SIAM J. CONTROL OPTIM.
Vol. 44, No. 1, pp. 328–348

STOCHASTIC APPROXIMATIONS AND DIFFERENTIAL INCLUSIONS*

MICHEL BENAÏM†, JOSEF HOFBAUER‡, AND SYLVAIN SORIN§

Abstract. The dynamical systems approach to stochastic approximation is generalized to the case where the mean differential equation is replaced by a differential inclusion. The limit set theorem of Benaïm and Hirsch is extended to this situation. Internally chain transitive sets and attractors are studied in detail for set-valued dynamical systems. Applications to game theory are given, in particular to Blackwell's approachability theorem and the convergence of fictitious play.

Key words. stochastic approximation, differential inclusions, set-valued dynamical systems, chain recurrence, approachability, game theory, learning, fictitious play

AMS subject classifications. 62L20, 34G25, 37B25, 62P20, 91A22, 91A26, 93E35, 34F05

DOI. 10.1137/S0363012904439301

1. Introduction.

1.1. Presentation. A powerful method for analyzing stochastic approximations or recursive stochastic algorithms is the so-called ODE (ordinary differential equation) method, which allows us to describe the limit behavior of the algorithm in terms of the asymptotics of a certain ODE,

$$\frac{dx}{dt} = F(x),$$

obtained by suitable averaging.

This method was introduced by Ljung [24] and extensively studied thereafter (see, e.g., the books by Kushner and Yin [23] or Duflo [14] for a comprehensive introduction and further references). However, until recently most works in this direction have assumed the simplest dynamics for F, for example, that F is linear or given by the gradient of a cost function. While this type of assumption makes perfect sense in engineering applications (where algorithms are often designed to minimize a cost function), there are several situations, including models of learning or adaptive behavior in games, for which F may have more complicated dynamics.

In a series of papers Benaïm [2, 3] and Benaïm and Hirsch [5] have demonstrated that the asymptotic behavior of stochastic approximation processes can be described with a great deal of generality beyond gradients and other simple dynamics. One of their key results is that the limit sets of the process are almost surely *compact connected attractor free* (or *internally chain transitive* in the sense of Conley [13]) for the deterministic flow induced by F.

*Received by the editors January 6, 2004; accepted for publication (in revised form) November 23, 2004; published electronically August 22, 2005. This research was partially supported by the Austrian Science Fund P15281 and the Swiss National Science Foundation grant 200021-1036251/1.
http://www.siam.org/journals/sicon/44-1/43930.html

†Institut de Mathematiques, Université de Neuchâtel, Rue Emile-Argand 11, Neuchâtel, Switzerland (michel.benaim@unine.ch).

‡Department of Mathematics, University College London, London WC1E 6BT, UK, and Institut für Mathematik, Universität Wien, Nordbergstrasse 15, 1090 Wien, Austria (jhofb@math.ucl.ac.uk).

§Laboratoire d'Econométrie, Ecole Polytechnique, 1 rue Descartes, 75005 Paris, France, and Equipe Combinatoire et Optimisation, UFR 929, Université P. et M. Curie - Paris 6, 175 Rue du Chevaleret, 75013 Paris, France (sorin@math.jussieu.fr).

The purpose of this paper is to show that such a dynamical system approach easily extends to the situation where the mean ODE is replaced by a differential inclusion. This is strongly motivated by certain problems arising in economics and game theory. In particular, the results here allow us to give a simple and unified presentation of Blackwell's approachability theorem, Smale's results on the prisoner's dilemma, and convergence of fictitious play in potential games. Many other applications[1] will be considered in a forthcoming paper, by Benaïm, Hofbauer, and Sorin [7], the present one being mainly devoted to theoretical issues.

The organization of the paper is as follows. Part 1 introduces the different notions of solutions, perturbed solutions, and stochastic approximations associated with a differential inclusion. Part 2 is devoted to the presentation of two classes of examples. Part 3 is a general study of the dynamical system defined by a differential inclusion. The main result (Theorem 3.6) on the limit set of a perturbed solution being internally chain transitive is stated. Then related notions—invariant and attracting sets, attractors, and Lyapunov functions—are analyzed. Part 4 contains the proof of the limit set theorem. Finally, Part 5 applies the previous results to two adaptive processes in game theory: approachability and fictitious play.

1.2. The differential inclusion. Let F denote a set-valued function mapping each point $x \in \mathbb{R}^m$ to a set $F(x) \subset \mathbb{R}^m$. We suppose throughout that the following holds.

Hypothesis 1.1 (standing assumptions on F).

(i) F is a closed set-valued map. That is,

$$\mathrm{Graph}(F) = \{(x, y) : y \in F(x)\}$$

is a closed subset of $\mathbb{R}^m \times \mathbb{R}^m$.

(ii) $F(x)$ is a nonempty compact convex subset of $\mathbb{R}^m$ for all $x \in \mathbb{R}^m$.

(iii) There exists $c > 0$ such that for all $x \in \mathbb{R}^m$

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

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

Definition I. *A* solution *for the differential inclusion*

$$\frac{d\mathbf{x}}{dt} \in F(\mathbf{x}) \tag{I}$$

with initial point $x \in \mathbb{R}^m$ is an absolutely continuous mapping $\mathbf{x} : \mathbb{R} \to \mathbb{R}^m$ such that $\mathbf{x}(0) = x$ and

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

for almost every $t \in \mathbb{R}$.

Under the above assumptions, it is well known (see Aubin and Cellina [1, Chapter 2.1] or Clarke et al. [12, Chapter 4.1]) that (I) admits (typically nonunique) solutions through every initial point.

[1]As pointed out to us by an anonymous referee, applications to resource sharing may be considered as in Buche and Kushner [11], where the dynamics are given by a differential inclusion. Possible applications to engineering include dry friction; see, e.g., Kunze [22].

Remark 1.2. Suppose that a differential inclusion is given on a compact convex set $C \subset \mathbb{R}^m$, of the form $F(x) = \Phi(x) - x$, such that $\Phi(x) \subset C$ for all $x \in C$ and Φ satisfies Hypothesis 1.1(i) and (ii), with $\mathbb{R}^m$ replaced by C. Then we can extend it to a differential inclusion defined on the whole space $\mathbb{R}^m$: For $x \in \mathbb{R}^m$ let $P(x) \in C$ denote the unique point in C closest to x, and define $F(x) = \Phi(P(x)) - x$. Then F satisfies Hypothesis 1.1.

1.3. Perturbed solutions. The main object of this paper is paths which are obtained as certain (deterministic or random) perturbations of solutions of (I).

DEFINITION II. *A continuous function* $\mathbf{y} : \mathbb{R}_+ = [0, \infty) \to \mathbb{R}^m$ *will be called a* perturbed solution *to* (I) *(we also say a perturbed solution to* F*) if it satisfies the following set of conditions* (II):

(i) $\mathbf{y}$ *is absolutely continuous.*

(ii) *There exists a locally integrable function* $t \mapsto U(t)$ *such that*

(a)

$$\lim_{t\to\infty} \sup_{0\le v\le T} \left\| \int_t^{t+v} U(s)\, ds \right\| = 0$$

for all $T > 0$*; and*

(b) $\frac{d\mathbf{y}(t)}{dt} - U(t) \in F^{\delta(t)}(\mathbf{y}(t))$ *for almost every* $t > 0$, *for some function* $\delta : [0, \infty) \to \mathbb{R}$ *with* $\delta(t) \to 0$ *as* $t \to \infty$. *Here* $F^\delta(x) := \{y \in \mathbb{R}^m : \exists z : \|z - x\| < \delta,\ d(y, F(z)) < \delta\}$ *and* $d(y, C) = \inf_{c\in C} \|y - c\|$.

The purpose of this paper is to investigate the long-term behavior of $\mathbf{y}$ and to describe its limit set

$$L(\mathbf{y}) = \bigcap_{t\ge 0} \overline{\{\mathbf{y}(s) : s \ge t\}}$$

in terms of the dynamics induced by F.

1.4. Stochastic approximations. As will be shown here, a natural class of perturbed solutions to F arises from certain stochastic approximation processes.

DEFINITION III. *A discrete time process* $\{x_n\}_{n\in\mathbb{N}}$ *living in* $\mathbb{R}^m$ *is a* solution for (III) *if it verifies a recursion of the form*

$$x_{n+1} - x_n - \gamma_{n+1} U_{n+1} \in \gamma_{n+1} F(x_n), \tag{III}$$

where the characteristics γ *and* U *satisfy*

- $\{\gamma_n\}_{n\ge 1}$ *is a sequence of nonnegative numbers such that*

$$\sum_n \gamma_n = \infty, \qquad \lim_{n\to\infty} \gamma_n = 0;$$

- $U_n \in \mathbb{R}^m$ *are (deterministic or random) perturbations.*

To such a process is naturally associated a continuous time process as follows.

DEFINITION IV. *Set*

$$\tau_0 = 0 \quad \text{and} \quad \tau_n = \sum_{i=1}^n \gamma_i \quad \text{for } n \ge 1,$$

and define the continuous time affine interpolated process $\mathbf{w} : \mathbb{R}_+ \to \mathbb{R}^m$ *by*

$$\mathbf{w}(\tau_n + s) = x_n + s\frac{x_{n+1} - x_n}{\tau_{n+1} - \tau_n}, \quad s \in [0, \gamma_{n+1}). \tag{IV}$$

1.5. From interpolated process to perturbed solutions. The next result gives sufficient conditions on the characteristics of the discrete process (III) for its interpolation (IV) to be a perturbed solution (II). If (U_i) are random variables, assumptions (i) and (ii) below have to be understood with probability one.

PROPOSITION 1.3. *Assume that the following hold:*

(i) *For all* $T > 0$

$$\lim_{n\to\infty} \sup\left\{\left\|\sum_{i=n}^{k-1} \gamma_{i+1} U_{i+1}\right\| : k = n+1, \ldots, m(\tau_n + T)\right\} = 0,$$

where

$$m(t) = \sup\{k \geq 0 : t \geq \tau_k\}; \tag{1.1}$$

(ii) $\sup_n \|x_n\| = M < \infty$.

Then the interpolated process $\mathbf{w}$ *is a perturbed solution of* F.

Proof. Let $\mathbf{U}, \gamma : \mathbb{R}_+ \to \mathbb{R}^m$ denote the continuous time processes defined by

$$\mathbf{U}(\tau_n + s) = U_{n+1}, \qquad \gamma(\tau_n + s) = \gamma_{n+1}$$

for all $n \in \mathbb{N}$, $0 \leq s < \gamma_{n+1}$.

Then, for any t,

$$\mathbf{w}(t) \in x_{m(t)} + (t - \tau_{m(t)})[\mathbf{U}(t) + F(x_{m(t)})];$$

hence

$$\dot{\mathbf{w}}(t) \in \mathbf{U}(t) + F(x_{m(t)}).$$

Let us set $\delta(t) = \|\mathbf{w}(t) - x_{m(t)}\|$. Then obviously

$$F(x_{m(t)}) \subset F^{\delta(t)}(\mathbf{w}(t)).$$

In addition,

$$\delta(t) \leq \gamma_{m(t)+1}[\|U_{m(t)+1}\| + c(1 + M)]$$

hence goes to 0, using hypothesis (i) of the statement of the proposition. It remains to check condition (ii)(a) of (II), but one has

$$\left\|\int_t^{t+v} \mathbf{U}(s)ds\right\| \leq \gamma_{m(t)+1}\|U_{m(t)+1}\| + \left\|\sum_{\ell=m(t)+1}^{m(t+v)-1} \gamma_{\ell+1} U_{\ell+1}\right\| + \gamma_{m(t+v)+1}\|U_{m(t+v)+1}\|,$$

and the result follows from condition (i). ∎

Sufficient conditions. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_n\}_{n\geq 0}$ a filtration of $\mathcal{F}$ (i.e., a nondecreasing sequence of sub-σ-algebras of $\mathcal{F}$). We say that a stochastic process $\{x_n\}$ given by (III) satisfies the *Robbins–Monro condition with martingale difference noise* (Kushner and Yin [23]) if its characteristics satisfy the following:

(i) $\{\gamma_n\}$ is a deterministic sequence.

(ii) $\{U_n\}$ is *adapted* to $\{\mathcal{F}_n\}$. That is, U_n is measurable with respect to $\mathcal{F}_n$ for each $n \geq 0$.

(iii) $\mathsf{E}(U_{n+1} \mid \mathcal{F}_n) = 0$.

The next proposition is a classical estimate for stochastic approximation processes. Note that F does not appear. We refer the reader to (Benaïm [3, Propositions 4.2 and 4.4]) for a proof and further references.

PROPOSITION 1.4. *Let $\{x_n\}$ given by* (III) *be a Robbins–Monro equation with martingale difference noise process. Suppose that one of the following condition holds:*

(i) *For some $q \geq 2$*

$$\sup_n \mathsf{E}(\|U_n\|^q) < \infty$$

and

$$\sum_n \gamma_n^{1+q/2} < \infty.$$

(ii) *There exists a positive number Γ such that for all $\theta \in \mathbb{R}^m$*

$$\mathsf{E}(\exp(\langle \theta, U_{n+1} \rangle) \mid \mathcal{F}_n) \leq \exp\left(\frac{\Gamma}{2}\|\theta\|^2\right)$$

and

$$\sum_n e^{-c/\gamma_n} < \infty$$

for each $c > 0$.
Then assumption (i) *of Proposition* 1.3 *holds with probability* 1.

Remark 1.5. Typical applications are

(i) U_n uniformly bounded in L^2 and $\gamma_n = \frac{1}{n}$,

(ii) U_n uniformly bounded and $\gamma_n = o(\frac{1}{\log n})$.

2. Examples.

2.1. A multistage decision making model. Let A and B be measurable spaces, respectively called the *action space* and the *states of nature*; $E \subset \mathbb{R}^m$ a convex compact set called the *outcomes space*; and $H : A \times B \to E$ a measurable function, called the *outcome* function.

At discrete times $n = 1, 2\ldots$ a decision maker (DM) chooses an action a_n from A and observes an outcome $H(a_n, b_n)$. We suppose the following.

(A) The sequence $\{a_n, b_n\}_{n\geq 0}$ is a random process defined on some probability space $(\Omega, \mathcal{F}, P)$ and adapted to some filtration $\{\mathcal{F}_n\}$. Here $\mathcal{F}_n$ has to be understood as the history of the process until time n.

(B) Given the history $\mathcal{F}_n$, DM and nature act independently:

$$\mathsf{P}((a_{n+1}, b_{n+1}) \in da \times db \mid \mathcal{F}_n) = \mathsf{P}(a_{n+1} \in da \mid \mathcal{F}_n)\mathsf{P}(b_{n+1} \in db \mid \mathcal{F}_n)$$

for any measurable sets $da \subset A$ and $db \subset B$.

(C) DM keeps track of only the cumulative average of the past outcomes,

$$x_n = \frac{1}{n}\sum_{i=1}^{n} H(a_i, b_i), \tag{2.1}$$

and his decisions are based on this average. That is,

$$\mathsf{P}(a_{n+1} \in da \mid \mathcal{F}_n) = Q_{x_n}(da),$$

where $Q_x(\cdot)$ is a probability measure over A for each $x \in E$, and $x \in E \mapsto Q_x(da) \in [0,1]$ is measurable for each measurable set $da \subset A$. The family $Q = \{Q_x\}_{x\in E}$ is called a *strategy* for DM.

Assumption (C) can be justified by considerations of limited memory and bounded rationality. It is partially motivated by Smale's approach to the prisoner's dilemma [27] (see also Benaïm and Hirsch [4, 5]), Blackwell's approachability theory ([8]; see also Sorin [28]), as well as fictitious play (Brown [10], Robinson [26]) and stochastic fictitious play (Benaïm and Hirsch [6], Fudenberg and Levine [15], Hofbauer and Sandholm [20]) in game theory (see the examples below).

For each $x \in E$ let

$$C(x) = \left\{ \int_{A\times B} H(a,b) Q_x(da)\nu(db) : \nu \in \mathcal{P}(B) \right\},$$

where $\mathcal{P}(B)$ denotes the set of probability measures over B. Then clearly

$$\mathsf{E}(H(a_{n+1}, b_{n+1}) \mid \mathcal{F}_n) \in C(x_n) \subset \overline{C}(x_n),$$

where $\overline{C}$ denote the smallest closed set-valued extension of C with convex values. More precisely, the graph of $\overline{C}$ is the intersection of all closed subsets $G \subset E \times E$ for which the fiber $G_x = \{y \in E : (x,y) \in G\}$ is convex and contains $C(x)$.

For $x \in \mathbb{R}^m$ let $P(x)$ denote the unique point in E closest to x. Extend $\overline{C}$ as in Remark 1.2 to a set-valued map on $\mathbb{R}^m$ by setting

$$\widehat{C}(x) = \overline{C}(P(x)).$$

Then the map

$$F(x) = -x + \overline{C}(P(x)) = -x + \widehat{C}(x) \tag{2.2}$$

clearly satisfies Hypothesis 1.1, and $\{x_n\}$ verifies the recursion

$$x_{n+1} - x_n = \frac{1}{n+1}(-x_n + H(a_{n+1}, b_{n+1})),$$

which can be rewritten as (see (III))

$$x_{n+1} - x_n \in \gamma_{n+1}[F(x_n) + U_{n+1}]$$

with $\gamma_n = \frac{1}{n}$ and $U_{n+1} = H(a_{n+1}, b_{n+1}) - \int_A H(a, b_{n+1}) Q_{x_n}(da)$. Hence, the conditions of Proposition 1.4 are satisfied and one deduces the following claim.

Proposition 2.1. *The affine continuous time interpolated process* (IV) *of the process* $\{x_n\}$ *given by* (2.1) *is almost surely a perturbed solution of* F *defined by* (2.2).

Example 2.2 (Blackwell's approachability theory). A set $\Lambda \subset E$ is said to be *approachable* if there exists a strategy Q such that $x_n \to \Lambda$ almost surely. Blackwell [8] gives conditions ensuring approachability. We will show in section 5.1 how Blackwell's results can be partially derived from our main results and generalized (Corollary 5.2) in certain directions.

2.2. Learning in games. The preceding formalism is well suited to analyzing certain models of learning in games.

Consider the situation where m players are playing a game over and over. Let A^i (for $i \in I = \{1, \dots, m\}$) be a finite set representing the actions (pure strategies) available to player i, and let X^i be the finite dimensional simplex of probabilities over A^i (the set of mixed strategies for player i). For $i \in I$ we let A^{-i} and X^{-i} respectively denote the actions and mixed strategies available to the opponents of i. The payoff function to player i is given by a function $U^i : A^i \times A^{-i} \to \mathbb{R}$. As usual, we extend U^i to a function (still denoted U^i) on $X^i \times X^{-i}$, by multilinearity.

Example 2.3 (fictitious and stochastic fictitious play). Consider the game from the viewpoint of player i so that the DM is player i, and "nature" is given by the other players. In fictitious or stochastic fictitious play the outcome space is the space $X^i \times X^{-i}$ of mixed strategies, and the outcome function is the "identity" function $H : A^i \times A^{-i} \to X^i \times X^{-i}$ mapping every profile of actions a to the corresponding profile of mixed strategy δ_a.

Let

$$BR^i(x^{-i}) = \operatorname*{Argmax}_{a^i \in A^i} U^i(a^i, x^{-i}) \subset A^i$$

be the set of best actions that i can play in response to x^{-i}.

Both classical fictitious play (Brown [10], Robinson [26]) and stochastic fictitious play (Benaïm and Hirsch [6], Fudenberg and Levine [15], Hofbauer and Sandholm [20]) assume that the strategy of player i, $Q^i = \{Q^i_x\}$, can be written as

$$Q^i_x(a^i) = q^i(a^i, x^{-i}),$$

where $q^i : A^i \times X^{-i} \to [0,1]$ is such that one of the following assumptions holds:

fictitious play assumption:

$$\sum_{a^i \in BR^i(x^{-i})} q^i(a^i, x^{-i}) = 1,$$

or *stochastic fictitious play assumption*, q^i is smooth in x^{-i} and

$$\sum_{a^i \in BR^i(x^{-i})} q^i(a^i, x^{-i}) \geq 1 - \delta$$

for some $0 < \delta \ll 1$.

In this framework, if a_ℓ denotes the profile of actions at stage ℓ, one has

$$x_n = \frac{1}{n} \sum_{\ell=1}^{n} a_\ell$$

and

$$x_{n+1} - x_n = \frac{1}{n+1}(a_{n+1} - x_n).$$

Thus for each i

$$\mathsf{E}(x^i_{n+1} - x^i_n \mid \mathcal{F}_n) \in \frac{1}{n+1}(\overline{BR}^i(x^{-i}_n) - x^i_n),$$

where $\overline{BR}^i(x^{-i}) \subset X^i$ is the convex hull of $BR^i(x^{-i})$ for the standard fictitious play, and $\overline{BR}^i(x^{-i}) = \sum_{a^i \in A^i} q^i(a^i, x^{-i})\delta_{a^i}$ for the stochastic fictitious play.

Thus the set-valued map F defined in (2.2) is given as

$$F^i(x) = -x + \overline{BR}^i(x^{-i}) \times X^{-i}.$$

Observe that if a subset $J \subset I$ of players plays a fictitious (or stochastic fictitious) play strategy, then F^i has to be replaced by

$$F^J(x) = \bigcap_{i \in J} F^i(x).$$

In particular, if all players play a fictitious play strategy, the differential inclusion induced by F is the best-response differential inclusion (Gilboa and Matsui [16], Hofbauer [19], Hofbauer and Sorin [21]), while if all play a stochastic fictitious play, F is a smooth best-response vector field (Benaïm and Hirsch [6], Fudenberg and Levine [15], Hofbauer and Sandholm [20]).

Example 2.4 (Smale approach to the prisoner's dilemma). We still consider the game from the viewpoint of player i, so that the DM is player i and nature the other players, but we take for H the payoff vector function

$$H : A^i \times A^{-i} \to E,$$
$$a \to U(a) = (U^1(a), \dots, U^m(a)),$$

where $E \subset \mathbb{R}^m$ is the convex hull of the payoff vectors $\{U(a)\}$.

This setting fits exactly with Smale's approach to the prisoner's dilemma [27] later revisited by Benaïm and Hirsch [4]. Details will be given in section 5.2, where Smale's approach will be reinterpreted in the framework of approachability.

3. Set-valued dynamical systems.

3.1. Properties of the trajectories of (I). Let $C^0(\mathbb{R}, \mathbb{R}^m)$ denote the space of continuous paths $\{\mathbf{z} : \mathbb{R} \to \mathbb{R}^m\}$ equipped with the topology of uniform convergence on compact intervals. This is a complete metric space for the distance $\mathbf{D}$ defined by

$$\mathbf{D}(\mathbf{x}, \mathbf{z}) = \sum_{k=1}^{\infty} \frac{1}{2^k} \min(\|\mathbf{x} - \mathbf{z}\|_{[-k,k]}, 1),$$

where $\|\cdot\|_{[-k,k]}$ stands for the supremum norm on $C^0([-k,k], \mathbb{R}^m)$.

Given a set $M \subset \mathbb{R}^m$, we let $S_M \subset C^0(\mathbb{R}, \mathbb{R}^m)$ denote the set of all solutions to (I) with initial conditions $x \in M$ ($S_M = \bigcup_{x \in M} S_x$), and $S_{M,M} \subset S_M$ the subset consisting of solutions $\mathbf{x}$ that remain in M (i.e., $\mathbf{x}(\mathbb{R}) \subset M$).

LEMMA 3.1. *Assume M compact. Then S_M is a nonempty compact set and $S_{M,M}$ is a compact (possibly empty) set.*

Proof. The first assertion follows from Aubin and Cellina [1, section 2.2, Theorem 1, p. 104]. The second easily follows from the first. ∎

3.2. Set-valued dynamical system induced by (I). The differential inclusion (I) induces a set-valued dynamical system $\{\Phi_t\}_{t \in \mathbb{R}}$ defined by

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

The family $\Phi = \{\Phi_t\}_{t \in \mathbb{R}}$ enjoys the following properties:

(a) $\Phi_0(x) = \{x\}$;
(b) $\Phi_t(\Phi_s(x)) = \Phi_{t+s}(x)$ for all $t, s \geq 0$;
(c) $y \in \Phi_t(x) \Rightarrow x \in \Phi_{-t}(y)$ for all $x, y \in \mathbb{R}^m$, $t \in \mathbb{R}$;
(d) $(x,t) \mapsto \Phi_t(x)$ is a closed set-valued map with compact values (i.e., $\Phi_t(x)$ is a compact set for each t and x).

Properties (a), (b), (c) are immediate to verify, and property (d) easily follows from Lemma 3.1.

For subsets $T \subset \mathbb{R}$ and $A \subset \mathbb{R}^m$ we will define

$$\Phi_T(A) = \bigcup_{t \in T} \bigcup_{x \in A} \Phi_t(x).$$

Invariant sets.

DEFINITION V. *A set $A \subset \mathbb{R}^m$ is said to be*

(i) strongly invariant *(for Φ) if $A = \Phi_t(A)$ for all $t \in \mathbb{R}$;*
(ii) quasi-invariant *if $A \subset \Phi_t(A)$ for all $t \in \mathbb{R}$;*
(iii) semi-invariant *if $\Phi_t(A) \subset A$ for all $t \in \mathbb{R}$;*
(iv) invariant *(for F) if for all $x \in A$ there exists a solution $\mathbf{x}$ to* (I) *with $\mathbf{x}(0) = x$ and such that $\mathbf{x}(\mathbb{R}) \subset A$.*

We call a set A *strongly positive invariant* if $\Phi_t(A) \subset A$ for all $t > 0$.

At first glance (at least for those used to ordinary differential equations) the good notion might seem to be the one defined by strong invariance. However, this notion is too strong for differential inclusions, as shown by the simple example below (Example 3.2), and the main notions that will really be needed here are invariance and strong positive invariance. We have included the definition of quasi invariance mainly because some of our later results may be related to a paper by Bronstein and Kopanskii [9] making use of this notion.[2] Observe, however, that by Lemma 3.3 below, quasi invariance coincides with invariance for compact sets.

Example 3.2. (a) Let F be the set-valued map defined on $\mathbb{R}$ by $F(x) = -\operatorname{sgn}(x)$ if $x \neq 0$ and $F(0) = [-1, 1]$. Then $\Phi_t(0) = \{0\}$ for $t \geq 0$, and $\Phi_t(0) = [t, -t]$ for $t < 0$. Hence $\{0\}$ is invariant and strongly positively invariant but is not strongly invariant.

(b) Let now $F(x) = x$ for $x < 0$, $F(x) = 1$ for $x > 0$, and $F(0) = [0, 1]$. Then $\Phi_t(0) = \{0\}$ for $t \leq 0$, and $\Phi_t(0) = [0, t]$ for $t \geq 0$. Hence $\{0\}$ is invariant but not strongly positively invariant.

LEMMA 3.3. *Every invariant set is quasi-invariant. Every compact quasi-invariant set is invariant.*

Proof. Suppose that A is invariant. Let $x \in A$ and $\mathbf{x}$ be a solution to (I) with $\mathbf{x}(0) = x$ and $\mathbf{x}(\mathbb{R}) \subset A$. For all $t \in \mathbb{R}$ we have $x \in \Phi_t(\mathbf{x}(-t))$. Hence A is quasi-invariant.

Conversely suppose that A is quasi-invariant and compact. Choose $x \in A$ and fix $N \in \mathbb{N}$. Then for every $p \in \mathbb{N}$ there exists, by quasi invariance and by gluing pieces of solutions together, a solution $\mathbf{x}_{p,N}$ to (I) such that $\mathbf{x}_{p,N}(0) = x$ and for all $q \in \{-2^p, \ldots, 2^p\}$, $\mathbf{x}_{p,N}(\frac{qN}{2^p}) \in A$. By Lemma 3.1, the sequence $\{\mathbf{x}_{p,N}\}_{p \in \mathbb{N}}$ is relatively compact in $C^0([-N, N], \mathbb{R}^m)$. Let $\mathbf{x}_N$ be a limit point of this sequence. Then for each dyadic point $t = \frac{qN}{2^p}$, where $q \in \{-2^p, \ldots, 2^p\}$, $\mathbf{x}_N(t) \in \overline{A}$. Continuity of $\mathbf{x}_N$ implies $\mathbf{x}_N([-N, N]) \subset \overline{A}$. Now let $\mathbf{x}$ be a limit point of the sequence $\{\mathbf{x}_N\}_{N \in \mathbb{N}}$ in $C^0(\mathbb{R}, \mathbb{R}^m)$. Then $\mathbf{x}(\mathbb{R}) \subset \overline{A}$ and $\mathbf{x}$ is a solution to (I). □

[2] Invariant sets in Bronstein and Kopanskii [9] coincide with what we define here as strongly invariant sets.

Remark 3.4. A invariant together with strong positive invariance implies $\Phi_t(A) = A$ for $t > 0$.

3.3. Chain-recurrence and the limit set theorem. Given a set $A \subset \mathbb{R}^m$ and $x, y \in A$, we write $x \hookrightarrow_A y$ if for every $\varepsilon > 0$ and $T > 0$ there exists an integer $n \in \mathbb{N}$, solutions $\mathbf{x}_1, \ldots, \mathbf{x}_n$ to (I), and real numbers $t_1, t_2, \ldots, t_n$ greater than T such that

(a) $\mathbf{x}_i(s) \in A$ for all $0 \leq s \leq t_i$ and for all $i = 1, \ldots, n$,
(b) $\|\mathbf{x}_i(t_i) - \mathbf{x}_{i+1}(0)\| \leq \varepsilon$ for all $i = 1, \ldots, n-1$,
(c) $\|\mathbf{x}_1(0) - x\| \leq \varepsilon$ and $\|\mathbf{x}_n(t_n) - y\| \leq \varepsilon$.

The sequence $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is called an (ε, T) chain (in A from x to y) for F.

DEFINITION VI. *A set $A \subset \mathbb{R}^m$ is said to be* internally chain transitive*, provided that A is compact and $x \hookrightarrow_A y$ for all $x, y \in A$.*

LEMMA 3.5. *An internally chain transitive set is invariant.*

Proof. Let A be such a set and $x \in A$. Let $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ be an (ε, T) chain from x to x. Set $\mathbf{y}_{\varepsilon,T}(t) = \mathbf{x}_1(t)$ for $0 \leq t \leq T$ and $\mathbf{z}_{\varepsilon,T}(t) = \mathbf{x}_n(t_n + t)$ for $-T \leq t \leq 0$. By Lemma 3.1 we can extract from $(\mathbf{y}_{1/p,T})_{p\in\mathbb{N}}$ and $(\mathbf{z}_{1/p,T})_{p\in\mathbb{N}}$ some subsequences converging, respectively, to $\mathbf{y}_T$ and $\mathbf{z}_T$, where $\mathbf{y}_T$ and $\mathbf{z}_T$ are solutions to (I), $\mathbf{y}_T(0) = x = \mathbf{z}_T(0)$, $\mathbf{y}_T([0, T]) \subset A$, and $\mathbf{z}_T([-T, 0]) \subset A$. The map $\mathbf{w}_T(t) = \mathbf{y}_T(t)$ for $t \geq 0$ and $\mathbf{w}_T(t) = \mathbf{z}_T(t)$ for $t \leq 0$ is then a solution to (I) with initial condition x and such that $\mathbf{w}_T([-T, T]) \subset A$. By Lemma 3.1, again we extract from $(\mathbf{w}_T)_{T\geq 0}$ a subsequence converging to a solution $\mathbf{w}$ whose range lies in A and with initial condition x. ∎

This notion of recurrence due to Conley [13] for classical dynamical systems is well suited to the description of the asymptotic behavior of a perturbed solution to (I), as shown by the following theorem.

THEOREM 3.6. *Let $\mathbf{y}$ be a bounded perturbed solution to* (I)*. Then, the limit set of $\mathbf{y}$,*

$$L(\mathbf{y}) = \bigcap_{t\geq 0} \overline{\{\mathbf{y}(s) : s \geq t\}},$$

is internally chain transitive.

This theorem is the set-valued version of the limit set theorem proved by Benaïm [2] for stochastic approximation and Benaïm and Hirsch [5] for asymptotic pseudotrajectories of a flow. We will deduce it from the more general results of section 4.

3.4. Limit sets. The set

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

is the ω-limit set of a point $x \in \mathbb{R}^m$. Note that $\omega_\Phi(x)$ contains the limit sets $L(\mathbf{x})$ of all solutions $\mathbf{x}$ with $\mathbf{x}(0) = x$ but is in general larger than the union of these.

In contrast to the limit set of a solution, the ω-limit set of a point need not be internally chain transitive.

Example 3.7. Let F be the set-valued map defined on $\mathbb{R}$ by $F(x) = 1 - x$ for $x > 0$ and $F(0) = [0, 1]$ and $F(x) = -x$ for $x < 0$. Then for every solution $\mathbf{x}$, one has $\lim_{t\to\infty} \mathbf{x}(t) = 0$ or 1. But $\omega_\Phi(0) = [0, 1]$ is not internally chain transitive.

More generally one defines

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

Definition VII. *A set* Y *is* forward precompact *if* $\overline{\Phi_{[t,\infty)}(Y)}$ *is compact for some* $t > 0$.

Lemma 3.8. (i) $\omega_\Phi(Y)$ *is the set of points* $p \in \mathbb{R}^m$ *such that*

$$p = \lim_{n\to\infty} \mathbf{y}_n(t_n)$$

for some sequence $\{\mathbf{y}_n\}$ *of solutions to* (I) *with initial conditions* $\mathbf{y}_n(0) \in Y$ *and some sequence* $\{t_n\} \in \mathbb{R}$ *with* $t_n \to \infty$.

(ii) $\omega_\Phi(Y)$ *is a closed invariant (possibly empty) set. If* Y *is forward precompact, then* $\omega_\Phi(Y)$ *is nonempty and compact.*

Proof. Point (i) is easily seen from the definition.

(ii) Let $p = \lim_{n\to\infty} \mathbf{y}_n(t_n) \in \omega_\Phi(Y)$. Set $\mathbf{z}_n(s) = \mathbf{y}_n(t_n + s)$ for all $s \in \mathbb{R}$. By Lemma 3.1 we may extract from $(\mathbf{z}_n)_{n\geq 0}$ a subsequence converging to some solution $\mathbf{z}$ with $\mathbf{z}(0) = p$ and $\mathbf{z}(s) = \lim_{n_k\to\infty} \mathbf{y}_{n_k}(t_{n_k} + s) \in \omega_\Phi(Y)$. This proves invariance. The rest is clear. ◻

Note that the limit set $\omega_\Phi(Y)$ is in general not strongly positively invariant (e.g., in Example 3.7 for $x < 0$, $\omega_\Phi(x) = \{0\}$).

3.5. Attracting sets and attractors. For applications it is useful to characterize $L(\mathbf{y})$ in terms of certain compact invariant sets for Φ, namely, *the attractors,* as defined below.

Given a closed invariant set L, the induced set-valued dynamical system Φ^L is the family of (set-valued) mappings $\Phi^L = \{\Phi_t^L\}_{t\in\mathbb{R}}$ defined on L by

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

Note that L is strongly invariant for Φ^L.

Definition VIII. *A compact set* $A \subset L$ *is called an* attracting set *for* Φ^L, *provided that there is a neighborhood* U *of* A *in* L *(i.e., for the induced topology) with the property that for every* $\varepsilon > 0$ *there exists* $t_\varepsilon > 0$ *such that*

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

for all $t \geq t_\varepsilon$. *Or, equivalently,* $\Phi^L_{[t_\varepsilon,\infty)}(U) \subset N^\varepsilon(A)$. *Here* $N^\varepsilon(A)$ *stands for the* ε*-neighborhood of* A.

If, additionally, A *is invariant, then* A *is called an* attractor *for* Φ^L.

The set U *is called a* fundamental neighborhood *of* A *for* Φ^L. *If* $A \neq L$ *and* $A \neq \emptyset$, *then* A *is called a* proper attracting set *(or* proper attractor*) for* Φ^L.

Furthermore, an attracting set (respectively, attractor) for Φ *is an attracting set (respectively, attractor) for* Φ^L *with* $L = \mathbb{R}^m$.

Example 3.9. Let F be the set-valued map from Example 3.2(a), i.e., defined on $\mathbb{R}$ by $F(x) = -\operatorname{sgn}(x)$ if $x \neq 0$ and $F(0) = [-1, 1]$. Then $\{0\}$ is an attractor and every compact set $A \subset \mathbb{R}$ with $0 \in A$ is an attracting set.

Proposition 3.10. *Let* A *be a nonempty compact subset of* L, *and* U *a neighborhood of* A *in* L. *Then the following hold:*

(i) A *is an attracting set for* Φ^L *with fundamental neighborhood* U *if and only if* U *is forward precompact and* $\omega_{\Phi^L}(U) \subset A$. *In this case* $\omega_{\Phi^L}(U)$ *is an attractor.*

(ii) A *is an attractor for* Φ^L *with fundamental neighborhood* U *if and only if* U *is forward precompact and* $\omega_{\Phi^L}(U) = A$.

Proof. (i) If A is an attracting set for Φ^L with fundamental neighborhood U, then $\omega_{\Phi^L}(U) \subset \bigcap_{\varepsilon>0} N^\varepsilon(A) \subset A$. Conversely, for t large enough $V_t = \overline{\Phi^L_{[t,\infty)}(U)}$ defines a decreasing family of compact sets converging to $\omega_{\Phi^L}(U) \subset A$. Hence for any $\varepsilon > 0$ there exists t_ε with $V_{t_\varepsilon} \subset N^\varepsilon(A)$ and A is an attracting set. In particular, $\omega_{\Phi^L}(U)$ itself is an attracting set, invariant by Lemma 3.8(ii).

(ii) If $A = \omega_{\Phi^L}(U)$, then A is an attractor by (i). Conversely, if A is an attractor with fundamental neighborhood U, then $\omega_\Phi(U) \subset A$ by (i). Let $x \in A$. Since A is invariant, there exists a solution $\mathbf{y}$ to (I) with $\mathbf{y}(0) = x$ and $\mathbf{y}(\mathbb{R}) \subset A$. Set $\mathbf{y}_n(t) = \mathbf{y}(t-n)$. Then $\mathbf{y}_n(n) = x$, proving that $x \in \omega_{\Phi^L}(U)$ (by Lemma 3.8(i)). □

PROPOSITION 3.11. *Every attractor is strongly positively invariant. (Example* 3.2(a) *provides an attractor that is not strongly invariant.)*

Proof. By invariance, $A \subset \Phi_T^L(A)$ for all $T > 0$. Hence, given $t > 0$,

$$\Phi_t^L(A) \subset \Phi_{t+T}^L(A) \subset \Phi_{t+T}^L(U) \subset \Phi_{[t+T,\infty)}^L(U)$$

for all $T > 0$. Thus $\Phi_t^L(A) \subset N^\varepsilon(A)$ for all $\varepsilon > 0$, and hence $\Phi_t^L(A) \subset A$ for all $t > 0$. □

Remark 3.12. In the family of attracting sets A with a given fundamental neighborhood U, there exists a minimal one, which is in addition invariant, strongly positively invariant, and independent of the set U used to define the family. It is also the largest positively quasi-invariant set included in U.

Any attractor $A \subset L$ can be written as $A = \omega_{\Phi^L}(U)$ for some U. Hence any fundamental neighborhood uniquely determines the attractor A. This implies, as in Conley [13], that Φ^L can have at most countably many attractors.

3.6. Attractors and stability.

DEFINITION IX. *A set $A \subset L$ is* asymptotically stable *for Φ^L if it satisfies the following three conditions:*

(i) *A is invariant.*

(ii) *A is* Lyapunov stable; *i.e., for every neighborhood U of A there exists a neighborhood V of A such that $\Phi_{[0,\infty)}(V) \subset U$.*

(iii) *A is* attractive; *i.e., there is a neighborhood U of A such that for every $x \in U : \omega_\Phi(x) \subset A$.*

Alternatively, instead of (iii) *one could ask for the following weaker requirement:*

(iii′) *There is a neighborhood U of A such that for every solution $\mathbf{x}$ with $\mathbf{x}(0) \in U$ one has $L(\mathbf{x}) \subset A$.*

We show now that for compact sets the concepts of attractor and asymptotic stability are equivalent. The proof of Corollary 3.18 below shows that it makes no difference whether one uses (iii) or (iii′) in the definition of asymptotic stability.

We start with an upper bound for entry times.

LEMMA 3.13. *Let V be an open set and K compact such that for all solutions $\mathbf{x}$ with $\mathbf{x}(0) \in K$ there is $t > 0$ with $\mathbf{x}(t) \in V$. Then there exists $T > 0$ such that for every solution $\mathbf{x}$ with $\mathbf{x}(0) \in K$ there is $t \in [0, T]$ with $\mathbf{x}(t) \in V$.*

Proof. Suppose that there is no such upper bound T for the entry times into V. Then for each $n \in \mathbb{N}$ there is $\mathbf{x}_n(0) = x_n \in K$ and a solution $\mathbf{x}_n$ such that $\mathbf{x}_n(t) \notin V$ for $0 \le t \le n$. Since K is compact, we can assume that $x_n \to x \in K$. And by Lemma 3.1 a subsequence of $\mathbf{x}_n$ converges to a solution $\mathbf{x}$ with $\mathbf{x}(0) = x$ and $\mathbf{x}(t) \notin V$ for all $t > 0$. □

LEMMA 3.14. *If a closed set A is Lyapunov stable, then it is strongly positively invariant.*

Proof. A is the intersection of a family of strongly positively invariant neighborhoods. □

Lemma 3.15. *If a compact set A satisfies* (ii) *and* (iii′)*, it is attracting.*

Proof. Let B be a compact neighborhood of A, included in the fundamental neighborhood U, and let W be a neighborhood of A. A being Lyapunov stable, there exists an open neighborhood V of A with $\Phi^L_{[0,\infty)}(V) \subset W$. For any $x \in B$ and any solution $\mathbf{x}$ with $\mathbf{x}(0) = x$, there exists $t > 0$ with $\mathbf{x}(t) \in V$. Applying Lemma 3.13 implies $\Phi^L_T(B) \subset \Phi^L_{[0,T]}(V)$; hence $\Phi^L_{[T,\infty)}(B) \subset W$ and A is attracting. □

Lemma 3.16. *If the set A is attracting and strongly positively invariant, then it is Lyapunov stable.*

Proof. Let A be attracting with fundamental neighborhood U, and V be any other (open) neighborhood of A. Then by definition there is $T > 0$ such that $\Phi^L_{[T,\infty)}(U) \subset V$. A being strongly positively invariant, $\Phi^L_{[0,T]}(A) \subset A$. Upper semicontinuity gives an $\varepsilon > 0$ such that $\Phi^L_{[0,T]}(N^\varepsilon(A)) \subset V$ and $N^\varepsilon(A) \subset U$. Hence $\Phi^L_{[0,\infty)}(N^\varepsilon(A)) \subset V$, which shows Lyapunov stability. □

Corollary 3.17. *For a compact set A, properties* (ii) *and* (iii′) *of Definition* IX*, together, are equivalent to attracting and strong positive invariance.*

Corollary 3.18. *A compact set A is an attractor if and only if it is asymptotically stable.*

We conclude with a simple useful condition ensuring that an open set contains an attractor.

Proposition 3.19. *Let U be an open set with compact closure. Suppose that $\Phi_T(\overline{U}) \subset U$ for some $T > 0$. Then U is a fundamental neighborhood of some attractor A.*

Proof. Since Φ has a closed graph, $\Phi_T(\overline{U})$ is compact. Therefore $\Phi_T(\overline{U}) \subset V \subset \overline{V} \subset U$ for some open set V. By upper semicontinuity of Φ_T (which follows from property (d) of a set-valued dynamical system) there exists $\varepsilon > 0$ such that $\Phi_t(\overline{U}) \subset V$ for $T - \varepsilon \le t \le T + \varepsilon$. Let $t_0 = T(T+1)/\varepsilon$. For all $t \ge t_0$ write $t = kT + r$ with $k \in \mathbb{N}$ and $r < T$. Hence $t = k(T + r/k)$ with $0 \le r/k < \varepsilon$. Thus

$$\Phi_t(\overline{U}) = \Phi_{T+r/k} \circ \cdots \circ \Phi_{T+r/k}(\overline{U}) \subset V.$$

Hence $\omega_\Phi(U) = \bigcap_{t \ge t_0} \overline{\Phi_{[t,\infty)}(U)} \subset \overline{V} \subset U$ is an attractor with fundamental neighborhood U. □

3.7. Chain transitivity and attractors.

Proposition 3.20. *Let L be internally chain transitive. Then L has no proper attracting set for Φ^L.*

Proof. Let $A \subset L$ be an attracting set. By definition, there exists a neighborhood U of A, and for all $\varepsilon > 0$ a number t_ε such that $\Phi^L_t(U) \subset N^\varepsilon(A)$ for all $t > t_\varepsilon$. Assume $A \ne L$ and choose ε small enough so that $N^{2\varepsilon}(A) \subset U$ and there exists $y \in L \setminus N^{2\varepsilon}(A)$. Then, for $T \ge t_\varepsilon$ and $x \in A$, there is no (ε, T) chain from x to y. In fact, $\mathbf{x}_1(0) \in N^{2\varepsilon}(A)$, and hence $\mathbf{x}_1(t_1) \in N^\varepsilon(A)$; by induction, $\mathbf{x}_i(t_i) \in N^\varepsilon(A)$ so that $\mathbf{x}_{i+1}(0) \in N^{2\varepsilon}(A)$ as well. Thus we arrive at a contradiction. □

Remark 3.21. This last proposition can also be deduced from Bronstein and Kopanskii [9, Theorem 1] combined with Lemma 3.1. Also the converse is true.

Recall that an attracting set (respectively, attractor) for Φ is an attracting set (respectively, attractor) for Φ^L with $L = \mathbb{R}^m$.

Lemma 3.22. *Let A be an attracting set for Φ and L a closed invariant set. Assume $A \cap L \ne \emptyset$. Then $A \cap L$ is an attracting set for Φ^L.*

Proof. The proof follows from the definitions. □

If A is a set, then

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

denotes its *basin of attraction.*

THEOREM 3.23. *Let A be an attracting set for Φ and L an internally chain transitive set. Assume $L \cap B(A) \neq \emptyset$. Then $L \subset A$.*

Proof. Suppose $L \cap B(A) \neq \emptyset$. Then there exists a solution $\mathbf{x}$ to (I) with $\mathbf{x}(0) = x \in B(A)$ and $\mathbf{x}(\mathbb{R}) \subset L$. Hence $d(\mathbf{x}(t), A) \to 0$ when $t \to \infty$, proving that L meets A. Proposition 3.20 and Lemma 3.22 imply that $L \subset A$. ☐

A *global attractor* for Φ is an attractor whose basin of attraction consists of all $\mathbb{R}^m$. If a global attractor exists, then it is unique and coincides with the maximal compact invariant set of Φ. The following corollary is an immediate consequence of Theorem 3.23 or even more easily of Lemma 3.5.

COROLLARY 3.24. *Suppose Φ has a global attractor A. Then every internally chain transitive set lies in A.*

3.8. Lyapunov functions.

PROPOSITION 3.25. *Let Λ be a compact set, $U \subset \mathbb{R}^m$ be a bounded open neighborhood of Λ, and $V : \overline{U} \to [0, \infty[$. Let the following hold:*

(i) *For all $t \geq 0$, $\Phi_t(U) \subset U$ (i.e., 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 $t > 0$, $V(y) < V(x)$;*

(iv) *V is upper semicontinuous, and for all $x \in \overline{U} \setminus \Lambda$, $y \in \Phi_t(x)$, and $t > 0$, $V(y) < V(x)$.*

(A) *Under* (i), (ii), *and* (iii), *Λ is a Lyapunov stable attracting set, and there exists an attractor contained in Λ whose basin contains U, and with $V^{-1}([0, r))$ as fundamental neighborhoods for small $r > 0$.*

(B) *Under* (i), (ii), *and* (iv), *there exists an attractor contained in Λ whose basin contains U.*

Proof. For the proof of (A), let $r > 0$ and $U_r = \{x \in U : V(x) < r\}$. Then $\{\overline{U}_r\}_{r>0}$ is a nested family of compact neighborhoods of Λ with $\bigcap_{r>0} \overline{U}_r = \Lambda$. Thus for $r > 0$ small enough, $\overline{U_r} \subset U$. Moreover, $\Phi_t(\overline{U}_r) \subset U_r$ for $t > 0$ by our hypotheses on U and V. Proposition 3.19 then implies the result.

For (B), let $A = \omega_\Phi(U)$, which is closed and invariant (by Lemma 3.8) and hence compact, since it is included in $\overline{U}$. Let $\alpha = \max_{y \in A} V(y)$ be reached at x, since V is upper semicontinuous. By invariance there exists a solution $\mathbf{x}$ and $t > 0$ with $z = \mathbf{x}(0) \in A$ and $\mathbf{x}(t) = x$. This contradicts (iv) unless $\alpha = 0$ and $A \subset \Lambda$. Thus U is a neighborhood of A, which is an attractor included in Λ. ☐

Remark 3.26. Given any attractor A, there exists a function V such that Proposition 3.25(iv) holds for $\Lambda = A$. Take $V(x) = \max\{d(y, A)g(t),\ y \in \Phi_t(x),\ t \geq 0\}$, where $d > g(t) > c > 0$ is any continuous strictly increasing function.

Let Λ be any subset of $\mathbb{R}^m$. A continuous function $V : \mathbb{R}^m \to \mathbb{R}$ is called a *Lyapunov function* for Λ if $V(y) < V(x)$ for all $x \in \mathbb{R}^m \setminus \Lambda$, $y \in \Phi_t(x)$, $t > 0$, and $V(y) \leq V(x)$ for all $x \in \Lambda$, $y \in \Phi_t(x)$, and $t \geq 0$. Note that for each solution $\mathbf{x}$, V is constant along its limit set $L(\mathbf{x})$.

The following result is similar to Benaïm [3, Proposition 6.4].

PROPOSITION 3.27. *Suppose that V is a Lyapunov function for Λ. Assume that $V(\Lambda)$ has empty interior. Then every internally chain transitive set L is contained in Λ and $V \mid L$ is constant.*

Proof. Let

$$v = \inf\{V(y) : y \in L\}.$$

Since L is compact and V is continuous, $v = V(x)$ for some point $x \in L$. Since L is invariant, there exists a solution $\mathbf{x}$ with $\mathbf{x}(t) \in L$ and $\mathbf{x}(0) = x$. Then $v = V(x) > V(\mathbf{x}(t))$, and thus is impossible for $t > 0$. Since $\mathbf{x}(t) \in \Phi_t(x)$, we conclude $x \in \Lambda$.

Thus v belongs to the range $V(\Lambda)$. Since $V(\Lambda)$ contains no interval, there is a sequence $v_n \notin V(\Lambda)$ decreasing to v. The sets $L_n = \{x \in L : V(x) < v_n\}$ satisfy $\Phi_t(\overline{L}_n) \subset L_n$ for $t > 0$. In fact, either $x \in \Lambda \cap \overline{L}_n$ and $V(y) \le V(x) < v_n$ or $V(y) < V(x) \le v_n$, for any $y \in \Phi_t(x)$, $t > 0$.

Thus, using Propositions 3.19 and 3.20, one obtains $L = \bigcap_n \overline{L}_n = \{x \in L : V(x) = v\}$. Hence V is constant on L. L being invariant, this implies, as above, $L \subset \Lambda$. ∎

COROLLARY 3.28. *Let V and Λ be as in Proposition 3.27. Suppose furthermore that V is C^m and Λ is contained in the critical points set of V. Then every internally chain transitive set lies in Λ and $V \mid L$ is constant.*

Proof. By Sard's theorem (Hirsch [18, p. 69]), $V(\Lambda)$ has empty interior and Proposition 3.27 applies. ∎

4. The limit set theorem.

4.1. Asymptotic pseudotrajectories for set-valued dynamics. The translation flow $\Theta : C^0(\mathbb{R}, \mathbb{R}^m) \times \mathbb{R} \to C^0(\mathbb{R}, \mathbb{R}^m)$ is the flow defined by

$$\Theta^t(\mathbf{x})(s) = \mathbf{x}(s + t).$$

A continuous function $\mathbf{z} : \mathbb{R}^+ \to \mathbb{R}^m$ is an asymptotic pseudotrajectory (APT) for Φ if

$$\lim_{t\to\infty} \mathbf{D}(\Theta^t(\mathbf{z}), S_{\mathbf{z}(t)}) = 0 \tag{4.1}$$

(or $\lim_{t\to\infty} \mathbf{D}(\Theta^t(\mathbf{z}), S) = 0$, where $S = \bigcup_{x\in\mathbb{R}^m} S_x$ denotes the set of all solutions of (I)).

Alternatively, for all T

$$\lim_{t\to\infty} \inf_{\mathbf{x}\in S_{\mathbf{z}}(t)} \sup_{0\le s\le T} \|\mathbf{z}(t+s) - \mathbf{x}(s)\| = 0.$$

In other words, for each fixed T, the curve

$$[0, T] \to \mathbb{R}^m : s \to \mathbf{z}(t + s)$$

shadows some Φ trajectory of the point $\mathbf{z}(t)$ over the interval $[0, T]$ with arbitrary accuracy for sufficiently large t. Hence $\mathbf{z}$ has a forward trajectory under Θ attracted by S. As usual, one extends $\mathbf{z}$ to $\mathbb{R}$ by letting $\mathbf{z}(t) = \mathbf{z}(0)$ for $t < 0$.

The next result is a natural extension of Benaïm and Hirsch [4], [5, Theorem 7.2].

THEOREM 4.1 (characterization of APT). *Assume $\mathbf{z}$ is bounded. Then there is equivalence between the following statements:*

(i) *$\mathbf{z}$ is an APT for Φ.*

(ii) *$\mathbf{z}$ is uniformly continuous, and any limit point of $\{\Theta^t(\mathbf{z})\}$ is in S.*

In both cases the set $\{\Theta^t(\mathbf{z});\ t \ge 0\}$ is relatively compact.

Proof. By hypothesis, $K = \overline{\{\mathbf{z}(t);\ t \ge 0\}}$ is compact.

For any $\varepsilon > 0$, there exists $\eta > 0$ such that $\|z - x\| < \varepsilon/2$, for any $x \in K$, any $z \in \Phi_s(x)$, and any $|s| < \eta$, using property (d) of the dynamical system.

$\mathbf{z}$ being an APT, there exists T such that $t > T$ implies

$$d(\mathbf{z}(t+s), \Phi_s(\mathbf{z}(t))) < \frac{\varepsilon}{2} \quad \forall |s| < \eta;$$

hence

$$\|\mathbf{z}(t+s) - \mathbf{z}(t)\| \leq \varepsilon$$

and $\mathbf{z}$ is uniformly continuous. Clearly any limit point belongs to S by the condition (4.1) above.

Conversely, if $\mathbf{z}$ is uniformly continuous, then the family of functions $\{\Theta^t(\mathbf{z});\ t \geq T\}$ is equicontinuous and hence (K being compact) relatively compact by Ascoli's theorem. Since any limit point belongs to S, property (4.1) follows. □

4.2. Perturbed solutions are APTs.

THEOREM 4.2. *Any bounded solution* $\mathbf{y}$ *of* (II) *is an APT of* (I).

Proof. Let us prove that $\mathbf{y}$ satisfies Theorem 4.1(ii). Set $v(t) = \dot{\mathbf{y}}(t) - U(t) \in F^{\delta(t)}(\mathbf{y}(t))$. Then,

$$\mathbf{y}(t+s) - \mathbf{y}(t) = \int_0^s v(t+\tau)d\tau + \int_t^{t+s} U(\tau)d\tau. \tag{4.2}$$

By assumption (iii) of (II), the second integral goes to 0 as $t \to \infty$. The boundedness of $\mathbf{y}$, $\mathbf{y}(\mathbb{R}) \subset M$, M compact (combined with the fact that F has linear growth) implies boundedness of v and shows that $\mathbf{y}$ is uniformly continuous. Thus the family $\Theta^t(\mathbf{y})$ is equicontinuous, and hence relatively compact. Let $\mathbf{z} = \lim_{t_n \to \infty} \Theta^{t_n}(\mathbf{y})$ be a limit point. Set $t = t_n$ in (4.2) and define $v_n(s) = v(t_n + s)$. Then, using the assumption (iii) on U, the second term in the right-hand side of this equality goes to zero uniformly on compact intervals when $n \to \infty$. Hence

$$\mathbf{z}(s) - \mathbf{z}(0) = \lim_{n \to \infty} \int_0^s v_n(\tau)d\tau.$$

Since (v_n) is uniformly bounded, it is bounded in $L^2[0,s]$, and by the Banach–Alaoglu theorem, a subsequence of v_n will converge weakly in $L^2[0,s]$ (or weak* in $L^\infty[0,s]$) to some function v with $v(t) \in F(\mathbf{z}(t))$, for almost every t, since $v_n(t) \in F^{\delta(t+t_n)}(\mathbf{y}(t+t_n))$ for every t. Here we use (ii) and that F is upper semicontinuous with convex values. In fact, by Mazur's theorem, a convex combination of $\{v_m, m \geq n\}$ converges almost surely to v and $\lim_{m\to\infty} \mathrm{Co}(\bigcup_{n \geq m} F^{\delta(t+t_n)}(\mathbf{y}(t+t_n))) \subset F(\mathbf{z}(t))$. Hence $\mathbf{z}(s) - \mathbf{z}(0) = \int_0^s v(\tau)d\tau$, proving that $\mathbf{z}$ is a solution of (I) and hence $\mathbf{z} \in S_{M,M}$. □

4.3. APTs are internally chain transitive.

THEOREM 4.3. *Let* $\mathbf{z}$ *be a bounded APT of* (I). *Then* $L(\mathbf{z})$ *is internally chain transitive.*

Proof. The set $\{\Theta^t(\mathbf{z}) : t \geq 0\}$ is relatively compact, and hence the ω-limit set of $\mathbf{z}$ for the flow Θ,

$$\omega_\Theta(\mathbf{z}) = \bigcap_{t \geq 0} \overline{\{\Theta^s(\mathbf{z}) : s \geq t\}},$$

is internally chain transitive. (By standard properties of ω-limit sets of bounded semiorbits, $\omega_\Theta(\mathbf{z})$ is a nonempty, compact, internally chain transitive set invariant under Θ; see Conley [13]; a short proof is also in Benaïm [3, Corollary 5.6].) By property (4.1), $\omega_\Theta(\mathbf{z}) \subset S$, the set of all solutions of (I).

Let $\Pi : (C^0(\mathbb{R}, \mathbb{R}^m), \mathbf{D}) \to (\mathbb{R}^m, \|\cdot\|)$ be the projection map defined by $\Pi(\mathbf{z}) = \mathbf{z}(0)$. One has $\Pi(\omega_\Theta(\mathbf{z})) = L(\mathbf{z})$. In fact if $p = \lim_{n\to\infty} \mathbf{z}(t_n)$, let $\mathbf{w}$ be a limit point of $\Theta^{t_n}(\mathbf{z})$. Then $\mathbf{w} \in \omega_\Theta(\mathbf{z})$ and $\Pi(\mathbf{w}) = p$.

It then easily follows that $L(\mathbf{z})$ is nonempty compact and invariant under Φ since $\omega_\Theta(\mathbf{z}) \subset S$. Since Π has Lipschitz constant 1, Π maps every (ε, T) chain for Θ to an (ε, T) chain for Φ. This proves that $L(\mathbf{z})$ is internally chain transitive for Φ. ◻

5. Applications.

5.1. Approachability. An application of Proposition 3.25 is the following result, which can be seen as a continuous asymptotic deterministic version of Blackwell's approachability theorem [8]. Note that one has no property on uniform speed of convergence.

Given a compact set $\Lambda \in \mathbb{R}^m$ and $x \in \mathbb{R}^m$, we let $\Pi_\Lambda(x) = \{y \in \Lambda : d^2(x, \Lambda) = \|x - y\|^2 = \langle x - y, x - y\rangle\}$.

Corollary 5.1. *Let $\Lambda \subset \mathbb{R}^m$ be a compact set, $r > 0$, and $U = \{x \in \mathbb{R}^m : d(x, \Lambda) < r\}$. Suppose that for all $x \in U \setminus \Lambda$ there exists $y \in \Pi_\Lambda(x)$ such that the affine hyperplane orthogonal to $[x, y]$ at y separates x from $x + F(x)$. That is,*

$$\langle x - y, x - y + v\rangle \le 0 \tag{5.1}$$

for all $v \in F(x)$. Then Λ contains an attractor for (I) *with fundamental neighborhood U.*

Proof. Set $V(x) = d(x, \Lambda)$. To apply Proposition 3.25 it suffices to verify condition (iii) of Proposition 3.25. Condition (i) will follow, and condition (ii) is clearly true.

Let $\mathbf{x}$ be a solution to (I) with initial condition $x \in U \setminus \Lambda$. Set $\tau = \inf\{t > 0 : \mathbf{x}(t) \in \Lambda\} \le \infty$, $g(t) = V(\mathbf{x}(t))$, and let $I \subset [0, \tau[$ be the set of $0 \le t < \tau$ such that $g'(t)$ and $\dot{\mathbf{x}}(t)$ exist and $\dot{\mathbf{x}}(t) \in F(\mathbf{x}(t))$. For all $t \in I$ and $y \in \Pi_\Lambda(\mathbf{x}(t))$

$$\begin{aligned} g(t+h) - g(t) &\le \|\mathbf{x}(t+h) - y\| - \|\mathbf{x}(t) - y\| \\ &= \|\mathbf{x}(t) + \dot{\mathbf{x}}(t)h - y\| - \|\mathbf{x}(t) - y\| + |h|\varepsilon(h), \end{aligned}$$

where $\lim_{h\to 0} \varepsilon(h) = 0$. Hence

$$\begin{aligned} g'(t) &\le \frac{1}{\|\mathbf{x}(t) - y\|}\langle \mathbf{x}(t) - y, \dot{\mathbf{x}}(t)\rangle \\ &= -g(t) + \frac{1}{\|\mathbf{x}(t) - y\|}\langle \mathbf{x}(t) - y, \mathbf{x}(t) - y + \dot{\mathbf{x}}(t)\rangle. \end{aligned}$$

Thus, $\dot{x} \in F(x)$ and (5.1) imply $g'(t) \le -g(t)$ for all $t \in I$. Since g and $\mathbf{x}$ are absolutely continuous, I has full measure in $[0, \tau[$. Hence $g(t) \le e^{-t} g(0)$ for all $t < \tau$. Therefore $V(\mathbf{x}(t)) < V(x)$ for all $0 < t < \tau$, which shows (iii). Finally, $V(\mathbf{x}(t)) \le e^{-t} V(x)$ shows that the sets $V^{-1}[0, r')$ (with $0 < r' \le r$) are fundamental neighborhoods of the attractor in Λ. ◻

In particular, if any point of E has a unique projection on Λ (for example, Λ convex), then $\overline{C} = C$, and one recovers exactly Blackwell's sufficient condition for approachability.

Corollary 5.2 (Blackwell's approachability theorem). *Consider the decision making process described in section* 2.1*, Example* 2.2*. Let $\Lambda \subset E$ be a compact set. Assume that there exists a strategy Q such that for all $x \in E \setminus \Lambda$ there exists $y \in \Pi_\Lambda(x)$ such that the hyperplane orthogonal to $[x, y]$ through y separates x from $\overline{C}(x)$. Then Λ is approachable.*

Proof. Let $L(x_n)$ denote the limit set of $\{x_n\}$. By Corollary 5.1, Λ is an attractor with fundamental neighborhood E, hence a global attractor. Thus Theorem 3.6 with Proposition 2.1 and Corollary 3.24 imply that $L(x_n)$ is almost surely contained in Λ. ∎

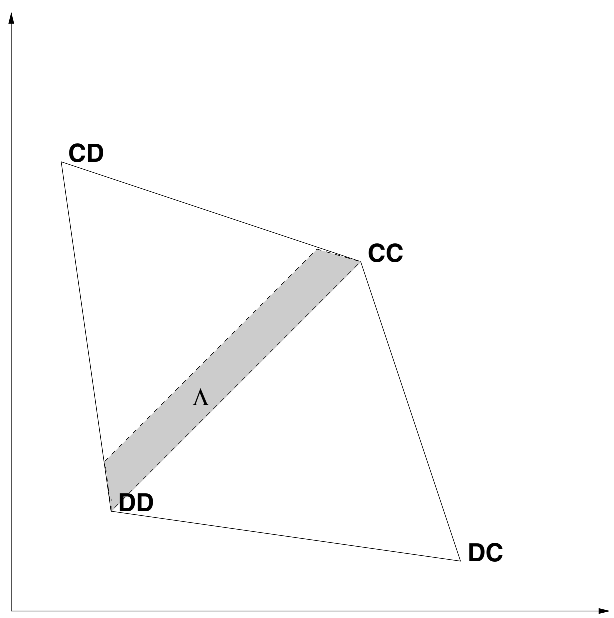
5.2. Smale's approach to the prisoner's dilemma. We develop here Example 2.4. Consider a 2×2 prisoner's dilemma game. Each player has two possible actions: cooperate (play C) or defect (play D). If both cooperate, each receives α; if both defect, each receives λ; if one cooperates and the other defects, the cooperator receives β and the defector γ. We suppose that $\gamma > \alpha > \lambda > \beta$, as is usual with a prisoner's dilemma game. We furthermore assume that

$$\gamma - \alpha < \alpha - \beta,$$

so that the outcome space E is the convex quadrilateral whose vertices are the payoff vectors

$$\mathbf{CD} = (\beta, \gamma), \qquad \mathbf{CC} = (\alpha, \alpha), \qquad \mathbf{DC} = (\gamma, \beta), \qquad \mathbf{DD} = (\lambda, \lambda);$$

see the figure below.



The outcome space E

Let δ be a nonnegative parameter. Adapting Smale [27] and Benaïm and Hirsch [4, 5], a δ-*good* strategy for player 1 is a strategy $Q^1 = \{Q^1_x\}$ (as defined in section 2.1) enjoying the following features:

$$Q^1_x(\text{play C}) = 1 \quad \text{if } x^1 > x^2$$

and

$$Q^1_x(\text{play C}) = 0 \quad \text{if } x^1 < x^2 - \delta.$$

The following result reinterprets the results of Smale [27] and Benaïm and Hirsch [4, 5] in the framework of approachability. It also provides some generalization (see Remark 5.4 below).

Theorem 5.3. (i) *Suppose that player* 1 *plays a* δ*-good strategy. Then the set*

$$\Lambda = \{x \in E : x^2 - \delta \le x^1 \le x^2\}$$

is approachable.

(ii) *Suppose that both players play a* δ*-good strategy and that at least one of them is continuous (meaning that the corresponding function* $x \to Q^i_x$*(play* C*) is continuous). Then*

$$\lim_{n\to\infty} x_n = \mathbf{CC}$$

almost surely.

Proof. (i) Let $x \in E \setminus \Lambda$. If $x^1 > x^2$, then

$$C(x) = \overline{C}(x) = [\mathbf{CC}, \mathbf{CD}],$$

and the line $\{u \in \mathbb{R}^2 : u^1 = u^2\}$ separates x from $\overline{C}(x)$. Similarly if $x^1 < x^2 - \delta$, then

$$C(x) = \overline{C}(x) = [\mathbf{DD}, \mathbf{DC}],$$

which is separated from x by the line $\{u \in \mathbb{R}^2 : u^1 = u^2 - \delta\}$. Assertion (i) then follows from Corollary 5.2.

(ii) If both play a δ-good strategy, then (i) and its analogue for player 2 imply that the diagonal

$$\Delta = \{x \in E : x^1 = x^2\}$$

is approachable. Thus $L(x_n) \subset \Delta$. Also (by Proposition 2.1, Theorem 3.6, and Lemma 3.5) $L(x_n)$ is invariant under the differential inclusion induced by

$$F(x) = -x + \overline{C}(x),$$

where $C(x) = C^1(x) \cap C^2(x)$ and $C^i(x)$ is the convex set associated with Q^i (the strategy of player i). Suppose that one player, say 1, plays a continuous strategy. Then $\overline{C}(x) \subset \overline{C^1}(x) = C^1(x)$ and for all $x \in \Delta$, $C^1(x) = [\mathbf{CD}, \mathbf{CC}]$. Now, there is only one subset of Δ which is invariant under $\dot{x} \in -x + [\mathbf{CD}, \mathbf{CC}]$; this is the point $\mathbf{CC}$. This proves that $L(x_n) = \mathbf{CC}$. ◻

Remark 5.4. (i) In contrast to Smale [27] and Benaïm and Hirsch [4, 5], observe that assertion (i) makes no hypothesis on player 2's behavior. In particular, it is unnecessary to assume that player 2 has a strategy of the form defined by section 2.1.

(ii) The regularity assumptions (on strategies) are much weaker than in Benaïm and Hirsch [4, 5].

(iii) A 0-good strategy makes the diagonal Δ approachable. However, if both players play a 0-good strategy, then $\overline{C}(x) = E$ for all $x \in \Delta$, and we are unable to predict the long-term behavior of $\{x_n\}$ on Δ.

5.3. Fictitious play in potential games. Here we generalize the result of Monderer and Shapley [25]. They prove convergence of the classical discrete fictitious play process, as defined in Example 2.3, for n-linear payoff functions. Harris [17] studies the best-response dynamics in this case but does not derive convergence of fictitious play from it. Our limit set theorem provides the right tool for doing this, even in the following, more general setting.

Let X^i, $i = 1, \dots, n$, be compact convex subsets of Euclidean spaces and $U\colon X^1 \times \cdots \times X^n \to \mathbb{R}$ be a C^1 function which is concave in each variable. U is interpreted as the common payoff function for the n players. We write $x = (x^i, x^{-i})$ and define $BR^i(x^{-i}) := \text{Argmax}_{x^i \in X^i} U(x)$ the set of maximizers. Then $x \mapsto BR(x) = (BR^1(x^{-1}), \dots, BR^n(x^{-n}))$ is upper semicontinuous (by Berge's maximum theorem, since U is continuous) with nonempty compact convex values. Consider the best response dynamics

$$\dot{\mathbf{x}} \in BR(\mathbf{x}) - \mathbf{x}. \tag{5.2}$$

Its constant solutions $\mathbf{x}(t) \equiv \hat{x}$ are precisely the Nash equilibria $\hat{x} \in BR(\hat{x})$; i.e., $U(\hat{x}) \geq U(x^i, \hat{x}^{-i})$ for all i and $x^i \in X^i$. Along a solution $\mathbf{x}(t)$ of (5.2), let $u(t) = U(\mathbf{x}(t))$. Then for almost all $t > 0$,

$$\dot{u}(t) = \sum_{i=1}^{n} \frac{\partial U}{\partial x^i}(\mathbf{x}(t))\dot{\mathbf{x}}^i(t) \tag{5.3}$$

$$\geq \sum_{i=1}^{n} [U(\mathbf{x}^i(t) + \dot{\mathbf{x}}^i(t), \mathbf{x}^{-i}(t)) - U(\mathbf{x}(t))] \tag{5.4}$$

$$= \sum_{i=1}^{n} \left[\max_{y^i \in X^i} U(y^i, \mathbf{x}^{-i}(t)) - U(\mathbf{x}(t))\right] \geq 0, \tag{5.5}$$

where from (5.3) to (5.4) we use the concavity of U in x^i, and (5.5) follows from (5.2) and the definition of BR^i. Since the function $t \mapsto u(t)$ is locally Lipschitz, this shows that it is weakly increasing. It is constant in a time interval T, if and only if $\mathbf{x}^i(t) \in BR^i(\mathbf{x}^{-i}(t))$ for all $t \in T$ and $i = 1, \dots, n$, i.e., if and only if $\mathbf{x}(t)$ is a Nash equilibrium for $t \in T$ (but $\mathbf{x}(t)$ may move in a component of the set of Nash equilibria (NE) with constant U).

THEOREM 5.5. *The limit set of every solution of* (5.2) *is a connected subset of NE, along which U is constant. If, furthermore, the set $U(NE)$ contains no interval in* $\mathbb{R}$, *then the limit set of every fictitious play path is a connected subset of NE along which U is constant.*

Proof. The first statement follows from the above. The second statement follows from Theorem 3.6 together with Proposition 3.27 with $V = -U$ and $\Lambda = NE$. ∎

Remark 5.6. The assumption that the set $U(NE)$ contains no interval in $\mathbb{R}$ follows via Corollary 3.28 if U is smooth enough (e.g., in the n-linear case) and if each X^i has at most countably many faces, by applying Sard's lemma to the interior of each face.

Example 5.7 (2×2 coordination game). The global attractor of (5.2) consists of three equilibria and two line segments connecting them. The internally chain transitive sets are the three equilibria. Hence every fictitious play process converges to one of these equilibria.

The case of (continuous concave-convex) two-person zero-sum games was treated in Hofbauer and Sorin [21], where it is shown that the global attractor of (5.2) equals the set of equilibria. In this case the full strength of Theorem 3.6 and the notion of chain transitivity are not needed; the invariance of the limit set of a fictitious play path implies that it is contained in the global attractor; compare Corollary 3.24.

Acknowledgments. This research was started during visits of Josef Hofbauer in Paris in 2002. Josef Hofbauer thanks the Laboratoire d'Econométrie, Ecole Polytechnique, and the D.E.A. OJME, Université P. et M. Curie - Paris 6, for financial support and Sylvain Sorin for his hospitality. Michel Benaïm thanks the Erwin Schrödinger Institute, and the organizers and participants of the 2004 Kyoto workshop on "game dynamics."

REFERENCES

[1] J.-P. Aubin and A. Cellina, *Differential Inclusions*, Springer, New York, 1984.

[2] M. Benaïm, *A dynamical system approach to stochastic approximations*, SIAM J. Control Optim., 34 (1996), pp. 437–472.

[3] M. Benaïm, *Dynamics of stochastic approximation algorithms*, in Séminaire de Probabilités XXXIII, Lecture Notes in Math. 1709, Springer, New York, 1999, pp. 1–68.

[4] M. Benaïm and M. W. Hirsch, *Stochastic Adaptive Behavior for Prisoner's Dilemma*, 1996, preprint.

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

[6] M. Benaïm and M. W. Hirsch, *Mixed equilibria and dynamical systems arising from fictitious play in perturbed games*, Games Econom. Behav., 29 (1999), pp. 36–72.

[7] M. Benaïm, J. Hofbauer, and S. Sorin, *Stochastic Approximations and Differential Inclusions: Applications*, Cahier du Laboratoire d'Econometrie, Ecole Polytechnique, 2005-011.

[8] D. Blackwell, *An analog of the minmax theorem for vector payoffs*, Pacific J. Math., 6 (1956), pp. 1–8.

[9] I. U. Bronstein and A. Ya. Kopanskii, *Chain recurrence in dynamical systems without uniqueness*, Nonlinear Anal., 12 (1988), pp. 147–154.

[10] G. Brown, *Iterative solution of games by fictitious play*, in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., Wiley, New York, 1951, pp. 374–376.

[11] R. Buche and H. J. Kushner, *Stochastic approximation and user adaptation in a competitive resource sharing system*, IEEE Trans. Automat. Control, 45 (2000), pp. 844–853.

[12] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[13] C. C. Conley, *Isolated Invariant Sets and the Morse Index*, CBMS Reg. Conf. Ser. in Math. 38, AMS, Providence, RI, 1978.

[14] M. Duflo, *Algorithmes Stochastiques*, Springer, New York, 1996.

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

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

[17] C. Harris, *On the rate of convergence of continuous time fictitious play*, Games Econom. Behav., 22 (1998), pp. 238–259.

[18] M. W. Hirsch, *Differential Topology*, Springer, New York, 1976.

[19] J. Hofbauer, *Stability for the Best Response Dynamics*, preprint, 1995.

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

[21] J. Hofbauer and S. Sorin, *Best response dynamics for continuous zero-sum games*, in Cahier du Laboratoire d'Econometrie, Ecole Polytechnique, 22002-2028.

[22] M. Kunze, *Non-Smooth Dynamical Systems*, Lecture Notes in Math. 1744, Springer, New York, 2000.

[23] H. J. Kushner and G. G. Yin, *Stochastic Approximations Algorithms and Applications*, Springer, New York, 1997.

[24] L. Ljung, *Analysis of recursive stochastic algorithms*, IEEE Trans Automat. Control, 22 (1977), pp. 551–575.

[25] D. Monderer and L. S. Shapley, *Fictitious play property for games with identical interests*, J. Econom. Theory, 68 (1996), pp. 258–265.

[26] J. Robinson, *An iterative method of solving a game*, Ann. Math., 54 (1951), pp. 296–301.

[27] S. Smale, *The prisoner's dilemma and dynamical systems associated to non-cooperative games*, Econometrica, 48 (1980), pp. 1617–1633.

[28] S. Sorin, *A First Course on Zero-Sum Repeated Games*, Springer, New York, 2002.