Dyn Games Appl (2011) 1: 172–207
DOI 10.1007/s13235-010-0006-z

Zero-Sum Repeated Games: Recent Advances and New Links with Differential Games

Sylvain Sorin

Published online: 16 October 2010

Abstract The purpose of this survey is to describe some recent advances in zero-sum repeated games and in particular new connections to differential games. Topics include: approachability, asymptotic analysis: recursive formula and operator approach, dual game and incomplete information, uniform approach.

Keywords Zero-sum two-person repeated games · Differential games · Approachability · Recursive formula · Asymptotic value · Uniform value

1 Introduction: Asymptotic and Uniform Approaches in RG, Quantitative and Qualitative DG

This survey is a sequel to Sorin [81], and covers some recent advances in the theory of two-person zero-sum repeated games (RG) and new connections with differential games (DG). Alternative developments in the theory of differential games within a much more general framework are treated in the companion paper by Buckdahn, Cardaliaguet and Quincampoix, [8] in this volume.

This section describes informally both models and some of the main issues. RG are played in discrete time. There is a state space M and action sets I for Player 1 and J for Player 2. At each stage n, the game is in some state $m_n \in M$, each player chooses an action ($i_n \in I$, $j_n \in J$) that, together with the current state, determines a stage payoff

This paper is dedicated to the memory of D. Blackwell. Comments from P. Cardaliaguet, R. Laraki, M. Oliu-Barton, J. Renault and G. Vigeral were very helpful. This research of was supported by grant ANR-08-BLAN- 0294-01 (France).

S. Sorin (✉)
Equipe Combinatoire et Optimisation, CNRS FRE 3232, Faculté de Mathématiques,
Université P. et M. Curie – Paris 6, Tour 15-16, 1° étage, 4 Place Jussieu, 75005 Paris, France
e-mail: sorin@math.jussieu.fr
url: http://www.ecp6.jussieu.fr/pageperso/sorin.html

S. Sorin
Laboratoire d'Econométrie, Ecole Polytechnique, 91128 Palaiseau, France

$g_n = g(m_n, i_n, j_n)$ and the joint law of the new state m_{n+1} and of the signals to the players (see Section 3 for a formal description). The basic structure of the game is stationary (i.e. independent of the stage n) and we are interested in the way the players can guide the process $\{g_n\}$ for large n.

DG are played in continuous time. There is a state space Z and control sets U for Player 1 and V for Player 2. 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. Notice that in the autonomous case, if the players use piecewise constant controls on intervals of size δ, the induced process is like a RG.

Clearly the above description extends to the N-player case as well. Let us specify the evaluation of the process in the zero-sum case.

1.1 RG Evaluations

For the RG framework there are basically three approaches.

1.1.1

We first introduce the *compact case*. For every probability distribution μ on the integers $n \geq 1$ ($\mu_n \geq 0$, $\sum_n \mu_n = 1$), one defines a game $G[\mu]$ with evaluation $\langle \mu, g \rangle = \sum_n g_n \mu_n$. Under standard assumptions on the basic data (for example if all sets involved in the definition of the game are finite), the natural product topology on plays specifies a game with compact strategy sets and continuous payoff function, hence the value $v[\mu]$ will exist.

The *asymptotic approach* studies the family of such games as the expected length (the mean of μ) goes to $+\infty$. (Alternatively one could consider these games has being played between time 0 and 1, the duration of stage n being μ_n.) Natural assumptions are that for each μ, μ_n is decreasing and $\mu_1 \to 0$ along the family.

The analysis concentrates on the corresponding family of values $\{v[\mu]\}$ and of (approximate) optimal strategies. Two typical examples correspond to:

- The *finite n-stage* game G_n with outcome given by the average of the first n stage payoffs:

$$\overline{g}_n = \frac{1}{n}\sum_{m=1}^{n} g_m.$$

- The *λ-discounted game* G_λ with outcome equal to the discounted sum of the payoffs:

$$\overline{g}_\lambda = \sum_{m=1}^{\infty} \lambda(1-\lambda)^{m-1} g_m.$$

The values of these games are denoted by v_n and v_λ respectively. The asymptotic analysis studies their behavior, as n goes to ∞ or λ goes to 0. The main issues are (i) to check whether the limit exists, (ii) is independent of the sequence of evaluations and (iii) to identify it.

Extensions consider games with random duration process where the weight μ_n is a random variable which law depends upon the previous path on a random duration tree, Neyman and Sorin [57], see Section 3.3.2. Note that the knowledge of the duration (i.e. of the evaluation process) by the players is crucial in the choice of the strategies.

The main tool for this analysis is the recursive formula, see Section 3.

1.1.2

An alternative analysis, called the *uniform approach*, considers the whole family of "long games" without specifying the exact duration. Hence one looks for strategies exhibiting asymptotic uniform properties in the following sense: they are almost optimal in any sufficiently long game. Explicitly:

Definition 1.1 $\underline{v}$ is the (uniform) *maxmin* if the two following conditions are satisfied:

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

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

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

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

A dual definition holds for the *minmax* $\overline{v}$. Whenever $\underline{v} = \overline{v}$, the game has a *uniform value*, denoted by v_∞. Note that the existence of v_∞ implies: $v_\infty = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$, and similarly for any random duration processes, Neyman and Sorin [57].

This analysis relies on an explicit construction of the strategies and is very sensitive to the information of the players: see e.g. stochastic games with known states and signals on the actions, Section 7.4. *and Zamir* [91].

1.1.3

A third approach specifies directly an outcome associated to the sequence $\{g_n\}$, like $\liminf \frac{1}{n}\sum_{m=1}^{n} g_m$ or a measurable function defined on plays (see Maitra and Sudderth [42, 43]). This describes an infinitely repeated game in normal form and the issue is the existence of a value. We will not cover this direction in this survey but let us mention that there are fascinating measurability issues and important connections with both the recursive formula and the uniform approach.

1.2 DG Evaluations

We now turn to some definitions of the payoff in DG, but we will be far to cover all cases.

1.2.1

First note that in DG, the fact that time is continuous and the hypothesis that each player knows the previous control of his opponent induce an issue in defining the strategies of the players, in such a way that the induced process (z_t, u_t, v_t) is well specified. One manner to proceed is as follows.

Let $\mathcal{U}$ (resp. $\mathcal{V}$) denote the sets of measurable control maps from $\mathbb{R}^+$ to U (resp. V). $\alpha \in \mathcal{A}'$ is a non anticipative strategy (NA) (resp. $\alpha \in \mathcal{A}$ is non anticipative strategy with delay (NAD)) if α maps $\mathbf{v} \in \mathcal{V}$ to $\alpha(\mathbf{v}) = \mathbf{u} \in \mathcal{U}$ such that if $\mathbf{v}_s = \mathbf{v}'_s$ on $[0, t]$ then $\alpha(\mathbf{v}) = \alpha(\mathbf{v}')$ on $[0, t]$, for all $t \in \mathbb{R}^+$ (resp. and there exists $\delta > 0$ 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 $t \in \mathbb{R}^+$). Note that a couple $(\alpha \in \mathcal{A}', \mathbf{v} \in \mathcal{V})$ or $(\mathbf{u} \in \mathcal{U}, \beta \in \mathcal{B}')$ induces a pair of control maps, hence a well defined dynamics on Z. This was the initial procedure to introduce upper and lower games and values, see e.g. Souganidis [84], Bardi and Capuzzo Dolcetta [3], Chapter VIII. It is more natural to work in a symmetric way with a normal form game defined on $\mathcal{A} \times \mathcal{B}$. A couple (α, β) defines a unique couple $(\mathbf{u}, \mathbf{v}) \in \mathcal{U} \times \mathcal{V}$ 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 [11], Cardaliaguet and Quincampoix [14].

1.2.2

A differential game is with fixed duration (or a "quantitative" DG) if the total evaluation is of the form

$$\Gamma(\alpha, \beta)(z_0) = \int_0^T \gamma_t \, dt + \overline{\gamma}(z_T) \tag{1}$$

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

$$\Gamma'(\alpha, \beta)(z_0) = \int_0^\infty \gamma_t \mu(dt)$$

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

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.

The basic approach is to prove that the maxmin (*resp. minmax*) *satisfies some dynamic programming inequality which leads to viscosity supersolution* (*resp. subsolution*) *of an Hamilton–Jacobi–Bellman equation and to use a comparison argument, see Appendix. Even if the framework seems quite different, this analysis is deeply related to the asymptotic approach for RG.*

1.2.3

Qualitative DG are concerned with asymptotic properties of the trajectories like $\mathbf{z}_t$ staying in some set C for all $t \in \mathbb{R}^+$ or from some time $T \geq 0$ on. By working on level sets for $\underline{v}$ this approach is very similar to the uniform approach for RG. Basic references are Krasovskii and Subbotin [32], Cardaliaguet [9], Cardaliaguet, Quincampoix and Saint-Pierre [15, 16].

1.2.4

Finally pursuit-evasion games or games of timing are more in the form of class 1.1.3.

1.3 Contents

We will first describe in Section 2 results related to approachability theory, since both asymptotic and uniform analysis are available and the link with DG is especially explicit and useful. The next Section 3 introduces the recursive structure and describes the extension of the recursive formula in terms of state space and payoff evaluation. Section 4 builds on Section 3 to develop the asymptotic approach. We recall the basic results, then present the operator approach and related tools based on the (generalized) Shapley operator. Section 5 deals with games with incomplete information and their dual, a tool that is fundamental both for RG and DG with incomplete information. Section 6, devoted to the uniform approach, recalls the fundamental properties and describes some recent achievements. In Section 7 we comment on several directions of research. Appendix is a brief presentation of basic results of quantitative DG.

2 Approachability

This section describes the exciting and productive interaction between RG and DG in a specific area: approachability, introduced and studied by Blackwell [6].

2.1 Definitions

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 $n = 1, 2, \dots$, 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. $\overline{g}_n = \frac{1}{n} \sum_{m=1}^{n} g_m$ is the average vector outcome up to stage n.

The aim of Player 1 is that $\overline{g}_n$ 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 $[v, +\infty)$.

$H_n = (I \times J)^n$ is the set of possible histories at stage $n + 1$ and $H_\infty = (I \times J)^\infty$ is the set of plays. Σ (resp. $\mathcal{T}$) is the set of strategies of Player 1 (resp. Player 2): mappings from $H = \bigcup_{n \geq 0} H_n$ to the sets of mixed actions $U = \Delta(I)$ (probabilities on I) (resp. $V = \Delta(J)$). At stage $n + 1$, given the history $h_n \in H_n$, Player 1 chooses an action $i_{n+1} \in I$ according to the probability distribution $\sigma(h_n) \in U$ (and similarly for Player 2). A couple (σ, τ) of strategies induces a probability on H_∞ and $\mathsf{E}_{\sigma,\tau}$ denotes the corresponding expectation.

The asymptotic notion is

Definition 2.1 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

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

where d_C stands for the distance to C.

If v_n is the value of the n-stage game with payoff $-E(d_C(\bar{g}_n))$, weak approachability means $v_n \to 0$.

The uniform notion is

Definition 2.2 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 $n \geq N$

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

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.

2.2 Preliminaries

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. At each stage $n = 1, 2, \dots$, 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. Accordingly, a strategy $\sigma^\star$ for Player 1 in $G^\star$ is a map from $H^\star = \bigcup_{n\geq 0} H_n^\star$ to U where $H_n^\star = (U \times V)^n$. A strategy $\tau^\star$ for Player 2 is defined similarly. A couple of strategies induces a play $\{(u_n, v_n)\}$ and a sequence of outcomes $\{g_n^\star\}$, and $\overline{g}_n^\star = \frac{1}{n}\sum_{m=1}^n g_m^\star$ denotes the average outcome up to stage n.

$G^\star$ is the game played in "mixed actions" or in expectation. Weak $\star$approachability, $v_n^\star$ and $\star$approachability are defined similarly.

2.3 Weak Approachability and Quantitative Differential Games

The next result is due to Vieille [87]. Recall that the aim is to obtain a good average outcome at stage n.

First consider the game $G^\star$. Use as state variable the accumulated payoff and consider the differential game Λ of fixed duration played on $[0, 1]$ starting from $z_0 = 0 \in \mathbb{R}^k$ with dynamics:

$$\dot{z}_t = u_t A v_t = f(z_t, u_t, v_t)$$

and terminal payoff $-d_C(\mathbf{z}(1))$. Thus the state variable is $z_t = \int_0^t \gamma_s\,ds$, γ_s being the payoff at time s. 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 $\xi \in \mathbb{R}^k$. $G_n^\star$ 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 $z \in \mathbb{R}^k$:

$$V_n\left(\frac{k}{n}, z\right) = \mathtt{val}_{U\times V} V_n\left(\frac{k+1}{n}, z + \frac{1}{n} u A v\right) \tag{2}$$

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 $z + \int_t^1 \gamma_s\,ds$). Then basic results from DG implies, see Appendix ($D\Phi$ is the gradient in z):

Theorem 2.1 (1) *$\Phi(z, t)$ is the unique viscosity solution on* $[0, 1] \times \mathbb{R}^k$ *of*

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

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 2.2

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

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 mth 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 $y_m^\star$ 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 $G_n^\star$ for large L, hence the result.

The last property implies the following:

Corollary 2.1 *Every set is weakly approachable or weakly excludable.*

2.4 Approachability and **B**-Sets

The main notion was introduced by Blackwell [6]:

Definition 2.3 A closed set C in $\mathbb{R}^k$ is a **B**-set for Player 1 (for a given 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:

$$\langle uAv - b, a - b\rangle \le 0, \quad \forall v \in V$$

where $\pi_C(a)$ denotes the set of closest points to a in C.

The basic result of Blackwell, [6] is

Theorem 2.3 *Let C be a* **B**-*set for Player* 1. *Then it is approachable in G and $\star$approachable in $G^\star$ by that player. An approachability strategy is given by $\sigma(h_n) = \hat{u}(\bar{g}_n)$ (resp. $\sigma^\star(h_n^\star) = \hat{u}(\bar{g}_n^\star)$).*

An important consequence of Theorem 2.3 is the next result.

Theorem 2.4 *A convex set C is either approachable or excludable.*

A further result due to Spinat [86], see also Hou [29], characterizes minimal approachable sets:

Theorem 2.5 *A set C is approachable iff it contains a* **B**-*set.*

2.5 Approachability and Qualitative Differential Games

We follow here as Soulamani, Quincampoix and Sorin [1].

To study $\star$approachability, consider an alternative differential game Γ where both the dynamics and the payoff function differ from the previous differential game Λ. The aim is to control asymptotically the average payoff, hence the discrete dynamics on the state variable 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 $\bar{\gamma}_t = \frac{1}{t}\int_0^t u_s A v_s \, ds$. A change of variable $z_t = \bar{\gamma}_{e^t}$ leads to

$$\dot{z}_t = u_t A v_t - z_t. \tag{3}$$

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

We recall the next definitions, following Cardaliaguet [9].

Definition 2.4 A nonempty closed set C in $\mathbb{R}^k$ is a *discriminating domain* for Player 1, given f if:

$$\forall a \in C,\ \forall p \in NP_C(a), \quad \sup_{v \in V} \inf_{u \in U} \big\langle f(a, u, v), p\big\rangle \le 0, \tag{4}$$

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 [9], states that Player 1 can ensure remaining in a discriminating domain as soon as he knows, at each time t, Player 2's control up to time t.

Theorem 2.6 *Assume that f satisfies Isaacs's 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 $t \geq 0$.*

We shall say that such a strategy α *preserves* the set C. The link with approachability is through the following result:

Theorem 2.7 *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* **B**-*set for Player 1.*

It is easy to deduce that starting from any point, not necessarily in C one has:

Theorem 2.8 *If a closed set $C \subset \mathbb{R}^k$ is a* **B**-*set for Player 1, there exists $\alpha \in \mathcal{A}'$, such that for every $\mathbf{v} \in \mathbf{V}$*

$$\forall t \geq 1 \quad d_C\big(\mathbf{z}\big[\alpha(\mathbf{v}), \mathbf{v}\big](t)\big) \leq Me^{-t}. \tag{5}$$

The main result is now:

Theorem 2.9 *A closed set C is $\star$approachable for Player 1 in $G^\star$ if and only if it contains a* **B**-*set for Player 1 (given A).*

The direct part follows from Blackwell's proof. The proof of converse implication is as follows: first one defines a map Ψ from strategies of Player 1 in $G^\star$ to nonanticipative strategies in Γ. In particular given $\varepsilon > 0$ and a strategy σ_ε that ε-approaches C in $G^\star$, its image is $\alpha_\varepsilon = \Psi(\sigma_\varepsilon)$. The next step is to show that the trajectories in the differential game Γ compatible with α_ε approach asymptotically C up to ε. Finally one proves that the ω-limit set of any trajectory compatible with some α is a discriminating domain. Explicitly, let $D(\alpha) = \bigcap_{\theta \geq 0} \mathrm{cl}\,\{\mathbf{x}[x_0, \alpha(\mathbf{w}), \mathbf{w}](t);\ t \geq \theta,\ \mathbf{w} \in \mathbf{V}\}$, where *cl* is the closure operator.

Lemma 2.1 *$D(\alpha)$ is a nonempty compact discriminating domain for Player 1 given f.*

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

2.6 On Strategies in the Differential and Repeated Games

This part describes explicitly the construction of an approachability strategy in the repeated game G starting from a preserving strategy in Γ. The idea of the construction is the following:

(a) Given a NA $\alpha' \in \mathcal{A}'$, construct an approximation in term of range by a NAD $\alpha \in \mathcal{A}$.
(b) When applied to α' preserving C (hence approaching C), this leads to $\alpha \in \mathcal{A}$ approaching C.

(c) This NAD α generates an $\star$approachability strategy in the repeated game $G^\star$.
(d) Finally $\star$approachability strategies in $G^\star$ induce approachability strategies in G.

For step (a), recall the definition of the range associated to a nonanticipative strategy $\alpha' \in \mathcal{A}'$:

$$R(\alpha', t) = \mathrm{cl}\{y \in \mathbb{R}^k \; \exists \mathbf{v} \in \mathbf{V}, \; y = \mathbf{x}[x_0, \alpha'(\mathbf{v}), \mathbf{v}](t)\}.$$

The next result is due to Cardaliaguet [10] and is inspired by the "extremal aiming" method of Krasovskii and Subbotin [32]. It is very much in the spirit of proximal normals and approachability.

Proposition 2.1 *For any $\varepsilon > 0$, $T > 0$ and any NA $\alpha' \in \mathcal{A}'$, there exists some NAD $\alpha \in \mathcal{A}$ such that, for all $t \in [0, T]$ and all $\mathbf{v} \in \mathbf{V}$:*

$$d_{R(\alpha',t)}\big(\mathbf{x}[x_0, \alpha(\mathbf{v}), \mathbf{v}](t)\big) \le \varepsilon.$$

The efurther result relies explicitly on the specific form (3) of the dynamics f in Γ and extends the approximation from a compact interval to $\mathbb{R}^+$.

Proposition 2.2 *Fix $z \in \mathbb{R}^k$. For any $\varepsilon > 0$ and any NA $\alpha' \in \mathcal{A}'$ in the game Γ, there exists some NAD $\alpha \in \mathcal{A}$ such that, for all $t \ge 0$ and all $\mathbf{v} \in \mathcal{V}$:*

$$d_{R(\alpha',t)}\big(\mathbf{z}_t[\alpha(\mathbf{v}), \mathbf{v}, z]\big) \le \varepsilon.$$

In particular this leads to step (b)

Proposition 2.3 *Let C be a **B**-set. For any $\varepsilon > 0$ there is some NAD α in the game Γ and some T such that for any $\mathbf{v}$ in $\mathcal{V}$*

$$d_C\big(\gamma[\alpha(\mathbf{v}), \mathbf{v}](t)\big) \le \varepsilon, \quad \forall t \ge T.$$

Step (c) is now to use the delay to define a strategy that depends only on the past moves. Hence α induces an ε-approachability strategy $\sigma^\star$ for C in $G^\star$.

The last step (d) is

Proposition 2.4 *Given $\sigma^\star$ a strategy that $\star$approach C up to $\varepsilon > 0$ in the game $G^\star$, there exists σ a strategy that approach C up to 2ε in the game G.*

and the idea is to use a martingale inequality to compare the trajectory and the trajectory "in law".

2.7 Remarks

(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 extension of approachability conditions for games with signals on the outcome, see Perchet [59] or on more general spaces, see Lehrer [39], Perchet and Quincampoix [60].

3 Recursive Structure of Compact Repeated Games and Shapley Operator

The simplest incarnation of the recursive formula for repeated games is the following equation, due to Shapley [72], involving the discounted value of finite stochastic game

$$v_\lambda(\omega) = \mathtt{val}_{X\times Y}\left\{\lambda g(\omega, x, y) + (1-\lambda)\sum_{\omega'} Q(\omega, x, y)[\omega']v_\lambda(\omega')\right\}, \quad \forall\omega \in \Omega \qquad (6)$$

where Q stands for the transition probability on the state space Ω, $X = \Delta(I)$ (the set of probabilities on X), $Y = \Delta(J)$ and for $h : I \times J \to \mathbb{R}$, $h(x, y) = \mathsf{E}_{x,y}h$. It expresses the value of the game today as a function of the current payoff and the value from tomorrow on. Two ingredients are concerned. First, the play of the game: the initial state ω is known by both players and also the new state will be, hence one can perform the analysis for each state separately: the "state" of the stochastic game is the natural "state variable" for the recursive formula. Second, the evaluation measure today and its decomposition between the weight today and the conditional distribution on stages from tomorrow on.

We will describe several extensions: first to general repeated game forms that correspond to point one above, then to general evaluation measures, point two.

3.1 Recursive Formula

We recall briefly that a recursive structure leading to an expression similar to (6) holds in general for two-person zero-sum repeated games described as follows:

M is a parameter space and g a function from $I \times J \times M$ to $\mathbb{R}$. For each $m \in M$ this defines a two-person zero-sum game with action spaces I and J for Player 1 and 2 respectively and payoff function $g(m, \cdot)$. (Again to simplify the presentation we will consider the case where all sets are finite, avoiding in particular measurability issues.)

The initial parameter m_1 is chosen at random and the players receive some initial information about it, say a_1 (resp. b_1) for Player 1 (resp. Player 2). This choice is performed according to some initial probability π on $M \times A \times B$, where A and B are the signal sets of both players.

At each stage n, Player 1 (resp. 2) chooses an action $i_n \in I$ (resp. $j_n \in J$). This determines a stage payoff $g_n = g(m_n, i_n, j_n)$, where m_n is the current value of the parameter. Then a new value of the parameter is selected and the players get some information. This is generated by a map Q from $M \times I \times J$ to probabilities on $M \times A \times B$. Hence at stage n a triple $(m_{n+1}, a_{n+1}, b_{n+1})$ is chosen according to the distribution $Q(m_n, i_n, j_n)$. The new parameter is m_{n+1}, and the signal 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 both the previous (m_n) and the new (m_{n+1}) values of the parameter.

Stochastic games correspond to public signals including the parameter, see Sorin [77], Neyman and Sorin [56], for a general presentation. *Incomplete information games* correspond to absorbing transition on the parameter and no further information (after the initial one) on the parameter (that remains fixed), see Aumann and Maschler [2], Sorin [77].

A play of the game induces a sequence $m_1, a_1, b_1, i_1, j_1, m_2, a_2, b_2, i_2, j_2, \dots$ while the information of Player 1 before his play at stage n is a 1-private history of the form $(a_1, i_1, a_2, i_2, \dots, a_n)$ and similarly for Player 2. The corresponding sequence of payoffs is $g_1, g_2, \dots$. Note that it is not known to the players except if included in the signals.

A strategy σ for Player 1 is a map from 1-private histories to $\Delta(I)$, the space of probabilities on the set I of actions: it defines the probability distribution of the stage action as a function of the past known to Player 1; τ is defined similarly for Player 2. Such a couple (σ, τ) induces, together with the components of the game, π and Q, a probability distribution on plays, hence on the sequence of payoffs.

The recursive structure relies on the construction of the universal belief space, Mertens and Zamir [52], that represents the infinite hierarchy of beliefs of the players: $\boldsymbol{\Omega} = M \times \Theta^1 \times \Theta^2$, where Θ^i, homeomorphic to $\Delta(M \times \Theta^{-i})$, is the type set of Player i. A consistent probability ρ on $\boldsymbol{\Omega}$ is such that the conditional probability induced by ρ at θ^i coincides with θ^i itself, as elements of $\Delta(M \times \Theta^{-i})$. The set of consistent probabilities is $\mathbb{P} \subset \Delta(\boldsymbol{\Omega})$. 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 the same for Player 2) that induces a consistent probability $\mathcal{P}_n \in \mathbb{P}$, see Mertens, Sorin and Zamir [49], Sections III.1, III.2, IV.3. This is referred to as the "entrance law". Taking into account the existence of a value for the repeated game, we suppose that the strategies used are announced to the players. The entrance law $\mathcal{P}_n$ and the (behavioral) strategies at stage n (say α_n and β_n), which are maps from type set to mixed action set, determine the current payoff and the new entrance law $\mathcal{P}_{n+1} = H(\mathcal{P}_n, \alpha_n, \beta_n)$. This updating rule is the basis of the recursive structure and $\mathbb{P}$ is the "state space" for the recursive structure. 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 is defined on the set of real bounded functions on $\mathbb{P}$ by:

$$\boldsymbol{\Psi}(f)(\mathcal{P}) = \sup_{\alpha} \inf_{\beta} \bigl\{ g(\mathcal{P}, \alpha, \beta) + f\bigl(H(\mathcal{P}, \alpha, \beta)\bigr) \bigr\}. \tag{7}$$

It is natural to introduce also the projective version

$$\overline{\boldsymbol{\Psi}}(f)(\mathcal{P}) = \sup_{\alpha} \inf_{\beta} \bigl\{ f\bigl(H(\mathcal{P}, \alpha, \beta)\bigr) \bigr\} \tag{8}$$

that will be crucial in the asymptotic analysis.

Then the usual relations hold, see Mertens, Sorin and Zamir, (1994) Section IV.3, for the finitely repeated game:

$$(n+1) v_{n+1}(\mathcal{P}) = \mathtt{val}_{\alpha \times \beta} \bigl\{ g(\mathcal{P}, \alpha, \beta) + n v_n\bigl(H(\mathcal{P}, \alpha, \beta)\bigr) \bigr\} \tag{9}$$

and the discounted game:

$$v_\lambda(\mathcal{P}) = \mathtt{val}_{\alpha \times \beta} \bigl\{ \lambda g(\mathcal{P}, \alpha, \beta) + (1-\lambda) v_\lambda\bigl(H(\mathcal{P}, \alpha, \beta)\bigr) \bigr\} \tag{10}$$

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 $\mathbb{P} \subset \Delta(\boldsymbol{\Omega})$. Hence to each compact repeated game G one can associate an auxiliary game Γ having the same values that satisfy the recursive equation. However the play, hence the strategies in both games differ.

In the framework of a stochastic game with state space Ω, this representation would correspond to the level of probabilities on the state space $\mathbb{P} \subset \Delta(\Omega)$. One recovers the initial Shapley formula (6) with $\mathcal{P}$ being the Dirac mass at ω, then (α, β) reduces 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 $v_\lambda(H(\mathcal{P}, \alpha, \beta)) = \mathsf{E}_{Q(\omega,x,y)} v_\lambda(.)$. The projective version is

$$\overline{\boldsymbol{\Psi}}(f)(\omega) = \mathtt{val}_{X \times Y} \left\{ \sum_{\omega'} Q(\omega, x, y)[\omega'] f(\omega') \right\}. \tag{11}$$

Let us describe the framework of repeated games with incomplete information (independent case with perfect monitoring). M is a product space $K \times L$, π is a product probability $p \otimes q$ with $p \in \Delta(K)$, $q \in \Delta(L)$ and in addition $a_1 = k$ 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 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 can be taken as follows: the "state space" M' is $\Delta(K) \times \Delta(L)$ and is interpreted as the space of beliefs on the true 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 $\mathbf{X} \times \mathbf{Y} \times M'$ by $g(p, q, x, y) = \sum_{k,\ell} p^k q^\ell g(k, \ell, x^k, y^\ell)$. Given (p, q, x, y), let $x(i) = \sum_k x_i^k p^k$ be the probability of action i and $p(i)$ be the conditional probability on K given the action i, explicitly $p^k(i) = \frac{p^k x_i^k}{x(i)}$ (and similarly for y and q). Since the actions are announced in the original game, and the strategy are known in the auxiliary game, these posterior probabilities are known by the players and we can work with $\boldsymbol{\Omega} = M'$ and take as $\mathbb{P}$ the set $\Delta(M')$. Finally the transition Q (from M' to $\Delta(M')$) is defined by the following procedure: $Q(p, q, x, y)(p', q') = \sum_{i,j;(p(i),q(j))=(p',q')} x(i)y(j)$. The resulting form of the Shapley operator is

$$\boldsymbol{\Psi}(f)(p,q) = \sup_{x\in\mathbf{X}} \inf_{y\in\mathbf{Y}} \left\{ \sum_{k,\ell} p^k q^\ell g\bigl(k,\ell,x^k,y^\ell\bigr) + \sum_{i,j} x(i)y(j) f\bigl(p(i),q(j)\bigr) \right\} \tag{12}$$

where with the previous notations

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

and the projective version writes

$$\overline{\boldsymbol{\Psi}}(f)(p,q) = \sup_{x\in\mathbf{X}} \inf_{y\in\mathbf{Y}} \left\{ \sum_{i,j} x(i)y(j) f\bigl(p(i),q(j)\bigr) \right\}. \tag{13}$$

The corresponding equations for v_n and v_λ are due to Aumann and Maschler, see [2], and Mertens and Zamir [50]. Recall that the auxiliary game Γ' is "equivalent" to the original one in terms of values but uses different strategy spaces. In fact in the original game the strategy of the opponent is unknown, hence the computation of the posterior distribution is not feasible.

3.2 General Partition and Games in Continuous Time

Similarly the recursive structure extends to arbitrarily evaluation of the stream of stage payoffs. First notice that the generalized Shapley operator can be used to describe any positive combination of weights between the past and the future like

$$\mathbf{S}[c,c'](f)(\mathcal{P}) = \sup_\alpha \inf_\beta \bigl\{ c g(\mathcal{P},\alpha,\beta) + c' f\bigl(H(\mathcal{P},\alpha,\beta)\bigr)\bigr\} \tag{14}$$

with $c > 0$, $c' > 0$, since one has

$$\mathbf{S}[c,c'](f) = c\boldsymbol{\Psi}\left(\frac{c'}{c} f\right)$$

and $\boldsymbol{\Psi} = \mathbf{S}[1,1]$. Note also that the projective operator $\overline{\boldsymbol{\Psi}} = \mathbf{S}[0,1]$ appears as the recession operator associated to $\boldsymbol{\Psi}$. Following this line, one has the classical fixed point formula for the discounted value

$$v_\lambda = \mathbf{S}[\lambda, 1-\lambda](v_\lambda) \tag{15}$$

and the recursive formula for the n stage value

$$v_n = \mathbf{S}\left[\frac{1}{n}, 1-\frac{1}{n}\right](v_{n-1}) \tag{16}$$

with obviously $v_0 = 0$.

Consider now an arbitrarily evaluation probability μ on $\mathbb{N}^\star$. We can approximate the corresponding value uniformly by considering measures with finite support. Then μ induces a finite partition Π of $[0,1]$ with $t_0 = 0$, $t_k = \sum_{m=1}^{k} \mu_m$, $t_N = 1$. Thus the repeated game is naturally represented as a game played between times 0 and 1, where the actions are constant on each subinterval (t_{k-1}, t_k): its length μ_k is the weight of stage k in the original game. Let v_Π be its value. The recursive equation is

$$v_\Pi = \mathtt{val}\left\{t_1 g + (1-t_1)\mathsf{E} v_{\Pi_{t_1}}\right\} = \mathbf{S}[t_1, 1-t_1](v_{\Pi_{t_1}})$$

where Π_{t_1} is the normalization on $[0,1]$ of the trace of the partition Π on the interval $[t_1, 1]$. Define now $V_\Pi(t_k)$ as the value of the game starting at time t_k, hence with $N-k$ stages and total weight $\sum_{m=k+1}^{N} \mu_m$. One obtains the alternative recursive formula

$$V_\Pi(t_k) = \mathtt{val}\left\{(t_{k+1}-t_k)g + \mathsf{E} V_\Pi(t_{k+1})\right\} = \mathbf{S}[t_{k+1}-t_k, 1]\big(V_\Pi(t_{k+1})\big). \tag{17}$$

The stationarity property of the game form induces time homogeneity

$$V_\Pi(t_k) = (1-t_k) V_{\Pi_{t_k}}(0) \tag{18}$$

where, as above, Π_{t_k} stands for normalization of Π restricted to the interval $[t_k, 1]$. By taking the linear extension we define this way for every finite partition Π, a function $V_\Pi(t)$ on $[0,1]$.

3.3 Further Extensions

3.3.1 Non Expansive Maps

Given a non expansive map $\mathbf{T}$ on a Banach space, one defines inductively

$$W_n = \mathbf{T}(W_{n-1}) = \mathbf{T}^n(0)$$

and for $\lambda \in]0,1[$:

$$W_\lambda = \mathbf{T}\big((1-\lambda)W_\lambda\big).$$

Then $\frac{W_n}{n}$ and λW_λ play the role of v_n and v_λ, see (16), (15).

3.3.2 *Random Duration* (Neyman and Sorin [57])

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 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 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 a non expansive map, Neyman [54]. Then a recursive formula analogous to (17) holds.

3.3.3 *Non Stationary Set Up*

Note that some of the results above like (17) have a natural extension to a non stationary set up and are very similar to the value of time discretization of a quantitative differential game.

4 Asymptotic Approach

We consider now the asymptotic behavior of v_n as n goes to ∞, or v_λ as λ goes to 0.

4.1 Basic Results

Concerning games with incomplete information on one side the first results proving the existence of $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ are due to Aumann and Maschler (1966), see [2], including in addition an identification of the limit as $\mathtt{Cav}_{\Delta(K)}u$. Here $u(p) = \mathtt{val}_{\Delta(I)\times\Delta(J)}\sum_k p^k g(k,x,y)$ is the value of the one shot non revealing game, where the informed player does not use his information and $\mathtt{Cav}_C$ is the concavification operator: given ϕ, a real bounded function defined on a convex set C, $\mathtt{Cav}_C(\phi)$ is the smallest function greater than ϕ and concave, on C.

Extensions of these results to games with lack of information on both sides were achieved by Mertens and Zamir [50]. In addition they identified the limit as the only solution of the system of implicit functional equations with unknown ϕ:

$$\begin{aligned}\phi(p,q) &= \mathtt{Cav}_{p\in\Delta(K)}\min\{\phi,u\}(p,q),\\ \phi(p,q) &= \mathtt{Vex}_{q\in\Delta(L)}\max\{\phi,u\}(p,q).\end{aligned}\tag{19}$$

Here again u stands for the value of the non revealing game: $u(p,q) = \mathtt{val}_{X\times Y}\sum_{k,\ell} p^k q^\ell g(k,\ell,x,y)$ and we will write $\mathbf{MZ}$ for the operator corresponding to (19)

$$\phi = \mathbf{MZ}(u).\tag{20}$$

Mertens and Zamir provided two proofs for this result: the first part is the same and shows by using the recursive structure and constructing sophisticated reply strategies that $h = \liminf v_n$ satisfies

$$h(p,q) \geq \mathtt{Cav}_{p\in\Delta(K)}\mathtt{Vex}_{q\in\Delta(L)}\max\{h,u\}(p,q).\tag{21}$$

Then one proof shows that Player 2 can achieve asymptotically any function satisfying (21). The second proof constructs inductively dual sequences of functions $\{c_n\}$ with $c_0 \equiv -\infty$ 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 $\{d_n\}$, that converges respectively to c and d satisfying

$$\begin{aligned} c(p,q) &= \mathtt{Cav}_{p\in\Delta(K)}\mathtt{Vex}_{q\in\Delta(L)} \max\{c, u\}(p,q), \\ d &= \mathtt{Vex}_{q\in\Delta(L)}\mathtt{Cav}_{p\in\Delta(K)} \min\{d, u\}(p,q). \end{aligned} \tag{22}$$

Now a *comparison principle* is used to deduce $c \geq d$. By contradiction, otherwise 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 (19) at (p_0, q_0) are trivial which implies $c(p_0, q_0) \geq u(p_0, q_0) \geq d(p_0, q_0)$.

Recall that in this framework the uniform value may not exists, see Section 6.1.

As for stochastic games, the existence of $\lim_{\lambda\to 0} v_\lambda$ in the finite case (Ω, I, J finite) is due to Bewley and Kohlberg [4] using algebraic arguments: the equation (6) can be written as a finite set of polynomial equalities and inequalities involving $\{x_\lambda^k, y_\lambda^k, v_\lambda(k), \lambda\}$ thus it defines a semi-algebraic set in some Euclidean space $\mathbb{R}^N$, hence by projection v_λ has an expansion in power series. The existence of $\lim_{n\to\infty} v_n$ is obtained by comparison, Bewley and Kohlberg [5], see Theorem (4.1).

4.2 Operator Approach

4.2.1 Non-Expansive Monotone Maps

As in Section 3.3.1 we define similar iterates for an operator $\mathbf{T}$ mapping $\mathcal{F}$ to itself, where $\mathcal{F}$ is a subset of the set $\mathcal{F}_0$ of real bounded functions on some set Ω. Assume:

(1) $\mathcal{F}$ is a convex cone, containing the constants and closed for the uniform norm.
(2) $\mathbf{T}$ is monotonic and translates the constants. (In particular $\mathbf{T}$ is non expansive.)

Define

$$V_n = \mathbf{T}^n[0], \qquad V_\lambda = \mathbf{T}[(1-\lambda)V_\lambda]$$

hence by normalizing $V_n = n v_n$, $v_\lambda = \lambda V_\lambda$ and introducing

$$\boldsymbol{\Phi}(\varepsilon, f) = \epsilon \mathbf{T}\left[\frac{1-\varepsilon}{\varepsilon} f\right]$$

one obtains as before

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

and we consider the asymptotic behavior of these families of functions which thus relies on the properties of $\boldsymbol{\Phi}(\varepsilon, \cdot)$, as ε goes to 0. Recall that in the case of a Shapley operator one has $\mathbf{S}[\varepsilon, 1-\varepsilon](f) = \boldsymbol{\Phi}(\varepsilon, f)$. Obviously any accumulation point w of the family v_n or v_λ will satisfy

$$w = \boldsymbol{\Phi}(0, w) \tag{23}$$

hence is a fixed point of the projective operator (8).

A general result in this framework is due to Neyman [54]:

Theorem 4.1 *If v_λ is of bounded variation in the sense that for any sequence λ_i decreasing to* 0

$$\sum_i \|v_{\lambda_{i+1}} - v_{\lambda_i}\| < \infty \tag{24}$$

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

Let us define sets of functions that will correspond to upper and lower bounds on the families of values following Rosenberg and Sorin [70], Sorin [80].

4.2.2 Uniform Domination

Let $\mathcal{L}^+$ be the set of functions $f \in \mathcal{F}$ that satisfy: there exists $R_0 \geq 0$ such that $R \geq R_0$ implies

$$\mathbf{T}(Rf) \leq (R+1)f. \tag{25}$$

$\mathcal{L}^+$ is defined in a dual way.

Theorem 4.2 *If $f \in \mathcal{L}^+$*, $\limsup_{n\to\infty} v_n$ *and* $\limsup_{\lambda\to 0} v_\lambda$ *are less than f.*

Note that the above condition is equivalent to

$$\boldsymbol{\Phi}(\varepsilon, f) \leq f$$

for $\varepsilon > 0$ small enough.

Corollary 4.1 *In particular 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.*

4.2.3 Pointwise Domination

More generally, when the set Ω is not finite, one can introduce the larger class $\mathcal{S}^+$ of functions satisfying

$$\theta^+(f)(\omega) = \limsup_{R\to\infty} \bigl\{\mathbf{T}(Rf)(\omega) - (R+1)f(\omega)\bigr\} \leq 0, \quad \forall \omega \in \Omega.$$

$\mathcal{S}^-$ is defined similarly with $\theta^-(f) \geq 0$.

Theorem 4.3 *Assume Ω compact. Let $\mathcal{S}_0^+$ (resp. $\mathcal{S}_0^-$) be the set of continuous functions in $\mathcal{S}^+$ (resp. $\mathcal{S}^-$). Then*

$$f^+ \geq f^-$$

for any $f^+ \in \mathcal{S}_0^+$ and $f^- \in \mathcal{S}_0^-$. Hence the intersection of the closures of $\mathcal{S}_0^+$ and $\mathcal{S}_0^-$ contains at most one point.

$\mathbf{T}$ has the recession property if $\lim_{\varepsilon\to 0} \boldsymbol{\Phi}(\varepsilon, f)(\omega) = \lim_{\varepsilon\to 0} \varepsilon \mathbf{T}(\frac{f}{\varepsilon})(\omega) = \mathbf{RT}(f)(\omega)$ exists. The next result is due to Vigeral [89].

Theorem 4.4 *Assume that $\mathbf{T}$ has the recession property and is convex. Then the family $\{v_n\}$ (resp. $\{v_\lambda\}$) has at most one accumulation point.*

The proof uses the inequality $\mathbf{RT}(x+y) \leq \mathbf{T}(x) + \mathbf{RT}(y)$ and relies on properties of the family of operators $\mathbf{T}_m$ defined by

$$\mathbf{T}_m(f) = \frac{1}{m}\mathbf{T}^m(mf).$$

4.2.4 Application to Games

The uniform domination property allows to prove existence and equality of $\lim v_n$ and $\lim v_\lambda$ in the following classes:

Theorem 4.5 (Rosenberg and Sorin [70]) *Absorbing games with compact action spaces.*

These are stochastic games where the state changes at most once. Notice that the algebraic approach cannot be used.

Theorem 4.6 (Sorin [79], Vigeral [90]) *Recursive games with finite state space and compact action spaces.*

These are stochastic games where the payoff is zero on non absorbing states, Everett [27].

The pointwise domination property is used to prove existence and equality of $\lim v_n$ and $\lim v_\lambda$ through the derived game, see Section 4.2.3, in the following cases:

Theorem 4.7 (Rosenberg and Sorin [70]) *Games with incomplete information on both sides.*

Recall that the first proof is due to Mertens and Zamir [50]. Any accumulation point w of the family v_λ (resp. v_n) as $\lambda \to 0$ belongs to the closure of $\mathcal{S}^+$, hence by symmetry the existence of a limit follows.

Theorem 4.8 (Rosenberg [65]) *Finite absorbing games with incomplete information on one side.*

This is the first proof in this area where both stochastic and information aspects are present.

Theorem 4.9 (Vigeral [89]) *Existence of* $\lim v_n$ (*resp.* $\lim v_\lambda$) *in all games where one player controls the transition and the family* $\{v_n\}$ (*resp.* $\{v_\lambda\}$) *is relatively compact.*

This follows from the property for convex $\mathbf{T}$ and applies in particular for (finite) dynamic programming, games with incomplete information on one side and mixture of those.

4.2.5 Derived Game (Rosenberg and Sorin [70])

Still dealing with the Shapley operator, one can use the existence of a pointwise limit:

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

$\varphi(f)(\omega)$ is the value of the "derived game" with payoff $g(\omega, x, y) - E_{(\omega,x,y)} f$, played on the product of the subsets of optimal strategies in the game $\boldsymbol{\Phi}(0, f)$.

In the setup of games with incomplete information on both sides (as well as in absorbing games, following an idea of Kohlberg [30]), any accumulation point of the sequence of values is close to a function f with $\varphi(f) \le 0$. This implies the existence, by Theorem 4.3, and the characterization of the asymptotic value as follows:

Let $\mathcal{E} f$ be the projection of the extreme points of the epigraph of f. Then $v = \lim v_n = \lim v_\lambda$ is a saddle continuous function satisfying both inequalities:

$$p \in \mathcal{E} v(\cdot, q) \quad \Rightarrow \quad v(p,q) \le u(p,q), \qquad q \in \mathcal{E} v(p, \cdot) \quad \Rightarrow \quad v(p,q) \ge u(p,q) \quad (26)$$

where u is the value of the non revealing game.

Then one shows that this recovers the characterization of Mertens and Zamir, (19), see also Laraki [33].

4.2.6 *Random Duration Process* (Neyman and Sorin [57])

The recursive formula for random duration processes implies that $v_\Theta = \frac{\mathbf{T}^{\Theta}(0)}{E(\theta)}$ has a limit, as $E(\theta)$ goes to ∞ either in finite stochastic games with signals and absorbing or recursive games with compact action sets.

The uniform domination Theorem 4.2 is true for v_Θ. Similarly $\lim_{E(\theta)\to\infty} v_\Theta$ exists and equals $\mathbf{MZ}(u)$ for games with lack of information on both sides. Finally, assume Θ monotonic in the sense that the conditional expected duration decreases with time. Then Theorem 4.1 still holds.

4.3 Comparison Principle

The operator approach using the generalized Shapley operator allows for an alternative proof of existence and characterization of the asymptotic value $\lim v_\lambda$ for games with incomplete information on both sides, due to Laraki [33]. The recursive equation is

$$\begin{aligned} \boldsymbol{\Phi}(\lambda, v_\lambda)(p,q) &= \sup_{x\in\mathbf{X}} \inf_{y\in\mathbf{Y}} \left\{ \lambda g(p,q,x,y) + (1-\lambda) \sum_{i,j} x(i) y(j) v_\lambda\big(p(i), q(j)\big) \right\} \\ &= v_\lambda(p,q). \end{aligned} \quad (27)$$

Remark that the family of functions $\{v_\lambda(p,q)\}$ is uniformly Lipschitz, hence relatively compact. To prove convergence it is enough to show that there is only one accumulation point. Note first that any accumulation point w satisfies

$$\boldsymbol{\Phi}(0, w) = w \quad (28)$$

i.e. is a fixed point of the projective operator (13). Assume now that w_1 and w_2, $w_1 \ge w_2$ are two different accumulation points and let (p_0, q_0) an extreme point of the (convex hull of) the set where the difference $w_1 - w_2$ is maximal. Using (28) and the fact that all functions involved are saddle, this implies that the set $\mathbf{X}(0, w_1)(p_0, q_0)$ of profile of mixed actions $x \in \mathbf{X}$ optimal in $\boldsymbol{\Phi}(0, w_1)$ is included in the set $NR_X(p_0)$ of non revealing actions at p_0 (meaning that for any move i having positive probability $p(i) = p_0$). Consider a family v_{λ_n} converging to w_1 and let x_n be optimal for $\boldsymbol{\Phi}(\lambda_n, v_{\lambda_n})(p_0, q_0)$. Jensen's inequality leads to

$$v_{\lambda_n}(p_0, q_0) \le \lambda_n g(p_0, q_0, x_n, j) + (1-\lambda_n) v_{\lambda_n}(p_0, q_0), \quad \forall j \in J$$

thus $v_{\lambda_n}(p_0, q_0) \le g(p_0, q_0, x_n, j)$ hence letting $\bar{x}$ being an accumulation point of the family $\{x_n\}$, thus in $NR_X(p_0)$ (by upper semi continuity) one obtains as λ_n goes to 0:

$$w_1(p_0, q_0) \le g(p_0, q_0, \bar{x}, j), \quad \forall j \in J$$

which implies $w_1(p_0, q_0) \le u(p_0, q_0)$. The dual property implies convergence. Moreover one recovers the characterization through the variational inequalities (26) hence one identifies the limit as $\mathbf{MZ}(u)$.

Consider the framework of splitting games, Sorin [77]. Recall that given H from $\Delta(K) \times \Delta(L)$ to $\mathbb{R}$ the corresponding Shapley operator is defined on real functions f on $\Delta(K) \times \Delta(L)$ by

$$\boldsymbol{\Psi}(f)(p, q) = \sup_{\mu \in M_p^K} \inf_{\nu \in M_q^L} \int_{\Delta(K) \times \Delta(L)} \big[H(p, q) + f(p, q)\big] \mu(dp) \nu(dq)$$

where M_p^K stands for the set of probabilities on $\Delta(K)$ with expectation p (and similarly for M_q^L). A procedure analogous to the previous one has been developed by Laraki [33, 34, 36]. It allows, by introducing a family of discounted games and identifying the limit of the values, to extend the range and the properties of the **MZ** operator, in particular to product of polytopes and Lipschitz functions.

4.4 The Limit Game

In addition to the convergence of the values, one could look for a normal form game $\mathcal{G}$ on $[0, 1]$ with strategy sets **U** and **V** and value w such that:

(1) the play at time t in $\mathcal{G}$ would be similar to the play at stage $[tn]$ in G_n (or at the fraction t of the total weight of the game for general evaluation)
(2) ε-optimal strategies in $\mathcal{G}$ would induce 2ε-optimal strategies in G_n, for large n.

Obviously then, $\lim v_n$ exists and is w.

One example was explicitly described (strategies and payoff) for the Big Match with incomplete information on one side in Sorin [74]. **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 piecewise constant approximation. **U** is the set of profiles 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 corresponding strategy of Player 1 in G_n has to satisfy the property that the probability of stopping the game before stage m is $\rho(\frac{m}{n})$. In the initial Big Match of Blackwell and Ferguson [7] (with complete information) one has $f(t) \equiv \frac{1}{2}$ and $\rho(t) = t$.

The auxiliary differential games introduced by Vieille [87] to study in weak approachability, Section 2.3 is also an example of a limit game.

The procedure is very similar to the approximation scheme of Souganidis [83] with $F(\rho, f) = \mathtt{val}\{\rho g + Ef\} = \rho \boldsymbol{\Phi}(\frac{f}{\rho})$, however the limit game is not given a priori and the operator is not smooth.

A recent example is in Laraki [37] and deals with absorbing games. Let $f(i, j)$ be the non absorbing payoff, $g(i, j)$ the absorbing payoff, $p(i, j)$ the probability of non absorption and $p^\star = 1 - p$. For $h : I \to \mathbb{R}$ define h on $\mathbb{R}^I$ by linear interpolation $h(\zeta) = \sum_{i \in I} \zeta(i) h(i)$. Given a stationary strategy $x \in X = \Delta(I)$ and a stationary pure strategy $j \in$

J the payoff in the discounted game is $r(\lambda, x, j) = \lambda f(x, j) + (1-\lambda)[p(x, j)r(\lambda, x, j) + \sum_i x(i)p^\star(i, j)g(i, j)]$. Define the absorbing part of the payoff $a(i, j) = p^\star(i, j)g(i, j)$ then

$$r(\lambda, x, j) = \frac{\lambda f(x, j) + (1-\lambda)a(x, j)}{\lambda p(x, j) + p^\star(x, j)}$$

and

$$v_\lambda = \max_{x\in\Delta(I)} \min_{j\in J} r(\lambda, x, j).$$

Let $w = \lim v_{\lambda_n}$ an accumulation point of the values and x_n an optimal stationary strategy of Player 1 in G_{λ_n}. Thus

$$v_{\lambda_n} \le \frac{\lambda_n f(x_n, j) + (1-\lambda_n)a(x_n, j)}{\lambda_n p(x_n, j) + p^\star(x_n, j)}, \quad \forall j \in J. \tag{29}$$

Assume that x_n converges to x.

If $p^\star(x, j) > 0$, going to the limit in (29) implies $w \le \frac{a(x,j)}{p^\star(x,j)}$.

Otherwise, letting $\alpha_n = \frac{x_n}{\lambda_n}$ one has

$$v_{\lambda_n} \le \frac{f(x_n, j) + (1-\lambda_n)a(\alpha_n, j)}{p(x_n, j) + p^\star(\alpha_n, j)}$$

hence for every ε, there exists $\alpha \in \mathcal{A} = (\mathbb{R}^+)^I$ such that

$$w \le \frac{f(x, j) + a(\alpha, j)}{1 + p^\star(\alpha, j)} + \varepsilon.$$

Thus

$$w \le \sup_{x\in X, \alpha\in\mathcal{A}} \min_{j\in J} \left[\frac{a(x, j)}{p^\star(x, j)} \mathbf{1}_{p^\star(x,j)>0} + \frac{f(x, j) + a(\alpha, j)}{1 + p^\star(\alpha, j)} \mathbf{1}_{p^\star(x,j)=0} \right] = W. \tag{30}$$

On the other hand let (x, α), ε-optimal in (30). Consider now $x[\lambda]$ proportional to $x + \lambda\alpha$, as a strategy of Player 1 in G_λ, for λ small enough and let j be a best reply, that one can take constant on a subsequence of λ_n. Then $v_{\lambda_n} \ge r(\lambda_n, x[\lambda_n], j)$ with

$$r\bigl(\lambda_n, x[\lambda_n], j\bigr) = \frac{\lambda_n f(x[\lambda_n], j) + (1-\lambda_n)a(x[\lambda_n], j)}{\lambda_n p(x[\lambda_n], j) + p^\star(x[\lambda_n], j)}$$

which is

$$r\bigl(\lambda_n, x[\lambda_n], j\bigr) = \frac{\lambda_n f(x, j) + (\lambda_n)^2 f(\alpha, j) + (1-\lambda_n)[a(x, j) + \lambda_n a(\alpha, j)]}{\lambda_n p(x, j) + (\lambda_n)^2 p(\alpha, j) + p^\star(x, j) + \lambda_n p^\star(\alpha, j)}.$$

Hence if $p^\star(x, j) > 0$, the limit of $r(\lambda_n, x[\lambda_n], j)$ is $\frac{a(x,j)}{p^\star(x,j)}$ and is otherwise $\frac{f(x,j)+a(\alpha,j)}{1+p^\star(\alpha,j)}$, if $p^\star(x, j) = 0$.

It follows that $w \ge W$.

Note that the proof shows that $\lim v_\lambda$ exists and is the value of the auxiliary game $\mathcal{G}$ with $\mathbf{U} = X \times \mathcal{A}$, $\mathbf{V} = Y \times \mathcal{B}$ and payoff function

$$L(x, \alpha, y, \beta) = \frac{a(x, y)}{p^\star(x, y)} \mathbf{1}_{p^\star(x,y)>0} + \frac{f(x, y) + a(\alpha, y) + a(x, \beta)}{1 + p^\star(\alpha, y) + p^\star(x, \beta)} \mathbf{1}_{p^\star(x,y)=0}.$$

Given a strategy (x, α) in $\mathcal{G}$, its image in G_λ is $x + \lambda\alpha$ (normalized).

4.5 Repeated Games and Evolution Equations

We follow Vigeral [88]. Consider again a non expansive mapping $\mathbf{T}$ from a Banach space X to itself.

The not-normalized values satisfy $V_n = \mathbf{T}^n(0)$ and

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

which can be considered as a discretization of the differential equation

$$\dot{x} = -Ax \tag{31}$$

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

The comparison between the iterates of $\mathbf{T}$ and the solution of (31) is as follows:

Theorem 4.10 (Chernoff's formula) *Let $U(t)$ be the solution of (31). Then*:

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

In particular with $U(0) = 0$ and $t = n$

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

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

$$\dot{x}(t) = \boldsymbol{\Phi}\big(\varepsilon(t), x(t)\big) - x(t) \tag{32}$$

where as usual $\boldsymbol{\Phi}(\varepsilon, x) = \varepsilon \mathbf{T}(\frac{1-\varepsilon}{\varepsilon} x)$. Notice that (32) is no longer autonomous.

Define the condition (C) by

$$\left\| \boldsymbol{\Phi}(\lambda, x) - \boldsymbol{\Phi}(\mu, x)\right\| \le |\lambda - \mu|\big(C + \|x\|\big).$$

Theorem 4.11 *Let $u(t)$ be the solution of (32), 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 (C).

(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 4.12 *Let $\bar{u}$ solution of (32) associated to $\bar{\varepsilon}$. Then $\|u(t) - \bar{u}(t)\| \to 0$ as soon as*

(i) *$\varepsilon(t) \sim \bar{\varepsilon}(t)$ as $t \to \infty$ or*
(ii) *$|\varepsilon - \bar{\varepsilon}| \in L^1$.*

5 The Dual of a Game with Incomplete Information

5.1 The Dual Game

Consider a two-person zero-sum game with incomplete information on one side defined by sets of actions S and T, a finite parameter space K, a probability $p \in P = \Delta(K)$ and for each k a real payoff function G^k on $S \times T$. Assume S and T convex and for each k, G^k bounded and bilinear on $S \times T$.

Note Obviously this covers the finite case where G^k is defined on $I \times J$, $S = \Delta(I), T = \Delta(J)$ and $G^k(s,t) = \mathsf{E}_{s,t} G^k$. However even if one starts with real payoff functions G^k on $A \times B$ where A and B are convex sets and $\mathtt{val}_{A\times B} G^k$ exists, there is a need for a mixed extension $\mathcal{A} = \Delta(A), \mathcal{B} = \Delta(B)$ with $G^k(\alpha, \beta) = \mathsf{E}_{\alpha,\beta}\, G^k$ on $\mathcal{A} \times \mathcal{B}$. Mixed actions are not only used for convexification of the action sets but also to have a linear structure on the payoff that allows to control the information (this aspect was known by Borel, see Sorin [78]).

The game is played as follows: $k \in K$ is selected according to p and told to Player 1 (the maximizer) while Player 2 only knows p. In normal form, Player 1 chooses $\mathbf{s} = \{s^k\}$ in S^K, Player 2 chooses t in T and the payoff is $G^p(\mathbf{s}, t) = \sum_k p^k G^k(s^k, t)$. 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 are concave in p on P, the first thanks to the splitting procedure, the second as an infimum of linear functions, see e.g. Sorin [77], Chapter 2.

Following De Meyer [22, 23], one introduces for each $z \in \mathbb{R}^k$, the "dual game" $G^*(z)$, where Player 1 chooses k and plays s in S while Player 2 plays 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. One has:

$$\underline{w}(z) = \sup_{\Delta(K)\times S^K} \inf_T \big[G^p(\mathbf{s}, t) - \langle p, z\rangle\big] = \sup_{\Delta(K)} \Big[\sup_{S^K} \inf_T \big(G^p(\mathbf{s}, t)\big) - \langle p, z\rangle\Big] \tag{33}$$

hence is convex in z. Similarly:

$$\overline{w}(z) = \inf_T \sup_{\Delta(K)\times S^K} \big[G^p(\mathbf{s}, t) - \langle p, z\rangle\big] = \inf_T \sup_S \Big[\sup_k \big(G^k(s, t) - z^k\big)\Big]. \tag{34}$$

Let $z = \sum_m \alpha^m z(m)$ be a barycentric combination in $\mathbb{R}^k$ and for $\varepsilon > 0$, $t(m)$ an ε-optimal strategy of Player 2 in $G^\star(z(m))$. Thus:

$$G^k\big(s, t(m)\big) - z^k(m) \le \underline{w}\big(z(m)\big) + \varepsilon, \quad \forall s \in S, k \in K.$$

Defining $t = \sum_m \alpha^m t(m)$ one obtains, by linearity:

$$G^k(s, t) - z^k \le \sum_m \alpha^m \underline{w}\big(z(m)\big) + \varepsilon, \quad \forall s \in S, k \in K$$

which implies $\underline{w}(z) \le \sum_m \alpha^m \underline{w}(z(m)) + \varepsilon$, hence $\underline{w}$ is also convex in z.

Obviously, since G is bounded $\underline{v}$ and $\overline{v}$ are Lipschitz in p and $\underline{w}, \overline{w}$ are 1-Lipschitz in z.

Theorem 5.1 *The following duality relations hold*:

$$\underline{w}(z) = \max_{p\in\Delta(K)} \{\underline{v}(p) - \langle p, z\rangle\}, \tag{35}$$

$$\underline{v}(p) = \inf_{z\in\mathbb{R}^K} \{\overline{w}(z) + \langle p, z\rangle\}, \tag{36}$$

$$\overline{w}(z) = \max_{p\in\Delta(K)} \{\overline{v}(p) - \langle p, z\rangle\}, \tag{37}$$

$$\overline{v}(p) = \inf_{z\in\mathbb{R}^K} \{\overline{w}(z) + \langle p, z\rangle\}. \tag{38}$$

Proof

$$\underline{w}(z) = \max_{\Delta(K)}\Big[\max_{S^K}\min_{T}\big(G^p(\mathbf{s},t)\big) - \langle p, z\rangle\Big] = \max_{\Delta(K)}\big[\underline{v}(p) - \langle p, z\rangle\big]$$

by definition, hence (35) and the dual equation (36) holds by Fenchel duality since $\underline{v}(p)$ is concave and continuous.

Let us now prove (38). Given $\varepsilon > 0$ and t an ε-optimal strategy of Player 2 in $G^\star(z)$ one has:

$$G^k(s,t) - z^k \le \overline{w}(z) + \varepsilon, \quad \forall s \in S, k \in K$$

which implies for all $p \in \Delta(K)$, $\mathbf{s} \in S^K$:

$$G^p(\mathbf{s},t) \le \underline{w}(z) + \varepsilon + \langle p, z\rangle$$

and in particular, for all $z \in \mathbb{R}^k$:

$$\overline{v}(p) \le \underline{w}(z) + \langle p, z\rangle$$

hence:

$$\overline{v}(p) \le \inf_{z\in\mathbb{R}^K} \{\underline{w}(z) + \langle p, z\rangle\}.$$

Let now t be ε-optimal in G^p and define $z(t)$ to be the vector payoff that t guarantees to Player 2: $z^k(t) = \max_S G^k(s,t)$. Optimality of t implies

$$\big\langle p, z(t)\big\rangle \le \overline{v}(p) + \varepsilon.$$

On the other hand, by playing t in $G^\star(z(t))$ Player 2 obtains at most 0; hence $\overline{w}(z(t)) \le 0$ which implies

$$\big\langle p, z(t)\big\rangle + \underline{w}\big(z(t)\big) \le \overline{v}(p) + \varepsilon$$

and equality in (38).

Finally (37) follows again by Fenchel duality since $\underline{w}(z)$ is convex and continuous. □

In terms of strategies one has the following correspondences:

Corollary 5.1 *Let $\varepsilon > 0$.*

(1) *Given z, let p achieve the maximum in (35) and $\mathbf{s}$ be ε-optimal in G^p: then $(p,\mathbf{s})$ is ε-optimal in $G^*(z)$.*
(2) *Given p, let z achieve the infimum up to ε in (37) and t be ε-optimal in $G^*(z)$: then t is also 2ε-optimal in G^p.*

5.2 The Recursive Equation for the Dual Game

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

$$(n+1)v_{n+1}(p) = \max_{x\in X^K}\min_{y\in Y}\left\{\sum_k p^k x^k G^k y + n\sum_i \hat{x}(i)v_n\big(p(i)\big)\right\} \tag{39}$$

with $\hat{x}(i) = \sum_k p^k x^k(i)$ and $p^k(i) = \text{Prob}\ (k|i)$. Note that, since Player 1 knows $p(i)$, this formula allows to construct inductively an optimal strategy for him in $G_n(p)$ (and in addition $p(i)$ will be a "state variable").

There are two issues here: first in the "true" game, Player 2 does not know $p(i)$, second the state variable is the martingale $p(i)$. The use of the dual game will be of interest for two purposes: construction of optimal strategies for the uninformed player and asymptotic analysis, De Meyer [23], De Meyer and Rosenberg [25].

Given G_n, let us consider the dual game $G_n^\star$ and its value w_n. From (35) or (37) one has

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

which, by using (39), leads to the recursive equation in the dual game:

$$(n+1)w_{n+1}(z) = \min_{y\in Y}\max_{i\in I} n w_n\left(\frac{n+1}{n}z - \frac{1}{n}G_i y\right). \tag{40}$$

In particular Player 2 has an optimal strategy in $G_{n+1}^\star(z)$ that depends only on z and the previous moves of the players: at stage 1 play y optimal in (40) and from stage 2 on, given the move i_1 of Player 1 at stage 1, play optimally in $G_n^\star(\frac{n+1}{n}z - \frac{1}{n}G_{i_1}y)$. Here z plays the role of a "state variable". Obviously a similar analysis is valid for G_λ and its dual.

5.3 The Associated Differential Game

The second advantage of dealing with (40) rather than with (39) is that the state variable evolves smoothly from z to $z + \frac{1}{n}(z - G_i y)$ while the martingale $p(i)$ could have jumps.

We follow Laraki [35] in considering w_n as the value of the time discretization with mesh $\frac{1}{n}$ of a differential game on [0, 1] with dynamic $\zeta(t) \in \mathbb{R}^K$ given by:

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

$x_t \in X$, $y_t \in Y$ and terminal payoff $\max_k \zeta^k(1)$. Basic results of differential games of fixed duration, see Appendix, 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 partial differential equation with boundary condition:

$$\frac{\partial\varphi}{\partial t} + u(D\varphi) = 0, \qquad \varphi(1,\zeta) = \max_k \zeta^k. \tag{41}$$

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

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

and finally $w(z) = \sup_{p\in\Delta(K)}\{u(p) - \langle p, z\rangle\}$. Hence $\lim v_\lambda = \lim v_n = \mathtt{Cav}_{\Delta(K)} u$, by taking the Fenchel conjugate. An alternative identification of the limit is through variational inequalities by translating in the primal the viscosity properties in the dual in terms of local sub- and super-differentials. This leads to the properties (26).

5.4 Differential Games with Incomplete Information

Similar tools have been recently introduced by Cardaliaguet [11] to study differential games of fixed duration and incomplete information on both sides $\Gamma(p, q)[\theta, t]$, see also the more general case of stochastic differential games with incomplete information developed in Cardaliaguet and Rainer [17], and in this issue by Buckdahn, Cardaliaguet and Quincampoix, [8].

K and L are finite sets and for each (k, ℓ) there is a differential game $\Gamma^{k\ell}$ on $[0, T]$ with control sets U and V (see Appendix for the hypotheses and a short reminder). The initial position is $z_0^{k\ell} \in Z^{k\ell}$, 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 $\bar{\gamma}^{k\ell}(z^{k\ell})$. $k \in K$ is chosen according to $p \in \Delta(K)$ and told to Player 1, similarly $\ell \in L$ is chosen according to $q \in \Delta(L)$ and told to Player 2. Then $\Gamma^{k\ell}$ is played. The corresponding game is $\Gamma(p, q)[z_0, 0]$. $\Gamma(p, q)[z, t]$ starting from $z = \{z^{k\ell}\}$ at time t is defined similarly. Note that the players use their information in choosing a strategy but in addition they have to use mixed strategies: $\alpha \in \bar{\mathcal{A}}$ is the choice at random of an element in $\mathcal{A}$. Hence a strategy for Player 1 is described by a profile $\hat{\alpha} = \{\alpha^k\} \in \bar{\mathcal{A}}^K$. The payoff induced by a couple 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).

Remark 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 in uninformed. In particular the analysis of Section 5.1 applies. Let us consider the minmax in $\Gamma(p, q)[z, t]$:

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

The dual game with respect to k and with parameter $\theta \in \mathbb{R}^K$ has a minmax that satisfies (34)

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

and (37)

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

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

Proposition 5.1

$$\overline{W}(\theta, q)[z, t] \le \inf_{\beta\in\mathcal{B}} \sup_{\alpha\in\mathcal{A}} \overline{W}\big(\theta(t+\delta), q\big)[\mathbf{z}_{t+\delta}, t+\delta] \tag{42}$$

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

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

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

Here $f(z,\cdot,\cdot,\cdot)$ stands for $\{f^{k\ell}(z^{k\ell},\cdot,\cdot,\cdot)\}$ 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 (37) to deduce from (42) the following property on $\overline{V}$:

Proposition 5.2 $\overline{V}$ *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*

$$\max\big\{\bar{L}\Phi(\bar{z},\bar{t},\bar{p}); \partial_t\Phi(\bar{z},\bar{t},\bar{p}) + H\big(\bar{z},\bar{t},D_z\Phi(\bar{z},\bar{t},\bar{p}),\bar{p},\bar{q}\big)\big\} \ge 0. \tag{43}$$

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

Finally a comparison principle extending Theorem A.3 proves the existence of a value V.

Theorem 5.2 *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 $F_1(\cdot,T,\cdot,\cdot)\le F_2(\cdot,T,\cdot,\cdot)$, then $F_1\le F_2$ on $Z\times[0,T]\times\Delta(K)\times\Delta(L)$.*

Using this characterization Souquière [85] 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 non revealing game and thus recovers Mertens–Zamir's result through differential games. This property does not hold in general, see examples in Cardaliaguet [12]. However one has the following approximation procedure. Given a finite partition Π of $[0,1]$ define inductively V_Π by:

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

where $\delta_{m+1} = t_{m+1} - t_m$. Then using results of Laraki [34, 36], Souquière [85] 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 [13], where moreover an algorithm for constructing approximate optimal strategies is provided. Hence the **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 non autonomous framework.

5.5 Continuous Time

In the same vein Cardaliaguet and Rainer [18] consider a continuous time game on $[0,T]$ with incomplete information on one side and payoff function $\gamma^k(t,u,v)$. Let $H(p,t) =$

$\mathtt{val}_{U\times V}\sum_k p^k\gamma^k(t,u,v)$. Then using the previous characterization they prove that the value is given by

$$V(t,p) = \sup_{\mathbf{p}\in M(p)} \mathsf{E}\int_t^T H\big(s,\mathbf{p}(s)\big)\,ds$$

where $M(p)$ is the set of càdlàg time martingales in $\Delta(K)$ starting from p. In addition they provide the construction of an optimal strategy for the informed player and explicit computations. (Compare with the splitting game, Section 4.3). More on this can be found in this issue, see Buckdahn, Cardaliaguet and Quincampoix, [8].

6 Uniform Approach

6.1 Basic Results

Concerning games with lack of information on one side, Aumann and Maschler (1966) proved the existence of a uniform value, see [2] and the famous formula $v(p) = \mathtt{Cav}_{p\in\Delta(K)}u(p)$. For games with lack of information on both sides, Aumann and Maschler (1967) proved that the maxmin and minmax exist, see [2], with moreover an explicit formula:

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

They also construct games without a value. For several extensions to the dependent case and signaling structure, mainly due to Mertens and Zamir, see Sorin [77].

In the framework of stochastic games with standard signaling (i.e. the moves are announced) the proof of the existence of a uniform value was obtained first for the "Big Match" by Blackwell and Ferguson [7], then for absorbing games by Kohlberg [30]. The main proof for general finite stochastic games is due to Mertens and Neyman [47]. This last result uses properties obtained by Bewley and Kohlberg [4] through their algebraic approach for v_λ 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 $\bar\lambda(L_n)$. In fact the result depends only on a property of the family $\{v_\lambda\}$ and allow one to extend the proof to absorbing (resp. recursive) games with compact action sets, Mertens, Neyman and Rosenberg [48] (resp. Vigeral [90]).

A first connection between incomplete information games and stochastic games is the so called "symmetric case". This corresponds to games where the state in M may not be known by the players but their information 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 the state process is very regular (martingale), Kohlberg and Zamir [31].

A collection of results proving the existence of the maxmin and/or the minmax for some classes of games includes Mertens and Zamir [51], Rosenberg and Vieille [71], Sorin [74, 75], see Sorin [77], Chapter 6 for a survey.

6.2 Dynamic Programming and MDP

In the framework of general dynamic programming (one person stochastic game with a state space Ω, a correspondence C from Ω to itself and a real bounded payoff g on Ω) Lehrer and Sorin [41] gave an example where $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ both exist and differ. They also proved that uniform convergence (on Ω) of v_n is equivalent to uniform convergence of v_λ and then the limits are the same. However this condition alone does not imply existence of the uniform value, v_∞, see Lehrer and Monderer [40], Monderer and Sorin [53].

Recent advances have been obtained by Renault [63] introducing new notions like the values v_{nm} (resp. ν_{nm}) of the game where the payoff is the average between stage $n+1$ and $n+m$ (resp. the minimum of all averages between stage $n+1$ and $n+\ell$ for $\ell \le m$).

Theorem 6.1 *Assume that the state space Ω is metric compact and the family of functions v_{nm} and ν_{nm} are uniformly equicontinuous. Then the uniform value v_∞ exits.*

Player 1 cannot get more than $\min_m \max_n \nu_{nm}$ and under the above conditions this quantity is also $\max_n \min_m \nu_{nm}$ (and the same with ν replaced by v). In particular when applied to Markov Decision Process (finite state space K, move space I, signal space A and transition from $K \times I$ to $K \times A$) the previous result implies:

Theorem 6.2 *General MDP processes with finite state space have a uniform value.*

This extends previous result by Rosenberg, Solan and Vieille [66]. For further development to the continuous time setup, see Quincampoix and Renault [61], Oliu-Barton and Vigeral [58].

6.3 Games with Transition Controlled by One Player

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

Again here Player 1 cannot get more than $\min_m \max_n \nu_{nm}$. One 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) and the finiteness assumption on the basic data implies

Theorem 6.3 (Renault [64]) *In the finite case, games with transition controlled by one player have a uniform value.*

The result extends previous work of Rosenberg, Solan and Vieille [69] and also the model of Markov game with lack of information on one side, Renault [62] for which explicit formulas are not yet available (Marino [44], Horner, Rosenberg, Solan and Vieille [28], Neyman [55]).

6.4 Stochastic Games with Signals

Consider a stochastic game and assume here that the signal to each player reveals the current stage 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 than in the standard signaling case since the state variable is not affected by the change in the information. However for example in the Big Match, when Player 1 has no information on Player 2's action the max min is 0 (Kohlberg [30]) and the uniform value does not exist.

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 6.4 *Maxmin and minmax exist in stochastic games with signals.*

This recent result, due to Coulomb [19], and Rosenberg, Solan and Vieille [67] is extremely involved and relies on the construction of two auxiliary games. Consider the maxmin and some discount factor λ. Introduce 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 2 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 the same analysis than in Mertens and Neyman [47] applies. It remains to show, for Player 1, that this auxiliary payoff indeed can be achieved in the real game. As 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 this on, without being detected.

An illuminating example, due to Coulomb [19], is as follows:

	α	β	γ
a	1^*	0^*	L
b	0	1	L

Payoffs (L large)

	α	β	γ
a	?	?	?
b	A	B	A

Signals to 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 a in the future is small enough.

For an overview, see Rosenberg, Solan and Vieille [68].

7 Comments, Conjectures and Open Problems

7.1 Comments

The above analysis shows that the same tools can be used for v_n and v_λ, even for general evaluation and presumably for all random duration processes … (Recall that G_n has finitely many stages while G_λ has infinitely many; that v_n satisfies a recurrence property while v_λ appears as a fixed point …) By considering the associated game on $[0, 1]$ one obtains a single point of view.

In a similar way, there is a unification in terms of structure of proofs from incomplete information games to stochastic games and mixture of those.

Moreover several concepts introduced in RG extend to the non autonomous framework, but at a differential level (on an interval of time on which the structure is essentially stationary).

Among the main relations between RG and DG (in addition to Section 2) one can note:

- The recursive formula which is a discrete version of the dynamic programming formula.
- RG with symmetric incomplete information on the state, Kohlberg and Zamir [31], are similar to DG on the Wasserstein space, Cardaliaguet and Quincampoix [14].
- The use of comparison principles. We can distinguish 3 types: (1) in DG there are two variational inequalities satisfied by the max min $\overline{v}(p)$ and by the min max $\underline{v}(p)$ and one

uses a comparison principle to prove that $\underline{v}(p) \geq \overline{v}(p)$; (2) in RG one has a family of functions satisfying a (parametrized) functional equation. Then the lim sup (resp. lim inf) of this family obeys a functional inequality and a comparison principle applies, see Mertens and Zamir [50]; (3) one shows uniqueness of the accumulation points of the family by checking a comparison principle at specified points for a class of functions (fixed points of the projective operator).

Formally the tools are very similar: in (43) the presence of $\bar{L}$ localizes the point and one works with the family of saddle functions.

The main differences are:

(1) In the RG framework one cannot directly go to the limit in time. The main obstacles are: (i) the difficulty to define the strategies in continuous time while taking into account the information structure (one has to find a finite dimensional summary, i.e. exhaustive statistics of the past history, see e.g. Mertens and Zamir [51]), (ii) the fact that the dynamics of the state variable may not be smooth (see e.g. the double scale in Sorin [76]).
(2) For RG one has equality in the recursive formula for a family of functions then one looks for the existence of a limit.
(3) In the DG framework the limit game is given but 2 equations are obtained for the maxmin and the minmax. The time discretization is used to obtain equality.

Notice finally that, while even restricted to the zero-sum case, this survey is far to be complete: among several very active domains not covered, let us mention stopping games where the results and tools are quite similar to those considered here: see Laraki and Solan [38] and the survey by Solan and Vieille [73].

7.2 Conjectures

Let us consider the general model of Section 3. Assume that all sets under consideration are finite: M, A, B, I and J.

The first conjecture is that the asymptotic value exists in the following strong sense:

(1) $\lim_{\lambda\to 0} v_\lambda$ exists, similarly
(2) $\lim_{n\to\infty} v_n$ exists, or more generally for any admissible family of random duration processes
(3) $\lim_{\mathbb{E}(\Theta)\to\infty} v_\Theta$ exists, and in addition the limit is the same
(4) $\lim_{\lambda\to 0} v_\lambda = \lim_{n\to\infty} v_n = \lim_{\mathbb{E}(T)\to\infty} v_T = \mathbf{v}$.

A second conjecture is that $\underline{v}$ and $\overline{v}$ exist in this framework. A third conjecture is that in games where the information of Player 1 contains the information of Player 2, the maxmin is equal to the asymptotic value: $\underline{v} = \mathbf{v}$, see Mertens [46], Mertens, Sorin and Zamir [49], Coulomb [20]. A last one is that the asymptotic value exists in stochastic games with finite state space, compact action sets and continuous payoff and transition function.

7.3 Open Problems

(1) Characterization of $\mathbf{v}$: A basic problem is to identify the asymptotic value, namely to find the analog of the **MZ** operator. A first class to consider is finite stochastic games and one looks for an operator proof of existence of $\lim_{\lambda\to 0} v_\lambda$.

(2) Construction of the limit game: Note that in games with incomplete information one can imagine 3 different limit games. The first one corresponds to the asymptotics of the auxiliary game used in the generalized Shapley operator. Then there are two other games related to each dual game.
When dealing with games with incomplete information on one side the auxiliary limit game is also the dual limit game for Player 1 and the optimal strategy of Player 1 is trivial: splitting at time 0 to realize the `Cav`. The extension to the non autonomous case is in Cardaliaguet and Rainer [18]. The dual limit game for Player 2 corresponds to the differential game introduced by Laraki [35]. Note that this is basically a game in continuous time.
Among the important questions are the similar constructions for games with lack of information on both sides, in particular to deduce properties of optimal strategies in long games, see De Meyer [24].
(3) Given a pair of stationary optimal strategies in a discounted stochastic game $x_\lambda(\omega)$ and $y_\lambda(\omega)$ the payoff in each state ω is determined, say $z_\lambda(\omega)$. Given the transition matrix $R_\lambda(\omega, \omega') = \sum_{ij} x^i_\lambda(\omega) y^j_\lambda(\omega) Q(i, j, \omega)(\omega')$, the occupation measure μ_λ starting from state ω is defined by

$$\mu_\lambda[\omega](\omega') = \sum_{n=1}^{\infty} \lambda(1-\lambda)^{n-1} R_\lambda^{n-1}(\omega, \omega')$$

and corresponds to the fraction of the length of the game spent in state ω'. Obviously one has $v_\lambda(\omega) = \langle \mu_\lambda[\omega](\cdot), z_\lambda(\cdot)\rangle$. If both state space and action spaces are finite all the above quantities are semi-algebraic functions of λ hence limits as λ goes to 0 exist and one obtains as well $\lim_{\lambda\to 0} v_\lambda(\omega) = v(\omega) = \langle \mu[\omega], z\rangle$.
Is it possible to identify directly these quantities through variational inequalities??
What is the relation with the limit game?
(4) The stationary aspect of the RG implies that, in the limit games, optimal strategies will be independent of the time: they will simply be a function of the current "state". On could also think that the corresponding stage payoff would be constant, hence equal to $\mathbf{v}$, see in this direction Sorin, Venel and Vigeral [82].
(5) Asymptotic approach with random duration process: The question is to find conditions on the non expansive operator $\mathbf{T}$ to deduce from the existence of an asymptotic value for deterministic evaluations the analog for the random case.
(6) Recursive structure with signals: Consider a game with lack of information on both sides and assume that the signals on the actions are independent of the state. However the natural belief space can be quite complex for example if Player 1 does not know the signal of Player 2. Mertens [45] introduced then a majorant and a minorant game and proved that they have the same asymptotic value. What would be the analog in terms of recursive formula?
(7) Extension to the dependent case: Consider an incomplete information game where the unknown parameter $k \in K$ is chosen according to p and the players have private information described by private partitions K^1 and K^2. Define the dual games and construct optimal strategies.
(8) Speed of convergence: For games with incomplete information the speed of convergence of v_n (resp. v_λ) to its limit is $\frac{1}{\sqrt{n}}$ (resp. $\sqrt{\lambda}$), Aumann and Maschler [2], Mertens and Zamir [50]. What is the relation with the approximation schemes having the same property in Souganidis [83]?

Appendix: Quantitative DG and Viscosity Solutions

We describe very briefly the main tools in the proof of existence of a value, due to Evans and Souganidis [26], but using NAD strategies.

Consider the case defined by (1) 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), $\bar{\gamma}$ (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 from z and let $\overline{v}[z,t]$ and $\underline{v}[z,t]$ denote the corresponding minmax and maxmin. Explicitly

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

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 $\mathbf{u}_s = \mathbf{u}_s(\alpha,\beta)$, $\mathbf{v}_s = \mathbf{v}_s(\alpha,\beta)$, $\mathbf{z}_s = \mathbf{z}_s(\alpha,\beta,z,t)$.
The first property is the following dynamic programming inequality:

Theorem A.1 *For* $0 \leq t \leq t+\delta \leq T$, $\overline{v}$ *satisfies*:

$$\overline{v}[z,t] \leq \inf_\beta \sup_\alpha \left\{\int_t^{t+\delta} \gamma\big(\mathbf{z}_s(\alpha,\beta,z,t), s, \mathbf{u}_s(\alpha,\beta), \mathbf{v}_s(\alpha,\beta)\big)\, ds + \overline{v}\big[\mathbf{z}_{t+\delta}(\alpha,\beta,z,t), t+\delta\big]\right\}. \tag{44}$$

In addition $\overline{v}$ is uniformly Lipschitz in z and t.
Property (44) 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:

$$\begin{aligned} \inf_\beta \sup_\alpha \bigg\{ & \frac{1}{\delta} \int_t^{t+\delta} \gamma\big(\mathbf{z}_s(\alpha,\beta,z,t), s, \mathbf{u}_s(\alpha,\beta), \mathbf{v}_s(\alpha,\beta)\big)\, ds \\ & + \frac{\Phi[\mathbf{z}_{t+\delta}(\alpha,\beta,z,t), t+\delta] - \Phi[t,z]}{\delta}\bigg\} \geq 0. \end{aligned} \tag{45}$$

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

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

which gives the differential inequality:

Theorem A.2

$$\partial_t \Phi[z,t] + H^+\big(D\Phi[z,t], z, t\big) \geq 0. \tag{46}$$

The fact that any smooth local majorant of $\overline{v}$ satisfies (46) can be express as: $\overline{v}$ 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 A.3 *Let W_1 be a viscosity subsolution and W_2 be a viscosity supersolution of*

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

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

to obtain finally:

Theorem A.4 *The differential game has a value*:

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

In fact the previous Theorem A.3 implies $\overline{v}[z,t] \le \underline{v}[z,t]$.

Note that the comparison Theorem A.3 is much more general and applies to W_1 s.c.s., W_2 s.c.i., H uniformly Lipschitz in p and satisfying: $|H(p,z_1,t_1) - H(p,z_2,t_2)| \le C(1 + \|p\|)\|(z_1,t_1) - (z_2,t_2)\|$. 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 [84] , Bardi and Capuzzo Dolcetta [3] and for viscosity solutions Crandall, Ishii and Lions [21].

References

1. 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
2. Aumann RJ, Maschler M (1995) Repeated games with incomplete information. MIT Press, Cambridge (with the collaboration of R. Stearns)
3. Bardi M, Capuzzo Dolcetta I (1996) Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations. Birkhauser, Basel
4. Bewley T, Kohlberg E (1976a) The asymptotic theory of stochastic games. Math Oper Res 1:197–208
5. Bewley T, Kohlberg E (1976b) The asymptotic solution of a recursion equation occurring in stochastic games. Math Oper Res 1:321–336
6. Blackwell D (1956) An analog of the minmax theorem for vector payoffs. Pac J Math 6:1–8
7. Blackwell D, Ferguson T (1968) The Big Match. Ann Math Stat 39:159–163
8. Buckdahn R, Cardaliaguet P, Quincampoix M (2010) Some recent aspects of differential game theory. Dyn Games Appl, this issue
9. Cardaliaguet P (1996) A differential game with two players and one target. SIAM J Control Optim 34:1441–1460
10. Cardaliaguet P (1997) Nonzero-sum differential games revisited. Working paper (unpublished)
11. Cardaliaguet P (2007) Differential games with asymmetric information. SIAM J Control Optim 46:816–838
12. Cardaliaguet P (2008) Representation formulas for differential games with asymmetric information. J Optim Theory Appl 138:1–16
13. Cardaliaguet P (2009) Numerical approximation and optimal strategies for differential games with lack of information on one side. In: Bernhard P, Gaitsgory V, Pourtalier O (eds) Advances in dynamic games and their applications. Annals of ISDG, vol 10. Birkhauser, Basel, pp 159–176
14. Cardaliaguet P, Quincampoix M (2008) Deterministic differential games under probability knowledge of initial condition. Int Game Theory Rev 10:1–16
15. Cardaliaguet P, Quincampoix M, Saint-Pierre P (1999) Numerical methods for differential games. In: Bardi M, Parthasarathy T, Raghavan TES (eds) Stochastic and differential games: theory and numerical methods. Annals of ISDG, vol 4. Birkhauser, Basel, pp 177–247
16. Cardaliaguet P, Quincampoix M, Saint-Pierre P (2007) Differential games through viability theory: old and recent results. In: Jorgensen S, Quincampoix M, Vincent T (eds) Advances in dynamic games theory. Annals of ISDG, vol 9. Birkhauser, Basel, pp 3–36
17. Cardaliaguet P, Rainer C (2009a) Stochastic differential games with asymmetric information. Appl Math Optim 59:1–36

18. Cardaliaguet P, Rainer C (2009b) On a continuous time game with incomplete information. Math Oper Res 34:769–794
19. Coulomb JM (2003) Stochastic games without perfect monitoring. Int J Game Theory 32:73–96
20. Coulomb JM (2003) Games with a recursive structure. In: Neyman A, Sorin S (eds) Stochastic games and applications. NATO science series C, vol 570. Kluwer Academic, Dordrecht, pp 427–442
21. Crandall MG, Ishii H, Lions P-L (1992) User's guide to viscosity solutions of second order partial differential equations. Bull Am Soc 27:1–67
22. De Meyer B (1996a) Repeated games and partial differential equations. Math Oper Res 21:209–236
23. De Meyer B (1996b) Repeated games, duality and the Central Limit theorem. Math Oper Res 21:237–251
24. De Meyer B (1999) From repeated games to Brownian games. Ann Inst H Poincaré, Probab Stat 35:1–48
25. De Meyer B, Rosenberg D (1999) "Cav u" and the dual game. Math Oper Res 24:619–626
26. Evans LC, Souganidis PE (1984) Differential games and representation formulas for solutions of Hamilton–Jacobi equations. Indiana Univ Math J 33:773–797
27. Everett H (1957) Recursive games. In: Dresher M, Tucker AW, Wolfe P (eds) Contributions to the theory of games, III. Annals of mathematical studies, vol 39. Princeton University Press, Princeton, pp 47–78
28. Horner J, Rosenberg D, Solan E, Vieille N (2010) On a Markov game with one-sided incomplete information. Oper Res 58:1107–1115
29. Hou T-F (1971) Approachability in a two-person game. Ann Math Stat 42:735–744
30. Kohlberg E (1974) Repeated games with absorbing states. Ann Stat 2:724–738
31. Kohlberg E, Zamir S (1974) Repeated games of incomplete information: the symmetric case. Ann Stat 2:1040
32. Krasovskii NN, Subbotin AI (1988) Game-theoretical control problems. Springer, Berlin
33. Laraki R (2001a) Variational inequalities, systems of functional equations and incomplete information repeated games. SIAM J Control Optim 40:516–524
34. Laraki R (2001b) The splitting game and applications. Int J Game Theory 30:359–376
35. Laraki R (2002) Repeated games with lack of information on one side: the dual differential approach. Math Oper Res 27:419–440
36. Laraki R (2004) On the regularity of the convexification operator on a compact set. J Convex Anal 11:209–234
37. Laraki R (2010a) Explicit formulas for repeated games with absorbing states. Int J Game Theory 39:53–69
38. Laraki R, Solan E (2005) The value of zero-sum stopping games in continuous time. SIAM J Control Optim 43:1913–1922
39. Lehrer E (2002) Approachability in infinite dimensional spaces. Int J Game Theory 31:253–268
40. Lehrer E, Monderer D (1994) Discounting versus averaging in dynamic programming. Games Econ Behav 6:97–113
41. Lehrer E, Sorin S (1992) A uniform Tauberian theorem in dynamic programming. Math Oper Res 17:303–307
42. Maitra A, Sudderth W (1996) Discrete gambling and stochastic games. Springer, Berlin
43. Maitra A, Sudderth W (2003) Borel stay-in-a-set games. Int J Game Theory 32:97–108
44. Marino A (2005) The value and optimal strategies of a particular Markov chain game. Preprint
45. Mertens J-F (1972) The value of two-person zero-sum repeated games: the extensive case. Int J Game Theory 1:217–227
46. Mertens J-F (1987) Repeated games. In: Proceedings of the international congress of mathematicians, Berkeley 1986. American Mathematical Society, Providence, pp 1528–1577
47. Mertens J-F, Neyman A (1981) Stochastic games. Int J Game Theory 10:53–66
48. Mertens J-F, Neyman A, Rosenberg D (2009) Absorbing games with compact action spaces. Math Oper Res 34:257–262
49. Mertens J-F, Sorin S, Zamir S (1994) Repeated games. CORE DP 9420-21-22
50. 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
51. Mertens J-F, Zamir S (1976) On a repeated game without a recursive structure. Int J Game Theory 5:173–182
52. Mertens J-F, Zamir S (1985) Formulation of Bayesian analysis for games with incomplete information. Int J Game Theory 14:1–29
53. Monderer D, Sorin S (1993) Asymptotic properties in dynamic programming. Int J Game Theory 22:1–11
54. Neyman A (2003) Stochastic games and nonexpansive maps. In: Neyman A, Sorin S (eds) Stochastic games and applications. NATO science series C, vol 570. Kluwer Academic, Dordrecht, pp 397–415

55. Neyman A (2008) Existence of optimal strategies in Markov games with incomplete information. Int J Game Theory 37:581–596
56. Neyman A, Sorin S (eds) (2003) Stochastic games and applications. NATO science series C, vol 570. Kluwer Academic, Dordrecht
57. Neyman A, Sorin S (2010) Repeated games with public uncertain duration process. Int J Game Theory 39:29–52
58. Oliu-Barton M, Vigeral G (2009) A uniform Tauberian theorem in optimal control. Preprint
59. Perchet V (2010) Approachability of convex sets in games with partial monitoring. Preprint
60. Perchet V, Quincampoix M (2010) Purely informative game: approachability in Wasserstein space. Preprint
61. Quincampoix M, Renault J (2009) On the existence of a limit value in some non expansive optimal control problems. Preprint
62. Renault J (2006) The value of Markov chain games with lack of information on one side. Math Oper Res 31:490–512
63. Renault J (2007) Uniform value in dynamic programming. Revised version: arXiv:0803.2758
64. Renault J (2009) The value of repeated games with an informed controller. Preprint
65. Rosenberg D (2000) Zero-sum absorbing games with incomplete information on one side: asymptotic analysis. SIAM J Control Optim 39:208–225
66. Rosenberg D, Solan E, Vieille N (2002) Blackwell optimality in Markov decision processes with partial observation. Ann Stat 30:1178–1193
67. Rosenberg D, Solan E, Vieille N (2003a) The maxmin value of stochastic games with imperfect monitoring. Int J Game Theory 32:133–150
68. Rosenberg D, Solan E, Vieille N (2003b) Stochastic games with imperfect monitoring. In: Haurie A, Muto S, Petrosjan LA, Raghavan TES (eds) Advances in dynamic games: applications to economics, management science, engineering, and environmental management. Annals of the ISDG, vol 8. pp 3–22
69. Rosenberg D, Solan E, Vieille N (2004) Stochastic games with a single controller and incomplete information. SIAM J Control Optim 43:86–110
70. Rosenberg D, Sorin S (2001) An operator approach to zero-sum repeated games. Isr J Math 121:221–246
71. Rosenberg D, Vieille N (2000) The Maxmin of recursive games with incomplete information on one side. Math Oper Res 25:23–35
72. Shapley LS (1953) Stochastic games. Proc Nat Acad Sci USA 39:1095–1100
73. Solan E, Vieille N (2004) Stopping games—recent results. In: Nowak AS, Szajowski K (eds) Advances in dynamic games. Annals of the ISDG, vol 7. Birkhauser, Basel, pp 235–245
74. Sorin S (1984) "Big Match" with lack of information on one side, Part I. Int J Game Theory 13:201–255
75. Sorin S (1985) "Big Match" with lack of information on one side, Part II. Int J Game Theory 14:173–204
76. Sorin S (1989) On repeated games without a recursive structure: existence of $\lim v_n$. Int J Game Theory 18:45–55
77. Sorin S (2002a) A first course on zero-sum repeated games. Springer, Berlin
78. Sorin S (2002b) Bluff and reputation. In: Schmidt C (ed) Game theory and economic analysis, Routledge. pp 57–73
79. 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, vol 570. Kluwer Academic, Dordrecht, pp 417–426
80. Sorin S (2004) Asymptotic properties of monotonic nonexpansive mappings. Discrete Events Dyn Syst 14:109–122
81. Sorin S (2005) New approaches and recent advances in two-person zero-sum repeated games. In: Nowak AS, Szajowski K (eds) Advances in dynamic games. Annals of the ISDG, vol 7. Birkhauser, Basel, pp 67–93
82. Sorin S, Venel X, Vigeral G (2010) Asymptotic properties of optimal trajectories in dynamic programming. Sankhya, Indian J Stat A 72:237–245
83. Souganidis PE (1985) Approximation schemes for viscosity solutions of Hamilton–Jacobi equations. J Differ Equ 17:781–791
84. Souganidis PE (1999) Two player zero sum differential games and viscosity solutions. In: Bardi M, Raghavan TES, Parthasarathy T (eds) Stochastic and differential games. Annals of the ISDG, vol 4. Birkhauser, Basel, pp 70–104
85. Souquière A (2010) Approximation and representation of the value for some differential games with asymmetric information. Int J Game Theory 39:699–722
86. Spinat X (2002) A necessary and sufficient condition for approachability. Math Oper Res 27:31–44
87. Vieille N (1992) Weak approachability. Math Oper Res 17:781–791

88. Vigeral G (2010a) Evolution equations in discrete and continuous time for non expansive operators in Banach spaces. ESAIM COCV, to appear
89. Vigeral G (2010b) Iterated monotonic nonexpansive operators and asymptotic properties of zero-sum stochastic games. Preprint
90. Vigeral G (2010c) Zero-sum recursive games with compact action spaces. Preprint
91. Zamir S (1973) On the notion of value for games with infinitely many stages. Ann Stat 1:791–796