Ergodic Properties of weak Asymptotic Pseudotrajectories for Set-valued Dynamical Systems

Mathieu Faure∗ and Gregory Roth†

Abstract: A successful method to describe the asymptotic behavior of various deterministic and stochastic processes such as asymptotically autonomous differential equations or stochastic approximation processes is to relate them to an appropriately chosen limit semiflow. Benaïm and Schreiber (2000) define a general class of such stochastic processes, which they call weak asymptotic pseudotrajectories and study their ergodic behavior. In particular, they prove that the weak∗ limit points of the empirical measures associated to such processes are almost surely invariant for the associated deterministic semiflow. Continuing a program started by Benaïm, Hofbauer and Sorin (2005), we generalize the ergodic results mentioned above to weak asymptotic pseudotrajectories relative to set-valued dynamical systems.

AMS 2000 subject classifications: Primary 34A60, 37L40, 37C50; secondary 62L20, 37B55.

Keywords and phrases: Set-valued dynamical systems, ergodic theory, invariant measures, stochastic approximations.

1. Introduction

Let \((E, d)\) be a complete metric space and consider a flow \(\phi\) on \(E\), namely \(\phi : E \times \mathbb{R} \to E\) is continuous and satisfies the following properties:

\(i\) for any \(x \in E\), \(\phi(x, 0) = x\),

\(ii\) for any \(t, s \in \mathbb{R}\) and \(x \in E\), \(\phi(\phi(x, t), s) = \phi(x, t + s)\).

In the sequel, we will prefer the notation \(\phi_t(x)\) instead of \(\phi(x, t)\). A continuous function \(X : \mathbb{R}_+ \to E\) is an asymptotic pseudotrajectory (APT) for the flow \(\phi\) if

\[
\lim_{t \to \infty} \sup_{s \in [0,T]} d(X(t + s), \phi_s(X(t))) = 0,
\]

for any \(T > 0\). Heuristically this means that, for any \(T > 0\), the curve joining \(X(t)\) to \(X(t + T)\) shadows the trajectory of the solution starting from \(X(t)\) with arbitrary accuracy, provided \(t\) is large enough. This concept has initially been introduced in Benaïm (1996) and Benaïm and Hirsch (1996), where the authors proved that the asymptotic behaviors of an APT can be described with a great deal of generality.

∗Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland.
†Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland. Tel. +41327182828. gregory.roth@unine.ch
through the study of the asymptotics of the flow \( \phi \). One of their main results is the characterization of the limit set of an APT, in the sense that it is *internally chain transitive*, i.e. compact, invariant and contains no proper attractor for the restricted flow (the terminology comes from the notion of *chain recurrence* introduced by Conley (1978), see also Bowen (1975)). Consequently, this result turns out to be a particularly useful tool for analyzing the long term behavior of a large class of "perturbed" systems, whose solutions are APTs relative to some "unperturbed" flow. For instance, given an asymptotically autonomous differential equation, its solution trajectories are APTs relative to the flow induced by its limit autonomous differential equation (see Benaïm and Hirsch (1996)). Also, under the right assumptions, the paths of a stochastic approximation process with decreasing step size are almost surely APTs for the flow induced by the mean ODE (see e.g. Benaïm (1999) or Pemantle (2007) for comprehensive overviews on the topic).

In Benaïm and Schreiber (2000), the authors investigate the ergodic or statistical behavior of APTs for a flow. In fact, they prove their main result for a more general class of stochastic processes that they call *weak asymptotic pseudotrajectories* (WAPT). Given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a nondecreasing family of sub-\(\sigma\)-algebras \((\mathcal{F}_t)_{t \geq 0}\), a process

\[
X : \mathbb{R}_+ \times \Omega \rightarrow E,
\]

is called a *weak asymptotic pseudotrajectory* (WAPT) for \( \phi \) provided that

(i) It is progressively measurable: \(X_{|[0,T] \times \Omega} \) is \(\mathcal{B}_{[0,T]} \times \mathcal{F}_T\) measurable for all \(T > 0\),

(ii) for each \(\alpha > 0\) and \(T > 0\),

\[
\lim_{t \to \infty} \mathbb{P}( \sup_{s \in [0,T]} d(X(t + s), \phi_s(X(t))) \geq \alpha | \mathcal{F}_t) = 0
\]

almost surely.

Notice that, if a random process \(X\) is almost surely an APT, namely for almost every \(\omega \in \Omega\), (1) holds, then \(X\) is a weak asymptotic pseudotrajectory. One should be aware that the characterization of limit sets no longer holds for a WAPT. By this we mean that they are not almost surely internally chain transitive in general. One of the main result of the paper quoted above is the following

**Theorem A** (Benaïm and Schreiber (2000), Theorem 1) *Given a WAPT \(X\) for a flow \(\phi\), the weak*\(^*\) *limit points of the empirical measures

\[
\mu_t(\omega) := \frac{1}{t} \int_0^t \delta_{X(s,\omega)} ds
\]

(where \(\delta_p\) is a Dirac measure at the point \(p\)) are almost surely invariant measures for \(\phi\). In particular, if the family of measures \(\{\frac{1}{t} \int_0^t \delta_{X(s)} ds\}_{t \geq 0}\) is tight, this implies that, with probability one, the process \(X(t)\) spends most of its time in any arbitrarily small neighborhood of the Birkhoff center of \(\phi\).*
In Benaïm, Hofbauer and Sorin (2005), the authors generalized the notion of asymptotic pseudotrajectory to set-valued dynamical systems $\Phi$ (induced for instance by a differential inclusion, see Section 2 for a general definition of set-valued dynamical systems and section 4 for the particular case of differential inclusions) and extended the characterization of limit sets. In this paper, we generalize the notion of WAPT to set-valued dynamical systems and extend Theorem A to these more general settings.

The paper is organized as follows. The first step (see Section 2) consists in defining properly the notions of invariant measure and Birkhoff center for set-valued dynamical systems. To this end, we heavily rely on Miller and Akin (1999). In this paper, the authors prove the equivalence between various definitions of an invariant measure for discrete time set-valued dynamical systems induced by closed relations. We give two equivalent definitions in the continuous time case (see Theorem 2.6). In the same framework, Aubin, Frankowska and Lasota (1991) prove a Poincaré recurrence Theorem. We give a topological version of this result in the continuous time case (see Theorem 2.10). In Section 3, we prove the main result about the ergodic behavior of WAPT (Theorem 3.2). We then give some examples of WAPT in Section 4, in particular stochastic approximation algorithms relative to a differential inclusion. The proofs of several technical results are postponed to the appendix to ease the reading.

2. Ergodic theory for set-valued dynamical systems

Set-valued dynamical systems, which are often referred to as general control systems or general dynamical systems, are generally used to describe multi-valued differential equations, including differential inclusions (see Bianchini and Zecca (1981), Li and Zhang (2002), Benaïm, Hofbauer and Sorin (2005, 2006) or Nieuwenhuis (2009)) and control systems (see Roxin (1965) or Kloeden (1975, 1978)). The literature on the subject is abundant and the terminology sometimes differs among authors.

2.1. Generalities

Let us first recall some classical notions. The Hausdorff distance between two nonempty closed sets $A$ and $B$ in $E$ is given by

$$D_H(A, B) := \max\{d_H(A, B), d_H(B, A)\},$$

where $d_H$ is the Hausdorff semidistance:

$$d_H(A, B) := \sup_{a \in A} d(a, B).$$

Let $\mathcal{C}(\mathbb{R}, E)$ denote the space of continuous $E$-valued applications, endowed with the topology of uniform convergence on compacts. This topological space is metrizable with the distance $D$, given by

$$D(x, y) := \sum_{k \in \mathbb{N}} \frac{1}{2^k} \min\{1, \sup_{t \in [-k, k]} d(x(t), y(t))\},$$
which makes it complete (since $E$ is complete).

**Definition 2.1.** A set-valued map $\Phi : \mathbb{R}_+ \times E \rightrightarrows E$ with nonempty and closed values is called a set-valued dynamical system (SVDS) on $E$ provided that

a) $\forall x \in E$, $\Phi_0(x) = \{x\}$,

b) $\forall x \in E$, $\forall s, t \in \mathbb{R}_+$, $\Phi_t(\Phi_s(x)) = \Phi_{t+s}(x)$,

c) for any $x \in E$, $t \mapsto \phi_t(x)$ is a continuous map for the Hausdorff distance,

d) for any $t \in \mathbb{R}_+$, the map $x \mapsto \phi_t(x)$ is upper semicontinuous, i.e. for any $x_0 \in E$, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $d_H(\phi_t(x), \phi_t(x_0)) < \varepsilon$ for any $x \in B(x_0, \delta)$.

**Definition 2.2.** A function $z : [0, T] \to E$ is a partial solution relative to the set-valued dynamical system $\Phi$ if it satisfies $z(t) \in \Phi_{t-s}(z(s))$ for $s, t \in [0, T]$, $s \leq t$. The set of such solutions is called $S_{\Phi}^{[0,T]}$. We call $S_{\Phi}^{[0,T]}(A)$ the set of partial solutions on $[0,T]$, starting in $A$. We denote by $S_{\Phi}^{+}$ the set of half solutions $S_{\Phi}^{[0,\infty)}$.

The partial solutions are continuous functions and, given $0 \leq t$ and two points $x, y \in E$ such that $y \in \Phi_t(x)$, there exists at least one partial solution $z$ on $[0, t]$ such that $z(0) = x$ and $z(t) = y$ (see Roxin (1965)). By the above remark, for all $x \in E$, $S_{\Phi}^{[0,t]}(x) \neq \emptyset$ which implies that $S_{\Phi}^{+}(x) \neq \emptyset$. A function $z : \mathbb{R} \to E$ which satisfies $z(t) \in \Phi_{t-s}(z(s))$, $\forall s, t \in \mathbb{R}$, $s \leq t$, is called an entire solution of $\Phi$. $S_{\Phi}$ is the set of all entire solutions, $S_{\Phi}(A)$ the subset of entire solutions starting from $A$. We will say that the set-valued dynamical system $\Phi$ is complete if, for any $x \in E$, there exists an entire solution with initial condition $x$, i.e., if $S_{\Phi}(x) \neq \emptyset$, $\forall x \in E$. For example, we will see in Section 4 that the SVDS induced by a standard differential inclusion is complete.

From now, let us assume that $E$ is compact and call it $M$ to avoid confusions. Therefore, we consider a SVDS $\Phi$ defined on $M$.

**Proposition 2.3.** The set of solutions $S_{\Phi}$ is a nonempty compact subset of $C(\mathbb{R}, M)$.

The compactness is a well-known consequence of Barbashin’s Theorem$^2$ (see for instance Barbashin (1948) or Aubin and Cellina (1984)). We prove the existence of at least one entire solution in the Appendix A.

**Remark 2.4.** Points c) and d) in Definition 2.1 imply that the map $(t, x) \mapsto \Phi_t(x)$ is jointly upper semicontinuous. Since $M$ is compact, upper semicontinuity is equivalent to saying that the graph of $\Phi$:

$$Gr(\Phi) := \{(t, x, y) : y \in \Phi_t(x)\}$$

is closed: if $(t_n, x_n, y_n) \to (t, x, y)$ and $y_n \in \Phi_{t_n}(x_n)$ then $y \in \Phi_t(x)$.

$^1$This means that $\Phi$ is a map from $\mathbb{R}_+ \times E$ to $2^E$. Set-valued maps are also called *relations* in the sequel.

$^2$Which states that, for any $t \geq 0$, the set $S_{\Phi}^{[0,t]}(A)$ is compact in $C([0,t], \mathbb{R})$. 
For our purpose, we need to give a proper definition of an invariant measure relative to set-valued dynamical systems. Recall that, if $\theta = \{\theta_t\}_{t \in \mathbb{R}}$ (resp. $\theta = \{\theta_t\}_{t \in \mathbb{R}_+}$) is a continuous flow (resp. semiflow) on a metric space $X$, a measure $\mu$ on $X$ is called $\theta$-invariant if $\mu(\theta_t^{-1}(A)) = \mu(A)$ for all Borel set $A \subset X$ and $t \in \mathbb{R}$ (resp. $t \in \mathbb{R}_+$).

Let $X$ and $X'$ be two metric spaces and $f$ be a Borel measurable map from $X$ to $X'$. We denote by $\mathcal{P}(X)$ the set of probability measures on $X$. Then we can define the map $f^* : \mathcal{P}(X) \rightarrow \mathcal{P}(X')$ by

$$f^*(\mu)(A') := \mu(f^{-1}(A')),$$

for any $\mu \in \mathcal{P}(X)$ and any Borel set $A'$ in $X'$. The support of $\mu \in \mathcal{P}(X)$, i.e. the smallest closed set $S \subset X$ satisfying $\mu(S) = 1$, is denoted by $\text{supp}(\mu)$. In the sequel, given a closed subset $S$ of $X$, we will sometimes assimilate a probability measure with support in $S$ to its restriction to the metric space $(S, d)$.

### 2.2. Invariant measures for a closed relation

Let $F : M \rightrightarrows M$ be a closed relation with nonempty values, which means that the graph of $F$,

$$\text{Gr}(F) := \{(x, y) \in M \times M \mid y \in F(x)\}$$

is closed. Let $M^\mathbb{Z}$ be the set of bi-infinite sequences in $M$. The relation $F$ induces a discrete time set-valued dynamical system on $M$, whose set of solutions is the nonempty set

$$S_F := \{x \in M^\mathbb{Z} : x_{i+1} \in F(x_i) \ \forall i \in \mathbb{Z}\}.$$

In order to define invariant measures in this discrete case, we follow Miller and Akin (1999). The shift homeomorphism $\tilde{\Theta} : M^\mathbb{Z} \rightarrow M^\mathbb{Z}$ is defined by $(\tilde{\Theta}(x))_i = x_{i+1}$, for $i \in \mathbb{Z}$. Notice that $M^\mathbb{Z}$, equipped with the product topology is metrizable via the following distance:

$$\delta(x, y) := \sup_{k \in \mathbb{Z}} \min \{d(x_k, y_k), 1/|k|\}.$$

The following theorem is due to Akin and Miller (see Miller and Akin (1999), Theorem 3.2).

**Theorem 2.5.** A probability measure $\mu \in \mathcal{P}(M)$ is called an invariant measure for $F$ if it satisfies the following equivalent conditions.

1. For every Borel set $A \subset M$

   $$\mu(A) \leq \mu(F^{-1}(A)).$$

2. There exists a Markov kernel $\kappa : M \rightarrow \mathcal{P}(M)$ satisfying

   $$x \in \text{supp}(\mu) \Rightarrow \text{supp}(\kappa(x, \cdot)) \subset F(x)$$

   and

   $$\mu(\cdot) = \kappa^*(\mu)(\cdot) := \int_M \kappa(x, \cdot) \mu(dx).$$
3. There exists $\tilde{\mu} \in \mathcal{P}(M \times M)$ the space of probability measures on $M \times M$ such that
\[ \text{supp}(\tilde{\mu}) \subset Graph(F) \]
and
\[ \mu = \pi_1^*(\tilde{\mu}) = \pi_2^*(\tilde{\mu}), \]
where $\pi_i : M \times M \to M$ is the $i$th projection.

4. There exists a probability measure $\nu$ on $M^\mathbb{Z}$ which is invariant with respect to the shift homeomorphism $\tilde{\Theta}$, satisfying
\[ \text{supp}(\nu) \subset S_F \]
and
\[ \mu = \tilde{\pi}_0^*(\nu), \]
where $\tilde{\pi}_0 : x \in M^\mathbb{Z} \mapsto x_0 \in M$.

The set $\mathcal{P}_F(M)$ of $F$-invariant measures is nonempty, compact and convex in $\mathcal{P}(M)$. In general, if $\mu \in \mathcal{P}_F(M)$ then
\[ \text{supp}(\mu) \subset \tilde{\pi}_0^*(S_F). \]

2.3. Invariant measures for set-valued dynamical systems

For a set-valued dynamical system $\Phi$, we now give two equivalent definitions of an invariant measure. Notice that, by definition, for any $t \in \mathbb{R}_+$, the set-valued map $\Phi_t : M \rightrightarrows M$ is a closed relation with nonempty images (hence its set of solutions $S_{\Phi_t}$ is nonempty). Let us introduce the Lipschitz map $\pi_0 : \mathcal{C}(\mathbb{R}, M) \to M$ defined by $\pi_0(y) = y(0)$ and the translation flow $\Theta : \mathbb{R} \times \mathcal{C}(\mathbb{R}, M) \to \mathcal{C}(\mathbb{R}, M)$, which associates to a real number $t$ and $y \in \mathcal{C}(\mathbb{R}, M)$ the translated map $\Theta_t(y)$, defined by
\[ \Theta_t(y)(s) = y(t + s). \]

**Theorem 2.6.** Given a probability measure $\mu \in \mathcal{P}(M)$, the two following statements are equivalent:

(i) for any $t \geq 0$, there exists a probability measure $\nu_t$ on $S_{\Phi_t}$ such that
\[ 1. \ \tilde{\pi}_0^*(\nu_t) = \mu, \]
\[ 2. \ \nu_t \text{ is } \tilde{\Theta}-\text{invariant}, \]

(ii) there exists a probability measure $\nu$ on $S_{\Phi}$ such that
\[ a) \ \pi_0^*(\nu) = \mu, \]
\[ b) \ \nu \text{ is } \Theta\text{-invariant}. \]

Such a probability measure will be called an invariant measure for the set-valued dynamical system $\Phi$. We call $\mathcal{P}_\Phi(M)$ (or $\mathcal{P}_\Phi$) the set of invariant measures for $\Phi$. It is a nonempty compact convex subset of $\mathcal{P}(M)$.
**Proof.** In order to prove $(i) \Rightarrow (ii)$, we define, for all $t > 0$, a new relation

$$K_t : Gr(\Phi_t) \Rightarrow S_{\Phi}^{[0,t]}$$

which associates, to $(x, y) \in Gr(\Phi_t)$, the set

$$K_t(x, y) = \{ z \in S_{\Phi}^{[0,t]} : z(0) = x ; z(t) = y \}.$$ 

Notice that $K_t(x, y)$ is nonempty and $K_t$ is a closed relation for all $t > 0$. Indeed, assume that the sequence $(x_n, y_n)_n$ converges to $(x, y) \in Gr(\Phi_t)$ and that, for all $n$, $z^n \in K_t(x_n, y_n)$ and $z^n \to z \in S_{\Phi}^{[0,t]}$. We easily have $z(0) = x$ and $z(t) = y$ and so $z \in K_t(x, y)$.

Since $S_{\Phi}^{[0,t]}$ is compact, closedness of $K_t$ is equivalent to upper semicontinuity, which can also be stated: for any $A$ closed in $S_{\Phi}^{[0,t]}$, $K_t^{-1}(A)$ is closed in $M \times M$. In particular, $K_t$ is measurable: for any closed $A \subset S_{\Phi}^{[0,t]}$, $K_t^{-1}(A)$ is a Borel set. By Theorem 8.1.3 of Aubin and Frankowska (2009) we can therefore choose, for all $t > 0$, a measurable selection of $K_t$,

$$\kappa_t : Gr(\Phi_t) \to S_{\Phi}^{[0,t]}.$$ 

Let $t > 0$ be fixed for now. We claim that there exists a measurable application $h_t : S_{\Phi_t} \to S_{\Phi}$ which conjugates the shift operators $\tilde{\Theta}$ and $\Theta_t$:

$$h_t \circ \tilde{\Theta} = \Theta_t \circ h_t.$$ (2)

In order to prove (2), we now define two sets:

$$A_t := \{ (x_n, y_n)_{n \in \mathbb{Z}} \in Gr(\Phi_t)^\mathbb{Z} , \ | \ y_n = x_{n+1} \text{ for all } n \};$$

and

$$B_t := \{ (z^n)_{n \in \mathbb{Z}} \in (S_{\Phi}^{[0,t]})^\mathbb{Z} \ | \ z^{n+1}(0) = z^n(t) \}.$$ 

Notice that $S_{\Phi_t}$ endowed with the metric $\delta$ (see previous section) is topologically equivalent to $A_t$, seen as a subset of the product space $Gr(\Phi_t)^\mathbb{Z}$, equipped with the induced product topology. Similarly, $(S_{\Phi}, D)$ is topologically equivalent to the set $B_t$, understood as a subset of the product space $(S_{\Phi}^{[0,t]})^\mathbb{Z}$, equipped with the product of the uniform convergence topology on $[0, t]$. We now construct a measurable function $\tilde{h}_t$ from $A_t$ to $B_t$ (to which can be associated a measurable function $h_t$ from $S_{\Phi_t}$ to $S_{\Phi}$) the following way: let $(x, y) = (x_n, y_n)_n$ be in $A_t$. Then $\tilde{h}_t(x, y)$ is given by $z = (z^n)_n$, where

$$z^n(s) = \kappa_t(x_n, y_n)(s), \ \forall s \in [0, t].$$

In other terms, $\tilde{h}_t$ is the countable product of the measurable map $\kappa_t$ and therefore is measurable.

Now the corresponding map $h_t : S_{\Phi_t} \to S_{\Phi}$ is also measurable. To understand why the conjugacy (2) holds, let us give some insights on the map $h_t$: given $(x_n)_n \in S_{\Phi_t}$,
we consider, for each couple \((x_n, x_{n+1})\) its image by \(\kappa_t\) (which is a partial solution curve of length \(t\)) and then build a solution \(z \in S_\Phi\) by joining together these partial solution curves. The conjugacy is a clear consequence of this construction.

By assumption, for any \(t \geq 0\), there exists a probability measure \(\nu_t\) on \(S_\Phi\) such that

1. \(\pi_0^\ast(\nu_t) = \mu\)
2. \(\nu_t\) is invariant for \(\tilde{\Theta}\).

Let \((t_n)_n\) be a strictly decreasing sequence, converging to 0 and \(\nu\) be a limit point of \((h_{t_n}^\ast(\nu_{t_n}))_n\) (such a point exists since the considered sequence is tight). Notice that \(\nu\) sits on \(S_\Phi\). We can assume, without loss of generality, that \(\nu = \lim_{n \to \infty} h_{t_n}^\ast(\nu_{t_n})\).

Since \(\pi_0 \circ h_t = \tilde{\pi}_0\), we have \(\pi_0^\ast(\nu) = \mu\). There remains to prove that \(\nu\) is \(\Theta\)-invariant. Let \(f\) be a continuous function on \(S_\Phi\) and \(T \geq 0\). It is sufficient to prove

\[
\int_{S_\Phi} f(z)\nu(dz) = \int_{S_\Phi} f(\Theta_T(z))\nu(dz). \tag{3}
\]

We call

\[
\xi_n = \int f(z)(h_{t_n}^\ast(\nu_{t_n}))(dz) \quad \text{and} \quad \xi'_n = \int f \circ \Theta_T(z)(h_{t_n}^\ast(\nu_{t_n}))(dz).
\]

The two members of equation (3) can be rewritten in the form

\[
\xi := \lim_{n \to \infty} \xi_n \quad \text{and} \quad \xi' := \lim_{n \to \infty} \xi'_n.
\]

Recall that \(\Theta_t \circ h_t = h_t \circ \tilde{\Theta}\) for all \(t \geq 0\). Call \(s_n = T - \frac{T}{t_n}t_n\). Then

\[
\Theta_T \circ h_{t_n} = \Theta_{s_n} \circ (\Theta_{t_n})^{[\frac{T}{t_n}]} \circ h_{t_n} = \Theta_{s_n} \circ h_{t_n} \circ \tilde{\Theta}^{[\frac{T}{t_n}]}.
\]

Since \(\nu_{t_n}\) is \(\tilde{\Theta}\)-invariant, we get:

\[
\xi'_n = \int f \circ \Theta_{s_n}(z)(h_{t_n}^\ast(\nu_{t_n}))(dz).
\]

Now we prove that \(|\xi_n - \xi'_n|\) converges to zero. Pick \(\varepsilon > 0\). Since \(S_\Phi\) is compact, \((t, z) \mapsto \Theta_t(z)\) is continuous and \(s_n \to 0\), there exists \(N\) large enough so that

\[
|f \circ \Theta_{s_n}(z) - f(z)| < \varepsilon,
\]

for all \(n \geq N\) and \(z \in S_\Phi\). Since \(\nu_{t_n}\) is a probability measure for all \(n\), we get the equation (3).

Conversely, assume that \((ii)\) holds. There exists a probability measure \(\nu\) which satisfies \(a)\) and \(b)\). For any \(t \geq 0\) we define the application \(g_t : S_\Phi \to S_{\Phi_t}\) which associates to \(z\), \(g_t(z) = (z(kt))_{k \in \mathbb{Z}}\) and another probability measure

\[
\nu_t := g_t^\ast(\nu).
\]
Since \( \tilde{\pi}_0 \circ g_t = \pi_0 \), we have \( \tilde{\pi}_0^*(\nu_t) = \tilde{\pi}_0^*(g_t^*(\nu)) \). Now we show that \( \nu_t \) is \( \tilde{\Theta} \)-invariant.

Since \( \tilde{\Theta} \circ g_t = g_t \circ \Theta_t \), we have
\[
\tilde{\Theta}_t^*(\nu_t) = \tilde{\Theta}_t^*(g_t^*(\nu)) = (g_t \circ \Theta_t)^*(\nu) = g_t^*((\Theta_t)^*(\nu)) = g_t^*(\nu) = \nu_t
\]

The set \( \mathcal{P}_{\Theta|\mathcal{S}_\Phi} \) of \( \Theta|\mathcal{S}_\Phi \)-invariant probability measures is a convex and compact nonempty set (by Krylov-Bogolubov Theorem). Since \( \mathcal{P}_\Phi \) is the image of \( \mathcal{P}_{\Theta|\mathcal{S}_\Phi} \) under \( \pi_0^* \), we easily check that it enjoys the same properties.  ■

**Remark 2.7.** If \( \Phi = \phi \) is a flow on \( M \) then \( \pi_0 \) restricts to an homeomorphism of \( S_\Phi \) to \( M \). Consequently, \( \mu \) is \( \phi \)-invariant if and only if it is the image of a \( \Theta \)-invariant measure (with its support in \( S_\Phi \)) under \( \pi_0^* \).

**Remark 2.8.** Another natural way to define invariant measures relative to a SVDS is to consider the set of half solutions \( S_\Phi^+ \) instead of \( S_\Phi \). Let us define the map \( \Theta^+ : \mathbb{R}_+ \times C(\mathbb{R}_+, M) \rightarrow C(\mathbb{R}_+, M) \) as the shift semi-flow defined similarly to \( \Theta \).

A probability measure \( \mu \) on \( M \) is then said to be a semi-invariant measure for the set-valued dynamical system \( \Phi \) if there exists a probability measure \( \nu^+ \) on \( C(\mathbb{R}_+, M) \) such that
\[
\begin{align*}
(i) & \quad \text{supp}(\nu^+) \subset S^+_\Phi, \\
(ii) & \quad \nu^+ \text{ is } \Theta^+-\text{invariant, and} \\
(iii) & \quad (\pi_0^+)^*(\nu^+) = \mu, \text{ where } \pi_0^+ : C(\mathbb{R}_+, M) \rightarrow M, \pi_0^+(\tilde{y}) = \tilde{y}(0).
\end{align*}
\]

Semi-invariance is, a priori, very similar to invariance (as defined in Theorem 2.6). We discuss the relationship between these two definitions in Appendix C.

### 2.4. Poincaré recurrence theorem for set-valued dynamical systems

As an application of the definition of an invariant measure, we shall state a topological version of the Poincaré recurrence theorem for set-valued dynamical systems. Roughly speaking, this theorem says that invariant measures sit on the closure of the set of recurrent points of the dynamic. The concept of recurrent point is closely related to the notion of \( \omega \)-limit set of a point \( x \in M \), defined by
\[
\omega_\Phi(x) := \bigcap_{t \geq 0} \Phi_{[t, \infty)}(x)
\]

It is characterized by the following: \( y \in \omega_\Phi(x) \) if and only if there exists \( (t_n)_n \uparrow \infty \), \( (z^n)_n \subset S_\Phi(x) \) such that \( z^n(t_n) \rightarrow y \) (see Benaïm, Hofbauer and Sorin (2005)). Notice that \( \omega_\Phi(x) \) contains the limit sets \( L(z) = \bigcap_{t \geq 0} z([t, \infty)) \) of all solutions \( z \) with \( z(0) = x \) but is in general larger than
\[
L(x) := \bigcup_{z \in S_\Phi(x)} L(z).
\]
(See Benaïm, Hofbauer and Sorin (2005))

In the classical framework of a flow, a point is recurrent provided it belongs to its own ω-limit set and the topological version of Poincaré recurrence theorem is stated as follows

**Theorem 2.9** (Poincaré). Let $(X, d)$ be a separable metric space and $\theta = \{\theta_t\}_{t \in \mathbb{R}}$ a flow on $X$. Let $\mu$ be an invariant measure for $\theta$. The closure of the set of all recurrent points,

$$R_\theta^\omega := \{x \in X : x \in \omega_\theta(x)\},$$

is called the Birkhoff center of $\theta$ and denoted $BC(\theta)$. Then

$$\mu(BC(\theta)) = 1.$$

The same definition of recurrence does not seem to fit in the set-valued framework. Intuitively, a point $x$ is recurrent if there exists an entire solution starting from $x$, whose limit set contains $x$. Thus it is more natural to define the set of recurrent points of $\Phi$ by

$$R_\Phi := \{x \in M : x \in L(x)\}.$$

Clearly we have $R_\Phi \subset R_\theta^\omega := \{x \in M : x \in \omega_\theta(x)\}$. The closure of $R_\Phi$ will be called the Birkhoff center of $\Phi$ and will be noted $BC(\Phi)$. Notice that, if $\Phi$ is actually a flow, then $L(x) = \omega(x)$ and consequently $R_\theta^\omega = R_\Phi$. The following statement is a Poincaré recurrence theorem for set-valued dynamical system.

**Theorem 2.10.** Let $\mu$ be an invariant measure for $\Phi$, then

$$\mu(BC(\Phi)) = 1.$$

**Proof.** Let $\mu$ be an invariant measure for $\Phi$ and $\nu$ be an invariant measure for $\Theta|_{S_\Phi}$ such that $\pi_0^\Theta(\nu) = \mu$. First of all, notice that

$$\pi_0(BC(\Theta)) \subset BC(\Phi).$$

Indeed, pick $z \in R_\theta^\omega|_{S_\Phi}$. There exists a sequence $t_n \uparrow +\infty$ such that $\Theta_{t_n}(z) \rightarrow_n z$. In particular, $\pi_0(z) = z(0) = \lim_n z(t_n)$, which means that $\pi_0(z) \in L(z(0))$. Using the last inclusion, we get

\[
\begin{align*}
\mu(BC(\Phi)) &\geq \mu(\pi_0(BC(\Theta))) \\
&= \nu(\pi_0^\Theta \circ \pi_0(BC(\Theta))) \\
&\geq \nu(BC(\Theta)).
\end{align*}
\]

The last quantity is equal to one by Theorem 2.9. ■
3. Ergodic properties of weak asymptotic pseudotrajectories

3.1. Definition

Let \((\Omega, \mathcal{F}, P)\) be a probability space and \(\{\mathcal{F}_t : t \geq 0\}\) a nondecreasing family of sub-\(\sigma\)-algebras. We define a process

\[ X : \mathbb{R}_+ \times \Omega \rightarrow M, \]

to be a weak asymptotic pseudotrajectory (WAPT) for the SVDS \(\Phi\) if it enjoys the following properties

(i) for almost every \(\omega\), the path \(X(\cdot, \omega)\) is uniformly continuous;
(ii) it is progressively measurable: \(X_{\lfloor 0,T \rfloor} \times \mathcal{F}_T\) measurable for all \(T > 0\);
(iii) for each \(\alpha > 0\) and \(T > 0\), we have

\[
\lim_{t \to \infty} P\left( \inf_{z \in S_{\Phi}, s \in [0,T]} d(X(t+s), z(s)) \geq \alpha \big| \mathcal{F}_t \right) = 0
\]

almost surely.

3.2. Ergodic behavior

Given a weak asymptotic pseudotrajectory \(X\) for \(\Phi\) and \(\omega \in \Omega\), let \(\mu_t(\omega)\) denote the empirical occupation measure of \(X(\cdot, \omega)\):

\[
\mu_t(\omega) := \frac{1}{t} \int_0^t \delta_{X(s, \omega)} ds.
\]

Remark 3.1. \(\mu_t(\omega)\) is defined as the unique Borel measure on \(M\) such that, for all continuous function \(f\) on \(M\),

\[
\frac{1}{t} \int_0^t f(X(s, \omega)) ds = \int_{x \in M} f(x) \mu_t(\omega)(dx).
\]

Let \(\mathcal{P}(X, \omega) \subset \mathcal{P}(M)\) denote the set of weak* limit points of \(\{\mu_t(\omega)\}_{t \geq 0}\). Notice that, since \(M\) is compact, \(\mathcal{P}(X, \omega)\) is nonempty and compact. We now state the main result of this section:

Theorem 3.2. Let \(X\) be a weak asymptotic pseudotrajectory for the set-valued dynamical system \(\Phi\). Then

\[
\mathcal{P}(X, \omega) \subset \mathcal{P}_\Phi, \text{ almost surely.}
\]

In particular,

\[
\bigcup_{\mu \in \mathcal{P}(X, \omega)} \text{supp}(\mu) \subset \text{BC}(\Phi), \text{ almost surely.}
\]
Before proving this theorem, we state some useful lemmas. First we introduce a family of probability measures \( \{ \nu_t \}_{t \geq 0} \) on \( \mathcal{C}(\mathbb{R}, M) \) related to the family \( \{ \mu_t \}_{t \geq 0} \) : given \( \omega \in \Omega \),

\[
\nu_t(\omega) := \frac{1}{t} \int_0^t \delta_{\Theta_s(X(\cdot, \omega))} ds,
\]

where we use the convention that an element \( X \in \mathcal{C}(\mathbb{R}_+, M) \) can also be seen as an element of \( \mathcal{C}(\mathbb{R}, M) \) with \( X(t) = X(0) \) for all \( t < 0 \).

**Lemma 3.3.**  The set \( \{ \Theta_t(\mathcal{C}(\cdot, \omega)) : t \geq 0 \} \) is almost surely relatively compact.

**Proof.** Hypothesis (i) in the definition of a WAPT guarantees that the set of functions \( \{ \Theta_t(\mathcal{C}(\cdot, \omega)) : t \geq 0 \} \) is almost surely equicontinuous. Since \( \mathcal{C}(\mathbb{R}, \omega) \subset M \) (which is compact), we can apply Ascoli-Arzela’s theorem to conclude. \( \blacksquare \)

**Lemma 3.4.**  The family \( \{ \nu_t(\omega) \}_{t \geq 0} \) is almost surely tight.

**Proof.**  By Lemma 3.3, we know that \( \{ \Theta_t(\mathcal{C}(\cdot, \omega)) : t \geq 0 \} \) is almost surely relatively compact. The tightness follows since the support of \( \nu_t \) is included in \( \{ \Theta_t(\mathcal{C}(\cdot, \omega)) : t \geq 0 \} \), for any \( t \geq 0 \). \( \blacksquare \)

The last lemma is a generalization (in the continuous case) of Theorem 6.9 in Walters (2000). A short proof is provided in appendix B for convenience.

**Lemma 3.5.**  Let \( (X, d) \) be a compact metric space, \( \theta = (\theta_t)_{t \in \mathbb{R}} \) be a flow on \( X \) and \( \{ \sigma_t \}_{t \geq 0} \) be a collection of probability measures on \( X \). Consider the family \( \{ \nu_t \}_{t \geq 0} \) of probability measures on \( X \), defined by

\[
\nu_t = \frac{1}{t} \int_0^t \theta_s^*(\sigma_t) ds.
\]

Then any limit point \( \nu \) of \( \{ \nu_t \}_{t \geq 0} \) is \( \theta \)-invariant.

We are now ready to prove our main result.

**Proof of Theorem 3.2.**

For all \( \omega \in \Omega \) and for all \( \mu \in \mathcal{P}(\mathcal{C}(\mathbb{R}, M)) \), there exist \( (t_j)_{j \geq 0} \) going to infinity and a probability measure \( \nu \) on \( \mathcal{C}(\mathbb{R}, M) \) such that

1. \( \mu_{t_j} \to \mu \) and \( \nu_{t_j} \to \nu \),
2. \( \nu \) is \( \Theta \)-invariant,
3. \( \pi_0^*(\nu) = \mu \).

The first point is a direct consequence of the definition of \( \mu \) and the tightness of \( \{ \nu_t \}_{t} \) (see Lemma 3.4), the second point is a consequence of Lemma 3.5 and the last point follows from the continuity of the map \( \pi_0 \) and the fact that \( \pi_0^*(\nu_t) = \mu_t \), for all \( t \in \mathbb{R}_+ \). The set of all such \( \nu \) will be called \( A(\omega, \mu) \):

\[
A(\omega, \mu) = \{ \nu \in \mathcal{P}(\mathcal{C}(\mathbb{R}, M)) : \exists t \to \infty \text{ such that } 1., 2. \text{ and } 3. \text{ hold} \}
\]

Let \( A(\omega) = \cup_{\mu \in \mathcal{P}(\mathcal{C}(\cdot, \omega))} A(\omega, \mu) \). We have \( A(\omega) \subset \mathcal{P}(\mathcal{C}(\mathbb{R}, M)) \), the set of \( \Theta \)-invariant probability measures on \( \mathcal{C}(\mathbb{R}, M) \). We now exhibit a set \( \tilde{\Omega} \subset \Omega \) of full measure such
that for all $\omega \in \tilde{\Omega}$ and for all $\nu \in A(\omega)$, $\text{supp}(\nu) \subset S_{\Phi}$. Let $\{C_k\}$ be the family of closed neighborhoods of $S_{\Phi}$ defined by

$$C_k = \{z \in \mathcal{C}(\mathbb{R}, M) : D(z, S_{\Phi}) \leq \frac{1}{k}\}.$$ 

It is sufficient to find, for all $k \geq 0$, a set $\Omega_k \subset \Omega$ of full measure such that for all $\omega \in \Omega_k$,

$$\bigcup_{\nu \in A(\omega)} \text{supp}(\nu) \subset C_k.$$

Let $k \in \mathbb{N}$ and $N \in \mathbb{N}$ (large) be fixed. First of all we choose $T > 0$ such that $T/N \in \mathbb{N}$ and $\sum_{j=T/N}^{\infty} \frac{1}{j^2} < \frac{1}{2k}$. Let $\delta > 0$ be small enough such that $\delta \sum_{j=0}^{T/N} \frac{1}{j^2} < \frac{1}{2k}$. With these choices of $T$ and $\delta$, we have for all $y \in \mathcal{C}(\mathbb{R}, M)$ and $s \geq 0$,

$$\sup_{u \in [-T/N,T/N]} d(\Theta_s(X)(u), y(u)) \leq \delta \Rightarrow D(\Theta_s(X), y) \leq \frac{1}{k}.$$ 

Therefore we have for all $i \geq 1$,

$$\sup_{u \in [0,T]} d(\Theta_{(i-1)T}(X)(u), y(u)) \leq \delta \Rightarrow D(\Theta_s(X), \Theta_{s-(i-1)T}(y)) \leq \frac{1}{k}, \quad (4)$$

for all $s \in [(i-1 + 1/N)T, (i-1/N)T]$. Since $S_{\Phi}$ is invariant for $\Theta$, (4) implies that the event

$$\left\{ \inf_{z \in S_{\Phi}} \sup_{u \in [0,T]} d(\Theta_{(i-1)T}(X)(u), z(u)) \leq \delta \right\}$$

is contained in the event

$$\left\{ \inf_{z \in S_{\Phi}} D(\Theta_s(X), z) \leq \frac{1}{k}, \forall s \in [(i-1 + 1/N)T, (i-1/N)T] \right\}. \quad (5)$$

For $n \geq 1$ set

$$U_n = I_{\{\inf_{z \in S_{\Phi}} \sup_{u \in [0,T]} d(\Theta_{(i-1)T}(X)(u), z(u)) > \delta\}}$$

and

$$M_n = \sum_{i=1}^{n} \frac{1}{i} \left( U_i - \mathbb{E}(U_i | \mathcal{F}_{(i-1)T}) \right).$$

Since $M_n$ is a martingale and $\sup_n \mathbb{E}(M_n^2) \leq 4 \sum_{i=1}^{\infty} \frac{1}{i^2}$, Doob’s convergence theorem implies that $(M_n)$ converges almost surely. Hence, by Kronecker lemma,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( U_i - \mathbb{E}(U_i | \mathcal{F}_{(i-1)T}) \right) = 0 \quad (6)$$
almost surely. In other terms, there exists $\Omega_{k,N} \subset \Omega$ (a subset of full measure) such that for all $\omega \in \Omega_{k,N}$, (6) holds. Now pick $\omega \in \Omega_{k,N}$ and $\nu \in A(\omega)$. Let $n_j = [t_j/T]$. Then

$$
\nu(C_k) \geq \lim_{j \to \infty} \frac{1}{n_jT} \sum_{i=1}^{n_j} \int_{(i-1/2N)T}^{(i-1/2N+1/2N)T} \delta_{\Theta_s(X)}(C_k)ds
$$

$$
\geq (1 - 2/N) \lim_{j \to \infty} \frac{1}{n_j} \sum_{i=1}^{n_j} \frac{1}{T(1-2/N)} \int_{(i-1/2N)T}^{(i-1/2N+1/2N)T} \delta_{\Theta_s(X)}(C_k)ds
$$

$$
\geq (1 - 2/N) \left(1 - \lim_{j \to \infty} \frac{1}{n_j} \sum_{i=1}^{n_j} U_i\right)
$$

$$
\geq (1 - 2/N) \left(1 - \lim_{j \to \infty} \frac{1}{n_j} \sum_{i=1}^{n_j} \left(U_i - \mathbb{E}(U_i|F_{(i-1)T})\right) + \frac{1}{n_jT} \sum_{i=1}^{n_j} \mathbb{E}(U_i|F_{(i-1)T})\right).
$$

The first term in the last equality is equal to zero. Additionally, by definition of a WAPT, we may assume without loss of generality that, on $\Omega_{k,N}$,

$$
\lim_{t \to \infty} \mathbb{P}\left(\inf_{\varepsilon \in S_\delta} \sup_{u \in [0,T]} d(X(t+u),z(u)) \geq \delta \right| F_t) = 0.
$$

Consequently, $\nu(C_k) \geq 1 - 2/N$ for all $\nu \in \cup_{\omega \in \Omega_{k,N}} A(\omega)$, which means that $\nu(C_k) = 1$ for all $\nu \in \cup_{\omega \in \Omega_{k,N}} A(\omega)$, with $\Omega_k := \cap_N \Omega_{k,N}$. We conclude the proof by setting $\tilde{\Omega} = \bigcap_k \Omega_k$. 

3.3. **A simple deterministic example**

Notice that the main result of previous section is not useless in the case of APTs. To understand why, let us consider a set-valued dynamical system on the circle $S^1$, defined by the differential inclusion $\frac{dx}{dt} \in F(x)$ with

$$
F(x) = \begin{cases} [0,1] & \text{if } x = 0 \\ 1-x & \text{if } 0 < x < 1. \end{cases}
$$

The dynamics have the following portrait
We easily check that the only two internally chain transitive sets are \( \{0\} \) and \( S^1 \). Let \( X \) be a bounded APT of this dynamics. Then by Theorem 4.3 of Benaïm, Hofbauer and Sorin (2005) we know that \( L(X) \) is either \( \{0\} \) or \( S^1 \). Moreover the only invariant measure for the dynamic is the Dirac measure on \( 0 \). Therefore by Theorem 3.2 we know that any limit point of the empirical measure of \( X \) is \( \delta_0 \). In particular, this means that \( X \) spends most of its time near \( 0 \).

4. Weak perturbed solution of a differential Inclusion

In this section, we give some examples of WAPT relative to a particular case of SVDS: the set-valued dynamical systems induced by a differential inclusion. In the whole section, we are in the particular case where \( E \) is the Euclidian space \( \mathbb{R}^m \).

**Definition 4.1.** A set-valued map \( F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m \) is said to be standard if it satisfies the following assumptions:

(i) for any \( x \in \mathbb{R}^m \), \( F(x) \) is a nonempty, compact and convex subset of \( \mathbb{R}^m \),

(ii) \( F \) is upper semicontinuous (see Definition 2.1),

(iii) there exists \( c > 0 \) such that

\[
\sup_{z \in F(x)} \|z\| \leq c(1 + \|x\|).
\]

Under the above assumptions (Definition 4.1), it is well known (see Aubin and Cellina (1984)) that the differential inclusion

\[
\frac{dz}{dt} \in F(z)
\]

admits at least one solution (i.e. an absolutely continuous mapping \( z : \mathbb{R} \rightarrow \mathbb{R}^m \) such that \( \dot{z}(t) \in F(z(t)) \) for almost every \( t \) through any initial point. To any \( x \in \mathbb{R}^m \) and \( t \in \mathbb{R}_+ \), we can therefore associate the nonempty set

\[
\Phi_t(x) := \{z(t) \mid z \text{ is a solution of } (8), \ z(0) = x\}.
\]

It is not hard to check that \( \Phi = (\Phi_t)_{t \in \mathbb{R}} \) is a complete set-valued dynamical system (see e.g. Benaïm, Hofbauer and Sorin (2005)).
Definition 4.2. A set $M \subset \mathbb{R}^m$ is invariant for $\Phi$ if, for every $x \in A$, there exists an entire solution curve $z$ such that $z(\mathbb{R}) \subset M$.

Let $M \subset \mathbb{R}^m$ be a compact and invariant subset (for $\Phi$). In the following we will consider the complete SVDS restricted to $M$, that we will also write $\Phi$ and that is defined, for all $x \in M$ and $t \in \mathbb{R}_+$, by

$$\Phi_t(x) := \{z(t) \mid z \text{ is a solution of (8), } z(0) = x, z(\mathbb{R}) \subset M\}.$$

Given a positive number $\delta$, let $F_\delta$ be the set-valued map defined by

$$F_\delta(x) := \{y \mid \exists z \in B(x, \delta) \text{ such that } d(y, F(z)) < \delta\}.$$

Definition 4.3. Given a function $\delta : [0, +\infty) \to [0, 1]$ decreasing to zero as $t$ goes to infinity and a locally integrable process $\overline{U} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$, we say that a process $Y : \mathbb{R}_+ \times \Omega \to M$ is a $(\delta, \overline{U})$-weak perturbed solution of the differential inclusion (8) provided

(i) $Y$ is absolutely continuous for all $\omega$,
(ii) for almost every $t > 0$,

$$\frac{dY(t)}{dt} - \overline{U}(t) \in F_\delta(Y(t)),$$

(iii) for any $T > 0$ and any $\gamma > 0$,

$$\lim_{t \to +\infty} P\left(\sup_{s \in [0, T]} \left\| \int_t^{t+s} \overline{U}(u)du \right\| \geq \gamma|\mathcal{F}_t\right) = 0,$$

almost surely.

Theorem 4.4. Assume that $Y$ is a $(\delta, \overline{U})$-weak perturbed solution of the differential inclusion (8) and that $\overline{U}$ is uniformly bounded by a positive constant $C$:

$$\sup_{\omega \in \Omega} \sup_{t \in [0, T]} \overline{U}(t, \omega) \leq C.$$ Then $Y$ is a weak asymptotic pseudotrajectory of $\Phi$.

Proof. Let $T > 0$ and define $\|F\| := \sup_{x \in M} \sup_{y \in F(x)} \|y\| < \infty$. Consider the compact set

$$K := \{y \in Lip([0, T], \mathbb{R}^m) \mid Lip(y) \leq \|F\| + C + 1, y(0) \in M\},$$

where $Lip([0, T], \mathbb{R}^m)$ denotes the set of Lipschitz functions on $[0, T]$ and $Lip(y)$ is the Lipschitz constant of $y$. The set $K$ is well adapted to our problem because it contains every solution curve of (8), restricted to an interval of length $T$ and every realization of any $(\delta, \overline{U})$-weak perturbed solution of the differential inclusion.

For $\delta \in [0, 1]$, let us define the set-valued application (with the convention $\Lambda^0 = \Lambda$):

$$\Lambda_\delta : K \Rightarrow K, \quad z \mapsto \Lambda_\delta(z),$$

(10)
where \( y \in \Lambda^\delta(z) \) if and only if there exists an integrable function \( h : [0,T] \to \mathbb{R}^m \) such that \( h(u) \in F^\delta(z(u)) \forall u \in [0,T] \) and
\[
y(\tau) = z(0) + \int_0^\tau h(u)du, \quad \forall \tau \in [0,T].
\]
Notice that \( \text{Fix}(\Lambda) := \{ z \in K \mid z \in \Lambda(z) \} \) is equal to \( S^\delta_{\Phi}[0,T] \), the set of partial solutions of \( \Phi \) on \([0,T]\).

By (ii), we have
\[
\frac{dY(t)}{dt} - U(t) \in F^\delta(t)(Y(t)), \quad \text{for almost every } t > 0.
\]

There exists an integrable function \( h : [0,T] \to \mathbb{R}^m \) such that \( h(u) \in F^\delta(t)(Y(t + u)) \forall u \in [0,T] \) and, for any \( \tau \in [0,T] \),
\[
Y(t + \tau) - \int_t^{t+\tau} U(u)du = Y(t) + \int_0^\tau h(u)du.
\]

Hence,
\[d_{[0,T]}(Y(t + \cdot), \Lambda^\delta(t)(Y(t + \cdot))) \leq \sup_{s \in [0,T]} \left| \int_t^{t+s} U(u)du \right|.
\]

Let \( \alpha > 0 \). The following statement is an immediate consequence of Corollary 4.11 in Faure and Roth (2010): there exists \( \gamma > 0 \) (which depends on \( T \) and \( \alpha \)) and \( \delta_0 > 0 \) such that, for any \( \delta < \delta_0 \)
\[
d_{[0,T]}(z, \Lambda^\delta(z)) < \gamma \Rightarrow d_{[0,T]}(z, S_{\Phi}) < \alpha.
\]

Consequently, for \( t \) large enough,
\[
d_{[0,T]}(Y(t + \cdot), S_{\Phi}) \geq \alpha \Rightarrow d_{[0,T]}(Y(t + \cdot), \Lambda^\delta(t)(Y(t + \cdot))) \geq \gamma.
\]

For these choices of \( t \) and \( \gamma \),
\[
P(d_{[0,T]}(Y(t + \cdot), S_{\Phi}) \geq \alpha \mid \mathcal{F}_t) \leq P(d_{[0,T]}(Y(t + \cdot), \Lambda^\delta(t)(Y(t + \cdot))) \geq \gamma \mid \mathcal{F}_t)
\]
\[
\leq P\left(\sup_{s \in [0,T]} \left| \int_t^{t+s} U(u)du \right| \geq \gamma \mid \mathcal{F}_t\right).
\]

By (iii), the last term tends to zero when \( t \) goes to infinity and the proof is complete. \( \blacksquare \)

### 4.1. Stochastic approximation algorithms

Stochastic approximation algorithms were born in the early 50s through the work of Robbins and Monro (1951) and Kiefer and Wolfowitz (1952). Let \( F : \mathbb{R}^m \to \mathbb{R}^m \) be a standard set-valued map and \( M \subset \mathbb{R}^m \) be a compact subset invariant for the set-valued dynamical system induced by the differential inclusion (8).

\footnote{we call \( d_{[0,T]} \) the uniform distance on \([0,T]\)}
Definition 4.5 (Weak generalized stochastic approximation process). Let \((U_n)_n\) be an uniformly bounded \(\mathbb{R}^m\)-valued random process and \((F_n)_n\) a sequence of set-valued maps on \(\mathbb{R}^m\). We say that \((x_n)_n\) is a generalized stochastic approximation process relative to the standard set-valued map \(F\) on \(M\) if the following assumptions are satisfied:

(i) we have the recursive formula

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

(ii) the step size \((\gamma_n)_n\) is deterministic and satisfies

\[ \sum_n \gamma_n = +\infty, \quad \lim_n \gamma_n = 0, \]

(iii) for all \(n \geq 0\), \(x_n \in M\),

(iv) for all \(T > 0\) and all \(\gamma > 0\),

\[ \lim_{t \to \infty} \mathbb{P} \left( \sup \left\{ \left\| \sum_{i=n}^{k-1} \gamma_{i+1}U_{i+1} \right\| \mid k \text{ such that } \sum_{i=n}^{k-1} \gamma_i \leq T \right\} \geq \gamma \mid F_n \right) = 0. \tag{11} \]

(v) for any \(\delta > 0\), there exists \(n_0 \in \mathbb{N}\) such that

\[ \forall n \geq n_0, \quad F_n(x_n) \subset F^\delta(x_n). \]

Remark 4.6. Let \((U_n)_n\) be a \(\mathbb{R}^m\)-valued random process adapted to the filtration \((\mathcal{F}_n)_n\) such that

(i) \(\mathbb{E}(U_{n+1}|\mathcal{F}_n) = 0\);

(ii) for all \(T > 0\), we have

\[ \lim_{R \to +\infty} \sup_n \mathbb{E} \left( \left\| U_{n+1} \right\| \mathbb{1}_{\left\{ \left\| U_{n+1} \right\| \geq R \right\}} \mid \mathcal{F}_n \right) = 0. \]

Then for all \(T > 0\) and all \(\gamma > 0\), Property (11) is satisfied (see Benaïm and Schreiber (2000)).

Consider a weak generalized stochastic approximation process \((x_n)_n\). Set \(\tau_n := \sum_{i=1}^n \gamma_i\) and \(m(t) := \sup \{ j \mid \tau_j \leq t \}\). We call \(X\) the continuous time affine interpolated process induced by \((x_n)_n\), \(\gamma\) the piecewise constant deterministic process induced by \((\gamma_n)_n\) and \(\mathcal{U}\) is the piecewise constant continuous time process associated to \((U_n)_n\):

\[ X(\tau_i + s) = x_i + s \frac{x_{i+1} - x_i}{\gamma_{i+1}}, \text{ for } s \in [0, \gamma_{i+1}] \]

\[ \gamma_{i+1} \text{ for } s \in [\gamma_i, \gamma_{i+1}], \]

and \(\mathcal{U}(t) := U_{n+1}, \text{ for } t \in [\tau_n, \tau_{n+1}]\).
Theorem 4.7. The interpolated process $X$ is a WAPT. Hence the conclusions of Theorem 3.2 hold.

Proof. By straightforward computations (see the proof of proposition 1.3 in Benaïm et al. Benaïm, Hofbauer and Sorin (2005)), it is not difficult to see that $(X(t))_t$ is a weak perturbed solution associated to $U$ and

$$
\delta(t) := \inf \left\{ \delta > 0 \mid \tau_n \geq t \Rightarrow F_n(x_n) \subset F_\delta(x_n) \right\} + \gamma(t) \left( U(t) + c \left( 1 + \sup_{x \in M} F(x) \right) \right),
$$

which converges to 0. Consequently $X$ is a WAPT relative to the SVDS induced by $F$ and the proof is complete. ■

Appendix A: Proof of Proposition 2.3

Recall that an element $\tilde{z} \in \mathcal{C}(\mathbb{R}_+, M)$ can also be seen as an element of $\mathcal{C}(\mathbb{R}, M)$, with the convention $\tilde{z}(t) = \tilde{z}(0)$ for all $t < 0$. Since $S_\Phi^+ \neq \emptyset$, the existence of at least one entire solution is a direct consequence of the following Proposition.

Proposition A.1. Let $(t_n) \uparrow \infty$ be a sequence of positive real numbers converging to infinity and $\tilde{z} \in S_\Phi^+$ be a solution. Then there exists a subsequence $(t_{n_k})_{k \geq 0}$ and an entire solution $z \in S_\Phi$ such that

$$
\lim_{k \to \infty} \Theta_{t_{n_k}}(\tilde{z}) = z.
$$

Proof. Pick some positive integer $N$. By Barbashin Theorem, the set $S_\Phi^{[-N,N]}$ is compact and, for $n \in \mathbb{N}$ large enough, $\Theta_{t_n}(\tilde{z}) \in S_\Phi^{[-N,N]}$. Therefore there exist an increasing sequence $\psi^N : \mathbb{N} \to \mathbb{N}$ and $z^N \in S_\Phi^{[-N,N]}$ such that

$$
\lim_{n \to \infty} \Theta_{t_{\psi^N(n)}}(\tilde{z}) = z^N,
$$

By the same arguments, there exist an extraction $\psi$ and $z^{N+1} \in S_\Phi^{[-N-1,N+1]}$ such that

$$
\lim_{n \to \infty} \Theta_{t_{\psi^{N+1}(n)}}(\tilde{z}) = z^{N+1},
$$

and then, in particular, $z^{N+1} \subset S_\Phi^{[-N,N]} = z^N$. We set $\psi^{N+1} := \psi \circ \psi^N$ and we iterate the process. In this way we construct an entire solution $z \in S_\Phi$ such that $z^{[\neg N,N]} = z^N \forall N$.

Let $(\delta_k) \downarrow 0$ be a decreasing sequence of positive real numbers. There exists a natural number $m_1$ such that for all $m \geq m_1$,

$$
\sup_{s \in [-1,1]} d(\Theta_{t_{\psi^1(m)}}(\tilde{z})(s), z(s)) < \delta_1,
$$

We set $n_1 := \psi^1(m_1)$. Now we define $n_k$ by induction. Fix $k > 1$. There exists a natural number $m_k \geq m_{k-1}$ such that for all $m \geq m_k$,

$$
\sup_{s \in [-k,k]} d(\Theta_{t_{\psi^k(m)}}(\tilde{z})(s), z(s)) < \delta_k,
$$

$$
\lim_{k \to \infty} \Theta_{t_{\psi^k(n_k)}}(\tilde{z}) = z^{N+1},
$$

and then, in particular, $z^{N+1} \subset S_\Phi^{[-N,N]} = z^N$. We set $\psi^{N+1} := \psi \circ \psi^N$ and we iterate the process. In this way we construct an entire solution $z \in S_\Phi$ such that $z^{[\neg N,N]} = z^N \forall N$.
We set \( n_k := \psi^k(m_k) \). Therefore, by construction, we have \( \lim_{k \to \infty} \Theta_{t_{n_k}}(z) = z \). ■

Appendix B: Proof of Lemma 3.5

Let \( \mu \) be a limit point of \( \{ \mu_t \}_{t \geq 0} \): there exists a sequence \( (t_n)_n \uparrow +\infty \) such that \( \mu = \lim_{n \to \infty} \mu_{t_n} \). We have to prove that, for any \( T \) and any continuous function \( f \),

\[
\int_X f(x) d\mu(x) = \int_X f(\theta_T(x)) d\mu(x). \tag{12}
\]

We have

\[
\int_X f(x) d\mu(x) = \lim_n \frac{1}{t_n} \int_0^{t_n} \int_X f(\theta_s(x)) d\sigma_{t_n}(x) ds
\]

and

\[
\int_X f(\theta_T(x)) d\mu(x) = \lim_n \frac{1}{t_n} \int_0^{t_n} \int_X f(\theta_{s+T}(x)) d\sigma_{t_n}(x) ds.
\]

By Fubini’s Theorem we can exchange the integral operators in both expressions. Consequently,

\[
\frac{1}{t_n} \left| \int_0^{t_n} \int_X f(\theta_s(x)) ds \right| d\sigma_{t_n}(x) - \int_0^{t_n} \int_X f(\theta_{s+T}(x)) ds \right| d\sigma_{t_n}(x) \\
\leq \frac{1}{t_n} \left| \int_X \left( \int_0^{t_n} f(\theta_s(x)) ds \right) d\sigma_{t_n}(x) \right| \\
\leq \frac{1}{t_n} \int_X \left( \int_0^{t_n} f(\theta_s(x)) ds \right) d\sigma_{t_n}(x) \\
\leq 2T \| f \|_{\infty}.
\]

Finally, taking the limit as \( n \) goes to infinity, we obtain (12). ■

Appendix C: Some remarks on semi-invariance

In this section we show (Proposition C.1) that every invariant measure for a SVDS \( \Phi \) is a semi-invariant measure for \( \Phi \). To do this, we start by proving two technical lemmas. Let us define the projection \( y^+ \) on \( \mathcal{C}(\mathbb{R}_+, M) \) of an element \( y \in \mathcal{C}(\mathbb{R}, M) \) by

\[
y^+(t) = y(t), \quad \forall t \geq 0,
\]

and, in the same way, \( A^+ := \{ y^+ : y \in A \} \) for a subset \( A \subset \mathcal{C}(\mathbb{R}, M) \). Be aware that the set \( (S_\Phi)^+ \) is contained in the set \( S_\Phi^+ \) of partial solutions, but not equal in general.

**Proposition C.1.** An invariant measure \( \mu \) on \( M \) for the set-valued dynamical system \( \Phi \) is a semi-invariant measure for \( \Phi \).
**Proof.** There exists a probability measure $\nu$ on $S_\Phi$ such that

a) $\pi^0_0(\nu) = \mu$,
b) $\nu$ is $\Theta$-invariant.

We need to construct a probability measure $\nu^+$ on $C(\mathbb{R}^+, M)$ which satisfies conditions (i), (ii) and (iii) of the definition of a semi-invariant measure given in Remark 2.8. A natural way to do this is to define $\nu^+$ as

$$\nu^+(A) := \nu(B_A), \text{ for all Borel sets } A \text{ of } C(\mathbb{R}^+, M),$$

where $B_A := \{z \in S_\Phi : z^+ \in A \}$. First, we have

$$\nu^+(S^+_\Phi) = \nu(\{z \in S_\Phi : z^+ \in S^+_\Phi \}) = \nu(S_\Phi) = 1.$$ 

which gives condition (i). Let $T > 0$ and $A \subset S^+_\Phi$ a Borel set. Since, for any $z \in S_\Phi$, we have $\Theta_T^+(z) = (\Theta_T(z))^+$,

$$B_{(\Theta_T^+)^{-1}(A)} = \{z \in S_\Phi : z^+ \in (\Theta_T^+)^{-1}(A) \} = \{z \in S_\Phi : (\Theta_T(z))^+ \in A \} = \Theta_T^{-1}(B_A).$$

Therefore, as $\nu$ sits on $S_\Phi$ and is $\Theta$-invariant, we have

$$\nu^+((\Theta_T^+)^{-1}(A)) = \nu \left( B_{(\Theta_T^+)^{-1}(A)} \right) = \nu (\Theta_T^{-1}(B_A)) = \nu (B_A) = \nu^+(A),$$

which gives condition (ii).

Let $D$ be a Borel subset of $M$. Notice that $B_{(\pi^0_0)^{-1}(D)} = \pi_0^{-1}(D) \cap S_\Phi = S_\Phi(D)$. Consequently

$$(\pi^0_0)^*(\nu^+)(D) = \nu^+((\pi^0_0)^{-1}(D)) = \nu(S_\Phi(D)) = \nu(\pi^{-1}_0(D)) = \mu(D),$$

and the result holds. ■

Whether the converse statement is also true is an open question.

**Acknowledgements** The authors thank Michel Benaïm for useful advices and discussions. They also acknowledge financial support from the Swiss National Science Foundation Grant 200020 130574.

**References**

AUBIN, J. P. and CELLINA, A. (1984). *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer-Verlag New York, Inc. Secaucus, NJ, USA.

AUBIN, J., FRANKOWSKA, H. and LASOTA, A. (1991). Poincareés recurrence theorem for set-valued dynamical systems. *Ann. Polon. Math* 54 85–91.
Aubin, J.-P. and Frankowska, H. (2009). *Set-valued analysis. Modern Birkhäuser Classics*. Birkhäuser Boston Inc., Boston, MA. Reprint of the 1990 edition [MR1048347]. MR2458436

Barbashin, E. (1948). On the theory of generalized dynamical systems. *Moskov. Gos. Ped. Inst. Ucen. Zap* 2 110–133.

Benaïm, M. (1996). A Dynamical System Approach to Stochastic Approximations. *SIAM Journal on Control and Optimization* 34 437.

Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. *Séminaire de probabilités de Strasbourg* 33 1–68.

Benaïm, M. and Hirsch, M. W. (1996). Asymptotic pseudotrajectories and chain recurrent flows, with applications. *Journal of Dynamics and Differential Equations* 8 141–176.

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

Benaïm, M., Hofbauer, J. and Sorin, S. (2006). Stochastic Approximations and Differential Inclusions. Part II: Applications. *Mathematics of Operations Research* 31 673–695.

Benaïm, M. and Schreiber, S. J. (2000). Ergodic properties of weak asymptotic pseudotrajectories for semiflows. *Journal of Dynamics and Differential Equations* 12 579–598.

Bianchini, R. and Zecca, P. (1981). Attainable Set at Time T for Multivalued Differential Equations. *Nonlinear Analy. Theory, Methods & Applic.* 5 1053–1059.

Bowen, R. (1975). Omega limit sets of Axiom A diffeomorphisms. *Journal of differential equations* 18 333–339.

Conley, C. C. (1978). *Isolated Invariant Sets and the Morse Index*. American Mathematical Society.

Faure, M. and Roth, G. (2010). Stochastic approximations of set-valued dynamical systems: convergence with positive probability to an attractor. *Mathematics of Operations Research* 35 624–640.

Kiefer, J. and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics* 23 462–466.

Kloeden, P. (1975). Asymptotic invariance and limit sets of general control systems. *Journal of Differential Equations* 19 91–105.

Kloeden, P. (1978). General control systems. *Mathematical Control Theory* 119–137.

Li, D. and Zhang, X. (2002). On the stability in general dynamical systems and differential inclusions. *J. Math. Anal. Appl* 274 705–724.

Miller, W. and Akin, E. (1999). Invariant measures for set-valued dynamical systems. *Transactions of the American Mathematical Society* 351 1203–1225.

Nieuwenhuis, J. (2009). Some remarks on set-valued dynamical systems. *The ANZIAM Journal* 22 308–313.

Pemantle, R. (2007). A survey of random processes with reinforcement. *Probability Surveys* 4 1–79.

Robbins, H. and Monro, S. (1951). A Stochastic Approximation Method. *The
Annals of Mathematical Statistics 400–407.

Roxin, E. (1965). Stability in general control systems. J. Differ. Equations 1 115–150.

Walters, P. (2000). An introduction to ergodic theory. Springer Verlag.