Journal of Dynamics and Games doi:10.3934/jdg.2020028

Volume **7**, Number **4**, October **2020** pp. **365–386**

REPLICATOR DYNAMICS: OLD AND NEW

Sylvain Sorin

Institut de Mathématiques Jussieu-PRG
Sorbonne Université, Campus P. & M. Curie
CNRS UMR 7586
4 Place Jussieu, 75005 Paris, France

(Communicated by Josef Hofbauer)

Abstract. We introduce the unilateral version associated to the replicator dynamics and describe its connection to on-line learning procedures, in particular to the multiplicative weight algorithm. We show the interest of handling simultaneously discrete and continuous time analysis.

We then survey recent results on extensions of this dynamics as maximization of the cumulative outcome with alternative regularization functions and variable weights. This includes no regret algorithms, time average version and link to best reply dynamics in two person games, application to equilibria and variational inequalities, convergence properties in potential and dissipative games.

Contents

2020 *Mathematics Subject Classification.* Primary: 91A22.

Key words and phrases. Replicator dynamics, multiplicative weight algorithm, discrete and continuous time, no regret, best reply dynamics, potential and dissipative games.

Part of this work was presented at “Journées Franco-Chiliennes d’Optimisation", Toulouse, July 2017, and dedicated to the memory of Felipe Alvarez.

This research was partially supported by a PGMO grant *COGLED*.

The author thanks Josef Hofbauer for many constructive comments and a referee for an extremely precise and helpful report.

1. **Replicator dynamics.**

1.1. **Basic properties.** We recall briefly the definition of the replicator dynamics, for an in-depth analysis see Hofbauer and Sigmund (1998, 2003) [48, 49], and then describe the unilateral version.

1.1.1. *One population.* The replicator dynamics was introduced by Taylor and Jonker (1978) [80].

Consider the evolution in time of the composition of a single large population with K types. The random interaction between her members occurs by couples and the result depends only on the types. It is thus represented by a $K \times K$ fitness matrix A: $A_{k\ell}$ is the outcome of type "k" facing type "ℓ" (amount of offsprings).

The corresponding discrete time dynamics, defined on the population size, is then:

$$N^k_{m+1} = N^k_m (1 + h\, e^k A x_m) \tag{1}$$

with the following notations:
N^k_m : number of members of type k at stage m,
$x^k_m = \frac{N^k_m}{\sum_{\ell \in K} N^\ell_m}$: proportion of type k at stage m, $x_m = (x^k_m)_{k \in K}$: vector of these proportions hence $x_m \in \Delta(K)$: simplex of $\mathbb{R}^K$,
e^k : k-th unit vector in $\mathbb{R}^K$,
h : time step size.

Letting $h \to 0$ leads to the continuous time version defined on the population composition. x^k_t is the proportion of type k at time t and the *Replicator Dynamics* on the simplex $\Delta(K)$ of $\mathbb{R}^K$ is given by:

$$\boxed{\dot{x}^k_t = x^k_t \left(e^k A x_t - x_t A x_t \right), \quad k \in K. \qquad (RD) \qquad (2)}$$

An alternative useful formulation is:

$$\frac{d}{dt} \log(x^k_t) = e^k A x_t - x_t A x_t, k \in K. \tag{3}$$

Note also that this defines a conservative dynamics: each face of the simplex is preserved.

This dynamics has strong links with *Evolutionary Stable Strategies*, Maynard Smith (1982) [55], see again the analysis in Hofbauer and Sigmund (1998, 2003) [48, 49].

1.1.2. *Two populations.* Elements of population one (having P types) are randomly matched with elements of population two (with Q types). The result of the interaction is specified by two $P \times Q$ matrices A and B: cross-matching between types (p, q) induces fitness A_{pq} (resp. B_{pq}) for type p in population one (resp. q in two).

Notice that the comparison between the fitness of a specific type and the average fitness in her population involves now the composition of both populations. The corresponding dynamics with $x_t \in \Delta(P), y_t \in \Delta(Q)$ is given by:

$$\begin{aligned} \dot{x}_t^p &= x_t^p[e^p A y_t - x_t A y_t], \qquad p \in P \\ \dot{y}_t^q &= y_t^q[x_t B e^q - x_t B y_t], \qquad q \in Q. \end{aligned} \tag{4}$$

1.1.3. *I populations.* We consider here the framework of a finite set I of non atomic populations, each with a finite set of types $S^i, i \in I$. The interaction is represented by a profile $s = (s^i; i \in I)$ of types and the outcome is specified by functions $F^i : S = \prod_{j\in I} S^j \to \mathbb{R}, i \in I$ (with multilinear extension to $\prod_j \Delta(S^j)$). $x^i \in \Delta(S^i)$ describes the composition of population $i \in I$, x^{ip} being the proportion of type $p \in S^i$ in it.

The dynamics for $x_t = (x_t^i; i \in I)$ is given by:

$$\boxed{\dot{x}_t^{ip} = x_t^{ip}[F^i(e^{ip}, x_t^{-i}) - F^i(x_t^i, x_t^{-i})] \qquad p \in S^i, i \in I. \quad (5)}$$

A natural alternative interpretation is to consider a game with a finite set of players I, each $i \in I$ having a finite set of choices S^i and a payoff function F^i.

Then $x_t^i \in \Delta(S^i)$ describes a mixed strategy of player i at time t and the replicator dynamics models an evolutionary behavior of each player as a function of her own past performance.

Remark that the set of rest points of the dynamics is:

$$\bigcup \{NE(T); T^i \subset S^i, i \in I\}$$

where $NE(T)$ is the set of equilibria in the game with pure strategy sets $(T^i; i \in I)$.

1.1.4. *Unilateral replicator dynamics.* This describes the evolution of the behavior of one agent facing an unknown environment. At each time t, she chooses at random an action k in a finite set K, its law x_t^k is thus a mixed strategy. Alternatively the choice set is the simplex $\Delta(K)$ and she controls the proportion of each k. The vector outcome is given by a bounded measurable process $u = \{u_t \in \mathbb{R}^K\}$.

The u-replicator dynamics $(u - RD)$ is specified by the following equation on $\Delta(K)$:

$$\boxed{\dot{x}_t^k = x_t^k[u_t^k - \langle x_t, u_t\rangle], \qquad k \in K. \qquad (u - RD) \quad (6)}$$

Notice that u_t may depend on the previous trajectory $\{x_s; s \leq t\}$.

Clearly the previous versions with one, two or I populations can be written in this form.

1.2. **Logit representation.** Recall that the *logit map* L from $\mathbb{R}^K$ to $\Delta(K)$ is defined by:

$$L^k(V) = \frac{\exp V^k}{\sum_{j\in K} \exp V^j} \quad . \tag{7}$$

Then the following explicit representation holds, Rustichini (1999) [66], Hofbauer, Sorin and Viossat (2009) [51]:

Proposition 1.1.

$$x_t = L(\int_0^t u_s ds) \quad \textit{follows} \quad (u - RD). \tag{8}$$

More generally, starting from x_0 at time 0 one has :

$$x_t^k = \frac{x_0^k \exp \int_0^t u_s^k ds}{\sum_{j \in K} x_0^j \exp \int_0^t u_s^j ds}.$$

Let $H(x) = \sum_k x^k \log x^k$ be the *entropy function* on $\Delta(K)$. Then the logit map satisfies the following maximization property:

Proposition 1.2.

$$L(V) \text{ is the argmax of } [\ \langle V, x\rangle - H(x);\ x \in \Delta(K)]. \tag{9}$$

2. On line learning.

2.1. Model and definitions.

$\{u_n; n \geq 1\}$ is a *discrete time* process of vectors in $\mathcal{U} = [-1, 1]^K$.

At each stage n, an agent having observed the past realizations of the vectors $u_1, ..., u_{n-1}$, chooses a component k_n in K.

The outcome at that stage is :

$$\omega_n = u_n^{k_n}$$

and the past history is given by:

$$h_{n-1} = (u_1, k_1, ..., u_{n-1}, k_{n-1}) \in H_{n-1}.$$

A strategy σ in this prediction problem is defined by the collection of vectors:

$$\sigma(h_{n-1}) \in \Delta(K), \qquad \forall h_{n-1} \in H_{n-1}, \forall n \geq 1,$$

where $\sigma(h_{n-1})$ denotes the probability distribution of the choice at stage n, k_n, given the past history h_{n-1}.

Note that here again u_n may depend on the past history h_{n-1} and on $\sigma(h_{n-1})$.

2.1.1. *External regret.*

The *External Regret*, ER, given $k \in K$ and $u \in \mathbb{R}^K$, is the vector $R(k, u) \in \mathbb{R}^K$ with components:

$$R^\ell(k, u) = u^\ell - u^k,\ \ell \in K.$$

The evaluation at stage n is given by $R_n = R(k_n, u_n)$ i.e. $R_n^\ell = u_n^\ell - \omega_n,\ \ell \in K$.

The average ER vector at stage n is $\overline{R}_n$, thus:

$$\overline{R}_n^\ell = \overline{u}_n^\ell - \overline{\omega}_n,\ \ell \in K.$$

(Given a sequence (u_m), $\overline{u}_n$ denotes the average: $\overline{u}_n = \frac{1}{n}\sum_{m=1}^n u_m$)

It compares, given a realization of the process (u_m), the actual (average) payoff - induced by the trajectory (k_m) in the set K - to the payoff corresponding to the choice of a constant component, see Hannan (1957) [34], Foster and Vohra (1999) [27], Fudenberg and Levine (1995) [29].

Definition 2.1. A strategy σ satisfies *external consistency* (or has no ER) if, for every process $\{u_m\}$:

$$\max_{k \in K} [\overline{R}_n^k]^+ \longrightarrow 0 \text{ a.s., as } n \to +\infty$$

where, as usual $v^+ = \max(v, 0)$; or equivalently

$$\sum_{m=1}^n (u_m^k - \omega_m) \leq o(n), \qquad \forall k \in K.$$

2.1.2. *Internal regret.* The *Internal Regret*, IR, given (k, u) is the $K \times K$ matrix $S(k, u)$ with components:

$$S^{j\ell}(k,u) = (u^\ell - u^j)\, \mathbf{I}_{\{j=k\}}.$$

where $\mathbf{I}_{\{j=.\}}$ denotes the indicator function.
The evaluation at stage n is $S_n = S(k_n, u_n)$, explicitly:

$$S_n^{k\ell} = \begin{cases} u_n^\ell - u_n^k & for\ k = k_n \\ 0 & otherwise. \end{cases}$$

The average IR $\overline{S}_n$ takes thus the form:

$$\overline{S}_n^{k\ell} = \frac{1}{n} \sum_{m=1, k_m=k}^{n} (u_m^\ell - u_m^k).$$

It compares, on average and for each component $k \in K$, the outcome computed *on the dates where k was played*, to the outcome for an alternative choice $\ell \in K$ on the same dates, see Foster and Vohra (1999), [27], Fudenberg and Levine (1999), [31]. (Note that one can ignore the moves played on a vanishing proportion of stages, and then the above quantities are rescaled averages computed on the pertinent dates.)

Definition 2.2. A strategy σ satisfies *internal consistency* (or has no IR) if, for every process $\{u_m\}$ and every couple k, ℓ:

$$[\overline{S}_n^{k\ell}]^+ \longrightarrow 0 \text{ a.s., as } n \to +\infty$$

2.1.3. *From ER to IR.* Note that one can construct from no ER procedures a no IR procedure: this involves several algorithms run in parallel with different inputs and at each stage the output implemented is some “invariant distribution".

Consider K parallel algorithms $(\phi(k), k \in K)$, that are externally consistent and generate each, at each stage m, a (row) vector $q_m(k) \in \Delta(K)$. Let Q_m be the $K \times K$ matrix whose k line is $q_m(k)$. Finally define the strategy σ at that stage m given this history, as a Q_m invariant measure $p_m \in \Delta(K)$, i.e. satisfying:

$$p_m = p_m Q_m.$$

Given the outcome $u_m \in \mathbb{R}^K$, let for each k, $p_m^k u_m$ be the entry of algorithm $\phi(k)$, at that stage, which then produces a new vector $q_{m+1}(k)$ and so on. Expressing the fact that $\phi(k)$ satisfies the no ER condition gives for all $j \in K$:

$$[\sum_{m=0}^{n} p_m^k u_m^j - \langle q_m(k), p_m^k u_m\rangle] \leq o(n)$$

Note that this corresponds to an “expected version” (see Section 2.2.1. below).
Remark that $\sum_k \langle q_m(k), p_m^k u_m\rangle = \sum_k \langle p_m^k q_m(k), u_m\rangle = \langle p_m, u_m\rangle$. Hence by summing over k, for any function $F : K \mapsto K$, inducing a potential competitor σ_F of σ with $j = F(k)$, the difference between the performances of σ_F and σ will satisfy as well:

$$[\sum_{m=0}^{n} \sum_{k} p_m^k u_m^{F(k)} - \langle p_m, u_m\rangle] \leq o(n)$$

which implies no IR or internal consistency.

Basic references are Stoltz and Lugosi (2005) [78], Blum and Mansour (2007) [12].

Note also that there are strong links between no IR and *calibration*, see e.g. Perchet (2014) [61].

2.2. **Exponential weight algorithm.**

2.2.1. *Conditional expectation.* Recall that the total regret at stage n that one wants to control is a vector of the form:

$$\sum_{m=1}^{n} u_m^k - \omega_m, \quad k \in K$$

where $\omega_m = u_m^{k_m}$ is the random payoff at stage m.

Let $x_m \in \Delta(K)$ be the law of k_m i.e. the mixed strategy given the history at stage m, then:

$$\mathsf{E}(\omega_m | h_{m-1}) = \langle u_m, x_m \rangle$$

so that $\omega_m - \langle u_m, x_m \rangle$ is a bounded martingale difference.

Hoeffding-Azuma's concentration inequality for a process (Z_n) of martingale differences with $|Z_n| \leq L$, see e.g. Cesa-Bianchi and Lugosi (2006) [19], states that:

$$\mathsf{P}\{|\overline{Z}_n| \geq \varepsilon\} \leq 2\exp(-\frac{n\,\varepsilon^2}{2L^2}).$$

In particular $\overline{Z}_n \to 0$ a.s. and the difference between the regret and its conditional expectation is controlled. Now we aim to bound quantities of the form:

$$\sum_{m=1}^{n} u_m^k - \langle u_m, x_m \rangle, \quad k \in K$$

or equivalently:

$$\sum_{m=1}^{n} \langle u_m, x \rangle - \langle u_m, x_m \rangle, \quad x \in \Delta(K)$$

by linearity.

The no ER condition becomes:

$$\boxed{ER_n(x) = \sum_{m=1}^{n} \langle u_m, x - x_m \rangle \leq o(n), \quad \forall x \in \Delta(K). \qquad (10)}$$

2.2.2. *Exponential weight algorithm in discrete time.* The strategy is defined as follows:

$$\boxed{\sigma^k(h_n) = x_{n+1}^k = \frac{\exp(A\sum_{m=1}^{n} u_m^k)}{\sum_{j\in K}\exp(A\sum_{m=1}^{n} u_m^j)} = L^k(A\sum\nolimits_{m=1}^{n} u_m) \qquad (11)}$$

where A is a positive parameter, recall (7).

This procedure, *exponential weight algorithm*, EW, was introduced by Vovk (1990) [84], see also Littlestone and Warmuth (1994) [54], Freund and Schapire (1999) [28]. A nice survey is Arora, Hazan and Kale (2012) [4].

The main result is the following:

Proposition 2.1. *Auer, Cesa-Bianchi, Freund, Shapire (1995)* [5]
For $A = 1/\sqrt{n}$, the exponential weight algorithm satisfies:

$$ER_n(x) \leq M\sqrt{n}.$$

2.2.3. *Continuous time approach.* Given a measurable process $\{u_t, t \geq 0\}$, with values in $[0,1]^K$, define the *continuous time exponential weight algorithm*, CTEW, Sorin (2009) [75], as the measurable process $x_t \in \Delta(K)$ satisfying:

$$\boxed{x_t^k = \frac{\exp \int_0^t u_s^k ds}{\sum_{j \in K} \exp \int_0^t u_s^j ds} \qquad (12)}$$

Note, using (8), that this corresponds to the unilateral replicator dynamics, defined by (6), with initial condition the barycenter of the simplex.

For similar continuous time approaches, see Cesa-Bianchi and Lugosi (2003) [18], Hart and Mas-Colell (2003) [39].

Proposition 2.2. *Sorin (2009)* [75]

Conditional expected external consistency holds for CTEW in the sense that:

$$\int_0^T \langle u_s, x - x_s \rangle ds \leq C.$$

A simple proof follows from (3) by integration, taking x_0 with full support, see Hofbauer, Sorin and Viossat (2009) [51]:

$$\int_0^T u_s^k - \langle u_s, x_s \rangle ds = \int_0^t \frac{\dot{x}_s^k}{x_s^k} ds = \log(\frac{x_t^k}{x_0^k}) \leq -\log(x_0^k).$$

2.2.4. *Application for discrete time process.* Given a discrete time process $\{u_m\}$ and a corresponding EW algorithm $\{x_m\}$ the aim is to get a bound on:

$$\frac{1}{n} \sum_{m=1}^n \langle u_m, x - x_m \rangle$$

from an evaluation of:

$$\frac{1}{T} \int_0^T \langle v_s, y - y_s \rangle ds$$

where $\{v_t\}$ is a continuous process constructed from $\{u_m\}$ and $\{y_t\}$ is the $CTEW$ algorithm associated to $\{v_t\}$.

This approach provides an alternative proof of the speed of convergence:

$$\frac{1}{n} \sum_{m=1}^n \langle u_m, x - x_m \rangle \leq M n^{-1/2}$$

as follows, see Sorin (2009) [75].

Given n, choose $T = \sqrt{n}$ so that:

- the bound in the continuous time version is of the order $1/T = 1/\sqrt{n}$

$$\frac{1}{T} \int_0^T \langle v_s, y - y_s \rangle ds \leq \frac{M_1}{\sqrt{n}}$$

- the error term with the discrete algorithm with step size $T/n = 1/\sqrt{n}$ is:

$$|\frac{1}{n} \sum_{m=1}^n \langle u_m, x_m \rangle - \frac{1}{T}(\int_0^T \langle v_t, y_t \rangle dt)| \leq \frac{M_2}{\sqrt{n}}.$$

2.3. **Extensions: Penalization and variable parameter.**

Recall from equations (8) and (9) that the replicator dynamics x_t maximizes on the simplex $\Delta(K)$ the amount:

$$\langle \int_0^t u_s ds, x\rangle - H(x)$$

which appears as a relaxation/regularization of argmax$\{\int_0^t u_s^k ds; k \in K\}$.

We follow the analysis in Kwon and Mertikopoulos (2017) [53] which:
- extends the analysis from $\Delta(K)$ to any compact convex set $X \subset \mathbb{R}^K$,
- replaces the entropy function by any bounded *strictly convex* lower semi-continuous penalization/regularization function F with domain X,
- uses time variable parameters

and keeps the consistency properties of the dynamics.

Let us define:

$$S_F(V) = \operatorname{argmax}\{\langle V, x\rangle - F(x);\ x \in \mathbb{R}^K\} = \operatorname{argmax}\{\langle V, x\rangle - F(x);\ x \in X\} \quad (13)$$

then the procedure is given by:

$$x_t = S_F(\eta_t \int_0^t u_s ds) \quad (14)$$

where $\{\eta_t\}$ is a positive weight process.

Alternatively, using the strict convexity of F, hence the differentiability of its Fenchel conjugate F^*, defined on $\mathbb{R}^K$ by:

$$F^*(v) = \sup_{y \in \mathbb{R}^K} [\langle v, y\rangle - F(y)]$$

see Rockafellar (1970) [64] (Sections 12 and 26), one has:

$$\boxed{x_t = \nabla F^*(\eta_t \int_0^t u_s ds). \quad (15)}$$

2.3.1. *Continuous time bound.* Assume η_t *decreasing* and let $r_X(F) = \sup_X F(x) - \inf_X F(x)$ be the range of F on X.

Proposition 2.3. *Kwon and Mertikopoulos (2017)* [53]

$$ER_t(x) = \int_0^t \langle u_s | x - x_s\rangle ds \leq \frac{1}{\eta_t} r_X(F).$$

For the case $\eta_t \equiv 1$ the proof proceeds as follows:

From (15) one deduces:

$$\frac{d}{dt} F^*(\int_0^t u_s\ ds) = \langle u_t, x_t\rangle$$

hence, with $W_t = \int_0^t u_s\ ds$, one obtains:

$$ER_t(x) = \int_0^t \langle u_s, x - x_s\rangle = \langle W_t, x\rangle - F^*(W_t) + F^*(0) \leq F(x) - \inf_{y\in X} F(y) \leq r_X(F)$$

using Fenchel inequality: $F(y) + F^*(v) \geq \langle v, y\rangle$.

The function $\alpha(t) = \langle W_t, x\rangle - F^*(W_t)$ plays the role of a potential.

2.3.2. *Discrete time bound.* Assume now that F is L *strongly convex* for some norm $\|.\|$ on $V = \mathbb{R}^K$ ($F - \frac{L}{2}\|.\|^2$ is convex).

Consider a *decreasing* sequence $\{\eta_k\}$ and a process $\{u_k\} \in V^*$. Let:

$$x_{k+1} = S_F(\eta_k \sum_{j=1}^{k} u_j).$$

Then one has the bound:

Proposition 2.4. *Kwon and Mertikopoulos (2017)* [53]

$$ER_k(x) = \sum_{j=1}^{k} \langle u_j | x - x_j \rangle \leq \frac{r_X(F)}{\eta_k} + \frac{\sum_{j=1}^{k} \eta_{j-1} \|u_j\|_*^2}{2L}$$

where the first term corresponds to a bound arising from the continuous time trajectory and the second to the approximation error between the continuous and discrete time processes.

For a precise analysis of the impact of the choice of the parameters, see again Kwon and Mertikopoulos (2017) [53] .

Notice the following, assuming ($\|u_j\|_*$) bounded:

i) $\eta_k = k^{-1/2}$ gives convergence of the mean regret ER_k/k to 0 with speed $O(k^{-1/2})$.

ii) $\eta_k = 1$ corresponds to the replicator dynamics with best continuous time convergence speed but bad discrete approximation: $\sum_{j=1}^{k} \eta_{j-1} \sim k$.

iii) $\eta_k = \frac{1}{k\varepsilon}$ gives ER_k/k of the order of ε.

It corresponds to the continuous time process:

$$x_t = \nabla F^*(\frac{1}{\varepsilon} \times \frac{1}{t} \int_0^t u_s ds)$$

and the discrete associated dynamics is *smooth fictitious play* since with $\overline{u}_k = \frac{1}{k} \sum_{j=1}^{k} u_j$, x_{k+1} maximizes:

$$\langle x, \overline{u}_k \rangle - \varepsilon F(x).$$

One recovers the bound of Fudenberg and Levine (1995) [29], see also Hofbauer, Sorin and Viossat (2009) [51].

2.4. **Comments.** The literature on algorithms satisfying no regret properties is huge and the approaches are diverse.

The precursors are Hannan (1957) [34] and Blackwell (1956) [13]. A good sample contains Cover (1991) [22], Foster and Vohra (1993, 1999)[26, 27], Fudenberg and Levine (1995, 1999) [29, 31], Hart and Mas Colell (2000, 2001, 2003) [37, 38, 39], Kalai and Vempala (2005) [52]; the books by Cesa-Bianchi and Lugosi (2006) [19], Fudenberg and Levine (1998) [30], Hart and Mas Colell (2013) [41]; and the surveys by Hart (2005) [36] and Perchet (2014) [61].

Benaim, Hofbauer and Sorin (2005, 2006) [9, 10] extends the tools of stochastic approximation, see Benaim (1999) [7], and apply them to prove no-regret properties, see also Benaim and Faure (2013), [8].

3. **Learning in finite games.**

3.1. **Finite games.** The framework is as follows: there is a finite set of players I having finite action spaces $(S^i, i \in I)$ and payoff functions $(G^i : S = S^i \times S^{-i} \to \mathbb{R}, i \in I)$ with multilinear extension to $\Delta(S)$.

We consider a repeated interaction in discrete time and with standard signalling: after each stage n all the players know the profile $s_n = (s_n^i, i \in I)$ used at that stage.

In addition each player i knows her own payoff function G^i.

Fix a player i and let $K = S^i$. From a unilateral point of view, player i knows:
- the stage payoff $\omega_n = G^i(k_n, s_n^{-i})$ as well as
- the vector payoff $u_n = G^i(., s_n^{-i}) \in \mathbb{R}^K$.
and the analysis of the previous Section 2 applies.

3.2. **Consistent procedures and plays.**

3.2.1. *External consistency and Hannan set.* Introduce the *empirical distribution on moves up to stage* n, $z_n = \frac{1}{n}\sum_{m=1}^n s_m \in \Delta(S)$, with $s_m = (s_m^j, j \in I)$.

Hence by linear extension, for each $k \in K$:

$$\begin{aligned}
\overline{R}_n^k &= \frac{1}{n}\sum_{m=1}^n u_m^k - \omega_m \\
&= \frac{1}{n}\sum_{m=1}^n \{G^i(k, s_m^{-i}) - G^i(s_m)\} \\
&= G^i(k, \frac{1}{n}\sum_{m=1}^n s_m^{-i}) - G^i(\frac{1}{n}\sum_{m=1}^n s_m) \\
&= G^i(k, z_n^{-i}) - G^i(z_n)
\end{aligned}$$

Thus σ^i, strategy of player i satisfies external consistency, definition 2.1, is equivalent to :

$$d(z_n, H^i) \to 0 \quad a.s.,$$

where d is the Euclidean distance in $\mathbb{R}^K$ and:

$$H^i = \{z \in \Delta(S); G^i(k, z^{-i}) - G^i(z) \leq 0, \forall k \in S^i\}$$

is Hannan's set for player i, Hannan (1957) [34].
Define $H = \cap_i H^i$ as the *Hannan's set.*

Proposition 3.1. *If each player follows some externally consistent procedure, the empirical distribution of moves converges a.s. to the Hannan set H.*

Note that no coordination among the players is required, the result follows from the properties of unilateral procedures.

Remark In the case of a zero-sum game, the set H exhibits an important property:

Let $z \in H$ with marginals z^1, z^2 and f be the payoff function. Then the external consistency property for player 1 ($z \in H^1$) implies:

$$f(z) \geq f(s^1, z^2), \quad \forall s^1 \in S^1.$$

Using the dual inequality for player 2 implies that $f(z)$ is equal to the *value* of the game and the marginals z^1, z^2 are *optimal strategies.*
Example For the zero-sum game:

0	1	−1
−1	0	1
1	−1	0

the distribution:

1/3	0	0
0	1/3	0
0	0	1/3

belongs to the Hannan set.

3.2.2. *Internal consistency and correlated equilibrium distributions.* Similarly to above, consider player i and introduce $\overline{S}_n = \mathbf{M}(z_n)$ with $\mathbf{M}$ being a $K \times K$ matrix defined on $\Delta(S)$ by:

$$\mathbf{M}^{k,j}(z) = \sum_{\ell \in S^{-i}} [G^i(j,\ell) - G^i(k,\ell)] z(k,\ell).$$

Then σ^i satisfies internal consistency, definition 2.2, is equivalent to $d(z_n, C^i) \to 0$ a.s. with:

$$C^i = \{z \in \Delta(S); \mathbf{M}^{k,j}(z) \leq 0, \forall k, j \in S^i\}.$$

Note that $C = \cap_i C^i$ is the set of *correlated equilibrium distributions*, Aumann (1974) [6].

Proposition 3.2. *If all players follow some internally consistent procedure, the empirical distribution of moves converges a.s. to C.*

This provides an alternative proof of existence of correlated equilibrium distributions through the existence of internally consistent procedures.
No similar property can be expected for Nash equilibria, see Hart and Mas Colell (2003) [40].

3.3. **Time average RD and perturbed best reply.** We extend here the analysis done in Hofbauer, Sorin and Viossat (2009) [51] for the replicator dynamics, see also Mertikopoulos and Sandholm (2016) [56].

3.3.1. *Normalization.* We use the notations of Section 2 and consider a procedure satisfying (15).

Recall that $S_F(V)$ denotes argmax$\{\langle V, x\rangle - F(x); x \in X\}$ and define the *payoff best reply correspondence* $\mathbf{br}$ defined on $\mathbb{R}^K$ by:

$$\mathbf{br}(V) = \text{argmax}\{\langle V, x\rangle; x \in X\}. \tag{16}$$

Introduce finally $\mathbf{br}^\varepsilon(V)$ as the $\text{argmax}_X \langle V, x\rangle - \varepsilon F(x)$. It is easy to see that for $\mathcal{U}$ compact, there exists a map $g : \mathbb{R}^+ \to \mathbb{R}^+$ with $g(r) \to 0$ as $r \to 0$ such that: above $\mathcal{U}$ the graph of $\mathbf{br}^\varepsilon$ is included in a $g(\varepsilon)$-neigborhood of the graph of $\mathbf{br}$ written $[\mathbf{br}]^{g(\varepsilon)}$.

Note also that $\mathbf{br}^\varepsilon(V) = S_F(V/\varepsilon)$.

Define $U_t = \frac{1}{t}\int_0^t u_s\, ds$. Since one has:

$$x_t = S_F(\eta_t \int_0^t u_s\, ds) = S_F(\eta_t t \times \frac{1}{t}\int_0^t u_s\, ds) = S_F(\eta_t t\, U_t) \tag{17}$$

remark that $\eta_t t \to \infty$ implies $d(x_t, \mathbf{br}(U_t)) \to 0$.

Introduce now the time average of the procedure:

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

Then one obtains:

$$\dot{X}_t = \frac{1}{t}[x_t - X_t]$$

thus finally from (17):

$$\dot{X}_t \in \frac{1}{t}[[\mathbf{br}]^{g(\frac{1}{\eta_t t})}(U_t) - X_t]. \tag{18}$$

3.3.2. *Two person games and best reply dynamics.* Consider a finite 2 person game with $P \times Q$ bimatrix payoff functions (A, B).

Use the previous results for player 1 with the $\{u_t\}$ process being defined by $u_t = Ay_t$ with $y_t \in Y = \Delta(Q)$, hence $U_t = AY_t$.
Introduce the *strategy best reply correspondence*, BR^1, defined on Y by:

$$BR^1(y) = \{x \in \Delta(S);\ xAy \geq x'Ay, \forall x' \in \Delta(S)\}.$$

Note the relation between payoff and strategy best replies:

$$\mathbf{br}(Ay) = BR^1(y), \quad \forall y \in Y.$$

Hence the previous equation (18) is:

$$\dot{X}_t \in \frac{1}{t}[[BR^1]^{g(\frac{1}{\eta_t t})}(Y_t) - X_t] \tag{19}$$

which is a perturbation of:

$$\dot{X}_t \in \frac{1}{t}[BR^1(Y_t) - X_t] \tag{20}$$

that corresponds to the best reply dynamics BRD, up to a change of time.

Explicitly the *best reply dynamics*, Gilboa and Matsui (1991)[33], is the differential inclusion on $M = X \times Y$ defined by:

$$\begin{array}{rcll} \dot{X}_t & \in & BR^1(Y_t) - X_t, & t \geq 0 \\ \dot{Y}_t & \in & BR^2(X_t) - Y_t, & t \geq 0. \qquad (BRD) \end{array} \tag{21}$$

Recall that this corresponds to a continuous time version (up to a change of time) of the discrete time *fictitious play* procedure, Brown (1949, 1951) [15, 16], Robinson (1951) [63], see e.g. Harris (1998) [35] Hofbauer and Sorin (2006) [50], Benaim, Hofbauer and Sorin (2005, 2006) [9, 10].

Notation. Let $\mathcal{P}$ be the family of procedures satisfying (15), induced by a penalization function F with moreover $\eta_t t \to \infty$ and $\mathcal{P}_1$ for $\eta_t = 1$.

The properties of stochastic approximation for differential inclusions, Benaim, Hofbauer and Sorin (2005, 2006) [9, 10], lead to the following results:

Proposition 3.3. *The limit set of every time average process $Z_t = (X_t, Y_t)$, with $(x_t, y_t) \in \mathcal{P}$ starting from an initial point $(x_0, y_0) \in M$, is a closed subset of M invariant and internally chain transitive under BRD.*

In particular this implies:

Proposition 3.4. *Let $\mathcal{A}$ be the global attractor (i.e., the maximal invariant set) of BRD. Then the limit set of every time average process $Z_t = (X_t, Y_t)$, with $(x_t, y_t) \in \mathcal{P}$ is a subset of $\mathcal{A}$.*

Recall that these similarities between RD and BRD were observed and the above properties conjectured in Gaunersdorfer and Hofbauer (1995) [32].

4. **Global dynamics in games.** Properties of dynamics in $\mathcal{P}$ have been studied in the framework of finite games by Coucheney, Gaujal and Mertikopoulos (2015) [21], Mertikopoulos and Sandholm (2016)[56] leading to several important extensions of the "Folk Theorem of evolutionary games". For population games see also Mertikopoulos and Sandholm (2018)[57] and Mertikopoulos and Zhou (2019) [58] for the continuous action case.

The replicator dynamics as well as dynamics in $\mathcal{P}$ can be extended to other configurations were equilibrium conditions take the form of variational inequalities.

We follow here Sorin and Wan (2016), [77].

4.1. **Variational inequalities.**

4.1.1. *Equilibrium and variational inequalities.* Three basic classes of games where equilibria are solutions of variational inequalities are as follows:

A) Finite games, see Section 3.1.

Let VG^i denote the vector payoff associated to G^i. Explicitly, $VG^{ip}(x^{-i}) = G^i(p, x^{-i})$, for all $p \in S^i, i \in I$. Hence $G^i(x) = \langle x^i, VG^i(x^{-i})\rangle$. Let $VG(x) = (VG^i(x^{-i}); i \in I)$.

An equilibrium is thus a solution of :

$$\langle VG(x), x - y\rangle = \sum_{i \in I} \langle VG^i(x^{-i}), x^i - y^i\rangle \geq 0, \quad \forall y \in X. \tag{22}$$

B) Concave $\mathcal{C}^1$ games.

Consider the case of I players with action sets $(X^i; i \in I)$ and payoff functions $(H^i; i \in I)$.

Assume that each X^i is a convex compact subset of a Hilbert space Y^i and that each $H^i : X = \prod_{j \in I} X^j \to \mathbb{R}$ is of class $\mathcal{C}^1$ and concave with respect to x^i.

An *equilibrium* is as usual a profile $x \in X$ satisfying:

$$H^i(x) \geq H^i(y^i, x^{-i}), \quad \forall y^i \in X^i, \forall i \in I. \tag{23}$$

which under our hypotheses is equivalent to:

$$\langle \nabla^i H^i(x), x^i - y^i\rangle \geq 0, \quad \forall y^i \in X^i, \forall i \in I. \tag{24}$$

where ∇^i is the gradient w.r.t. x^i.

C) Population games. Consider a finite set I of non atomic populations, each with a finite set S^i of types. A configuration is a vector $x = (x^i; i \in I)$ where each $x^i \in X^i = \Delta(S^i)$ describes the composition of population $i \in I$. The payoffs are defined by a family of continuous functions $(g^{ip}, i \in I, p \in S^i)$, all from $X = \prod_i X^i$ to $\mathbb{R}$, where $g^{ip}(x)$ is the outcome of a member in population i choosing p, given the configuration x. A *Wardrop equilibrium*, Wardrop (1952)[85], is a profile $x \in X$ satisfying:

$$x^{ip} > 0 \Rightarrow g^{ip}(x) \geq g^{iq}(x), \quad \forall p, q \in S^i, \forall i \in I \tag{25}$$

meaning that if p is used in population i, it is a best choice. An equivalent characterization of (25) is through the solutions of the variational inequality:

$$\langle g^i(x), x^i - y^i\rangle \geq 0, \quad \forall y^i \in X^i, \forall i \in I,$$

or alternatively:

$$\langle g(x), x - y\rangle = \sum_{i \in I} \langle g^i(x), x^i - y^i\rangle \geq 0, \quad \forall y \in X.$$

This formulation is standard in transportation, see Dafermos (1980) [23], Dupuis and Nagurney (1993), [24], Smith (1979) [73].

4.1.2. *Dynamics for variational inequalities.* Note that VG, ∇H and g play similar roles in the three frameworks above.

We call them *evaluation functions* and denote them by $\Phi = \{\Phi^{ip}\}$. By extension one will speak of the game $\Gamma(\Phi)$ when a game has evaluation functions Φ. Assume that X^i is a convex compact subset of $\mathbb{R}^{S^i}$ for each $i \in I$, $X = \prod_i X^i$ and that $\Phi^{ip} : X = \prod_{j\in I} X^j \longrightarrow \mathbb{R}$ is continuous for each $i \in I, p \in S^i$. $NE(\Phi)$, set of equilibria of $\Gamma(\Phi)$, is the set of profiles $x \in X$ satisfying:

$$\langle \Phi(x), x - y\rangle = \sum_i \langle \Phi^i(x), x^i - y^i\rangle \geq 0, \qquad \forall y \in X. \tag{26}$$

An equivalent representation is given by:

$$\Pi_X[x + \Phi(x)] = x. \tag{27}$$

where Π_X is the projection operator on X.

Note that under our hypotheses, Brouwer's fixed point theorem implies the existence of a solution.

Assume $X^i = \Delta(S^i), \forall i \in I$, then the replicator dynamics has the form:

$$\boxed{\dot{x}_t^{ip} = x_t^{ip}[\Phi_t^{ip}(x_t) - \overline{\Phi}^i(x_t)], \quad p \in S^i, i \in I, \qquad (28)}$$

where:

$$\overline{\Phi}^i(x) = \langle x^i, \Phi^i(x)\rangle = \sum_{p\in S^i} x^{ip}\Phi^{ip}(x)$$

which allows for arbitrary continuous (not multilinear) functions Φ.

More generally for a strategy of player i in the class $\mathcal{P}_1$ the corresponding $\{u_t\}$ process is $\{\Phi^i(x_t)\}$.

Note that if each of the players in I use any procedure in $\mathcal{P}_1$ and $\{x_t\}$ converges to x^*, the property:

$$\int_0^t \langle \Phi(x_s), x - x_s\rangle ds \leq o(t), \, \forall x \in X$$

implies by taking the average that:

$$\langle \Phi(x^*), x - x^*\rangle \leq 0, \, \forall x \in X$$

hence $x^* \in NE(\Phi)$.

Definition 4.1. A dynamics $\dot{x}_t = \mathcal{B}_\Phi(x_t)$ satisfies *positive correlation (PC)*, Sandholm (2010) [69], if:

$$\mathcal{B}_\Phi^i(x) \neq 0 \Rightarrow \langle \mathcal{B}_\Phi^i(x), \Phi^i(x)\rangle > 0, \quad \forall i \in I, \forall x \in X. \tag{29}$$

This corresponds to *myopic adjustment dynamics*, Swinkels (1993) [79] and extends the gradient-like property.

Proposition 4.1. *Replicator dynamics as well as dynamics in $\mathcal{P}_1$ satisfies (PC).*

Note that if F^* is $\mathcal{C}^2$ with Hessian $\nabla^2 F^*$ one has from (15):

$$\dot{x}_t^i = (\nabla^2 F^*(\int_0^t \Phi^i(x_s)ds))\Phi^i(x_t)$$

hence

$$\langle \dot{x}^i_t, \Phi^i(x_t)\rangle = \nabla^2 F^*(\int_0^t \Phi^i(x_s)ds)[\Phi^i(x_t), \Phi^i(x_t)]$$

(where $A[x,x]$ denotes the bilinear form xAx) which is non negative by convexity and positive if $\dot{x}^i_t \neq 0$

4.2. **Potential games.**

Definition 4.2. A real valued function W, of class $\mathcal{C}^1$ on a neighborhood Ω of X, is a *potential* for Φ if for each $i \in I$, there exists a strictly positive function $\mu^i(x)$ defined on X such that:

$$\boxed{\left\langle \nabla^i W(x) - \mu^i(x)\Phi^i(x), y^i \right\rangle = 0, \quad \forall x \in X, \forall y^i \in X^i_0, \forall i \in I, \qquad (30)}$$

where $X^i_0 = \{y \in \mathbb{R}^{|S^i|}, \sum_{p\in S^i} y_p = 0\}$ is the tangent space to X^i.

The game $\Gamma(\Phi)$ is then called a *potential game*, Monderer and Shapley (1996) [59], Sandholm (2001, 2009), [67, 68].

The potential W allows to find equilibria since one has:

Proposition 4.2. *Let $\Gamma(\Phi)$ be a game with potential W.*
1. Every local maximum of W is an equilibrium of $\Gamma(\Phi)$.
2. If W is concave on X, then any equilibrium of $\Gamma(\Phi)$ is a global maximum of W on X.

Moreover the potential enjoys also interesting dynamical properties:

Proposition 4.3. *Consider a potential game $\Gamma(\Phi)$ with potential function W.*
If the dynamics $\dot{x} = \mathcal{B}_\Phi(x)$ satisfies (PC), then W is a strict Lyapunov function for $\mathcal{B}_\Phi$. All ω-limit points are rest points of $\mathcal{B}_\Phi$.

In fact, if x_t is not a rest point:

$$\frac{d}{dt}W(x_t) = \sum_i \langle \nabla^i W(x_t), \dot{x}^i_t\rangle = \sum_i \mu^i(x_t)\langle \Phi^i(x_t), \dot{x}^i_t\rangle > 0. \tag{31}$$

In particular this applies to RD and dynamics in $\mathcal{P}_1$.

4.3. **Dissipative games.**

Definition 4.3. The game $\Gamma(\Phi)$ is *dissipative* if Φ satisfies:

$$\langle \Phi(x) - \Phi(y), x - y\rangle \leq 0, \qquad \forall\ (x,y) \in X \times X. \tag{32}$$

In the framework of population games, Hofbauer and Sandholm (2002) [47] studied this class under the name *stable games*.

Recall that this class extends zero-sum games, Rockafellar (1970) [65].

Let $SNE(\Phi)$ be the set of $x \in X$ satisfying:

$$\langle \Phi(y), x - y\rangle \geq 0, \qquad \forall y \in X. \tag{33}$$

Proposition 4.4. *If $\Gamma(\Phi)$ is dissipative*

$$SNE(\Phi) = NE(\Phi).$$

in particular $NE(\Phi)$ is convex.

Results similar to potential games hold for dissipative games and several evolutionary dynamics but with ad hoc Lyapunov functions.

Proposition 4.5. *Consider a dissipative game* $\Gamma(\Phi)$ *and let* $\bar{x} \in NE(\Phi)$.

For the replicator dynamics, H *is a Lyapunov function with:*

$$H(x) = \sum_{i \in I} \sum_{p \in supp(\bar{x}^i)} \bar{x}^i_p \log \frac{\bar{x}^i_p}{x^i_p}.$$

More generally for dynamics in $\mathcal{P}_1$ *with smooth* F

$$H(x) = \langle \nabla F(x), x - \bar{x} \rangle - F(x)$$

or in the dual space of outcomes:

$$G(W) = F^*(W) - \langle W, \bar{x} \rangle$$

with $W^i_t = \int_0^t \Phi^i(x_s) ds$, *are Lyapunov functions.*

In fact one has:

$$\begin{aligned} \frac{d}{dt} G(W_t) &= \sum_i \langle \Phi^i(x_t), \nabla F^*(W^i_t) \rangle - \langle \Phi^i(x_t), \bar{x}^i \rangle \\ &= \sum_i \langle \Phi^i(x_t), x^i_t - \bar{x}^i \rangle \leq 0 \end{aligned}$$

by using (33).

5. **Smooth case: Hessian Riemannian metrics.** Consider the initial one population model with in addition symmetric interaction: $A = {}^t A$.

We follow Akin (1979) [1], Hofbauer and Sigmund (1998) [48].

Ax derives from the potential $W(x) = \frac{1}{2} x A x$ and the replicator dynamics is a *gradient for the Shahshahani metric*, Shahshahani (1979) [71], $(.|.)_x$, defined on the tangent space, for x in the interior of $\Delta(K)$, by:

$$(u|v)_x = \sum_k \frac{1}{x^k} u^k v^k.$$

This means that RD writes:

$$\dot{x}_t = \operatorname*{grad}_{x_t} W(x_t) \tag{34}$$

In fact:

$$\operatorname*{grad}_{x_t} W(x_t) = \{x^k_t [e^k A x_t - x_t A x_t]; k \in K\}.$$

satisfies:

$$(\operatorname*{grad}_{x_t} W(x_t) | v)_{x_t} = \sum_k \frac{1}{x^k_t} x^k_t [e^k A x_t - x_t A x_t] v^k = \langle A x_t, v \rangle = DW(x_t).v$$

where DW stands for the differential.

More generally for a dynamics in $\mathcal{P}_1$, assume F $\mathcal{C}^2$ strictly convex with $\|\nabla F(x)\| \to \infty$ as $x \to \partial X$.

Then from:

$$x_t = \nabla F^*(\int_0^t u_s ds)$$

one obtains:

$$\nabla F(x_t) = \int_0^t u_s ds$$

and:

$$\nabla^2 F(x_t) \dot{x}_t = u_t$$

finally :

$$\dot{x}_t = (\nabla^2 F(x_t))^{-1} u_t$$

which corresponds when: $u_t = \nabla f(x_t)$ to the dynamics introduced and studied in Alvarez, Bolte and Brahic (2004) [3]. $\nabla^2 F(x_t)$ defines a Riemannian metric on the interior of X and if u_t is a gradient, x_t is the corresponding Riemannian gradient. For a precise analysis in the framework of population games, see Mertikopoulos and Sandholm (2018) [57].

A similar property holds for potential games as defined in Section 4.2.

6. Concluding remarks.

6.1. Comments.

Starting from evolution of populations, based on realized fitness, one obtains a class of dynamics $\{x_t\}$ or $\{x_n\}$ adapted to a process $\{u_t\}$ or $\{u_n\}$, satisfying:
- no regret properties in on-line learning,
- convergence to Hannan set for finite games,
- convergence to equilibria for some classes of games (potential, dissipative).

Moreover the same mechanism applies:
- for general unknown process $\{u_t\}$ (the changing unpredictable environment),
- for continuous vector field $\{u_t\} = g(x_t)$,
- for games when equilibrium conditions are expressed as variational inequalities.

The analysis shows the analogy between Exponential Weights and Replicator Dynamics (discrete/continuous time) similar to Fictitious Play versus Best Reply Dynamics, and the connection between the two groups through time averaging.

All the properties emerge from independent/autonomous or uncoupled dynamics. The link between the individual dynamics of the players at time t is through the past trajectory of moves $\{x_s; s < t\}$ that impact their outcomes $\{\Phi(x_s); s < t\}$. In particular the knowledge of the other players payoff function is not assumed. As a consequence a player i is unable to check whether a profile is an equilibrium. Even, in discrete time, one could consider a sequence of players i (with the same characteristics) where player i_n acting at stage n only knows the sequence of $\{\Phi^i(x_m); m < n\}$ and not the sequence of previous moves $\{x^i_m; m < n\}$: the behavior at stage m is only a function of the past vector outcomes up to stage m, and does not depend on the past behavior up to that stage.

Convergence to the set of equilibria occurs in two classes:
a) potential games where the dynamics essentially mimick gradient dynamics (even if the players do not know it),
b) dissipative games, that extend results for 0-sum games.

It is interesting to see that two properties of gradient of convex functions are used for the evaluation Φ: gradient in case a), monotone operator in case b).

In fact the relation with minimization of convex functions is a very wide and active area of research, see e.g. the recent books: Bubeck (2015) [17], Hazan (2011) [42], Hazan (2019) [43], Shalev-Shwartz (2012) [72].

This corresponds to the case $u_t = -\nabla f(x_t)$ for f convex.

A last remark is that the analysis of Replicator Dynamics can be performed at two levels:
either using (6) which corresponds to the differential approach and an incremental discrete dynamics: $x_{k+1} - x_k = \delta_k T(u_k)$
or using (8) which is the integral approach and is linked to a cumulative discrete dynamics of the form $x_{k+1} = S(\alpha_k \sum_{j=1}^{k} u_j)$.

Using strategies in the class $\mathcal{P}$ follows the second formulation and is related to the dual averaging procedures introduced by Nesterov (2009) [60].

6.2. **Research directions and open problems.**

6.2.1. *Correlated equilibria.* Recall that one can build from a family of no ER procedures a no IR procedure. It would be interesting to see the impact of this construction on the underlying dynamics, for example starting with Replicator Dynamics applied to several processes and then "mixed" through some invariant distribution.

In fact RD satisfies Hannan's property but in general does not converge to the set of correlated equilibria, Viossat (2007) [81] gives an example of a game with a single (pure) correlated equilibrium and where replicator dynamics attributes a vanishing weight to it, see also Viossat [82], Viossat and Zapechelnyuk (2013) [83].

6.2.2. *Continuum of actions.* We describe briefly the extension of the previous approach in the "simplex case" $X^i = \Delta(S^i)$ from S^i finite to S^i compact. A first study in this direction goes back to Bomze (1990) [14].

- Population game

Each $A^i, i \in I$ is a compact subset of some euclidean space. Let $Z^i = \Delta(A^i), Z = \prod_i Z^i$. Assume that each $F^i : A^i \times Z \to \mathbb{R}, i \in I$ is continuous (for the weak* topology on Z). $F^i(a^i, z)$ is the payoff of a member of population $i \in I$ with type a^i given the configuration z.

The equilibrium condition is:

$$\sum_i \int_{A^i} F^i(u,z)[z^i(du) - \zeta^i(du)] \geq 0, \qquad \forall \zeta \in Z \tag{35}$$

so that the evaluation function is: $\Phi^i(z) = F^i(., z)$.

- I "atomic splitting" players

Let $G^i : Z \to \mathbb{R}$ be the payoff function of player i. The evaluation function Φ^i corresponds to the Gateaux derivative $\delta^i G^i$ of G^i with respect to z^i. $\Phi^i(z)$ is a continuous function f^i on A^i such that:

$$\lim_{t \to 0} \frac{G^i(z^i + t(\zeta^i - z^i), z^{-i}) - G^i(z)}{t} = \int_{A^i} f^i(u)[\zeta^i(du) - z^i(du)]$$

- Variational inequalities and dynamics

The equilibria $z = \{z^i\} \in Z$ are solutions of the varIational inequalities between continuous functions $\Phi(z)$ and probabilities ζ on A:

$$[\Phi(z), z - \zeta] = \sum_{i \in I} \int_{A^i} \Phi^i(z)(u)(z^i(du) - \zeta^i(du)) \geq 0.$$

A potential function W on Z satisfies $\delta^i W = \Phi^i, i \in I$ on the tangent space.

The dynamics are now defined on the set of measures, for example for RD:

$$\dot{z}_t^i(B) = \int_B \Phi^i(u, z_t) z_t^i(du) - z_t^i(B) \int_{A^i} \Phi^i(u, z_t) z_t^i(du).$$

See e.g. the analysis in Cheung (2016) [20] and the references therein.

6.2.3. *Convergence.* The convergence properties, either to the Hannan set, to correlated equilibria distribution or through a Lyapounov function are, in general, convergence to a set.

In the framework of Section 5, more precise results are available, implying convergence of the trajectories:

-for the replicator dynamics, in the one population symmetric case, Akin and Hofbauer (1982) [2],

- for the Hessian Riemannian gradient flows, Alvarez, Bolte and Brahic (2004) [3].

A natural research direction is to look for extensions.

REFERENCES

[1] E. Akin, *The Geometry of Population Genetics*, Springer, 1979.

[2] E. Akin and J. Hofbauer, Recurrence of the unfit, *Mathematical Biosciences*, **61** (1982), 51–62.

[3] F. Alvarez, J. Bolte and O. Brahic, Hessian Riemannian gradient flows in convex programming, *SIAM Journal on Control and Optimization*, **43** (2004), 477–501.

[4] S. Arora, E. Hazan and S. Kale, The multiplicative weights update method: A meta algorithm and applications, *Theory of Computing*, **8** (2012), 121–164.

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

[6] R. J. Aumann, Subjectivity and correlation in randomized strategies, *Journal of Mathematical Economics*, **1** (1974), 67–96.

[7] M. Benaim, Dynamics of stochastic approximation algorithms, *Séminaire de Probabilités, XXXIII*, **1709** (1999), 1–68.

[8] M. Benaim and M. Faure, Consistency of vanishingly smooth fictitious play, *Mathematics of Operations Research*, **38** (2013), 437–450.

[9] M. Benaim, J. Hofbauer and S. Sorin, Stochastic approximations and differential inclusions, *SIAM J. Opt. and Control*, **44** (2005), 328–348.

[10] M. Benaim, J. Hofbauer and S. Sorin, Stochastic approximations and differential inclusions. Part II: Applications, *Mathematics of Operations Research*, **31** (2006), 673–695.

[11] M. Benaim, J. Hofbauer and S. Sorin, Perturbations of set–valued dynamical systems, with applications to game theory, *Dynamic Games and Applications*, **2** (2012), 195–205.

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

[13] A. Blum and Y. Mansour, From external to internal regret, *Journal of Machine Learning Reserach*, **8** (2007), 1307–1324.

[14] I. M. Bomze, Dynamic aspects of evolutionary stability, *Monatshefte Math.*, **110** (1990), 189–206.

[15] G. W. Brown, Some notes on computation of games solutions, Report P–78, The Rand Corporation, 1949.

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

[17] S. Bubeck, Convex optimization: Algorithms and complexity, *Fondations and Trends in Machine Learning*, **8** (2015), 231–357.

[18] N. Cesa–Bianchi and G. Lugosi, Potential–based algorithms in on–line prediction and game theory, *Computational Learning Theory (Amsterdam, 2001),* 48–64, Lecture Notes in Comput. Sci., 2111, Lecture Notes in Artificial Intelligence, Springer, Berlin, 2001.

[19] N. Cesa–Bianchi and G. Lugosi, *Prediction, Learning and Games*, Cambridge University Press, 2006.

[20] M.–W. Cheung, Imitative dynamics for games with continuous strategy space, *Games and Economic Behavior*, **99** (2016), 206–223.

[21] P. Coucheney, B. Gaujal and P. Mertikopoulos, Penalty–regulated dynamics and robust learning procedures in games, *Mathematics of Operations Research*, **40** (2015), 611–633.

[22] T. Cover, Universal portfolios, *Math. Finance*, **1** (1991), 1–29.

[23] S. C. Dafermos, Traffic equilibrium and variational inequalities, *Transportation Sci.*, **14** (1980), 42–54.

[24] P. Dupuis and A. Nagurney, Dynamical systems and variational inequalities, *Ann. Oper. Res.*, **44** (1993), 9–42.

[25] M. Faure, P. Gaillard, B. Gaujal and V. Perchet, Online learning and game theory. A quick overview with recent results and applications, *ESAIM: Proceedings and Surveys*, **51** (2015), 246–271.

[26] D. Foster and R. Vohra, A randomization rule for selecting forecasts, *Operations Research*, **41** (1993), 704–707.

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

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

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

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

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

[32] A. Gaunersdorfer and J. Hofbauer, Fictitious play, Shapley polygons and the replicator equation, *Games and Economic Behavior*, **11** (1995), 279–303.

[33] I. Gilboa and A. Matsui, Social stability and equilibrium, *Econometrica*, **59** (1991), 859–867.

[34] J. Hannan, Approximation to Bayes risk in repeated plays, *Contributions to the Theory of Games, III*, Drescher M., A.W. Tucker and P. Wolfe eds., Princeton University Press, (1957), 97–139.

[35] C. Harris, On the rate of convergence of continuous time fictitious play, *Games and Economic Behavior*, **22** (1998), 238–259.

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

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

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

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

[40] S. Hart and A. Mas–Colell, Uncoupled dynamics do not lead to Nash equilibria, *American Economic Review*, **93** (2003), 1830–1836.

[41] S. Hart and A. Mas Colell, Simple Adaptive Strategies: From Regret–Matching to Uncoupled Dynamics, World Scientific Publishing, 2013.

[42] E. Hazan, The convex optimization approach to regret minimization, *Optimization for machine learning*, S. Sra, S. Nowozin, S. Wright eds, MIT Press, (2011), 287–303.

[43] E. Hazan, *Optimization for Machine Learning*, `https://arxiv.org/pdf/1909.03550.pdf`, 2019.

[44] J. Hofbauer, From Nash and Brown to Maynard Smith: Equilibria, dynamics and ESS, *Selection*, **1**(2000), 81—-88.

[45] J. Hofbauer, Deterministic evolutionary game dynamics, in *Evolutionary Game Dynamics*, K. Sigmund ed., Proceedings of Symposia in Applied Mathematics, A.M.S., **69** (2011), 61–79.

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

[47] J. Hofbauer and W. H. Sandholm, Stable games and their dynamics, *Journal of Economic Theory*, **144** (2009), 1665–1693, 1693.e4.

[48] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge U.P., 1998.

[49] J. Hofbauer and K. Sigmund, Evolutionary game dynamics, *Bulletin of the A.M.S.*, **40** (2003), 479–519.

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

[51] J. Hofbauer, S. Sorin and Y. Viossat, Time average replicator and best reply dynamics, *Mathematics of Operations Research*, **34** (2009), 263–269.

[52] A. Kalai and S. Vempala, Efficient algorithms for online decision problems, *Journal of Computer and System Sciences*, **71** (2005), 291–307.

[53] J. Kwon and P. Mertikopoulos, A continuous time approach to on–line optimization, *Journal of Dynamics and Games*, **4** (2017), 125–148.

[54] N. Littlestone and M. K. Warmuth, The weighted majority algorithm, *Information and Computation*, **108** (1994), 212–261.
[55] J. Maynard Smith, *Evolution and the Theory of Games*, Cambridge U.P., 1982.
[56] P. Mertikopoulos and W. H. Sandholm, Learning in games via reinforcement and regularization, *Mathematics of Operations Research*, **41** (2016), 1297–1324.
[57] P. Mertikopoulos and W. H. Sandholm, Riemannian game dynamics, *Journal of Economic Theory*, **177** (2018), 315–364.
[58] P. Mertikopoulos and Z. Zhou, Learning in games with continuous action sets and unknown payoff functions, *Mathematical Programming*, **173** (2019), 465–507.
[59] D. Monderer and L. S. Shapley, Potential games, *Games Econom. Behav.*, **14** (1996), 124–143.
[60] Y. Nesterov, Primal–dual subgradient methods for convex problems, *Mathematical Programming*, **120** (2009), 221–259.
[61] V. Perchet, Approachability, regret and calibration: Implications and equivalences, *Journal of Dynamics and Games*, **1** (2014), 181–254.
[62] V. Perchet, Exponential weight approachability, applications to calibration and regret minimization, *Dynamic Games and Applications*, **5** (2015), 136–153.
[63] J. Robinson, An iterative method of solving a game, *Annals of Mathematics*, **54** (1951), 296–301.
[64] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
[65] R. T. Rockafellar, Monotone operators associated with saddle–functions and minmax problems, *Nonlinear Functional Analysis, F. Browder, ed.*, Proceedings of Symposia in Pure Math, AMS, **18** (1970), 241–250.
[66] A. Rustichini, Optimal properties of stimulus–response learning models, *Games and Economic Behavior*, **29** (1999), 244–273.
[67] W. H. Sandholm, Potential games with continuous player sets, *Journal of Economic Theory*, **97** (2001), 81–108.
[68] W. H. Sandholm, Large population potential games, *Journal of Economic Theory*, **144** (2009), 1710–1725.
[69] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, 2010.
[70] W. H. Sandholm, *Population Games and Deterministic Evolutionary Dynamics,* Economic Learning and Social Evolution. MIT Press, Cambridge, MA, 2010.
[71] S. Shahshahani, A new mathematical framework for the study of linkage and selection, *Memoirs of the American Mathematical Society*, **17** (1979), ix+34 pp.
[72] S. Shalev–Shwartz, Online Learning and Online Convex Optimization, *Foundations and Trends in Machine Learning*, **4** (2012), 107–194.
[73] M. J. Smith, The existence, uniqueness and stability of traffic equilibria, *Transportation Res., Part B*, **13** (1979), 295–304.
[74] M. J. Smith, The stability of a dynamic model of traffic assignment – an application of a method of Lyapunov, *Transportation Sci.*, **18** (1984), 245–252.
[75] S. Sorin, Exponential weight algorithm in continuous time, *Mathematical Programming*, Ser. B, **116** (2009), 513–528.
[76] S. Sorin, On some global and unilateral adaptive dynamics, *Evolutionary Game Dynamics*, K. Sigmund (ed.), Proceedings of Symposia in Applied Mathematics, **69** (2011), 81–109.
[77] S. Sorin and C. Wang, Finite composite games: Equilibria and dynamics, *Journal of Dynamics and Games*, **3** (2016), 101–120.
[78] G. Stoltz and G. Lugosi, Internal regret in on–line portfolio selection, *Machine Learning*, **59** (2005), 125–159.
[79] J. M. Swinkels, Adjustment dynamics and rational play in games, *Games and Economic Behavior*, **5** (1983), 455–484.
[80] P. Taylor and L. Jonker, Evolutionary stable strategies and game dynamics, *Mathematical Biosciences*, **40** (1978), 145–156.
[81] Y. Viossat, The replicator dynamics does not lead to correlated equilibria, *Games and Economic Behavior*, **59** (2007), 397–407.
[82] Y. Viossat, Evolutionary dynamics may eliminate all strategies used in correlated equilibrium, *Mathematical Social Sciences*, **56** (2008), 27–43.
[83] Y. Viossat and A. Zapechelnyuk, No–regret dynamics and fictitious play, *Journal of Economic Theory*, **148** (2013), 825–842.
[84] V. Vovk, Aggregating strategies, *Proceedings of the 3rd Annual Conference on Computational Learning Theory*, (1990), 371–383.

[85] G. Wardrop, Some theoretical aspects of road traffic research, *Proc. Inst. Civ. Eng.*, **1** (1952), 325–362.

Received for publication October 2019.

E-mail address: sylvain.sorin@imj-prg.fr