MATHEMATICS OF OPERATIONS RESEARCH
Vol. 34, No. 2, May 2009, pp. 263–269
ISSN 0364-765X | EISSN 1526-5471 | 09 | 3402 | 0263

DOI 10.1287/moor.1080.0359

Time Average Replicator and Best-Reply Dynamics

Josef Hofbauer
Department of Mathematics, University of Vienna, A-1090 Vienna, Austria,
josef.hofbauer@univie.ac.at

Sylvain Sorin
Equipe Combinatoire et Optimisation, CNRS FRE3232 Faculté de Mathématiques, UPMC-Paris 6, 75013 Paris, France, and
Laboratoire d'Econométrie, Ecole Polytechnique, 75005 Paris, France,
sorin@math.jussieu.fr

Yannick Viossat
Ceremade, Université Paris-Dauphine, F-75016 Paris, France,
viossat@ceremade.dauphine.fr

Using an explicit representation in terms of the logit map, we show, in a unilateral framework, that the time average of the replicator dynamics is a perturbed solution of the best-reply dynamics.

Key words: replicator dynamics; best-reply dynamics; logit map; perturbed differential inclusion; internally chain transitive set; attractor
MSC2000 subject classification: Primary: 91A22
OR/MS subject classification: Primary: Noncooperative games
History: Received June 30, 2007; revised July 10, 2008. Published online in *Articles in Advance* April 24, 2009.

1. Presentation. The two prime examples of deterministic evolutionary game dynamics are the *replicator dynamics* (RD) and the *best-response dynamics* (BRD).

In the framework of a symmetric two-person game with $K \times K$ payoff matrix A played within a single population, the *replicator equation* is given by

$$\dot{x}_t^k = x_t^k(e^k A x_t - x_t A x_t), \quad k \in K \quad \text{(RD)}, \tag{1}$$

where K is the set of pure strategies, $\Delta = \Delta(K)$ is the set of mixed strategies (the simplex on K), e^k is the kth unit vector in Δ, $x_t \in \Delta$ is the composition of the population at time t with x_t^k denoting the frequency of strategy k. It was introduced by Taylor and Jonker [25] as the basic selection dynamics for the evolutionary games of Maynard Smith [20]; see Hofbauer and Sigmund [16] for a summary. The interpretation is that in an infinite population of replicating players, the per-capita growth rate of the frequencies of pure strategies is linearly related to their payoffs.

In the same framework, the *best-reply dynamics* is defined by the differential inclusion on Δ,

$$\dot{z}_t \in \mathrm{BR}(z_t) - z_t, \quad t \geq 0 \quad \text{(BRD)}. \tag{2}$$

It was introduced by Gilboa and Matsui [10] and studied further by Hofbauer [14], Hofbauer and Sigmund [16], and Cressman [5]. Here $\mathrm{BR}(z) \subset \Delta$ denotes the set of all pure and mixed best replies to the strategy profile $z \in \Delta$. The interpretation is that in an infinite population of players, in each small time interval, a small fraction of players revises their strategies and changes to a best reply against the present population distribution. It is the prototype of a population model of rational (but myopic) behaviour.

(BRD) is closely related to the fictitious play process introduced by Brown [4], which models repeated decision making by a single decision maker in each role of the game. Consider a bimatrix game played repeatedly and let (x_n, y_n) denote the strategies at step n. The (discrete-time) fictitious play process satisfies for all $n > 0$, $x_{n+1} \in \mathrm{BR}_1(Y_n)$, $y_{n+1} \in \mathrm{BR}_2(X_n)$ where $X_n = (1/n)\sum_{s=1}^n x_s$, $Y_n = (1/n)\sum_{s=1}^n y_s$, and BR_i is the set of best replies for player i. In a continuous-time setting, letting (x_t, y_t) denote the strategies at time t, the process satisfies for all $t > 0$, $x_t \in \mathrm{BR}_1(Y_t)$, $y_t \in \mathrm{BR}_2(X_t)$ where $X_t = (1/t)\int_0^t x_s\, ds$ and $Y_t = (1/t)\int_0^t y_s\, ds$. This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}(\mathrm{BR}(Z_t) - Z_t), \quad t > 0 \quad \text{(CFP)} \tag{3}$$

where $\mathrm{BR}(Z_t) = \mathrm{BR}_1(Y_t) \times \mathrm{BR}_2(X_t)$. This is equivalent to (BRD) via the change in time $Z_{e^s} = z_s$.

Despite the different interpretations of (RD) and (BRD) and the different dynamic characters, there are amazing similarities in the long-run behaviour of these two dynamics. This has been summarized in the following heuristic principle; see Gaunersdorfer and Hofbauer [9] and Hofbauer [14]:

For many games, the long-run behaviour ($t \to \infty$) *of the time averages* $X_t = (1/t)\int_0^t x_s\,ds$ *of the trajectories* x_t *of the replicator equation is the same as for the BR trajectories.*

In this paper we will provide a rigorous statement that largely explains this heuristic. We show that for any interior solution of (RD), for every $t > 0$, the solution x_t at time t is an approximate best reply against its time-average X_t, and the approximation gets better as $t \to \infty$. This implies that X_t is a perturbed solution of (BRD) and hence the limit set of X_t has the same properties as a limit set of a true orbit of (BRD), i.e., it is invariant and internally chain transitive under (BRD); these terms will be explained in §5. The main tool to prove this result is the logit map, which is a canonical smoothing of the best-response correspondence. We show that x_t equals the logit approximation at X_t with error rate $1/t$.

2. Unilateral processes. We are interested in consequences for games, but it is instructive to consider the point of view of one player without hypotheses on the behaviour of the others. This gives rise to the unilateral process defined below. There are two interpretations: The first one is that a single decision maker repeatedly chooses a mixed strategy, receives a corresponding stream of payoffs, and adapts his behaviour accordingly, without necessarily knowing whether he is facing Nature or other decision makers. Alternatively, we may think of a game played between several populations (one per role in the game) and consider the evolution of behaviour in one population according to a given dynamics without hypotheses on the evolution of other populations; in particular, without assuming that the other populations evolve according to the same dynamics.

In both cases, the player—or the population of players—is facing a (measurable) outcome process $\mathcal{U} = \{U_t, t \ge 0\}$, with values in the cube $C = [-c, c]^K \subset \mathbb{R}^K$ where K is his finite set of actions and c is some positive constant. The outcome U_t is a vector; its component U_t^k is the payoff received at time t if k is the action played at that time. The constant c is a bound on the payoffs. The cumulative vector-valued outcome up to stage t is $S_t = \int_0^t U_s\,ds$, and its time average is denoted by $\bar{U}_t = (1/t)S_t$.

Note that **br** denotes the (payoff-based) best-reply correspondence from C to the simplex Δ on K, defined by

$$\mathbf{br}(U) = \left\{ x \in \Delta;\ \langle x, U\rangle = \max_{y\in\Delta}\langle y, U\rangle \right\}.$$

The $\mathcal{U}$-fictitious play process (FPP) is defined on Δ by the differential inclusion

$$\dot{X}_t \in \frac{1}{t}[\mathbf{br}(\bar{U}_t) - X_t] \quad \text{(FPP)}. \tag{4}$$

The $\mathcal{U}$-replicator process (RP) is specified by the following equation on Δ:

$$\dot{x}_t^k = x_t^k[U_t^k - \langle x_t, U_t\rangle], \quad k \in K \quad \text{(RP)}. \tag{5}$$

Explicitly, in the framework of an N-player game with finite strategy sets K^i for each player $i \in N$ and payoff for player i defined by a function G^i from $\prod_{i\in N} K^i$ to $\mathbb{R}$ one has, considering player 1, $U_t^k = G^1(k, x_t^{-1})$, where $x_t = (x_t^1, x_t^{-1}) \in \prod_{i\in N}\Delta^i$, with $\Delta^i = \Delta(K^i)$ denoting the simplex on K^i. This describes the vector-valued payoff process that player 1 is facing.

If all players follow a (payoff-based) fictitious play dynamics, each time average strategy satisfies (4). For $N = 2$ this is (CFP).

If all players follow the replicator process, then (5) yields the N-player replicator equation on $\prod_{i\in N}\Delta^i$,

$$\dot{x}_t^{ik} = x_t^{ik}[G^i(k, x_t^{-i}) - G^i(x_t)] \quad k \in K^i, \ \ i \in N. \tag{6}$$

Finally, in the framework of a symmetric two-person game with payoff matrix A played within a single population, $U_t = Ax_t$, $U_t^k = e^k A x_t$ and (5) yields (RD).

3. Logit rule and perturbed best reply. Define the logit map L from $\mathbb{R}^K$ to Δ by

$$L^k(V) = \frac{\exp V^k}{\sum_j \exp V^j}. \tag{7}$$

Given $\eta > 0$, let $[\mathbf{br}]^\eta$ be the correspondence from C to Δ with graph being the η-neighborhood for the uniform norm of the graph of **br**.

The L map and the **br** correspondence are related as follows:

PROPOSITION 3.1. *For any $U \in C$ and $\varepsilon > 0$*

$$L(U/\varepsilon) \in [\mathbf{br}]^{\eta(\varepsilon)}(U)$$

with $\eta(\varepsilon) \to 0$ as $\varepsilon \to 0$.

PROOF. Given $\eta > 0$, define the correspondence D^η from C to Δ by

$$D^\eta(U) = \left\{ x \in \Delta;\, (U^k + \eta < \max_{j\in K} U^j \Rightarrow x^k \le \eta),\ \forall\, k \in K \right\}$$

and note that $D^\eta \subset [\mathbf{br}]^\eta$.

Let $\varepsilon(\eta)$ satisfy

$$\exp(-\eta/\varepsilon(\eta)) = \eta.$$

By definition of L, one has for all (j, k)

$$L^k(U/\varepsilon) = \frac{\exp((U^k - U^j)/\varepsilon)}{1 + \sum_{l\neq j} \exp((U^l - U^j)/\varepsilon)},$$

and it follows that $\varepsilon \le \varepsilon(\eta)$ implies

$$L(U/\varepsilon) \in D^\eta(U).$$

Finally, define $\eta(\varepsilon)$ to be the inverse function of ε to get the result. □

REMARKS. L is also given by

$$L(V) = \operatorname*{arg\,max}_{x\in\Delta} \left\{ \langle x, V\rangle - \sum_k x^k \log x^k \right\};$$

see, e.g., Rockafellar [23, p. 148]. Hence, introducing the (payoff-based) perturbed best reply $\mathbf{br}^\varepsilon$ from C to Δ defined by

$$\mathbf{br}^\varepsilon(U) = \operatorname*{arg\,max}_{x\in\Delta} \left\{ \langle x, U\rangle - \varepsilon \sum_k x^k \log x^k \right\},$$

one has

$$L(U/\varepsilon) = \mathbf{br}^\varepsilon(U),$$

and Proposition 3.1 follows also from Berge's maximum theorem (Berge [3, p. 116]). The map $\mathbf{br}^\varepsilon$ is the logit ε-approximation of the **br** correspondence.

4. Explicit representation of the replicator process.

4.1. CEW. The following procedure has been introduced in discrete time in the framework of online algorithms under the name "multiplicative weight algorithm" (Freund and Schapire [7], Littlestone and Warmuth [19]). We use here the name (CEW) (continuous exponential weight) for the process defined, given $\mathcal{U}$, by

$$x_t = L\left(\int_0^t U_s\, ds\right) \quad \text{(CEW)}. \tag{8}$$

4.2. Properties of CEW. The main property of (CEW) that will be used is that it provides an explicit solution of (RD). In fact, applying the logit map to the cumulative outcome stream $\int_0^t U_s\, ds$ generates a replicator process for the current outcome stream U_t.

PROPOSITION 4.1. (CEW) *satisfies* (RP).

PROOF. Taking the derivative of $\log x_t^k$ leads to

$$\frac{\dot{x}_t^k}{x_t^k} = U_t^k - \sum_j \frac{U_t^j \exp \int_0^t U_v^j\, dv}{\sum_m \exp \int_0^t U_v^m\, dv},$$

which is

$$\dot{x}_t^k = x_t^k [U_t^k - \langle x_t, U_t\rangle],$$

which hence gives the previous (RP) Equation (5). □

Note that (CEW) specifies the solution starting from the barycenter of Δ. The link with the best-reply correspondence is the following.

PROPOSITION 4.2. *CEW satisfies*

$$x_t \in [\mathbf{br}]^{\delta(t)}(\bar{U}_t)$$

with $\delta(t) \to 0$ *as* $t \to \infty$.

PROOF. Write

$$x_t = L\left(\int_0^t U_s\, ds\right) = L(t\bar{U}_t) \in [\mathbf{br}]^{\eta(1/t)}(\bar{U}_t)$$

by Proposition 3.1, with $U = \bar{U}_t$ and $\varepsilon = 1/t$. Let $\delta(t) = \eta(1/t)$. □

4.3. Time average. We describe here the consequences for the time-average behavior process. Define

$$X_t = \frac{1}{t}\int_0^t x_s\, ds.$$

PROPOSITION 4.3. *If* x_t *is* (CEW), *then* X_t *satisfies*

$$\dot{X}_t \in \frac{1}{t}([\mathbf{br}]^{\delta(t)}(\bar{U}_t) - X_t)$$

with $\delta(t) \to 0$ *as* $t \to \infty$.

PROOF. Because

$$\dot{X}_t = \frac{1}{t}(x_t - X_t),$$

the result follows from Proposition 4.2. □

4.4. Initial conditions. The solution of (RP) starting from $x_0 \in \operatorname{int}\Delta$ is given by $x_t = L(U_0 + \int_0^t U_s\, ds)$ with $U_0^k = \log x_0^k$. The average process satisfies

$$\dot{X}_t \in \frac{1}{t}([\mathbf{br}]^{\delta(t)}(U_0/t + \bar{U}_t) - X_t), \tag{9}$$

which can be written as

$$\dot{X}_t \in \frac{1}{t}([\mathbf{br}]^{\alpha(t)}(\bar{U}_t) - X_t), \tag{10}$$

with $\alpha(t) \to 0$ as $t \to \infty$.

5. Consequences for games. Consider a two-person (bimatrix) game (A, B). If the game is symmetric, this gives rise to the single-population replicator dynamics (RD) and best-reply dynamics (BRD) as defined in §1. Otherwise, we consider the two population replicator dynamics

$$\begin{aligned} \dot{x}_t^k &= x_t^k(e^k A y_t - x_t A y_t), \quad k \in K^1 \\ \dot{y}_t^k &= y_t^k(x_t B e^k - x_t B y_t), \quad k \in K^2 \end{aligned} \tag{11}$$

and the (BRD) dynamics corresponding to (2). Let M be the state space (a simplex Δ or a product of simplices $\Delta^1 \times \Delta^2$).

We now use the previous results with the $\mathcal{U}$ process being defined by $U_t = Ay_t$ for player 1; hence, $\bar{U}_t = AY_t$. Note that[1] $\mathbf{br}(AY) = \mathrm{BR}_1(Y)$.

PROPOSITION 5.1. *The limit set of every replicator time-average process* X_t *starting from an initial point* $x_0 \in \operatorname{int} M$ *is a closed subset of* M *that is invariant and internally chain transitive* (ICT) *under* (BRD).

[1] The following result is not true for general N person games, with $N \geq 3$, because of the nonlinearity of their payoff functions. However, Proposition 5.1 still holds for N person games with linear incentives.

Here the *limit set* of a process X_t is the set of all accumulation points of X_t as $t \to \infty$. A set A is *invariant* under a (set-valued) flow Φ_t generated by a differential inclusion such as (BRD), if for every $x \in A$ there exists a solution $\mathbf{x}$, defined for all $t \in \mathbb{R}$, with $\mathbf{x}(0) = x$ and $\mathbf{x}(t) \in A$ for all $t \in \mathbb{R}$. For compact sets A this is equivalent to $A \subset \Phi_t(A)$ for all $t \in \mathbb{R}$; see Benaïm et al. [1, Lemma 3.3]. A set A is *internally chain transitive* (ICT) if any two points $x, y \in A$ can be connected by finitely many arbitrarily long pieces of orbits lying completely within A with arbitrarily small jumps between them. For the precise, definition, see Benaïm et al. [1, §3.3].

Proof of Proposition 5.1. Equation (10) implies that X_t satisfies a perturbed version of (CFP); hence, X_{e^t} is a perturbed solution to the differential inclusion (BRD), according to Definition II of Benaïm et al. [1]. Now apply Theorem 3.6 of that paper. □

A consequence of Proposition 5.1 is the following.

Proposition 5.2. *Let $\mathcal{A}$ be the global attractor* (*i.e., the maximal invariant set*) *of* (BRD). *Then the limit set of every replicator time-average process X_t starting from an initial point $x_0 \in \operatorname{int} M$ is a subset of $\mathcal{A}$.*

We now discuss some consequences and special cases.

(1) *If the time average of an interior orbit of the replicator dynamics converges, then the limit is a Nash equilibrium.* Indeed, by Proposition 5.1 the limit is a singleton invariant set of (BRD), and hence a Nash equilibrium. In particular, the time average of a periodic orbit in $\operatorname{int} M$ is an interior Nash equilibrium. (This statement is wrong for three-person games; see Plank [21] for a counterexample.) As another consequence, one obtains: If an interior orbit of the replicator dynamics converges, then the limit is a Nash equilibrium. (For a direct proof that works also for N-person games, see Hofbauer and Sigmund [16, Theorem 7.2.1].)

(2) For *two-person zero-sum games*, the global attractor of (BRD) equals the (convex) set of Nash equilibria (this is a strengthened version of Brown and Robinson's convergence result for fictitious play (Robinson [22]), stemming from Hofbauer and Sorin [17]). Therefore, by Proposition 5.2 the time averages of (RD) converge to the set of Nash equilibria as well. For a direct proof (in the special case when an interior equilibrium exists), see Hofbauer and Sigmund [16, 11.2.6]. Note that orbits of (RD) in general do not converge, but oscillate around the set of Nash equilibria, as in the matching pennies game.

(3) In *potential games* the only ICT sets of (BRD) are (connected subsets of) components of Nash equilibria; see Benaïm et al. [1, Theorem 5.5]. Hence, by Proposition 5.1 time averages of (RD) converge to such components. In fact, orbits of (RD) themselves converge, because the common payoff function is an increasing Ljapunov function; see Hofbauer and Sigmund [16, Theorem 11.2.2].

(4) For *games with a strictly dominated strategy*, the global attractor of (BRD) is contained in a face of M with no weight on this strategy. Hence, time averages of (RD) converge to this face, i.e., the strictly dominated strategy is eliminated on the average. In fact, the frequency of a strictly dominated strategy under (RD) vanishes; see Hofbauer and Sigmund [16, Theorem 8.3.2].

(5) Consider now the *rock–paper–scissors game* with payoff matrix

$$A = \begin{pmatrix} 0 & -b_2 & a_3 \\ a_1 & 0 & -b_3 \\ -b_1 & a_2 & 0 \end{pmatrix}, \quad a_i, b_i > 0$$

in a single population setting. There are two cases, see Gaunersdorfer and Hofbauer [9]. If $a_1a_2a_3 \geq b_1b_2b_3$, then the NE $\hat{x}$ is the global attractor of (BRD). Hence, Proposition 5.2 implies that the time averages of (RD) converge to $\hat{x}$ as well. Note that in case of equality, $a_1a_2a_3 = b_1b_2b_3$, the orbits of (RD) oscillate around $\hat{x}$ and hence do not converge; only their time averages do. If $a_1a_2a_3 < b_1b_2b_3$, then there are two ICT sets under (BRD), $\hat{x}$ and the Shapley triangle; see Gaunersdorfer and Hofbauer [9]. Then Proposition 5.1 implies that time averages of (RD) converge to one of these, whereas the limit set of all nonconstant orbits of (RD) is the boundary of M. However, our results do not show that for most orbits, the time average converges to the Shapley triangle. This still requires a more direct argument, as in Gaunersdorfer and Hofbauer [9].

(6) *If $\hat{x} \in \operatorname{int} M$ is the global attractor of* (BRD)*, then time averages of* (RD) *converge to $\hat{x}$*. In the literature on (RD) the following sufficient condition for the convergence of its time averages is known: If the (RD) is *permanent*, i.e., all interior orbits have their ω-limit set contained in a compact set in $\operatorname{int} M$, then the time averages of (RD) converge to the unique interior equilibrium $\hat{x}$. (See Hofbauer and Sigmund [16, Theorem 13.5.1].) It is tempting to conjecture that, for generic payoff matrices A, permanence of (RD) is equivalent to the global attractor of (BRD) being equal to the unique interior equilibrium.

6. External consistency. Another property related to the average outcome process and (CEW) is *external consistency* (sometimes called "no regret").

6.1. Definition. A procedure satisfies external consistency if for each process $\mathcal{U}$ with values in $\mathbb{R}^K$, it produces a process $\{x_t\} \in \Delta$, such that for all k

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle]\, ds \le C_t = o(t).$$

This property says that the (expected) average payoff induced by $\{x_t\}$ along the play is asymptotically not less than the payoff obtained by any fixed choice $k \in K$; see Fudenberg and Levine [8].

6.2. CEW. We recall this result from Sorin [24], where the aim is to compare discrete- and continuous-time procedures.

Proposition 6.1. (CEW) *satisfies external consistency.*

Proof. Define $W_t = \sum_{k\in K} \exp S_t^k$. Then

$$\dot{W}_t = \sum_k \exp(S_t^k) U_t^k = \sum_k W_t x_t^k U_t^k = \langle x_t, U_t \rangle W_t.$$

Hence,

$$W_t = W_0 \exp\left(\int_0^t \langle x_s, U_s \rangle\, ds\right).$$

Thus, $W_t \ge \exp(S_t^k)$ for every k, implies:

$$\int_0^t \langle x_s, U_s \rangle\, ds \ge \int_0^t U_s^k\, ds - \log W_0. \quad \square$$

6.3. RP. In fact, a direct and simpler equivalent proof is available; see Hofbauer [15].

Proposition 6.2. (RP) *satisfies external consistency.*

Proof. By integrating Equation (5), one obtains, on the support of x_0:

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle]\, ds = \int_0^t \frac{\dot{x}_s^k}{x_s^k}\, ds = \log\left(\frac{x_t^k}{x_0^k}\right) \le -\log x_0^k. \quad \square$$

Remark. This proof shows, in fact, more: For any accumulation point $\bar{x}$ of x_t, one component $\bar{x}^k$ will be positive, hence, the corresponding asymptotic average difference in payoffs will be 0. In fact, if $x_{t_n}^k \to \bar{x}^k$, then

$$\frac{1}{t_n} \int_0^{t_n} [U_s^k - \langle x_s, U_s \rangle]\, ds \to 0 \quad \text{as } t_n \to +\infty.$$

Back to a game framework this implies that if player 1 follows (RP), then the set of accumulation points of the empirical correlated distribution process will belong to her reduced Hannan set; see Fudenberg and Levine [8], Hannan [11], Hart [12]:

$$\bar{H}^1 = \{\theta \in \Delta(S);\, G^1(k, \theta^{-1}) \le G^1(\theta),\, \forall\, k \in S^1, \text{ with equality for at least one component}\}.$$

6.4. Internal consistency. A procedure satisfies *internal consistency* (or "conditional no regret") if for each process $\mathcal{U}$ with values in $\mathbb{R}^K$, it produces a process $\{x_t\} \in \Delta$, such that for all k and all j,

$$\int_0^t x_s^j [U_s^k - U_s^j]\, ds \le C_t' = o(t).$$

In a discrete-time context, this property says that the average payoff on periods where j was played is asymptotically not less than the payoff that would have been received on these periods by any fixed choice $k \in K$ (Foster and Vohra [6]). In a game context, if all players use a procedure satisfying internal consistency, then the set of accumulation points of the empirical correlated distribution process will belong to the set of correlated equilibria (Hart and Mas-Colell [13]).

The example from Viossat [26] of a game where the limit set for the replicator dynamics is disjoint from the unique correlated equilibrium shows that (RP) does not satisfy internal consistency.

7. Comments. We can now compare several processes in the spirit of (payoff-based) fictitious play. The original fictitious play process (I) is defined by

$$x_t \in \mathbf{br}(\bar{U}_t).$$

The corresponding time average satisfies (CFP).
With a smooth best-reply process (see Hopkins [18]) one has (II)

$$x_t = \mathbf{br}^{\varepsilon}(\bar{U}_t)$$

and the corresponding time average satisfies a smooth fictitious play process.
Finally, the replicator process (III) satisfies

$$x_t = \mathbf{br}^{1/t}(\bar{U}_t),$$

and the time average follows a time-dependent perturbation of the fictitious play process.

In (I), the process $\{x_t\}$ follows exactly the best-reply correspondence, but does not have good unilateral properties.

On the other hand for (II), $\{x_t\}$ satisfies a weak form of external consistency, with an error term $\alpha(\varepsilon)$ vanishing with ε (Fudenberg and Levine [8], Benaïm et al. [2]).

In contrast, (III) satisfies exact external consistency because of both a smooth and vanishing approximation of **br**.

Acknowledgments. The second author gratefully acknowledges the support of the ANR BLAN 06-3146280 "MAEV," and the third author the support of the ANR, project "Croyances," and of the Fondation du Risque, Chaire Groupama, "Les particuliers face au risque."

References

[1] Benaïm, M., J. Hofbauer, S. Sorin. 2005. Stochastic approximations and differential inclusions. *SIAM J. Control Optim.* **44**(1) 328–348.
[2] Benaïm, M., J. Hofbauer, S. Sorin. 2006. Stochastic approximations and differential inclusions. Part II: Applications. *Math. Oper. Res.* **31** 673–695.
[3] Berge, C. 1997. *Topological Spaces*. Dover Publications, Mineola, NY.
[4] Brown, G. W. 1951. Iterative solution of games by fictitious play. T. C. Koopmans, ed. *Activity Analysis of Production and Allocation*. Vol. 13. Wiley, New York, 374–376.
[5] Cressman, R. 2003. *Evolutionary Dynamics and Extensive Form Games*. M.I.T. Press, Cambridge, MA.
[6] Foster, D. P., R. Vohra. 1999. Regret in the on-line decision problem. *Games Econom. Behav.* **29**(1–2) 7–35.
[7] Freund, Y., R. E. Schapire. 1999. Adaptive game playing using multiplicative weights. *Games Econom. Behav.* **29**(1–2) 79–103.
[8] Fudenberg, D., D. K. Levine. 1995. Consistency and cautious fictitious play. *J. Econom. Dynam. Control* **19**(5–7) 1065–1089.
[9] Gaunersdorfer, A., J. Hofbauer. 1995. Fictitious play, Shapley polygons, and the replicator equation. *Games Econom. Behav.* **11**(2) 279–303.
[10] Gilboa, I., A. Matsui. 1991. Social stability and equilibrium. *Econometrica* **59**(3) 859–867.
[11] Hannan, J. 1957. Approximation to Bayes risk in repeated play. M. Drescher, A. W. Tucker, P. Wolfe, eds. *Contributions to the Theory of Games, III*, Vol. 3. Princeton University Press, Princeton, NJ, 97–139.
[12] Hart, S. 2005. Adaptive heuristics. *Econometrica* **73**(5) 1401–1430.
[13] Hart, S., A. Mas-Colell. 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* **68**(5) 1127–1150.
[14] Hofbauer, J. 1995. *Stability for the Best Response Dynamics*. Preprint.
[15] Hofbauer, J. 2004. *Time Averages of the Replicator Dynamics and Correlated Equilibria*. Preprint.
[16] Hofbauer, J., K. Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK.
[17] Hofbauer, J., S. Sorin. 2006. Best response dynamics for continuous zero-sum games. *Discrete Contin. Dynam. Systems Ser. B* **6**(1) 215–224.
[18] Hopkins, E. 2002. Two competing models of how people learn in games. *Econometrica* **70**(6) 2141–2166.
[19] Littlestone, N., M. K. Warmuth. 1994. The weighted majority algorithm. *Inform. Comput.* **108** 212–261.
[20] Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.
[21] Plank, M. 1997. Some qualitative differences between the replicator dynamics of two player and *n* player games. *Nonlinear Anal.* **30**(3) 1411–1417.
[22] Robinson, J. 1951. An iterative method of solving a game. *Ann. Math.* **54** 296–301.
[23] Rockafellar, R.-T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.
[24] Sorin, S. 2009. Exponential weight algorithm in continuous time. *Math. Programming Ser. B* **116**(1–2) 513–528.
[25] Taylor, P., L. Jonker. 1978. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **40**(2) 145–156.
[26] Viossat, Y. 2007. The replicator dynamics does not lead to correlated equilibria. *Games Econom. Behav.* **59**(2) 397–407.