudderth, W.D., 1992. An operator solution of stochastic games. Israel J. Math. 78, 33–49.
Maitra, A.P., Sudderth, W.D., 1993. Borel stochastic games with lim sup payoff. Ann. Probab. 21, 861–885.
Maitra, A.P., Sudderth, W.D., 1996. Discrete Gambling and Stochastic Games. Springer.
Maitra, A.P., Sudderth, W.D., 1998. Finitely additive stochastic games with Borel measurable payoffs, Int. J. Game Theory 27, 257–267.
Maitra, A.P., Sudderth, W.D., 2003. Borel stay-in-a-set games. Int. J. Game Theory 32, 97–108.
Marino, A., 2005. The value and optimal strategies of a particular Markov chain game. Preprint.
Martin, D.A., 1975. Borel determinacy. Ann. Math. 102, 363–371.
Martin, D.A., 1998. The determinacy of Blackwell games. J. Symbolic Logic 63, 1565–1581.
Mertens, J.-F., 1972. The value of two-person zero-sum repeated games: the extensive case. Int. J. Game Theory 1, 217–227.
Mertens, J.-F., 1987. Repeated games. Proceedings of the International Congress of Mathematicians, Berkeley 1986, American Mathematical Society, pp. 1528–1577.
Mertens, J.-F., 1998. The speed of convergence in repeated games with incomplete information on one side. Int. J. Game Theory 27, 343–359.
Mertens, J.-F., 2002. Stochastic games. In: Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory, vol. 3. Elsevier, pp. 1809–1832.
Mertens, J.-F., Neyman, A., 1981. Stochastic games. Int. J. Game Theory 10, 53–66.
Mertens, J.-F., Zamir, S., 1971. The value of two-person zero-sum repeated games with lack of information on both sides. Int. J. Game Theory 1, 39–64.
Mertens, J.-F., Zamir, S., 1976a. On a repeated game without a recursive structure. Int. J. Game Theory 5, 173–182.
Mertens, J.-F., Zamir, S., 1976b. The normal distribution and repeated games. Int. J. Game Theory 5, 187–197.
Mertens, J.-F., Zamir, S., 1985. Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29.
Mertens, J.-F., Zamir, S., 1995. Incomplete information games and the normal distribution. CORE D.P. 9520.
Mertens, J.-F., Neyman, A., Rosenberg, D., 2009. Absorbing games with compact action spaces. Math. Operat. Res. 34, 257–262.
Mertens, J.-F., Sorin, S., Zamir, S., 1994. Repeated Games, CORE D.P. 9420-21-22, to appear Cambridge University Press 2014.
Milman, E., 2006. Approachable sets of vector payoffs in stochastic games. Games Econ. Behav. 56, 135–147.
Monderer, D., Sorin, S., 1993. Asymptotic properties in dynamic programming. Int. J. Game Theory 22, 1–11.
Neveu, J., 1975. Discrete-Parameter Martingales. North-Holland, Amsterdam.
Neyman, A., 2003. Stochastic games and nonexpansive maps. In: Neyman, A., Sorin, S., (Eds.), Stochastic Games and Applications, NATO Science Series C 570, Kluwer Academic Publishers, pp. 397–415.
Neyman, A., 2008. Existence of optimal strategies in Markov games with incomplete information. Int. J. Game Theory 37, 581–596.
Neyman, A., 2012a. The value of two-person zero-sum repeated games with incomplete information and uncertain duration. Int. J. Game Theory 41, 195–207.
Neyman, A., 2012b. Continuous-time stochastic games. DP 616, CSR Jerusalem.
Neyman, A., 2013. Stochastic games with short-stage duration. Dyn. Games Appl. 3, 236–278.
Neyman, A., Sorin, S., 1998. Equilibria in repeated games of incomplete information: the general symmetric case. Int. J. Game Theory 27, 201–210.

Neyman, A., Sorin, S., (Eds.) 2003. Stochastic Games and Applications, NATO Science Series, vol. C 570. Kluwer Academic Publishers.
Neyman, A., Sorin, S., 2010. Repeated games with public uncertain duration process. Int. J. Game Theory 39, 29–52.
Oliu-Barton, M., 2013. The asymptotic value in finite stochastic games. Math. Operat. Res. to appear.
Oliu-Barton, M., Vigeral, G., 2013. A uniform Tauberian theorem in optimal control. In: Cardaliaguet, P., Cressman, R., (Eds.), Advances in Dynamic Games, Annals of the International Society of Dynamic Games, vol. 12. pp. 199–215.
Perchet, V., 2009. Calibration and internal no-regret with random signals. In: Gavalda, R., Lugosi, G., Zegmann, T., Zilles, S., (Eds.), Algorithmic Learning Theory, vol. LNAI 5809, Springer, pp. 68–82.
Perchet, V., 2011a. Approachability of convex sets in games with partial monitoring. J. Optim. Theory Appl. 149, 665–677.
Perchet, V., 2011b. Internal regret with partial monitoring calibration-based optimal algorithm. J. Mach. Learn. Res. 12, 1893–1921.
Perchet, V., Quincampoix, M., 2012. On an unified framework for approachability in games with or without signals. Preprint hal- 00776676.
Peres, Y., Schramm, O., Sheffield, S., Wilson, D., 2009. Tug of war and the infinity Laplacian. J. Am. Math. Soc. 22, 167–210.
Quincampoix, M., Renault, J., 2011. On the existence of a limit value in some non expansive optimal control problems. SIAM J. Control Optim. 49, 2118–2132.
Radzik, T., Raghavan, T.E.S., 1994. Duels. In: Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory, vol. 2, pp. 761–768.
Renault, J., 2006. The value of Markov chain games with lack of information on one side. Math. Operat. Res. 31, 490–512.
Renault, J., 2011. Uniform value in dynamic programming. J. Eur. Math. Soc. 13, 309–330.
Renault, J., 2012. The value of repeated games with an informed controller. Math. Operat. Res. 37, 154–179.
Renault, J., 2013. General limit value in dynamic programming. Preprint hal 00769763.
Renault, J., Venel, X., 2012. A distance for probability spaces, and long-term values in Markov decision processes and repeated games. Preprint hal-00674998.
Rosenberg, D., 1998. Duality and Markovian strategies. Int. J. Game Theory 27, 577–597.
Rosenberg, D., 2000. Zero-sum absorbing games with incomplete information on one side: asymptotic analysis. SIAM J. Control Optim. 39, 208–225.
Rosenberg, D., Sorin, S., 2001. An operator approach to zero-sum repeated games. Israël J. Math. 121, 221–246.
Rosenberg, D., Vieille, N., 2000. The maxmin of recursive games with incomplete information on one side. Math. Operat. Res. 25, 23–35.
Rosenberg, D., Solan, E., Vieille, N., 2001. Stopping games with randomized strategies. Prob. Theory Rel. Fields 119, 433–451.
Rosenberg, D., Solan, E., Vieille, N., 2002. Blackwell optimality in Markov decision processes with partial observation. Ann. Stat. 30, 1178–1193.
Rosenberg, D., Solan, E., Vieille, N., 2003a. The maxmin value of stochastic games with imperfect monitoring. Int. J. Game Theory 32, 133–150.
Rosenberg, D., Solan, E., Vieille, N., 2003b. Stochastic games with imperfect monitoring. In: Haurie, A., Muto, S., Petrosjan, L.A., Raghavan, T.E.S., (Eds.), Advances in Dynamic Games: Applications to Economics, Management Science, Engineering, and Environmental Management. Annals of the ISDG, vol. 8, pp. 3–22.
Rosenberg, D., Solan, E., Vieille, N., 2004. Stochastic games with a single controller and incomplete information. SIAM J. Control Optim. 43, 86–110.
Scarf, H.E., Shapley, L.S., 1957. Games with partial information. In: Dresher, M., Tucker, A.W., Wolfe, P., (Eds.), Contributions to the Theory of Games, III, Annals of Mathematical Studies, vol. 39. Princeton University Press, pp. 213–229.
Shapley, L.S., 1953. Stochastic games. Proc. Nat. Acad. Sci. U.S.A 39, 1095–1100.

Shmaya, E., 2011. The determinacy of infinite games with eventual perfect monitoring. Proc. AMS 139, 3665–3678.
Shmaya, E., Solan, E., 2004. Zero-sum dynamic games and a stochastic variation of Ramsey's theorem. Stoch. Process. Appl. 112, 319–329.
Solan, E., 2003. Continuity of the value of competitive Markov decision processes. J. Theoret. Probab. 16, 831–845.
Solan, E., Vieille, N., 2002. Uniform value in recursive games. Ann. Appl. Probab. 12, 1185–1201.
Solan, E., Vieille, N., 2004. Stopping games -recent results. In: Nowak, A.S., Szajowski, K., (Eds.), Advances in Dynamic Games, Annals of the ISDG, vol. 7. Birkhauser, pp. 235–245.
Sorin, S., 1984. "Big Match" with lack of information on one side (Part I). Int. J. Game Theory 13, 201–255.
Sorin, S., 1985a. "Big Match" with lack of information on one side, Part II. Int. J. Game Theory 14, 173–204.
Sorin, S., 1985b. On repeated games with state dependent matrices. Int. J. Game Theory 14, 249–272.
Sorin, S., 1989. On repeated games without a recursive structure: existence of $\lim v_n$. Int. J. Game Theory 18, 45–55.
Sorin, S., 1992. Repeated games with complete information. In: Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory, vol. 1. Elsevier, pp. 71–107.
Sorin, S., 2002. A First Course on Zero-Sum Repeated Games. Springer.
Sorin, S., 2003. The operator approach to zero-sum stochastic games. In: Neyman, A., Sorin, S., (Eds.), Stochastic Games and Applications, NATO Science Series C 570, Kluwer Academic Publishers, pp. 417–426.
Sorin, S., 2004. Asymptotic properties of monotonic nonexpansive mappings. Discrete Events Dyn. Syst. 14, 109–122.
Sorin, S., Vigeral, G., 2012. Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems. J. Opim. Theory Appl. 157, 564–576.
Sorin, S., Vigeral, G., 2013. Reversibility and oscillations in zero-sum discounted stochastic games. Preprint hal 00869656.
Sorin, S., Venel, X., Vigeral, G., 2010. Asymptotic properties of optimal trajectories in dynamic programming. Sankhya 72A, 237–245.
Souganidis, P.E., 1985. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. J. Differ. Equat. 17, 781–791.
Souganidis, P.E., 1999. Two Player zero sum differential games and viscosity solutions. In: Bardi, M., Raghavan, T.E.S., Parthasarathy, T., (Eds.), Stochastic and Differential Games, Annals of the ISDG, vol. 4. Birkhauser, pp. 70–104.
Souquière, A., 2010. Approximation and representation of the value for some differential games with asymmetric information. Int. J. Game Theory 39, 699–722.
Spinat, X., 2002. A necessary and sufficient condition for approachability. Math. Operat. Res. 27, 31–44.
van den Dries, L., 1998. Tame Topology and o-Minimal Structures. London Mathematical Society Lecture Notes Series. vol. 248. Cambridge University Press, Cambridge.
Vieille, N., 1992. Weak approachability. Math. Operat. Res. 17, 781–791.
Vieille, N., 2002. Stochastic games: recent results. In: Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory, vol. 3. Elsevier, pp. 1834–1850.
Vigeral, G., 2010a. Evolution equations in discrete and continuous time for non expansive operators in Banach spaces. ESAIM COCV 16, 809–832.
Vigeral, G., 2010b. Iterated monotonic nonexpansive operators and asymptotic properties of zero-sum stochastic games. Preprint.
Vigeral, G., 2010c. Zero-sum recursive games with compact action spaces. Preprint.
Vigeral, G., 2013. A zero-sum stochastic game with compact action sets and no asymptotic value. Dyn. Games Appl. 3, 172–186.

Zachrisson, L.E., 1964. Markov Games. In: Dresher, M., Shapley, L.S., Tucker, A.W., (Eds.), Advances in Game Theory, Annals of Mathematical Studies, vol. 52. Princeton University Press, pp. 210–253.
Zamir, S., 1973. On the notion of value for games with infinitely many stages. Ann. Stat. 1, 791–796.
Zamir, S., 1992. Repeated games of incomplete information: zero-sum. In: Aumann, R.J., Hart, S., (Eds.), Handbook of Game Theory, vol. 1. Elsevier, pp. 109–154.
Ziliotto, B., 2013. Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture $maxmin = lim v_n$. Preprint hal-00824039.
Ziliotto, B., 2014. General limit value in stochastic games. Preprint.