Stochastic Approximations and Differential Inclusions. Part II: Applications *

Michel Benaïm
Institut de Mathematiques, Université de Neuchâtel
Rue Emile-Argand 11, Neuchâtel, Suisse
michel.benaim@unine.ch

Josef Hofbauer
Department of Mathematics, University College London
London WC1E 6BT, U.K.
j.hofbauer@ucl.ac.uk

Sylvain Sorin
Equipe Combinatoire et Optimisation, UFR 921,
Université P. et M. Curie - Paris 6,
175 Rue du Chevaleret, 75013 Paris
and Laboratoire d'Econométrie, Ecole Polytechnique,
1 rue Descartes, 75005 Paris, France
sorin@math.jussieu.fr

revised December 2005

*We acknowledge financial support from the Swiss National Science Foundation, grant 200021-1036251/1 and from UCL's Centre for Economic Learning and Social Evolution (ELSE).

Abstract

We apply the theoretical results on "stochastic approximations and differential inclusions" developed in Benaïm, Hofbauer and Sorin (2005) to several adaptive processes used in game theory including: classical and generalized approachability, no-regret potential procedures (Hart and Mas-Colell), smooth fictitious play (Fudenberg and Levine).

Keywords: Stochastic approximation, differential inclusions, set valued dynamical systems, approachability, no regret, consistency, smooth fictitious play.

1 Introduction

The first paper of this series (Benaïm, Hofbauer and Sorin, 2005), henceforth referred to as BHS, was devoted to the analysis of the long term behavior of a class of continuous paths called *perturbed solution* that are obtained as certain perturbations of trajectories solutions to a *differential inclusion* in $\mathbb{R}^m$

$$\dot{\mathbf{x}} \in M(\mathbf{x}). \tag{1}$$

A fundamental and motivating example is given by (continuous time linear interpolation of) discrete stochastic approximations of the form

$$X_{n+1} - X_n = a_{n+1} Y_{n+1} \tag{2}$$

with

$$\mathsf{E}(Y_{n+1} \mid \mathcal{F}_n) \in M(X_n)$$

where $n \in \mathbb{N}$, $a_n \geq 0, \sum_n a_n = +\infty$ and $\mathcal{F}_n$ is the σ-algebra generated by $(X_0, \ldots, X_n)$, under conditions on the increments $\{Y_n\}$ and the coefficients $\{a_n\}$. For example if:

(i) $\sup_n \mid Y_{n+1} - \mathsf{E}(Y_{n+1} \mid \mathcal{F}_n) \mid < \infty$, and

(ii) $a_n = o(\frac{1}{\log(n)})$

the interpolation of a process $\{X_n\}$ satisfying (2) is almost surely a perturbed solution of (1).

Following the dynamical system approach to stochastic approximations initiated by Benaïm and Hirsch (Benaïm (1996, 1999); Benaïm and Hirsch (1996, 1999)) it was shown in BHS that the set of limit points of a perturbed solution is a *compact invariant attractor free set* for the set-valued dynamical system induced by (1).

From a mathematical viewpoint, this type of property is a natural generalization of Benaïm and Hirsch's previous results.[1] In view of applications, it is strongly motivated by a large class of problems, especially in game theory, where the use of differential inclusions is unavoidable since one deals with unilateral dynamics where the strategies chosen by a player's opponents (or nature) are unknown to this player.
In BHS a few applications where given: 1) in the framework of approachability theory (where one player aims at controlling the asymptotic behavior of the Cesaro mean of a sequence of vector payoffs corresponding to the outcomes of a repeated game) and 2) for the study of fictitious play (where each player uses, at each stage of a repeated game, a move which is a best reply to the past frequencies of moves of the opponent).

The purpose of the current paper is to explore much further the range of possible applications of the theory and to convince the reader that it provides a unified and powerful approach to several questions such as approachability or consistency (no regret). The price to pay is a bit of theory, but as a reward we obtain neat and simpler (sometime much simpler...) proofs of numerous results arising in different contexts.

The general structure for the analysis of such discrete time dynamics relies on the identification of a state variable for which the increments satisfies an equation like (2). This requires in particular vanishing step size (for example the state variable will be a time average—of payoffs or moves—) and a Markov property for the conditional law of the increments (the behavioral strategy will be a function of the state variable).

The organization of the paper is as follows. Section 2 summarizes the results of BHS that will be needed here. In section 3 we first consider generalized approachability, where the parameters are a correspondence N and a potential function Q adapted to a set C and extend results obtained by Hart and Mas-Colell (2001a). In Section 4 we deal with (external) consistency (or no regret): the previous set C is now the negative orthant and an approachability strategy is constructed explicitely through a potential function P, following Hart and Mas-Colell (2001a). A similar approach (Section 5) allows also to recover conditional (or internal) consistency properties via generalized approachability. The next section 6 shows analogous results for an alternative dynamics: smooth fictitious play. This allows to retrieve and extend certain properties obtained by Fudenberg and Levine (1995, 1999) on consistency and conditional consistency. Section 7 deals with several extensions of the previous to the case where the

[1] Benaïm and Hirsch's analysis was restricted to asymptotic pseudo trajectories (perturbed solutions) of differential equations and flows.

information available to a player is reduced, and section 8 applies to results recently obtained by Benaïm and BenArous (2003).

2 General framework and previous results

Consider the differential inclusion (1). All the analysis will be done under the following condition, which corresponds to Hypothesis 1.1 in BHS:

Hypothesis 2.1 (Standing assumptions) M is an upper semi continuous correspondence from $\mathbb{R}^m$ to itself, with compact convex non–empty values and which satisfies the following growth condition. There exists $c > 0$ such that for all $x \in \mathbb{R}^m$

$$\sup_{z \in M(x)} \|z\| \leq c\,(1 + \|x\|).$$

Here $\|\cdot\|$ denotes any norm on $\mathbb{R}^m$.

Remark These conditions are quite standard and such correspondences are sometimes called Marchaud maps (see Aubin (1991), p. 62). Note also that in most of our applications one has $M(x) \subset K_0$ where K_0 is a given compact set, so that the growth condition is automatically satisfied.

In order to state the main results of BHS that will be used here, we first recall some definitions and notation.
The *set–valued dynamical system* $\{\Phi_t\}_{t\in\mathbb{R}}$ induced by (1) is defined by

$$\Phi_t(x) = \{\mathbf{x}(t) : \mathbf{x} \text{ is a solution to (1) with } \mathbf{x}(0) = x\}$$

where a solution to the differential inclusion (1) is an absolutely continuous mapping $\mathbf{x} : \mathbb{R} \to \mathbb{R}^m$ satisfying

$$\frac{d\mathbf{x}(t)}{dt} \in M(\mathbf{x}(t))$$

for almost every $t \in \mathbb{R}$.
Given a set of times $T \subset \mathbb{R}$ and a set of positions $V \subset \mathbb{R}^m$

$$\Phi_T(V) = \bigcup_{t\in T} \bigcup_{v\in V} \Phi_t(v)$$

denote the set of possible values, at some time in T, of trajectories being in V at time 0.

Given a point $x \in \mathbb{R}^m$ we let

$$\omega_\Phi(x) = \bigcap_{t \geq 0} \overline{\Phi_{[t,\infty)}(x)}$$

denote its ω–*limit set* (where as usual the bar stands for the closure operator). The corresponding notion for a set Y, denoted $\omega_\Phi(Y)$, is defined similarly with $\Phi_{[t,\infty)}(Y)$ instead of $\Phi_{[t,\infty)}(x)$.
A set A is said *invariant* if, for all $x \in A$ there exists a solution $\mathbf{x}$ with $\mathbf{x}(0) = x$ such that $\mathbf{x}(\mathbb{R}) \subset A$, and *strongly positively invariant* if $\Phi_t(A) \subset A$ for all $t > 0$.
A non–empty compact set A is called an *attracting* set if there exists a neighborhood U of A and a function $\mathbf{t}$ from $(0, \varepsilon_0)$ to $\mathbb{R}^+$ with $\varepsilon_0 > 0$ such that

$$\Phi_t(U) \subset A^\varepsilon$$

for all $\varepsilon < \varepsilon_0$ and $t \geq \mathbf{t}(\varepsilon)$, where A^ε stands for the $\varepsilon-$neighborhood of A. This corresponds to a strong notion of attraction, uniform with respect to the initial conditions and the feasible trajectories.
If additionally A is invariant, then A is called an *attractor*.
Given an attracting set (resp. attractor) A, its *basin of attraction* is the set

$$B(A) = \{x \in \mathbb{R}^m : \omega_\Phi(x) \subset A\}.$$

When $B(A) = \mathbb{R}^m$, we call A a *globally* attracting set (resp. a global attractor).

Remark The following terminology is sometimes used in the literature. A set A is said *asymptotically stable* if it is

(i) invariant,

(ii) *Lyapounov stable*, i.e., for every neighborhood U of A there exists a neighborhood V of A such that its forward image $\Phi_{[0,\infty)}(V)$ satisfies $\Phi_{[0,\infty)}(V) \subset U$, and

(iii) *attractive*, i.e., its basin of attraction $B(A)$ is a neighborhood of A.

However, as shown in (BHS, Corollary 3.18) attractors and compact asymptotically stable sets coincide.

Given a closed invariant set L, the induced dynamical system Φ^L is defined on L by

$$\Phi_t^L(x) = \{\mathbf{x}(t) : \mathbf{x} \text{ is a solution to (1) with } \mathbf{x}(0) = x \text{ and } \mathbf{x}(\mathbb{R}) \subset L\}.$$

A set L is said *attractor free* if there exists no proper subset A of L which is an attractor for Φ^L.

We now turn to the discrete random perturbations of (1) and consider, on a probability space $(\Omega, \mathcal{F}, P)$, random variables $X_n, n \in \mathbb{N}$, with values in $\mathbb{R}^m$, satisfying the difference inclusion

$$X_{n+1} - X_n \in a_{n+1}[M(X_n) + U_{n+1}] \tag{3}$$

where the coefficients a_n are nonnegative numbers with

$$\sum_n a_n = +\infty.$$

Such a process $\{X_n\}$ is a *Discrete Stochastic Approximation* (DSA) of the differential inclusion (1) if the following conditions on the perturbations $\{U_n\}$ and the coefficients $\{a_n\}$ hold:

(i) $\mathsf{E}(U_{n+1} \mid \mathcal{F}_n) = 0$ where $\mathcal{F}_n$ is the σ-algebra generated by $(X_1, \cdots, X_n)$,

(ii) **(a)** $\sup_n \mathsf{E}(\|U_{n+1}\|^2) < \infty$ and $\sum_n a_n^2 < +\infty$ or
(b) $\sup_n \|U_{n+1}\| < K$ and $a_n = o(\frac{1}{\log(n)})$.

Remark More general conditions on the characteristics (a_n, U_n) can be found in (BHS, Proposition 1.4).

A typical example is given by equations of the form (2) by letting

$$U_{n+1} = Y_{n+1} - \mathsf{E}(Y_{n+1} \mid \mathcal{F}_n).$$

Given a trajectory $\{X_n(\omega)\}_{n\geq 1}$, its set of accumulation points is denoted by $L(\omega) = L(\{X_n(\omega)\})$. The *limit set* of the process $\{X_n\}$ is the random set $L = L(\{X_n\})$.

The principal properties established in BHS express relations between limit sets of DSA and attracting sets through the following results involving *internally chain transitive* (ICT) sets. (We do not define ICT sets here, see BHS Section 3.3, since we only use the fact that they satisfy Properties 2 and 4 below).

Property 1 *The limit set L of a bounded DSA is almost surely an ICT set.*

This result is in fact stated in BHS for the limit set of the continuous time interpolated process but under our conditions both sets coïncide.

Properties of the limit set L will then be obtained through the next result (BHS, Lemma 3.5, Proposition 3.20 and Theorem 3.23):

Property 2 **(i)** *ICT sets are non–empty, compact, invariant and attractor free.*

(ii) *If A is an attracting set with $B(A) \cap L \neq \emptyset$ and L is ICT, then $L \subset A$.*

Some useful properties of attracting sets or attractors are the two following (BHS, Propositions 3.25 and 3.27).

Property 3 (Strong Lyapounov) *Let $\Lambda \subset \mathbb{R}^m$ be compact with a bounded open neighborhood U and $V : \overline{U} \to [0, \infty[$. Assume the following conditions:*

(i) *U is strongly positively invariant,*

(ii) $V^{-1}(0) = \Lambda$,

(iii) *V is continuous and for all $x \in U \setminus \Lambda, y \in \Phi_t(x)$ and $t > 0$, $V(y) < V(x)$.*

Then Λ contains an attractor whose basin contains U.
The map V is called a strong Lyapounov function associated to Λ.

Let $\Lambda \subset \mathbb{R}^m$ be a set and $U \subset \mathbb{R}^m$ an open neighborhood of Λ. A continuous function $V : U \to \mathbb{R}$ is called a *Lyapunov function* for $\Lambda \subset \mathbb{R}^m$ if $V(y) < V(x)$ for all $x \in U \setminus \Lambda, y \in \Phi_t(x), t > 0$; and $V(y) \leq V(x)$ for all $x \in \Lambda, y \in \Phi_t(x)$ and $t \geq 0$.

Property 4 (Lyapounov) *Suppose V is a Lyapunov function for Λ. Assume that $V(\Lambda)$ has an empty interior. Then every internally chain transitive set $L \subset U$ is contained in Λ and $V|L$ is constant.*

3 Generalized approachability: a potential approach

We follow here the approach of Hart and Mas-Colell (2001a, 2003). Throughout this section C is a closed subset of $\mathbb{R}^m$ and Q is a 'potential function' that attains its minimum on C. Given a correspondence N we consider a dynamical system defined by

$$\dot{\mathbf{w}} \in N(\mathbf{w}) - \mathbf{w}. \tag{4}$$

We provide two sets of conditions on N and Q that imply convergence of the solutions to (4) and of the corresponding DSA to the set C. When applied in the approachability framework (Blackwell, 1956) this will extend Blackwell's property.

Hypothesis 3.1 Q is a $\mathcal{C}^1$ function from $\mathbb{R}^m$ to $\mathbb{R}$ such that

$$Q \geq 0 \text{ and } C = \{Q = 0\}$$

and N is a correspondence satisfying the standard hypothesis 2.1.

3.1 Exponential convergence

Hypothesis 3.2 There exists some positive constant B such that for $w \in \mathbb{R}^m \setminus C$

$$\langle \nabla Q(w), N(w) - w \rangle \leq -B\ Q(w),$$

meaning $\langle \nabla Q(w), w' - w \rangle \leq -B\ Q(w)$ for all $w' \in N(w)$.

Theorem 3.3 *Let $\mathbf{w}(t)$ be a solution of (4). Under hypotheses 3.1 and 3.2, $Q(\mathbf{w}(t))$ goes to zero at exponential rate and the set C is a globally attracting set.*

Proof: If $\mathbf{w}(t) \notin C$

$$\frac{d}{dt} Q(\mathbf{w}(t)) = \langle \nabla Q(\mathbf{w}(t)), \dot{\mathbf{w}}(t) \rangle,$$

hence

$$\frac{d}{dt} Q(\mathbf{w}(t)) \leq -B\ Q(\mathbf{w}(t)),$$

so that

$$Q(\mathbf{w}(t)) \leq Q(\mathbf{w}(0)) e^{-Bt}.$$

This implies that, for any $\varepsilon > 0$, any bounded neighborhood V of C satisfies $\Phi_t(V) \subset C^{\varepsilon}$, for t large enough.
Alternatively, Property 3 applies to the forward image $W = \Phi_{[0,\infty)}(V)$. ∎

Corollary 3.4 *Any bounded DSA of (4) converges a.s. to C.*

Proof: Being a DSA implies Property 1. C is a global attracting set, thus Property 2 applies. Hence the limit set of any DSA is a.s. included in C. ∎

3.2 Application: approachability

Following again Hart and Mas-Colell (2001a) and (2003) and assuming hypothesis 3.2, we show here that the above property extends Blackwell's approachability theory (Blackwell, 1956; Sorin, 2002) in the convex case. (A first approach can be found in BHS, §5.)

Let I and L be two finite sets of moves. Consider a two-person game with vector payoffs described by an $I \times L$ matrix A with entries in $\mathbb{R}^m$. At each stage $n+1$, knowing the previous sequence of moves $h_n = (i_1, \ell_1, ..., i_n, \ell_n)$ player 1 (resp. 2) chooses i_{n+1} in I (resp. ℓ_{n+1} in L). The corresponding stage payoff is $g_{n+1} = A_{i_{n+1}, \ell_{n+1}}$ and $\bar{g}_n = \frac{1}{n}\sum_{m=1}^{n} g_m$ denotes the average of the payoffs until stage n. Let $X = \Delta(I)$ denote the simplex of mixed moves (probabilities on I) and similarly $Y = \Delta(L)$. $\mathcal{H}_n = (I \times L)^n$ denotes the space of all possible sequences of moves up to time n. A *strategy* for player 1 is a map

$$\sigma : \bigcup_n \mathcal{H}_n \to X, \qquad h_n \in \mathcal{H}_n \to \sigma(h_n) = (\sigma_i(h_n))_{i \in I}$$

and similarly $\tau : \bigcup_n \mathcal{H}_n \to Y$ for player 2. A pair of strategies (σ, τ) for the players specifies at each stage $n+1$ the distribution of the current moves given the past according to the formulae:

$$P(i_{n+1} = i, \ell_{n+1} = \ell \mid \mathcal{F}_n)(h_n) = \sigma_i(h_n)\tau_\ell(h_n),$$

where $\mathcal{F}_n$ is the σ-algebra generated by h_n. It then induces a probability on the space of sequences of moves $(I \times L)^{\mathbb{N}}$ denoted $\mathsf{P}_{\sigma,\tau}$.
For x in X we let xA denote the convex hull of the family $\{xA_\ell = \sum_{i \in I} x_i A_{i\ell}; \ell \in L\}$. Finally $d(., C)$ stands for the distance to the closed set C: $d(x, C) = \inf_{y \in C} d(x, y)$.

Definition 3.5 Let N be a correspondence from $\mathbb{R}^m$ to itself. A function $\tilde{x}$ from $\mathbb{R}^m$ to X is N-*adapted* if

$$\tilde{x}(w)A \subset N(w), \quad \forall w \notin C.$$

Theorem 3.6 *Assume hypotheses 3.1, 3.2 and that $\tilde{x}$ is N-adapted. Then any strategy σ of player 1 that satisfies $\sigma(h_n) = \tilde{x}(\bar{g}_n)$ at each stage n, whenever $\bar{g}_n \notin C$, approaches C: explicitly, for any strategy τ of player 2,*

$$d(\bar{g}_n, C) \to 0 \qquad \mathsf{P}_{\sigma,\tau} \ a.s.$$

Proof: The proof proceeds in 2 steps.
First we show that the discrete dynamics associated to the approachability process is a DSA of (4), like in BHS §2 and §5. Then we apply the previous Corollary 3.4. Explicitly, the sequence of outcomes satisfy:

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

By the choice of player 1's strategy $E_{\sigma,\tau}(g_{n+1} \mid \mathcal{F}_n) = \gamma_n$ belongs to $\tilde{x}(\bar{g}_n)A \subset N(\bar{g}_n)$, for any strategy τ of player 2. Hence one writes

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

which shows that $\{\bar{g}_n\}$ is a DSA of (4) (with $a_n = 1/n$ and $Y_{n+1} = g_{n+1} - \bar{g}_n$, so that $E(Y_{n+1} \mid \mathcal{F}_n) \in N(\bar{g}_n) - \bar{g}_n$). Then Corollary 3.4 applies. ■

Remark The fact that $\tilde{x}$ is N-adapted implies that the trajectories of the deterministic continuous time process when player 1 follows $\tilde{x}$ are always feasible under N — while N might be much more regular and easier to study.

Convex case

Assume C convex. Let us show that the above analysis covers Blackwell (1956)'s original framework . Recall that Blackwell's sufficient condition for approachability states that, for any $w \notin C$, there exists $x(w) \in X$ with:

$$\langle w - \Pi_C(w), x(w)A - \Pi_C(w)\rangle \leq 0 \tag{5}$$

where $\Pi_C(w)$ denotes the projection of w on C.
Convexity of C implies the following property:

Lemma 3.7 *Let $Q(w) = \|w - \Pi_C(w)\|_2^2$, then Q is C^1 with $\nabla Q(w) = 2(w - \Pi_C(w))$.*

Proof: We simply write $\|w\|^2$ for the square of the L^2 norm.

$$\begin{aligned} Q(w + w') - Q(w) &= \|w + w' - \Pi_C(w + w')\|^2 - \|w - \Pi_C(w)\|^2 \\ &\leq \|w + w' - \Pi_C(w)\|^2 - \|w - \Pi_C(w)\|^2 \\ &= 2\langle w', w - \Pi_C(w)\rangle + \|w'\|^2. \end{aligned}$$

Similarly

$$\begin{aligned} Q(w+w') - Q(w) &\geq \|w+w' - \Pi_C(w+w')\|^2 - \|w - \Pi_C(w+w')\|^2 \\ &= 2\langle w', w - \Pi_C(w+w')\rangle + \|w'\|^2. \end{aligned}$$

C being convex, Π_C is continuous (1 Lipschitz), hence there exists two constants c_1 and c_2 such that

$$c_1\|w'\|^2 \leq Q(w+w') - Q(w) - 2\langle w', w - \Pi_C(w)\rangle \leq c_2\|w'\|^2.$$

Thus Q is C^1 and $\nabla Q(w) = 2(w - \Pi_C(w))$. ∎

Proposition 3.8 *If player 1 uses a strategy σ which, at each position $\bar{g}_n = w$, induces a mixed move $x(w)$ satisfying Blackwell's condition (5), then approachability holds: for any strategy τ of player 2,*

$$d(\bar{g}_n, C) \to 0 \qquad \mathsf{P}_{\sigma,\tau} \ a.s.$$

Proof: Let $N(w)$ be the intersection of $\mathbf{A}$, the convex hull of the family $\{A_{i\ell}; i \in I, \ell \in L\}$, with the closed half space $\{\theta \in \mathbb{R}^m; \langle w - \Pi_C(w), \theta - \Pi_C(w)\rangle \leq 0\}$. Then N is u.s.c. by continuity of Π_C, and (5) makes x N-adapted. Furthermore, the condition

$$\langle w - \Pi_C(w), N(w) - \Pi_C(w)\rangle \leq 0$$

can be rewritten as

$$\langle w - \Pi_C(w), N(w) - w\rangle \leq -\|w - \Pi_C(w)\|^2$$

which is

$$\langle \frac{1}{2}\nabla Q(w), N(w) - w\rangle \leq -Q(w)$$

with $Q(w) = \|w - \Pi_C(w)\|^2$, by the previous Lemma 3.7. Hence hypotheses 3.1 and 3.2 hold and Theorem 3.6 applies. ∎

Remark

(i) The convexity of C was used to get the property of Π_C, hence of Q ($\mathcal{C}^1$) and of N (u.s.c.).

Define the support function of C on $\mathbb{R}^m$ by:

$$w_C(u) = \sup_{c \in C}\langle u, c\rangle.$$

The previous condition of hypothesis 3.2 holds in particular if Q satisfies:

$$\langle \nabla Q(w), w \rangle - w_C(\nabla Q(w)) \geq B.Q(w), \tag{6}$$

and N fulfills the following inequality:

$$\langle \nabla Q(w), N(w) \rangle \leq w_C(\nabla Q(w)) \qquad \forall w \in \mathbb{R}^m \setminus C \tag{7}$$

which are the original conditions of Hart and Mas-Colell (2001a, p. 34).

(ii) Blackwell (1956) obtains also a speed of convergence of $n^{-1/2}$ for the expectation of the distance: $\rho_n = \mathsf{E}(d(\bar{g}_n, C))$. This corresponds to the exponential decrease $\rho_t^2 = Q(\mathbf{x}(t)) \leq Le^{-t}$ since in the DSA, stage n ends at time $t_n = \sum_{m \leq n} 1/m \sim \log(n)$.

(iii) BHS proves results very similar to Proposition 3.8 (Corollaries 5.1 and 5.2 in BHS) for arbitrary (i.e non necessarily convex) compact sets C but under a stronger separability assumption.

3.3 Slow convergence

We follow again Hart and Mas-Colell (2001a) in considering a hypothesis weaker than 3.2.

Hypothesis 3.9 Q and N satisfy, for $w \in \mathbb{R}^m \setminus C$:

$$\langle \nabla Q(w), N(w) - w \rangle < 0.$$

Remark This is in particular the case if C is convex, inequality (7) holds, and whenever $w \notin C$:

$$\langle \nabla Q(w), w \rangle > w_C(\nabla Q(w)) \tag{8}$$

(A closed half space with exterior normal vector $\nabla Q(w)$ contains C and $N(w)$ but not w, see Hart and Mas-Colell (2001a) p.31).

Theorem 3.10 *Under hypotheses 3.1 and 3.9, Q is a strong Lyapounov function for (4).*

Proof: Using hypothesis 3.9, one obtains if $\mathbf{w}(t) \notin C$:

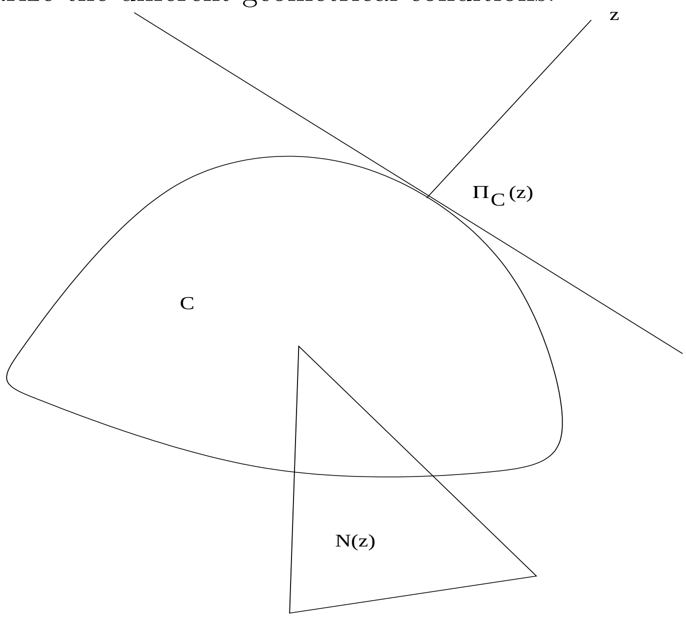
$$\frac{d}{dt} Q(\mathbf{w}(t)) = \langle \nabla Q(\mathbf{w}(t)), \dot{\mathbf{w}}(t) \rangle = \langle \nabla Q(\mathbf{w}(t)), N(\mathbf{w}(t)) - \mathbf{w}(t) \rangle < 0.$$

■

Corollary 3.11 *Assume hypotheses 3.1 and 3.9. Then any bounded DSA of (4) converges a.s. to C.*
Furthermore, theorem 3.6 applies when hypothesis 3.2 is replaced by hypothesis 3.9.

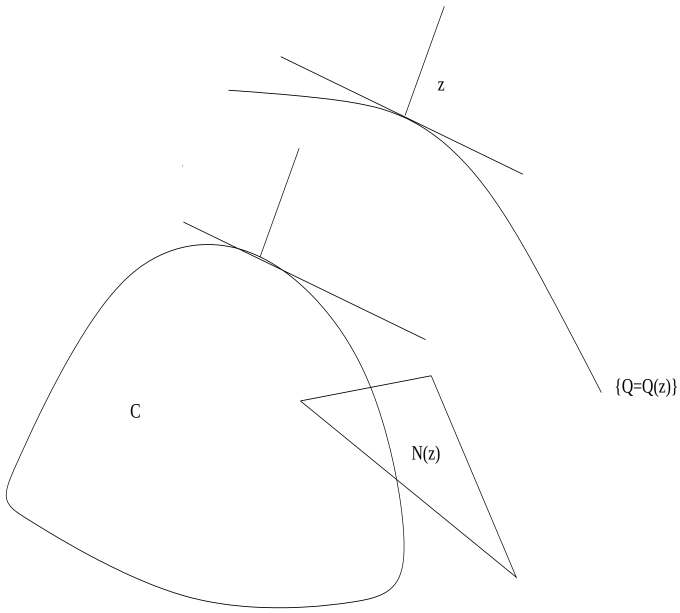
Proof: Follows from Properties 1, 2 and 3. The set C contains a global attractor, hence the limit set of a bounded DSA is included in C. ■

We summarize the different geometrical conditions:



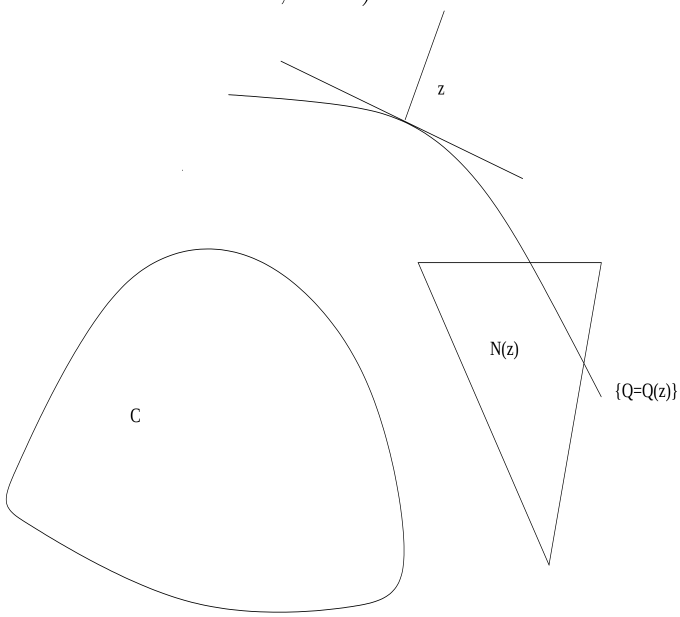
Condition (5)

The hyperplane through $\Pi_C(z)$ orthogonal to $z - \Pi_C(z)$ separates z and $N(z)$ (Blackwell, 1956).



Conditions (7) and (8)

The supporting hyperplane to C with orthogonal direction $\nabla Q(z)$ separates $N(z)$ from z (Hart and Mas-Collel, 2000).



Condition of hypothesis 3.9

$N(z)$ belongs to the interior of the half space defined by the exterior normal vector $\nabla Q(z)$ at z.

4 Approachability and consistency

We consider here a framework where the previous set C is the negative orthant and the vector of payoffs describes the vector of regrets in a strategic game, see Hart and Mas-Colell (2001a), (2003). The consistency condition amounts to the convergence of the average regrets to C. The interest of the approach is that the same function P will be used to play the rôle of the function Q on one hand and to define the strategy and hence the correspondence N on the other hand. Also the procedure can be defined on the payoff space as well as on the set of correlated moves.

4.1 No regret and correlated moves

Consider a finite game in strategic form. There are finitely many players labeled $a = 1, 2, \ldots, A$. We let S^a denote the finite moves set of player a, $S = \prod_a S^a$, and $Z = \Delta(S)$ the set of probabilities on S (correlated moves).
Since we will consider everything from the view point of player 1 it is convenient to set $S^1 = I, X = \Delta(I)$ (mixed moves of player 1), $L = \prod_{a \neq 1} S^a$, and $Y = \Delta(L)$ (correlated mixed moves of player 1's opponents) hence $Z = \Delta(I \times L)$.
Throughout, $X \times Y$ is identified with a subset of Z through the natural embedding $(x, y) \to x \times y$, where $x \times y$ stands for the product probability of x and y. As usual, I (L, S) is also identified with a subset of X (Y, Z) through the embedding $k \to \delta_k$. We let $U : S \to \mathbb{R}$ denote the payoff function of player 1 and we still denote by U its linear extension to Z, and its bilinear extension to $X \times Y$.
Let m be the cardinality of I and $R(z)$ denote the m-dimensional vector of regrets for player 1 at z in Z, defined by

$$R(z) = \{U(i, z^{-1}) - U(z)\}_{i \in I}$$

where z^{-1} stands for the marginal of z on L. (Player 1 compares his payoff using a given move i to his actual payoff, assuming the other players' behavior, z^{-1}, given.)
Let $D = \mathbb{R}^m_-$ be the closed negative orthant associated to the set of moves of player 1.

Definition 4.1 *H (for Hannan's set) is the set of probabilities in Z satisfying the no-regret condition for player 1. Formally:*

$$H = \{z \in Z : U(i, z^{-1}) \leq U(z), \forall i \in I\} = \{z \in Z : R(z) \in D\}.$$

Definition 4.2 P is a *potential function* for D if it satisfies the following set of conditions:

(i) P is a $\mathcal{C}^1$ nonnegative function from $\mathbb{R}^m$ to $\mathbb{R}$,

(ii) $P(w) = 0$ iff $w \in D$,

(iii) $\nabla P(w) \geq 0$,

(iv) $\langle \nabla P(w), w \rangle > 0, \forall w \notin D$.

Definition 4.3 Given a potential P for D, the *P-regret-based dynamics* for player 1 is defined on Z by

$$\dot{\mathbf{z}} \in N(\mathbf{z}) - \mathbf{z} \tag{9}$$

where

(i) $N(z) = \varphi(R(z)) \times Y \subset Z$, with

(ii) $\varphi(w) = \frac{\nabla P(w)}{|\nabla P(w)|} \in X$ whenever $w \notin D$ and $\varphi(w) = X$ otherwise.

Here $|\nabla P(w)|$ stands for the L^1 norm of $\nabla P(w)$.

Remark This corresponds to a process where only the behavior of player 1, outside of H, is specified. Note that even the dynamics is truly independent among the players ("uncoupled" according to Hart and Mas Collel, see Hart (2005)) the natural state space is the set of correlated moves (and not the product of the sets of mixed moves) since the criteria involves the actual payoffs and not only the marginal empirical frequencies.

The associated discrete process is as follows. Let $s_n \in S$ be the random variable of profile of actions at stage n, and $\mathcal{F}_n$ the σ-algebra generated by the history $h_n = (s_1, \ldots, s_n)$. The average $\bar{z}_n = \frac{1}{n}\sum_{m=1}^{n} s_m$ satisfies:

$$\bar{z}_{n+1} - \bar{z}_n = \frac{1}{n+1}[s_{n+1} - \bar{z}_n]. \tag{10}$$

Definition 4.4 A *P-regret-based strategy* for player 1 is specified by the conditions:

(i) For all $(i, \ell) \in I \times L$

$$\mathsf{P}(i_{n+1} = i, \ell_{n+1} = \ell | \mathcal{F}_n) = \mathsf{P}(i_{n+1} = i | \mathcal{F}_n) \mathsf{P}(\ell_{n+1} = \ell | \mathcal{F}_n),$$

(ii) $\mathsf{P}(i_{n+1} = i | \mathcal{F}_n) = \varphi_i(R(\bar{z}_n))$ whenever $R(\bar{z}_n) \notin D$, where $\varphi(\cdot) = \{\varphi_i(\cdot)\}_{i \in I}$ is like in definition 4.3.

The corresponding discrete time process (10) is called a *P-regret-based discrete dynamics.*

Clearly, one has

Proposition 4.5 *The P-regret-based discrete dynamics (10) is a DSA of (9).*

The next result is obvious but crucial.

Lemma 4.6 *Let $z = x \times y \in X \times Y \subset Z$, then*

$$\langle x, R(z) \rangle = 0.$$

Proof:
One has

$$\sum_{i \in I} x_i [U(i, y) - U(x \times y)] = 0.$$

■

4.2 Blackwell's framework

Given $w \in \mathbb{R}^m$, let w^+ be the vector with components $w_k^+ = \max(w_k, 0)$. Define $Q(w) = \sum_k (w_k^+)^2$. Note that $\nabla Q(w) = 2w^+$, hence Q satisfies the conditions $(i) - (iv)$ of definition 4.2. If Π denotes the projection on D one has $w - \Pi(w) = w^+$ and $\langle w^+, \Pi(w) \rangle = 0$.
In the game with vector payoff given by the regret of player 1, the set of feasible expected payoffs corresponding to xA (cf. §3.2), when player 1 uses θ, is $\{R(z); z = \theta \times z^{-1}\}$. Assume that player 1 uses a Q-regret-based strategy. Since at $w = \bar{g}_n$, $\theta(w)$ is proportional to $\nabla Q(w)$, hence to w^+, Lemma 4.6 implies that condition (5): $\langle w - \Pi w, xA - \Pi w \rangle \leq 0$ is satisfied; in fact, this quantity reduces to: $\langle w^+, R(y) - \Pi w \rangle$ which equals 0.
Hence a Q-regret-based strategy approaches the orthant D.

4.3 Convergence of P-regret-based dynamics

The previous dynamics in Section 3 were defined on the payoff space. Here, we take the image by R (which is linear), of the dynamical system (9) and obtain the following differential inclusion in $\mathbb{R}^m$:

$$\dot{\mathbf{w}} \in \hat{N}(\mathbf{w}) - \mathbf{w} \tag{11}$$

where

$$\hat{N}(w) = R(\varphi(w) \times Y).$$

The associated discrete dynamics to (10) is given as

$$\bar{w}_{n+1} - \bar{w}_n = \frac{1}{n+1}(w_{n+1} - \bar{w}_n) \tag{12}$$

with $w_n = R(z_n)$.

Theorem 4.7 *The potential P is a strong Lyapounov function associated to the set D for (11), and similarly $P \circ R$ to the set H for (9). Hence, D contains an attractor for (11) and H contains an attractor for (9).*

Proof: Remark that $\langle \nabla P(w), \hat{N}(w) \rangle = 0$: in fact $\nabla P(w) = 0$ for $w \in D$, and for $w \notin D$ use Lemma 4.6. Hence for any $\mathbf{w}(t)$ solution to (11)

$$\frac{d}{dt} P(\mathbf{w}(t)) = \langle \nabla P(\mathbf{w}(t)), \dot{\mathbf{w}}(t) \rangle = -\langle \nabla P(\mathbf{w}(t)), \mathbf{w}(t) \rangle < 0$$

and P is a strong Lyapounov function associated to D, in view of conditions $(i) - (iv)$ of definition 4.2. The last assertion follows from Property 3. ∎

Corollary 4.8 *Any P-regret-based discrete dynamics (10) approaches D in the payoff space, hence H in the action space.*

Proof: D (resp. H) contains an attractor for (11) whose basin of attraction contains $R(Z)$ (resp. Z) and the process (12) (resp. (10)) is a bounded DSA, hence Properties 1, 2 and 3 apply. ∎

Remark A direct proof is available as follows :
Let $\mathbf{R}$ the range of R and define, for $w \notin D$,

$$N(w) = \{w' \in \mathbb{R}^m; \langle w', \nabla P(w) \rangle = 0\} \cap ?\mathbf{R}.$$

Hypotheses (3.1) and 3.9 are satisfied and Corollary 3.11 applies.

5 Approachability and conditional consistency

We keep the framework of Section 4 and the notation introduced in 4.1, and follow Hart and Mas-Colell (2000), (2001a), (2003) in studying conditional (or internal) regrets. One constructs again an approachability strategy from an associate potential function P. Like in the previous Section 4 the dynamics can be defined either in the payoff space or in the space of correlated moves.

We still consider only player 1 and denote by U his payoff.
Given $z = (z_s)_{s \in S} \in Z$, introduce the family of m comparison vectors of dimension m (testing k against j with $(j,k) \in I^2$)

$$C(j,k)(z) = \sum_{\ell \in L} [U(k,\ell) - U(j,\ell)] z_{(j,\ell)}.$$

(This corresponds to the change in the expected gain of Player 1 at z when replacing move j by k.) Remark that if one let $(z \mid j)$ denote the conditional probability on L induced by z given $j \in I$ and z^1 the marginal on I, then

$$\{C(j,k)(z)\}_{k \in I} = z^1_j R((z \mid j))$$

where we recall that $R((z \mid j))$ is the vector of regrets for player 1 at $(z \mid j)$.

Definition 5.1 The set of no conditional regret (for player 1) is

$$C^1 = \{z; C(j,k)(z) \leq 0, \forall j,k \in I\}.$$

It is obviously a subset of H since

$$\sum_j \{C(j,k)(z)\}_{k \in I} = R(z).$$

Property The intersection over all players a of the sets C^a is the set of correlated equilibria of the game.

5.1 Discrete standard case

Here we will use approachability theory to retrieve the well known fact (see Hart and Mas-Colell (2000)) that player 1 has a strategy such that the vector $C(\bar{z}_n)$ converges to the negative orthant of $\mathbb{R}^{m^2}$, where $\bar{z}_n \in Z$ is the average (correlated) distribution on S.
Given $s \in S$ define the auxiliary "vector payoff" $B(s)$ to be the $m \times m$ real valued matrix where, if $s = (j,\ell) \in I \times L$, hence j is the move of player 1, the

only non-zero line is line j with entry on column k being $U(k,\ell) - U(j,\ell)$. The average payoff at stage n is thus a matrix B_n with coefficient

$$B_n(j,k) = \frac{1}{n} \sum_{r, i_r = j} (U(k,\ell_r) - U(j,\ell_r)) = C(j,k)(\bar{z}_n)$$

which is the test of k versus j on the dates, up to stage n, where j was played. Consider the Markov chain on I with transition matrix

$$M_n(j,k) = \frac{B_n(j,k)^+}{b_n},$$

for $j \neq k$ where $b_n > \max_j \sum_k B_n(j,k)^+$. By standard results on finite Markov chains, M_n admits (at least) one invariant probability measure. Let $\mu_n = \mu(B_n)$ be such a measure. Then (dropping the subscript n)

$$\mu_j = \sum_k \mu_k M(k,j) = \sum_{k \neq j} \mu_k \frac{B(k,j)^+}{b} + \mu_j (1 - \sum_{k \neq j} \frac{B(j,k)^+}{b}).$$

Thus b disappears and the condition writes

$$\sum_{k \neq j} \mu_k B(k,j)^+ = \mu_j \sum_{k \neq j} B(j,k)^+.$$

Theorem 5.2 *Any strategy of player 1 satisfying $\sigma(h_n) = \mu_n$ is an approachability strategy for the negative orthant of $\mathbb{R}^{m^2}$. Namely*

$$\forall j,k \qquad \lim_{n \to \infty} B_n(j,k)^+ = 0 \qquad a.s.$$

Equivalently, $(\bar{z}_n)$ approaches the set of no conditional regret for player 1 :

$$\lim_{n \to \infty} d(\bar{z}_n, C^1) = 0.$$

Proof: Let Ω denote the closed negative orthant of $\mathbb{R}^{m^2}$. In view of proposition 3.8 it is enough to prove that inequality (5)

$$\langle b - \Pi_\Omega(b), b' - \Pi_\Omega(b) \rangle \leq 0, \quad \forall b \notin \Omega$$

holds for every regret matrix b', feasible under $\mu = \mu(b)$.
As usual, since the projection is on the negative orthant Ω, $b - \Pi_\Omega(b) = b^+$ and $\langle b - \Pi_\Omega(b), \Pi_\Omega(b) \rangle = 0$. Hence it remains to evaluate

$$\sum_{j,k} B(j,k)^+ \mu_j [U(k,\ell) - U(j,\ell)]$$

but the coefficient of $U(j,\ell)$ is precisely

$$\sum_k B^+(k,j) \mu_k - \mu_j \sum_k B^+(j,k) = 0$$

by the choice of $\mu = \mu(b)$. ∎

5.2 Continuous general case

We first state a general property (compare Lemma 4.6):

Lemma 5.3 *Given* $a \in \mathbb{R}^{m^2}$, *let* $\mu \in X$ *satisfy :*

$$\sum_{k:\, k\neq j} \mu_k a(k,j) = \mu_j \sum_{k:\, k\neq j} a(j,k), \forall j \in I$$

then

$$\langle a, C(\mu \times y)\rangle = 0, \quad \forall y \in Y.$$

Proof: As above one computes:

$$\sum_j \sum_k a(j,k)\mu_j[U(k,y) - U(j,y)]$$

but the coefficient of $U(j,y)$ is precisely

$$\sum_k a(k,j)\mu_k - \mu_j \sum_k a(j,k) = 0.$$

■

Let P be a potential function for Ω the negative orthant of $\mathbb{R}^{m^2}$, for example $P(w) = \sum_{ij}(w_{ij}^+)^2$, as in the standard case above.

Definition 5.4 The *P-conditional regret dynamics* in continuous time is defined on Z by:

$$\dot{\mathbf{z}} \in \mu(\mathbf{z}) \times Y - \mathbf{z} \tag{13}$$

where $\mu(z)$ is the set of $\mu \in X$ that are solution to:

$$\sum\nolimits_{k\in S} \mu_k \nabla P_{kj}(C(z)) = \mu_j \sum\nolimits_k \nabla P_{jk}(C(z))$$

whenever $C(z) \notin \Omega$ (∇P_{jk} denotes the jk component of the gradient of P). In particular $\mu(z) = X$ whenever $C(z) \in \Omega$.

The associated process in $\mathbb{R}^{m^2}$ is the image under C:

$$\dot{\mathbf{w}} \in C(\nu(\mathbf{w}) \times Y) - \mathbf{w} \tag{14}$$

where $\nu(w)$ is the set of $\nu \in X$ with

$$\sum\nolimits_{k\in S} \nu_k \nabla P_{kj}(w) = \nu_j \sum\nolimits_k \nabla P_{jk}(w).$$

Theorem 5.5 *The processes (13) and (14) satisfy:*

$$C^{+}(j,k)(\mathbf{z}(t)) = \mathbf{w}^{+}_{jk}(t) \to_{t\to\infty} 0.$$

Proof: Apply Theorem 3.10 with:

$$N(w) = \{w' \in (\mathbb{R}^m)^2 : \langle \nabla P(w), w' \rangle = 0\} \cap \mathbf{C}$$

where $\mathbf{C}$ is the range $C(Z)$ of C. Since $\mathbf{w}(t) = C(\mathbf{z}(t))$ the previous lemma 5.3 implies that $\dot{\mathbf{w}}(t) \in N(\mathbf{w}(t)) - \mathbf{w}(t)$. ∎

The discrete processes corresponding to (13) and (14) are respectively in Z

$$\bar{z}_{n+1} - \bar{z}_n = \frac{1}{n+1}[\mu_{n+1} \times z^{-1}_{n+1} - \bar{z}_n + (z_{n+1} - \mu_{n+1} \times z^{-1}_{n+1})] \tag{15}$$

where μ_{n+1} satisfies:

$$\sum\nolimits_{k\in S} \mu^k_{n+1} \nabla P_{kj}(C(\bar{z}_n)) = \mu^j_{n+1} \sum\nolimits_k \nabla P_{jk}(C(\bar{z}_n))$$

and in $\mathbb{R}^{m^2}$

$$\bar{w}_{n+1} - \bar{w}_n = \frac{1}{n+1}[C(\mu_{n+1} \times z^{-1}_{n+1}) - \bar{w}_n + (w_{n+1} - C(\mu_{n+1} \times z^{-1}_{n+1})]. \tag{16}$$

Corollary 5.6 *The discrete processes (15) and (16) satisfy:*

$$C^{+}(j,k)(\bar{z}_n) = \bar{w}^{jk,+}_n \to_{t\to\infty} 0 \qquad a.s.$$

Proof: (15) and (16) are bounded DSA of (13) and (14) and Properties 1, 2, and 3 apply. ∎

Corollary 5.7 *If all players follow the above procedure, the empirical distribution of moves converges a.s. to the set of correlated equilibria.*

6 Smooth fictitious play and consistency

We follow the approach of Fudenberg and Levine concerning consistency (1995) and conditional consistency (1999) and deduce some of their main results (see Theorems 6.6, 6.12 below) as corollaries of dynamical properties. Basically the criteria are similar to the ones studied in Section 4 and 5 but the procedure is different and based only on the previous behavior of the opponents. Like in sections 4 and 5 we continue to adopt the point of view of player 1.

6.1 Consistency

Let

$$V(y) = \max_{x \in X} U(x, y).$$

The *average regret evaluation* along $h_n \in \mathcal{H}_n$ is

$$e(h_n) = e_n = V(\overline{y}_n) - \frac{1}{n}\sum\nolimits_{m=1}^{n} U(i_m, \ell_m).$$

where as usual $\bar{y}_n$ stands for the time average of (ℓ_m) up to time n. (This corresponds to the maximal component of the regret vector $R(\bar{z}_n)$).

Definition 6.1 *(Fudenberg and Levine, 1995)* Let $\eta > 0$. A strategy σ for player 1 is said η-consistent if for any opponents strategy τ

$$\limsup_{n \to \infty} e_n \leq \eta \qquad \mathsf{P}_{\sigma,\tau} \text{ a.s.}$$

6.2 Smooth fictitious play

A *smooth perturbation* of the payoff U is a map

$$U^\varepsilon(x, y) = U(x, y) + \varepsilon\rho(x),\ 0 < \varepsilon < \varepsilon_0$$

such that:

(i) $\rho : X \to \mathbb{R}$ is a $\mathcal{C}^1$ function with $\|\rho\| \leq 1$,

(ii) $\mathsf{argmax}_{x \in X} U^\varepsilon(., y)$ reduces to one point and defines a continuous map

$$\mathbf{br}^\varepsilon : Y \to X$$

called a *smooth best reply function*,

(iii) $D_1 U^\varepsilon(\mathbf{br}^\varepsilon(y), y).D\mathbf{br}^\varepsilon(y) = 0$ (for example $D_1 U^\varepsilon(., y)$ is 0 at $\mathbf{br}^\varepsilon(y)$. This occurs in particular if $\mathbf{br}^\varepsilon(y)$ belongs to the interior of X).

Remark A typical example is

$$\rho(x) = -\sum_k x_k \log x_k. \tag{17}$$

which leads to

$$\mathbf{br}_i^\varepsilon(y) = \frac{\exp(U(i, y)/\varepsilon)}{\sum_{k \in I} \exp(U(k, y)/\varepsilon)} \tag{18}$$

as shown by Fudenberg and Levine (1995, 1999).

Let

$$V^\varepsilon(y) = \max_x U^\varepsilon(x, y) = U^\varepsilon(\mathbf{br}^\varepsilon(y), y).$$

Lemma 6.2 *(Fudenberg and Levine (1999))*

$$DV^\varepsilon(y)(h) = U(\mathbf{br}^\varepsilon(y), h).$$

Proof: One has

$$DV^\varepsilon(y) = D_1U^\varepsilon(\mathbf{br}^\varepsilon(y), y).D\mathbf{br}^\varepsilon(y) + D_2U^\varepsilon(\mathbf{br}^\varepsilon(y), y)$$

The first term is zero by condition (iii) above. For the second term one has

$$D_2U^\varepsilon(\mathbf{br}^\varepsilon(y), y) = D_2U(\mathbf{br}^\varepsilon(y), y)$$

which, by linearity of $U(x, .)$ gives the result. ∎

Definition 6.3 A smooth fictitious play strategy for player 1 associated to the smooth best response function $\mathbf{br}^\varepsilon$ (in short a $\mathsf{SFP}(\varepsilon)$ strategy) is a strategy σ^ε such that

$$\mathsf{E}_{\sigma^\varepsilon,\tau}(i_{n+1} \mid \mathcal{F}_n) = \mathbf{br}^\varepsilon(\bar{y}_n)$$

for any τ.

There are two classical interpretations of $\mathsf{SFP}(\varepsilon)$ strategies. One is that player 1 chooses to randomize his moves. Another one called *stochastic fictitious play* (Fudenberg and Levine (1998), Benaïm and Hirsch (1999)) is that payoffs are perturbed in each period by random shocks and that player 1 plays the best reply to the empirical mixed strategy of its opponents. Under mild assumptions on the distribution of the shocks it was shown by Hofbauer and Sandholm (2002) (Theorem 2.1) that this can always be seen as a $\mathsf{SFP}(\varepsilon)$ strategy for a suitable ρ.

6.3 SFP and consistency

Fictitious play was initially used as a global dynamics (i.e. the behavior of each player is specified) to prove convergence of the empirical strategies to optimal strategies (see Brown (1951) and Robinson (1951) and for recent results BHS Section 5.3 and Hofbauer and Sorin (2006)).
Here we deal with unilateral dynamics and consider the consistency property. Hence the state space cannot be reduced to the product of the sets of mixed

moves but has to incorporate the payoffs.
Explicitly, the discrete dynamics of averaged moves is

$$\overline{x}_{n+1} - \overline{x}_n = \frac{1}{n+1}[i_{n+1} - \overline{x}_n], \qquad \overline{y}_{n+1} - \overline{y}_n = \frac{1}{n+1}[\ell_{n+1} - \overline{y}_n]. \tag{19}$$

Let $u_n = U(i_n, \ell_n)$ be the payoff at stage n and $\overline{u}_n$ be the average payoff up to stage n so that

$$\overline{u}_{n+1} - \overline{u}_n = \frac{1}{n+1}[u_{n+1} - \overline{u}_n]. \tag{20}$$

Lemma 6.4 *Assume that player 1 plays a* $\mathsf{SFP}(\varepsilon)$ *strategy. Then the process* $(\bar{x}_n, \bar{y}_n, \bar{u}_n)$ *is a DSA of the differential inclusion*

$$\dot{\omega} \in N(\omega) - \omega \tag{21}$$

where $\omega = (x, y, u) \in X \times Y \times \mathbb{R}$ *and*

$$N(x, y, u) = \{(\mathbf{br}^\varepsilon(y), \beta, U(\mathbf{br}^\varepsilon(y), \beta)) : \beta \in Y\}.$$

Proof: To shorten notation we write $\mathsf{E}(. \mid \mathcal{F}_n)$ for $\mathsf{E}_{\sigma^\varepsilon,\tau}(. \mid \mathcal{F}_n)$ where τ is any opponents strategy. By assumption $\mathsf{E}(i_{n+1} \mid \mathcal{F}_n) = \mathbf{br}^\varepsilon(\bar{y}_n)$. Set $\mathsf{E}(\ell_{n+1} \mid \mathcal{F}_n) = \beta_n \in Y$. Then, by conditional independence of i_{n+1} and ℓ_{n+1}, one gets that $\mathsf{E}(u_{n+1} \mid \mathcal{F}_n) = U(\mathbf{br}^\varepsilon(\bar{y}_n), \beta_n)$. Hence $\mathsf{E}((i_{n+1}, \ell_{n+1}, u_{n+1}) \mid \mathcal{F}_n) \in N(x_n, y_n, u_n)$. ∎

Theorem 6.5 *The set* $\{(x, y, u) \in X \times Y \times \mathbb{R} : V^\varepsilon(y) - u \leq \varepsilon\}$ *is a global attracting set for (21). In particular, for any* $\eta > 0$*, there exists* $\bar{\varepsilon}$ *such that for* $\varepsilon \leq \bar{\varepsilon}$*,* $\limsup_{t\to\infty} V^\varepsilon(\mathbf{y}(t)) - \mathbf{u}(t) \leq \eta$ *(i.e. continuous* $\mathsf{SFP}(\varepsilon)$ *satisfies* η*-consistency.)*

Proof: Let $\mathbf{w}^\varepsilon(t) = V^\varepsilon(\mathbf{y}(t)) - \mathbf{u}(t)$. Taking time derivative one obtains, using Lemma 6.2 and (21):

$$\begin{aligned}
\dot{\mathbf{w}}^\varepsilon(t) &= DV^\varepsilon(\mathbf{y}(t)).\dot{\mathbf{y}}(t) - \dot{\mathbf{u}}(t) \\
&= U(\mathbf{br}^\varepsilon(\mathbf{y}(t)), \beta(t)) - U(\mathbf{br}^\varepsilon(\mathbf{y}(t)), \mathbf{y}(t)) - U(\mathbf{br}^\varepsilon(\mathbf{y}(t)), \beta(t)) + \mathbf{u}(t) \\
&= \mathbf{u}(t) - U(\mathbf{br}^\varepsilon(\mathbf{y}(t)), \mathbf{y}(t)) \\
&= -\mathbf{w}^\varepsilon(t) + \varepsilon\rho(\sigma^\varepsilon(\mathbf{y}(t))).
\end{aligned}$$

Hence

$$\dot{w}^\varepsilon(t) + w^\varepsilon(t) \leq \varepsilon$$

so that $w^\varepsilon(t) \leq \varepsilon ? + Ke^{-t}$ for some constant K and the result follows. ∎

Theorem 6.6 *For any $\eta > 0$, there exists $\bar{\varepsilon}$ such that for $\varepsilon \leq \bar{\varepsilon}$, $\mathsf{SFP}(\varepsilon)$ is η-consistent.*

Proof: The assertion follows from lemma 6.4, Property 1, Property 2 (*ii*) and Theorem 6.5.

■

6.4 Remarks and Generalizations

The definition given here of a $\mathsf{SFP}(\varepsilon)$ strategy can be extended in some interesting directions. Rather than developing a general theory we focus on two particular examples.

1. Strategies based on pairwise comparison of payoffs: Suppose that ρ is given by (17). Then, playing a $\mathsf{SFP}(\varepsilon)$ strategy requires for player 1 the computation of $\mathbf{br}^{\varepsilon}(\bar{y}_n)$ given by (18) at each stage. In case where the cardinality of S^1 is very large (say 2^N with $N \geq 10$) this computation is not feasible! An alternative feasible strategy is the following:
Assume that I is the set of vertices of a graph. Write $i \sim j$ when i and j are neigbhours in this graph. Assume furthermore that the graph is *symmetric* ($\sim$ is a symmetric relation) and *connected* (given any two points $i, j \in I$ there exists a finite sequence $i_1 = i, i_2, \ldots, i_m = j$ such that $i_l \sim i_{l+1}$ for $l = 1, \ldots, m-1$). Let $N(i) = \{j \in I \setminus \{i\} : \ i \sim j\}$. The strategy is as follows: Let i be the action chosen at time n (i.e. $i_n = i$). At time $n+1$, player 1 picks an action j at random in $N(i)$. He then switches to j (i.e. $i_{n+1} = j$) with probability

$$R(i, j, \bar{y}_n) = \min\left[1, \frac{|N(i)|}{|N(j)|} \exp\left(\frac{1}{\varepsilon}\left(U(j, \bar{y}_n) - U(i, \bar{y}_n)\right)\right)\right]$$

and keeps i (i.e $i_{n+1} = i$) with the complementary probability $1 - R(i, j, \bar{y}_n)$. Here $|N(i)|$ stands for the cardinal of $N(i)$.
Note that this strategy only involves at each step the computation of the payoffs difference $(U(j, \bar{y}_n) - U(i, \bar{y}_n))$. While this strategy is not an $\mathsf{SFP}(\varepsilon)$ strategy, one still has:

Theorem 6.7 *For any $\eta > 0$, there exists $\bar{\varepsilon}$ such that, for $\varepsilon \leq \bar{\epsilon}$, the strategy described above is η-consistent.*

Proof: For fixed $y \in Y$, let $Q(y)$ be the Markov transition matrix given by $Q(i, j, y) = \frac{1}{|N(i)|} R(i, j, y)$ for $j \in N(i)$, $Q(i, j, y) = 0$ for $j \notin N(i) \cup \{i\}$, and $Q(i, i, y)) = 1 - \sum_{j \neq i} Q(i, j, y)$. Then $Q(y)$ is an irreducible Markov matrix

having $\mathbf{br}^{\varepsilon}(y)$ as unique invariant probability: this is easily seen by checking that $Q(y)$ is reversible with respect to $\mathbf{br}^{\varepsilon}(y)$. That is $\mathbf{br}_i^{\varepsilon}(y)Q(i,j,y) = \mathbf{br}_j^{\varepsilon}(y)Q(j,i,y)$.
The discrete time process (19), (20) is not a DSA (as defined here) to (21) because $\mathsf{E}(i_{n+1} \mid \mathcal{F}_n) \neq \mathbf{br}^{\varepsilon}(\bar{y}_n)$. However, the conditional law of i_{n+1} given $\mathcal{F}_n$ is $Q(x_n, \cdot, \bar{y}_n)$ and using the techniques introduced by Métivier and Priouret (1992) to deal with Markovian perturbations (see e.g. Duflo 1996, Chapter 3.IV) it can still be proved that the assumptions of Proposition 1.3 in BHS are fulfilled, from which it follows that the interpolated affine process associated to (19), (20) is a perturbed solution (see BHS, for a precise definition) to (21). Hence Property 1 applies and the end of the proof is similar to the proof of Theorem 6.6. ■

2. Convex sets of actions: Suppose that X and Y are two convex compact subsets of finite dimensional Euclidean spaces. U is a bounded function with $U(x,.)$ linear on Y. The discrete dynamics of averaged moves is

$$\overline{x}_{n+1} - \overline{x}_n = \frac{1}{n+1}[x_{n+1} - \overline{x}_n], \qquad \overline{y}_{n+1} - \overline{y}_n = \frac{1}{n+1}[y_{n+1} - \overline{y}_n]. \tag{22}$$

with $x_{n+1} = \mathbf{br}^{\varepsilon}(\bar{y}_n)$. Let $u_n = U(x_n, y_n)$ be the payoff at stage n and $\overline{u}_n$ be the average payoff up to stage n so that

$$\overline{u}_{n+1} - \overline{u}_n = \frac{1}{n+1}[u_{n+1} - \overline{u}_n]. \tag{23}$$

Then the results of the previous section 6.3 still hold.

6.5 SFP and conditional consistency

We keep here the framework of Section 4 but extend the analysis from consistency to conditional consistency (which is like studying external regrets (Section 4) and then internal regrets (Section 5)). Given $z \in Z$, recall that we let $z^1 \in X$ denote the marginal of z on I. That is

$$z^1 = (z_i^1)_{i \in I} \text{ with } z_i^1 = \sum_{\ell \in L} z_{i\ell}.$$

Let $z[i] \in \mathbb{R}^L$ be the vector with components $z[i]_\ell = z_{i\ell}$. Note that $z[i]$ belongs to tY for some $0 \leq t \leq 1$. A conditional probability on L induced by z given $i \in I$ satisfies

$$z \mid i = (z \mid i)_{\ell \in L} \text{ with } (z \mid i)_\ell z_i^1 = z_{i\ell} = z[i]_\ell.$$

Let $[0,1].Y = \{ty : \ 0 \le t \le 1, \ y \in Y\}$. Extend U to $X \times ([0,1] \times Y)$ by $U(x, ty) = tU(x, y)$ and similarly for V. The conditional evaluation function at $z \in Z$ is

$$ce(z) = \sum_{i \in I} V(z[i]) - U(i, z[i]) = \sum_{i \in I} z_i^1 [V(z \mid i) - U(i, z \mid i)] = \sum_{i \in I} z_i^1 V(z|i) - U(z).$$

with the convention that $z_i^1 V(z \mid i) = z_i^1 U(i, z \mid i) = 0$ when $z_i^1 = 0$.

Like in Section 5, conditional consistency means consistency with respect to the conditional distribution given each event of the form "i was played". In a discrete framework the conditional evaluation is thus

$$ce_n = ce(\bar{z}_n)$$

where as usual $\bar{z}_n$ stands for the empirical correlated distribution of moves up to stage n.

Conditional consistency is defined like consistency but with respect to (ce_n). More precisely:

Definition 6.8 A strategy σ for player 1 is said η-conditionally consistent if for any opponents strategy τ

$$\limsup_{n \to \infty} ce_n \le \eta \qquad \mathsf{P}_{\sigma,\tau} \ a.s.$$

Given a smooth best reply function $\mathbf{br}^\varepsilon : Y \to X$, let us introduce a correspondence $\mathbf{Br}^\varepsilon$ defined on $[0,1] \times Y$ by $\mathbf{Br}^\varepsilon(ty)= \mathbf{br}^\varepsilon(y)$ for $0 < t \le 1$ and $\mathbf{Br}^\varepsilon(0) = X$. For $z \in Z$, let $\mu^\varepsilon(z) \subset X$ denote the set of all $\mu \in X$ that are solution to the equation

$$\sum_{i \in I} \mu_i b^i = \mu \tag{24}$$

for some vectors family $\{b^i\}_{i \in I}$ such that $b^i \in \mathbf{Br}^\varepsilon(z[i])$.

Lemma 6.9 *μ^ε is an u.s.c correspondence with compact convex non–empty values.*

Proof: For any vectors family $\{b^i\}_{i \in I}$ with $b^i \in X$ the function $\mu \to \sum_{i \in I} \mu_i b^i$ maps continuously X into itself. It then has fixed points by Brouwer's fixed point theorem, showing that $\mu^\varepsilon(z) \neq \emptyset$. Let $\mu, \nu \in \mu^\varepsilon(z)$. That is $\mu = \sum_i \mu_i b^i$ and $\nu = \sum_i \nu^i c^i$ with $b^i, c^i \in \mathbf{Br}^\varepsilon(z[i])$. Then for any $0 \le t \le 1$ $t\mu + (1-t)\nu = \sum_i (t\mu_i + (1-t)\nu_i) d^i$ with $d^i = \frac{t\mu_i b^i + (1-t)\nu_i c^i}{(t\mu_i + (1-t)\nu_i)}$. By convexity of $\mathbf{Br}^\varepsilon(z[i]), d^i \in \mathbf{Br}^\varepsilon(z[i])$. Thus $t\mu + (1-t)\nu \in \mu^\varepsilon(z)$ proving convexity of $\mu^\varepsilon(z)$.Using the fact that $\mathbf{Br}^\varepsilon$ has a closed graph, it is easy to show that μ^ε has a closed graph, from which it will follow that it is u.s.c with compact values. Details are left to the reader.

∎

Definition 6.10 A conditional smooth fictitious play strategy for player 1 associated to the smooth best response function $\mathbf{br}^\varepsilon$ (in short a $\mathsf{CSFP}(\varepsilon)$ strategy) is a strategy σ^ε such that $\sigma^\varepsilon(h_n) \in \mu^\varepsilon(\bar{z}_n)$.

The random discrete process associated to $\mathsf{CSFP}(\varepsilon)$ is thus defined by:

$$\overline{z}_{n+1} - \overline{z}_n = \frac{1}{n+1}[z_{n+1} - \overline{z}_n] \tag{25}$$

where the conditional law of $z_{n+1} = (i_{n+1}, \ell_{n+1})$ given the past up to time n is a product law $\sigma^\varepsilon(h_n) \times \tau(h_n)$. The associated differential inclusion is

$$\dot{\mathbf{z}} \in \mu^\varepsilon(\mathbf{z}) \times Y - \mathbf{z}. \tag{26}$$

Extend $\mathbf{br}^\varepsilon$ to a map, still denoted $\mathbf{br}^\varepsilon$, on $[0,1] \times Y$ by choosing a non–empty selection of $\mathbf{Br}^\varepsilon$ and define

$$V^\varepsilon(z[i]) = U(\mathbf{br}^\varepsilon(z[i]), z[i]) - \varepsilon z_i^1 \rho(\mathbf{br}^\varepsilon(z[i]))$$

(so that if $z_i^1 > 0$ $V^\varepsilon(z[i]) = z_i^1 V^\varepsilon(z \mid i)$ and $V^\varepsilon(0) = 0$). Let

$$ce^\varepsilon(z) = \sum_i (V^\varepsilon(z[i]) - U(z[i])) = \sum_i V^\varepsilon(z[i]) - U(z).$$

The evaluation along a solution $t \to z(t)$ to (26) is

$$\mathbf{W}^\varepsilon(t) = ce^\varepsilon(\mathbf{z}(t)).$$

The next proof is in spirit similar to Section 6.3 but technically heavier. Since we are dealing with smooth best reply to conditional events there is a discontinuity at the boundary and the analysis has to take care of this aspect.

Theorem 6.11 *The set $\{z \in Z : \ ce^\varepsilon(z) \leq \varepsilon\}$ is an attracting set for (26) whose basin is Z. In particular, conditional consistency holds for continuous* $\mathsf{CSFP}(\varepsilon)$.

Proof: We shall compute

$$\dot{\mathbf{W}}^\varepsilon(t) = \frac{d}{dt}\sum\nolimits_i V^\varepsilon(\mathbf{z}[i](t)) - \frac{d}{dt}U(\mathbf{z}(t)).$$

The last term is

$$\frac{d}{dt}U(\mathbf{z}(t)) = U(\mu^\varepsilon(t), \beta(t)) - U(\mathbf{z}(t))$$

by linearity, with $\beta(t) \in Y$ and $\mu^\varepsilon(t) \in \mu^\varepsilon(\mathbf{z}(t))$. We now pass to the first term. First observe that

$$\frac{d}{dt}\mathbf{z}_i^1 \in \mu_i^\varepsilon(\mathbf{z}) - \mathbf{z}_i^1 \geq -\mathbf{z}_i^1.$$

Hence $\mathbf{z}_i^1(t) > 0$ implies $\mathbf{z}_i^1(s) > 0$ for all $s \geq t$. It then exists $\tau_i \in [0, \infty]$ such that $\mathbf{z}_i^1(s) = 0$ for $s \leq \tau_i$ and $\mathbf{z}_i^1(s) > 0$ for $s > \tau_i$. Consequently the map $t \to V^\varepsilon(\mathbf{z}[i](t))$ is differentiable everywhere but possibly at $t = \tau_i$ and is zero for $t \leq \tau_i$. If $t > \tau_i$, then

$$\begin{aligned} \frac{d}{dt}V^\varepsilon(\mathbf{z}[i](t)) &= \frac{d}{dt}U^\varepsilon(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)), \mathbf{z}[i](t)) - \varepsilon \mathbf{z}_i^1(t)\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t))) \\ &= U^\varepsilon(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)), \dot{\mathbf{z}}[i](t)) - \dot{\mathbf{z}}_i^1(t)\varepsilon\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t))) \end{aligned} \tag{27}$$

by Lemma 6.2. If now $t < \tau_i$, both $\dot{\mathbf{z}}[i](t)$ and $\frac{d}{dt}V^\varepsilon(\mathbf{z}[i](t))$ are zero, so that equality (27) is still valid.
Finally, using $\frac{d}{dt}\mathbf{z}_{ij}(t) = \mu^\varepsilon{}_i(t)\beta_j(t) - \mathbf{z}_{ij}(t)$, we get that

$$\dot{\mathbf{W}}^\varepsilon(t) = \sum_i U^\varepsilon(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)), \mu_i^\varepsilon(t)\beta(t) - \mathbf{z}[i](t))$$

$$+ \sum_i (\mu_i^\varepsilon(t) - \mathbf{z}_i^1(t))\varepsilon\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t))) - U(\mu^\varepsilon(t), \beta(t)) + U(\mathbf{z}(t))$$

for all (but possibly finitely many) $t \geq 0$. Replacing gives

$$\dot{\mathbf{W}}^\varepsilon(t) = -\mathbf{W}^\varepsilon(t) + \mathbf{A}(t)$$

where

$$\mathbf{A}(t) = -U(\mu^\varepsilon(t), \beta(t))$$
$$+ \sum_i U^\varepsilon(\mathbf{br}^\varepsilon(\mathbf{z}[i]((t)), \mu_i^\varepsilon(t)\beta(t)) + \sum_i \mu_i^\varepsilon(t)\varepsilon\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t))).$$

Thus one obtains:

$$\mathbf{A}(t) = -U(\mu^\varepsilon(t), \beta(t)) + \sum\nolimits_i \mu_i^\varepsilon(t) \left[U(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)), \beta(t)) + \varepsilon\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)))\right].$$

Now equation (24) and linearity of $U(., y)$ implies

$$U(\mu^\varepsilon(t), \beta(t)) = \sum\nolimits_i \mu_i^\varepsilon(t) U(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)), \beta(t))).$$

Thus

$$\mathbf{A}(t) = \varepsilon \sum_i \mu_i^\varepsilon(t)\rho(\mathbf{br}^\varepsilon(\mathbf{z}[i](t)))$$

so that

$$\dot{\mathbf{W}}^{\varepsilon}(t) \leq -\mathbf{W}^{\varepsilon}(t) + \varepsilon$$

for all (but possibly finitely many) $t \geq 0$. Hence

$$\mathbf{W}^{\varepsilon}(t) \leq e^{-t}(\mathbf{W}^{\varepsilon}(0) - \varepsilon) + \varepsilon$$

for all $t \geq 0$. ■

Theorem 6.12 *For any $\eta > 0$, there exists $\bar{\varepsilon} > 0$ such that, for $\varepsilon \leq \bar{\varepsilon}$, a* CSFP$(\varepsilon)$ *strategy is η-consistent.*

Proof: Let $\mathcal{L} = \mathcal{L}(\bar{z}_n)$ be the limit set of $(\bar{z}_n)$ defined by (25). Since $(\bar{z}_n)$ is a DSA to (26) and $\{z \in Z : ce^{\varepsilon}(z) \leq \varepsilon\}$ is an attracting set for (26) whose basin is Z (Theorem 6.11), it suffices to apply Property 2 (*ii*).

■

7 Extensions

We study in this section extensions of the previous dynamics in the case where the information of player 1 is reduced: either he does not recall his past moves, or he does not know the other players moves sets, or he is not told their moves.

7.1 Procedure in law

We consider here procedures where player 1 is uninformed of his previous sequences of moves, but know only its law (team problem).

The general framework is as follows. A discrete time process $\{w_n\}$ is defined through a recursive equation by:

$$w_{n+1} - w_n = a_{n+1} V(w_n, i_{n+1}, \ell_{n+1}) \tag{28}$$

where $(i_{n+1}, \ell_{n+1}) \in I \times L$ are the moves[2] of the players at stage $n+1$ and $V : \mathbb{R}^m \times I \times L \to \mathbb{R}^m$ is some bounded measurable map.

A typical example is given, in the framework of approachability (see section 3.2), by

$$V(w, i, \ell) = -w + A_{i\ell} \tag{29}$$

where $A_{i\ell}$ is the vector valued payoff corresponding to (i, ℓ) and $a_n = 1/n$. In such case $w_n = \bar{g}_n$ is the average payoff.

[2] For convenience, we keep the notation used for finite games but it is unnecessary to assume here that the move spaces are finite.

Assume that player 1 uses a strategy (as defined in section 3.2) of the form

$$\sigma(h_n) = \psi(w_n)$$

where for each w, $\psi(w)$ is some probability over I. Hence w plays the rôle of a state variable for player 1 and we call such σ a $\psi-$strategy. Let $V_\psi(w)$ be the range of V under σ at w, namely the convex hull of

$$\{\int_I V(w,i,\ell)\psi(w)(di);\ \ell \in L\}.$$

Then the associated continuous time process associated to (28) is

$$\dot{\mathbf{w}} \in V_\psi(\mathbf{w}). \tag{30}$$

We consider now another discrete time process where, after each stage n, player 1 is not informed upon his realized move i_n but only upon ℓ_n. Define by induction the new input at stage $n+1$:

$$w^*_{n+1} - w^*_n = a_{n+1} \int_I V(w^*_n, i, \ell_{n+1})\psi(w^*_n)(di). \tag{31}$$

Remark that the range of V under $\psi(w^*)$ at w^* is $V_\psi(w^*)$ so that the continuous time process associated to (31) is again (30). Explicitely (28) and (31) are DSA of the same differential inclusion (30).

Definition 7.1 A ψ-procedure in law is a strategy σ of the form $\sigma(h_n) = \psi(w^*_n)$ where for each w, $\psi(w)$ is some probability over I and $\{w^*_n\}$ is given by (31).

The key observation is that a procedure in law for player 1 is independent on the moves of player 1 and only requires the knowledge of the map V and the observation of the opponents moves. The interesting result is that such a procedure will in fact induce, under certain assumptions (see hypothesis 7.2 below), the same asymptotic behavior in the original discrete process.

Suppose that player 1 uses a ψ-procedure in law. Then the coupled system (28, 31) is a DSA to the differential inclusion

$$(\dot{w}, \dot{w}^*) \in V^2_\psi(w, w^*) \tag{32}$$

where $V^2_\psi(w, w^*)$ is the convex hull of

$$\{(\int_I V(w,i,\ell)\psi(w^*)(di), \int_I V(w^*,i,\ell)\psi(w^*)(di));\ \ell \in L\}.$$

We shall assume, from now on, that (32) meets the standing hypothesis 2.1. We furthermore assume that

Hypothesis 7.2 The map V satisfies one of the two following conditions:

(i) There exists a norm $||\cdot||$ such that $w \to w + V(w, i, \ell)$ is contracting uniformly in $s = (i, \ell)$. That is

$$||w + V(w, s) - (u + V(u, s))|| \leq \rho ||w - u||$$

for some $\rho < 1$.

(ii) V is C^1 in w and there exists $\alpha > 0$ such that all eigenvalues of the symmetric matrix

$$\frac{\partial V}{\partial w}(w, s) + \frac{{}^t\partial V}{\partial w}(w, s)$$

are bounded by $-\alpha$.

(^{t} stands for the transpose). Remark that hypothesis 7.2 holds trivially for (29). Under this later hypothesis one has the following result.

Theorem 7.3 *Assume that $\{w_n, w_n^*\}$ is a bounded sequence. Under a $\psi-$procedure in law the limit sets of $\{w_n\}$ and $\{w_n^*\}$ coincide, and this limit set is an ICT set of the differential inclusion (30). Under a $\psi-$strategy the limit set of $\{w_n\}$ is also an ICT set of the same differential inclusion.*

Proof: Let $\mathcal{L}$ be the limit set of $\{w_n, w_n^*\}$. By properties 1 and 2, $\mathcal{L}$ is compact and invariant. Choose $(w, w^*) \in \mathcal{L}$ and let $t \to (\mathbf{w}(t), \mathbf{w}^*(t))$ denote a solution to (32) that lies in $\mathcal{L}$ (by invariance) with initial condition (w, w^*). Let $\mathbf{u}(t) = \mathbf{w}(t) - \mathbf{w}^*(t)$.

Assume condition (i) in hypothesis 7.2. Let $Q(t) = ||\mathbf{u}(t)||$. Then for all $0 \leq s \leq 1$

$$Q(t+s) = ||\mathbf{u}(t) + \dot{\mathbf{u}}(t)s + o(s)|| = ||(1-s)\mathbf{u}(t) + (\dot{\mathbf{u}}(t) + \mathbf{u}(t))s + o(s)||$$

$$\leq (1-s)Q(t) + s||\dot{\mathbf{u}}(t) + \mathbf{u}(t)|| + o(s).$$

Now $\dot{\mathbf{u}}(t) + \mathbf{u}(t)$ can be written as

$$\mathbf{w}(t) - \mathbf{w}^*(t) + \int_{I \times L} [V(\mathbf{w}(t), i, \ell) - V(\mathbf{w}^*(t), i, \ell)]\, \psi(\mathbf{w}^*(t))(di)d\nu(\ell)$$

for some probability measure ν over L. Thus by condition (i)

$$Q(t+s) \leq (1-s)Q(t) + s\rho Q(t) + o(s),$$

from which it follows that

$$\dot{Q}(t) \leq (\rho - 1)Q(t)$$

for almost every t. Hence, for all $t \geq 0$:

$$Q(0) \leq e^{(\rho-1)t}Q(-t) \leq e^{(\rho-1)t}K$$

for some constant K. Letting $t \to +\infty$ shows that $Q(0) = 0$. That is $w = w^*$.

Assume now condition (ii). Let $||\cdot||$ denote the Euclidean norm on $\mathbb{R}^m$ and $\langle \cdot, \cdot \rangle$ the associated scalar product. Then

$$\langle V(w,s) - V(w^*,s), w - w^* \rangle = \int_0^1 \langle \partial_w V(w^* + u(w-w^*), s).(w-w^*), w - w^* \rangle du$$

$$\leq -\frac{\alpha}{2}||w - w^*||^2.$$

Therefore

$$\frac{d}{dt}Q^2(t) = 2\langle \mathbf{w}(t) - \mathbf{w}^*(t), \dot{\mathbf{w}}(t) - \dot{\mathbf{w}}^*(t) \rangle \leq -\alpha Q^2(t)$$

from which it follows (like previously) that $Q(0) = 0$.
We then have proved that given hypothesis 7.2, $\{w_n\}$ and $\{w_n^*\}$ have the same limit set under a $\psi-$procedure in law. Since $\{w_n^*\}$ is a DSA to (30), this limit set is ICT for (30) by Property 1. The same property holds for $\{w_n\}$ under a $\psi-$strategy.

■

Remark Let $\mathcal{R}$ denote the set of chain-recurrent points for (28). Hypothesis 7.2 can be weakened to the assumption that conditions (i) or (ii) are satisfied for V restricted to $\mathcal{R} \times I \times L$.

The previous result applies to the framework of Sections 4 and 5 and show that the discrete regret dynamics will have the same properties when based on the (conditional) expected stage regret $E_x R(s)$ or $E_x C(s)$.

7.2 Best prediction algorithm

Consider a situation where at each stage n an unknown vector U_n ($\in [-1, +1]^I$) is selected and a player chooses a component $i_n \in I$. Let $\omega_n = U_n^{i_n}$. Assume

that U_n is announced after stage n.
Consistency is defined trough the evaluation vector V_n with $V_n^i = \bar{U}_n^i - \bar{\omega}_n, i \in I$, where, as usual, $\bar{U}_n$ is the average vector and $\bar{\omega}_n$ the average realization.
Conditional consistency is defined through the evaluation matrix W_n with $W_n^{jk} = (1/n)(\sum_{m,i_m=j} U_m^k - \omega_m)$.
This formulation is related to on line algorithms, see Foster and Vohra (1999) for a general presentation. In the previous framework, the vector U_n is $U(., \ell_n)$ where ℓ_n is the choice of players other than 1 at stage n. The claim is that all previous results go through (V_n or W_n converges to the negative orthant) when dealing with the dynamics expressed on the payoffs space. This means that player 1 does not need to know the payoff matrix, nor the set of moves of the other players; only a compact range for the payoffs is requested. A sketch of proofs is as follow.

7.2.1 Approachability: consistency

We consider the dynamics of section 4. The regret vector R^* if i is played, is $R^*(i) = \{U^j - U^i\}_{j \in I}$. Lemma 4.6 is now for $\theta \in \Delta(I)$

$$\langle \theta, R^*(\theta) \rangle = 0$$

since $R^*(\theta)$, the expectation of R^* under θ is

$$R^*(\theta) = \sum_{i \in I} \theta(i) R^*(i) = \{U^j - \langle \theta, U \rangle\}_j$$

hence the properties of the P-regret based dynamics on the payoff space $\mathbb{R}^m$ still hold (Theorem 4.7 and Corollary 4.8).

7.2.2 Approachability: conditional consistency

The content of Section 5 extends as well. The $I \times I$ regret matrix is defined, at stage n, given the move i_n, by all lines being 0 except line i_n which is the vector $\{U_n^j - U_n^i\}_{j \in J}$. Then the analysis is identical and the convergence of the regret to the negative orthant holds for P-conditional regret dynamics as in Theorem 5.5 and Corollary 5.6.

7.2.3 SFP: consistency

In the framework of Section 6, the only hypothesis used on the set Y was that it was convex compact, hence one can take $L = [-1, +1]^I$ and $U(x, \ell) = \langle x, \ell \rangle$. Then all computations go through.

7.2.4 SFP: conditional consistency

For the analog of Section 6.5 let us define the $I \times I$ evaluation matrix M_n at stage n and given the move i_n, by all lines equal to 0 except line i_n being the vector U_n. Its average at stage n is $\bar{M}_n$. μ_n is an invariant measure for the Markov matrix defined by the family $\mathrm{BR}^\epsilon(\bar{M}_n^i)$, where $(\bar{M}_n^i)$ denotes the i-line of $(\bar{M}_n)$.

7.3 Partial information

We consider here the framework of section 7.2 but where only ω_n is observed by player 1, not the vector U_n. In a game theoretical framework, this means that the move of the opponent at stage n is not observed by player 1 but only the corresponding payoff $U(i_n, \ell_n)$ is known.
This problem has been studied in Auer and alii (1995), Foster and Vohra (1997), Fudenberg and Levine (1999), Hart and Mas-Colell (2001b) and in a game theoretical framework by Banos (1968) and Megiddo (1980) (note that working in the framework of 7.2 is more demanding than finding an optimal strategy in a game, since the payoffs can actually vary stage after stage).
The basic idea is to generate, from the actual history of payoffs and moves $\{\omega_n, i_n\}$ and the knowledge of the strategy σ a sequence of pseudo-vectors $\tilde{U}_n \in \mathbb{R}^S$ to which the previous procedures applies.

7.3.1 Consistency

We follow Auer and alii (1995) and define $\tilde{U}_n$ by

$$\tilde{U}_n^i = \frac{\omega_n}{\sigma_n^i} \mathbf{1}_{\{i=i_n\}}$$

where as usual i_n is the component chosen at stage n and σ_n^i stands for $\sigma(h_{n-1})(i)$. The associated pseudo-regret vector is $\{\tilde{R}_n^i = \tilde{U}_n^i - \omega_n\}_{i \in I}$. Notice that

$$E(\tilde{R}_n^i | h_{n-1}) = U_n^i - \langle \sigma_n, U_n \rangle$$

hence, in particular

$$\langle \sigma_n, E(\tilde{R}_n | h_{n-1}) \rangle = 0.$$

To keep $\tilde{U}_n$ bounded one defines first τ_n adapted to the vector $\tilde{U}_n$ as in Section 7.2, namely proportional to $\nabla P(\frac{1}{n-1} \sum_{m=1}^{n-1} \tilde{R}_m)$, see section 4, then σ is specified by

$$\sigma_n^i = (1 - \delta)\tau_n^i + \delta / K$$

for $\delta > 0$ small enough and K being the cardinality of the set I.
The discrete dynamics is thus

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

The corresponding dynamics in continuous time satisfies:

$$\dot{\mathbf{w}}(t) = \alpha(t) - \mathbf{w}(t)$$

with $\alpha(t) = U_t - \langle p(t), U_t\rangle$ for some measurable process U_t with values in $[-1, 1]$ and $p(t) = (1-\delta)q(t) + \delta/K$ with

$$\nabla P(w(t)) = \|\nabla P(w(t))\| q(t).$$

Define the condition

$$\langle \nabla P(w), w\rangle \geq B\|\nabla P(w)\|\|w^+\| \tag{33}$$

on $\mathbb{R}^S \setminus D$ for some positive constant B (satisfied for example by $P(w) = \sum_s (w_s^+)^2$).

Proposition 7.4 *Assume that the potential satisfies in addition (33). Then consistency holds for the continuous process $\tilde{R}_t$ and both discrete processes $\tilde{R}_n$ and R_n.*

Proof: One has

$$\begin{aligned} \frac{d}{dt}P(w(t)) &= \langle \nabla P(w(t)), \dot{w}(t)\rangle \\ &= \langle \nabla P(w(t)), \alpha(t) - w(t)\rangle. \end{aligned}$$

Now

$$\begin{aligned} \langle \nabla P(w(t)), \alpha(t)\rangle &= \|\nabla P(w(t))\|\langle q(t), \alpha(t)\rangle \\ &= \|\nabla P(w(t))\|\langle \frac{1}{1-\delta}p_t - \frac{\delta}{(1-\delta)K}, \alpha(t)\rangle \\ &\leq \|\nabla P(w(t))\|\frac{\delta}{(1-\delta)K}R \end{aligned}$$

for some constant R since $\langle p(t), \alpha(t)\rangle = 0$ and the range of α is bounded. It follows, using (33), that given $\varepsilon > 0$, $\delta > 0$ small enough and $\|w^+(t)\| \geq \varepsilon$ implies

$$\begin{aligned} \frac{d}{dt}P(w(t)) &\leq \|\nabla P(w(t))\|(\frac{\delta}{(1-\delta)K}R - B\|w^+(t)\|) \\ &\leq -\|\nabla P(w(t))\|B\varepsilon/2. \end{aligned}$$

Now $\langle \nabla P(w), w\rangle > 0$ for $w \notin D$ implies $\|\nabla P(w)\| \geq a > 0$ on $\|w^+\| \geq \varepsilon$. Let $\beta > 0$, $A = \{P \leq \beta\}$ and choose $\varepsilon > 0$ such that $\|w^+\| \leq \varepsilon$ is included A. Then the complement of A is an attracting set and consistency holds for the process $\tilde{R}_t$, hence as in section 4, for the discrete time process $\tilde{R}_n$. The result concerning the actual process R_n with $R_n^k = U_n^k - \omega_n$ finally follows from another application of Theorem 7.3 since both processes have same conditional expectation. ■

7.3.2 Conditional consistency

A similar analysis holds in this framework. The pseudo regret matrix is now defined by

$$\tilde{C}_n(i,j) = \frac{\sigma_n^i}{\sigma_n^j} U_n^j \mathbf{1}_{\{j=i_n\}} - U_n^i \mathbf{1}_{\{i=i_n\}}$$

hence

$$E(\tilde{C}_n(i,j)|h_{n-1}) = \sigma_n^i (U_n^j - U_n^i)$$

and this relation allows to invoke ultimately Theorem 7.3, hence to work with the pseudo process. The construction is similar to subsection 5.2, in particular equation (A6). $\mu(w)$ is a solution of

$$\sum_k \mu^k(w) \nabla_{kj} P(w) = \mu^j(w) \sum_k \nabla_{jk} P(w)$$

and player 1 uses a perturbation $\nu(t) = (1-\delta)\mu(w(t)) + \delta u$ where u is uniform. Then the analysis is as above and leads to

Proposition 7.5 *Assume that the potential satisfies in addition (33). Then consistency holds for the continuous process $\tilde{C}_t$ and both discrete processes $\tilde{C}_n$ and C_n.*

8 A learning example

We consider here a process analyzed by Benaïm and Ben Arous (2003). Let $S = \{0, \ldots, K\}$,

$$X = \Delta(S) = \{x \in \mathbb{R}^{K+1} : x_k \geq 0, \sum_{k=0}^{K} x_k = 1\}$$

be the K dimensional simplex, and $f = \{f_k\}, k \in S$ a family of bounded real valued functions on X. Suppose that a "player" has to choose an infinite sequence $x_1, x_2, \ldots \in S$ (identified with the extreme points of X) and is rewarded at time $n+1$ by

$$y_{n+1} = f_{x_{n+1}}(\bar{x}_n)$$

where

$$\bar{x}_n = \frac{1}{n} \sum_{1 \leq m \leq n} x_m.$$

Let

$$\bar{y}_n = \frac{1}{n} \sum_{1 \leq m \leq n} y_m$$

denote the average payoff at time n. The goal of the player is thus to maximize its long term average payoff $\liminf \bar{y}_n$. In order to analyze this system note that the average discrete process satisfies

$$\begin{aligned} \bar{x}_{n+1} - \bar{x}_n &= \frac{1}{n}(x_{n+1} - \bar{x}_n), \\ \bar{y}_{n+1} - \bar{y}_n &= \frac{1}{n}(f_{x_{n+1}}(\bar{x}_n) - \bar{y}_n). \end{aligned}$$

Therefore, it is easily seen to be a DSA of the following differential inclusion

$$(\dot{\mathbf{x}}, \dot{\mathbf{y}}) \in -(\mathbf{x}, \mathbf{y}) + N(\mathbf{x}, \mathbf{y}) \tag{34}$$

where $(x, y) \in X \times [\alpha_-, \alpha_+]$, $\alpha_- = \inf_{S,X} f_k(x), \alpha_+ = \sup_{S,X} f_k(x)$ and N is defined as

$$N(x, y) = \{(\theta, \langle \theta, f(x) \rangle) : \theta \in X\}.$$

Definition 8.1 *f has a gradient structure if, letting*

$$g_k(x_1, \ldots, x_K) = f_0(1 - \sum_{k=1}^{K} x_k, x_1, \ldots, x_K) - f_k(1 - \sum_{k=1}^{K} x_k, x_1, \ldots, x_K)$$

there exists a C^1 function V, defined in a neighborhood of

$$Z = \{z \in \mathbb{R}^K, z = \{z_k\}, k = 1, \ldots, K, \text{ with } (x_0, z) \in X \text{ for some } x_0 \in [0,1]\},$$

satisfying

$$\nabla V(z) = g(z).$$

Theorem 8.2 *Assume that f has a gradient structure. Then every compact invariant set of (34) meets the graph*

$$S = \{(x, y) \in X \times [\alpha_-, \alpha_+] : y = \langle f(x), x\rangle\}.$$

Proof: We follow the computation in Benaïm and Ben Arous (2003). Note that (34) can be rewritten as

$$\begin{aligned} &\dot{x} + x \in X \\ &\dot{y} = \langle x + \dot{x}, f(x)\rangle - y. \end{aligned}$$

Hence

$$\begin{aligned} \frac{y(s+t) - y(s)}{t} &= \frac{1}{t}\int_s^{s+t} \dot{y}(u)du \\ &= \frac{1}{t}[\int_s^{s+t} ?\langle f(x(u)), x(u)\rangle - y(u)du + \int_s^{s+t} ?\langle f(x(u)), \dot{x}(u)\rangle du] \end{aligned}$$

but $x(u) \in X$ implies

$$\begin{aligned} \langle f(x(u)), \dot{x}(u)\rangle &= \sum\nolimits_{k=0}^{K} f_k(x(u))\dot{x}_k(u) \\ &= \sum\nolimits_{k=1}^{K} [-f_0(x(u)) + f_k(x(u))]\dot{x}_k(u) \\ &= -\sum\nolimits_{k=1}^{K} g_k(z(u))\dot{z}_k(u) \\ &= -\frac{d}{dt}V(z(u)) \end{aligned}$$

where $z(u) \in \mathbb{R}^m$ is defined by $z_k(u) = x_k(u)$. So that

$$\frac{1}{t}\int_s^{s+t} (\langle f(x(u)), x(u)\rangle - y(u))du = \frac{(y(s+t) + V(z(s+t)) - (y(s) + V(z(s))}{t}$$

and the right hand term goes to zero uniformly (in s, y, z) as $t \to \infty$. Let now $\mathcal{L}$ be a compact invariant set. Replacing $\mathcal{L}$ by one of its connected components we can always assume that $\mathcal{L}$ is connected. Suppose that $\mathcal{L} \cap S = \emptyset$. Then $(\langle f(x), x\rangle - y)$ has constant sign on $\mathcal{L}$ (say > 0) and, by compactness, is bounded below by a positive number δ. Thus for any trajectory $t \to (x(t), y(t))$ contained in $\mathcal{L}$

$$\frac{1}{t}\int_s^{s+t} (\langle f(x(u)), x(u)\rangle - y(u))du \geq \delta.$$

A contradiction. ∎

Corollary 8.3 *The limit set of* $\{(\bar{x}_n, \bar{y}_n)_n\}$ *meets* S. *In particular*

$$\liminf \bar{y}_n \leq \sup_{x \in X} \langle x, f(x) \rangle.$$

If, furthermore (x_n) *is such that* $\lim_{n\to\infty} \bar{x}_n = x^*$ *then*

$$\lim_{n\to\infty} \bar{y}_n = \sup_{x \in X} \langle x^*, f(x^*) \rangle.$$

Proof : One uses the fact that the discrete process is a DSA hence the limit set is invariant, being ICT by Property 2. The second part of the corollary follows from the proof part (a) of Theorem 4 in Benaïm and Ben Arous (2003).

∎

9 Concluding remarks

The main purpose of the paper was to show that stochastic approximation tools are extremely effective for analyzing several game dynamics and that the use of differential inclusions is needed. Note that certain discrete dynamics do not enter this framework: one example is the procedure of Hart and Mas-Colell (2001a) which depends both on the average regret and on the last move. The corresponding continuous process generates in fact a differential equation of order 2. Moreover, as shown in Hart and Mas-Colell (2003), see also Cahn (2004), this continuous process has regularity properties not shared by the discrete counterpart.
Among the open problems not touched upon in the present work are the questions related to the speed of convergence and to the convergence to a subset of the approachable set.

References

[1] Aubin J.-P., (1991) *Viability Theory*, Birkhäuser.

[2] Auer P., Cesa-Bianchi N., Freund Y. and R.E. Shapire, (1995) Gambling in a rigged casino: the adversarial multi-armed bandit problem, *Proceedings of the 36 th Annual IEEE Symposium on Foundations of Computer Science*, 322-331.

[3] Auer P., Cesa-Bianchi N., Freund Y. and R.E. Shapire, (2002) The non-stochastic multiarmed bandit problem, *SIAM J. Comput.*, **32**, 48-77.

[4] Banos A., (1968) On pseudo-games, *Annals of Mathematical Statistics*, **39**, 1932-1945.

[5] Benaïm M., (1996) A dynamical system approach to stochastic approximation, *SIAM Journal on Control and Optimization*, **34**, 437-472.

[6] Benaïm M., (1999) Dynamics of stochastic approximation algorithms, *Séminaire de Probabilités XXXIII*, Lecture Notes in Math. **1709**, 1–68, Springer.

[7] Benaïm M. and G. Ben Arous, (2003) A two armed bandit type problem, *International Journal of Game Theory*, **32**, 3-16.

[8] Benaïm M. and M.W. Hirsch, (1996) Asymptotic pseudotrajectories and chain recurrent flows, with applications, *J. Dynam. Differential Equations*, **8**, 141-176.

[9] Benaïm M. and M.W. Hirsch, (1999) Mixed equilibria and dynamical systems arising from fictitious play in perturbed games, *Games and Economic Behavior*, **29**, 36-72.

[10] Benaïm M., J. Hofbauer and S. Sorin, (2005) Stochastic approximations and differential inclusions, *SIAM Journal on Control and Optimization*, **44**, 328-348.

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

[12] Brown G., (1951) Iterative solution of games by fictitious play, in Koopmans T.C. (ed.) *Activity Analysis of Production and Allocation* , Wiley, 374-376.

[13] Cahn A., (2004) General procedures leading to correlated equilibria, *International Journal of Game Theory*, **33**, 21-40.

[14] Duflo M., (1996) *Algorithmes Stochastiques*, Springer.

[15] Foster D. and R. Vohra, (1997) Calibrated learning and correlated equilibria, *Games and Economic Behavior*, **21**, 40-55.

[16] Foster D. and R. Vohra, (1998) Asymptotic calibration, *Biometrika*, **85**, 379-390.

[17] Foster D. and R. Vohra, (1999) Regret in the on-line decision problem, *Games and Economic Behavior*, **29**, 7-35.

[18] Freund Y. and R.E. Shapire, (1999) Adaptive game playing using multiplicative weights, *Games and Economic Behavior*, **29**, 79-103.

[19] Fudenberg D. and D. K. Levine, (1995) Consistency and cautious fictitious play, *Journal of Economic Dynamics and Control*, **19**, 1065-1089.

[20] Fudenberg D. and D. K. Levine, (1998) *The Theory of Learning in Games*, MIT Press.

[21] Fudenberg D. and D. K. Levine, (1999) Conditional universal consistency, *Games and Economic Behavior*, **29**, 104-130.

[22] Hannan J., (1957) Approximation to Bayes risk in repeated plays, *Contributions to the Theory of Games, III*, Dresher M., A.W. Tucker and P. Wolfe Eds., 97-139, Princeton U.P.

[23] Hart S., (2005) Adaptive heuristics, *Econometrica*, **73**, 1401-1430.

[24] Hart S. and A. Mas-Colell, (2000) A simple adaptive procedure leading to correlated equilibria, *Econometrica*, **68**, 1127-1150.

[25] Hart S. and A. Mas-Colell, (2001a) A general class of adaptive strategies, *Journal of Economic Theory*, **98**, 26-54.

[26] Hart S. and A. Mas-Colell, (2001b) A reinforcement procedure leading to correlated equilibria, in *Economic Essays: A Festschrift for Wener Hildenbrandt*, Debreu G., W. Neuefeind and W. Trockel Eds., Springer, 181-200.

[27] Hart S. and A. Mas-Colell, (2003) Regret-based continuous time dynamics, *Games and Economic Behavior*, **45**, 375-394.

[28] Hofbauer J. and W.H. Sandholm, (2002) On the global convergence of stochastic fictitious play, *Econometrica*, **70**, 2265-2294.

[29] Hofbauer J. and S. Sorin, (2006) Best response dynamics for continuous zero–sum games, *Discrete and Continuous Dynamical Systems*, **B, 6**, 215-224.

[30] Megiddo N., (1980) On repeated games with incomplete information played by non-Bayesian players, *International Journal of Game Theory*, **9**, 157-167.

[31] Métivier, M. and P. Priouret, (1992) Théorèmes de convergence preque-sûre pour une classe d'algorithmes stochastiques à pas décroissants, *Prob. Theor. Relat. Fields*, **74**, 403-438.

[32] Robinson J., (1951) An iterative method of solving a game, *Annals of Mathematics*, **54**, 296-301.

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