THE ROYAL
SWEDISH
ACADEMY OF
SCIENCES

INSTITUT
MITTAG-LEFFLER

Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

Continuous Limit in Dynamics with Choice

L. Kapitanski and S. Zivanovic Gonzalez

REPORT No. 10, 2012/2013, fall

ISSN 1103-467X
ISRN IML-R- -10-12/13- -SE+fall

Continuous Limit in Dynamics with Choice

Lev Kapitanski [1] and Sanja Živanović Gonzalez[2]

[1]*Department of Mathematics, University of Miami, Coral Gables, FL 33124, USA*
[2]*Department of Mathematics and Computer Science, Barry University, Miami Shores, FL 33161, USA*

December 7, 2012

Abstract

We are interested in time evolution of systems that switch their modes of operation at discrete moments of time. The intervals between switching may, in general, vary. The number of modes may be finite or infinite. The mathematical setting for such systems is variable time step dynamics with choice. We have used this setting before to study the long term behavior of such systems. In this paper, we define and study the continuous time dynamics whose trajectories are limits of trajectories of discrete systems as time step goes to zero. The limit dynamics is multivalued. In the special case of a switched system, when the dynamics is generated by switching between solutions of a finite number of systems of ODEs, we show that our continuous limit solution set coincides with the solution set of the relaxed differential inclusion.

Acknowledgements. This work has been completed during the fall 2012 program "Hamiltonians in Magnetic Field" at the Institut Mittag-Leffler (Djursholm, Sweden). LK thanks the organizers of the program and the leadership and staff of the Institut Mittag-Leffler for their kind support and hospitality.

1 Introduction

Many if not most of natural (real life) and man-made (engineered) systems switch their regimes of operation at discrete moments of time. The intervals between switching – the dwell times – may, in general, vary. The number of regimes may be finite or infinite. It is important to understand the workings of such systems. In control theory, the systems with both continuous and discrete behavior are known as hybrid systems. There is a large literature on control, stability, etc. of such systems, see [4, 2, 10, 11]. A special class of hybrid systems is formed by the so-called switched systems, [9]. A switched system (with time-dependent switching) is governed by a finite set of (systems of) ordinary differential equations:

$$\dot{x} = F_n(x) \quad \text{on the interval } [t_n, t_{n+1}) , \quad j = 0, 1, \dots , \tag{1}$$

where each F_n is taken from a finite set of (vector-)functions $\{f_1, \dots, f_N\}$, see [9]. If we know in advance the sequence of dwell times, $\tau_0, \tau_1, \dots, \tau_n, \dots$ and the sequence of the corresponding regimes, $j_0, j_1, \dots, j_n, \dots \in \{1, 2, \dots, N\}$, then the values of $x(t)$ at the switching instances, $x_n = x(t_n)$, can be determined successively starting from $x(0) = x_0$ as

$$x_{n+1} = S_{j_n}^{\tau_n}(x_n) , \tag{2}$$

where S_j^τ is the evolution map for the equation in the regime number j, i.e., $S_j^\tau(\eta)$ is the value at time $t = \tau$ of the solution $y(t)$ of the jth constituent equation $\dot{y} = f_j(y)$ with the initial condition $y(0) = \eta$.

Control theorists also study the discrete-time switched systems where the continuous phase is missing or not important. The discrete-time switched systems are described directly by equation (2) where the maps S_j^τ are given in advance. Most of the studies of such systems deal with the (already difficult) case when all S_j^τ are constant matrices, see e.g. [11]. However, the nonlinear equations (2) appear naturally in a quantized setting when discretized equations (1) are used (or being implemented),

$$x_{n+1} = x_n + \tau_n \, f_{j_n}(x_n) . \tag{3}$$

In this example, $S_j^\tau(x) = x + \tau \, f_j(x)$, and the "dwell-time" τ is interpreted as the (variable) time-step. The occurrence of τ_n may be more intricate as in the following example:

$$x_{n+1} = x_n + \sin(\tau_n \, x_n) . \tag{4}$$

There is only one regime, while τ_n may vary over a finite set or some interval, and the evolution map depends on τ: $S^\tau(x) = x + \sin(\tau x)$.

In [7] we propose a convenient mathematical setting to work with systems like (2). Let X be the state space, let $\mathcal{J}$ be the set of labels for all possible regimes, and let $\mathcal{I}$ be the set of possible dwell-times. Suppose that to every choice of $j \in \mathcal{J}$ and $\tau \in \mathcal{I}$ corresponds a map $S_j^\tau : X \to X$. This makes it possible to consider the switched trajectories generated as in (2). We have trajectories, but we don't have dynamical system. However, it is not hard to build a dynamical system of which (2) is a part. Denote by $\Sigma_{\mathcal{J}}$ the set of all infinite one-sided strings (sequences) of symbols from $\mathcal{J}$, and denote by $\Sigma_{\mathcal{I}}$ the set of all infinite one-sided strings of numbers from $\mathcal{I}$. In the lingo of symbolic dynamics, $\Sigma_{\mathcal{J}}$ and $\Sigma_{\mathcal{I}}$ are the full one-sided shifts over the alphabets $\mathcal{J}$ and $\mathcal{I}$. The strings from $\Sigma_{\mathcal{J}}$ encode the sequences of regimes, and the strings from $\Sigma_{\mathcal{I}}$ encode the sequences of the dwell-times during which the corresponding regimes are active. It is convenient to identify the infinite strings in $\Sigma_{\mathcal{J}}$ with the maps $w : [0, 1, 2, \dots] \to \mathcal{J}$ and the infinite strings in $\Sigma_{\mathcal{I}}$ with the maps $h : [0, 1, 2, \dots] \to \mathcal{I}$. Thus, $w(n-1)$ is the label of the nth regime in the sequence w, and $h(n-1)$ is the nth time interval in the sequence h. Often we think of w as a sequence and write down the consecutive symbols as $w(0)w(1)w(2)\dots$. On the sets $\Sigma_{\mathcal{J}}$ and $\Sigma_{\mathcal{I}}$ acts the shift operator, σ, that erases the first symbol in sequences. Thus, $\sigma(w) = w(1)w(2)\dots$, or we can write $\sigma(w)(n) = w(n+1)$; in the same fashion σ acts on $\Sigma_{\mathcal{I}}$. Now, to include the trajectories (2) into a semi-dynamical system, we use the skew-product construction. Consider the product space $\mathfrak{X} = X \times \Sigma_{\mathcal{J}} \times \Sigma_{\mathcal{I}}$ and define the map $\mathfrak{S} : \mathfrak{X} \to \mathfrak{X}$ acting by the rule

$$\mathfrak{S}(x, w, h) = (S_{w(0)}^{h(0)}(x), \sigma(w), \sigma(h))\,.$$

The iterations of the map $\mathfrak{S}$ define on the space $\mathfrak{X}$ a semi-dynamical system which we call the *variable time-step dynamics with choice* associated with the maps S_j^τ. The switched trajectories (2) appear as the first coordinates of the trajectories of the points in $\mathfrak{X}$ under the iterations of the map $\mathfrak{S}$.

We used this setting in [7] to study the long-term behavior in variable time-step dynamics with choice. We were interested, in particular, in global compact attractors for such systems. Our assumptions on the spaces X, $\mathcal{J}$, $\mathcal{I}$, and on the maps S_j^τ were quite general: X could be any complete metric space, $\mathcal{J}$ was a metric compact, and $\mathcal{I}$ was a compact subset of $(0, +\infty)$. The assumptions on the maps S_j^τ were basically minimal to guarantee that for fixed j and τ the discrete-time system (S_j^τ, X) would possess a global compact attractor.

In this paper we use the same setting but for a completely different purpose. We want to understand what happens with the dynamics (2) when the intervals between switching the regimes shrink to 0.

Prior to [7] we studied global attractors in dynamics with choice when the time-step, or the dwell-time, is fixed, [5]. One of the numerical examples involved the following discrete version of the Ross-Macdonald type model of malaria transmission:

$$\begin{aligned} x_{n+1} &= x_n + \tau\,[a_n\, y_n(1 - x_n) - r_n\, x_n] \\ y_{n+1} &= y_n + \tau\,[b_n\, x_n(1 - y_n) - m_n\, y_n\,.] \end{aligned} \tag{5}$$

For computational purposes, the coefficients (a_n, b_n, r_n, m_n) were chosen from a set of two quadruples, $(4, 6, 1, 2)$ and $(2, 10, 3, 2)$ (just two different regimes, $\mathcal{J} = \{1, 2\}$). We have observed that the appearance of the global attractor depends on the size of τ. If τ is relatively large, a certain part of the boundary of the attractor looks like the feathery end of a brush stroke. For smaller τ, the same part of the boundary becomes smoother, approaching a smooth curve, it seems, as τ goes to zero. It was then that we started thinking about the limit $\tau \to 0$ in dynamics with choice.

The question of continuous limit itself suggests that $\tau = 0$ should be included in the set of admissible dwell-times, so we take $\mathcal{I} = [0, \tau_*]$, some $\tau_* > 0$. It is natural to assume that S_j^0 is the identity map for every j. However, we do not assume that S_j^τ form a semigroup with respect to τ, because we want to include examples (3) and (4). To handle various limits, we need some continuity and compactness from the maps S_j^τ. However, since the identity map is among them, we are forced to impose the assumption on X, that closed, bounded subsets of X are compact. Since we always assume that X is a complete metric space, the latter restriction still permits X to be a finite-dimensional Riemannian manifold, or a Frechét-Montel vector space. Our other assumptions on the maps S_j^τ refer to reasonable properties of the switched trajectories, and in examples can be checked with *a priori* estimates. We try to keep the number of assumptions to the minimum.

The main results can be stated informally as follows. The continuous limit dynamics associated with the maps S_j^τ is inherently multi-valued. For every initial state x_0, the set $\mathcal{F}_t(x_0)$ of the limits of the values at time t of the switched trajectories with the maximal dwell-time going to 0 is compact. For any closed, bounded set $B \subset X$, the set $\mathcal{F}_t(B) = \cup_{x \in B} \mathcal{F}_t(x)$ is compact. On the hyperspace $X^\sharp$ of all closed, bounded (= compact) subsets of X with the Hausdorff metric, the map $\mathcal{F}_t : X^\sharp \to X^\sharp$ is continuous, satisfies the semigroup property, $\mathcal{F}_{t_2} \circ \mathcal{F}_{t_1} = \mathcal{F}_{t_1+t_2}$, and thus defines a semi-dynamical system. If we assume in addition that the points of X that are close to each other can be connected by a unique "shortest curve" (with reasonable properties), then each set

$$\bigcup_{0 \le t \le T} \mathcal{F}_t(B)$$

is lined (filled) by the integral curves. We say that $\gamma(t)$ is an integral curve of the continuous limit dynamics if $\gamma : [0, T] \to X$ is a continuous map and $\gamma(t_1+t_2) \in \mathcal{F}_{t_2}(\gamma(t_1))$ for all $t_1, t_2 \in [0, T]$ with $t_1 + t_2 \in [0, T]$.

This result can be viewed as an extension of the Filippov-Ważewski relaxation theorem. An application to the switched system (1) shows that in that case the continuous limit dynamics is the same as the dynamics generated by the differential inclusion

$$\dot{x} \in \overline{co}\,(f_1(x), \ldots, f_N(x))\ . \tag{6}$$

The continuous limit dynamics for the Euler approximations (3) is also given by (6). However, our abstract results are much more general and we can apply them in a much greater variety of situations.

The paper is organized as follows. In Section 2 we set the notation for the variable time-step dynamics with choice. In Section 3 we prove all the abstract results on the properties of the maps $\mathcal{F}_t$ and define the Continuous Limit Dynamics. In Section 4 we show that the Continuous Limit Dynamics for the switched system (1) is governed by the differential inclusion (6) thus making a link with the Filippov-Ważewski theorem.

2 Abstract Setting

State space. Throughout this paper (X, d) is a complete metric space with the Heine-Borel property (i.e., every closed, bounded subset of X is compact). For example, X can be a finite-dimensional Riemannian manifold, or X can be a Montel space.

The distance from a point, x, to a set, B, is $d(x, B) = \inf_{y \in B} d(x, y)$. We denote by $e(A, B) = \sup_{x \in A} d(x, B)$ the excess of the set A over the set B. The Hausdorff distance between the sets A and B is the maximum of the two numbers, $e(A, B)$ and $e(B, A)$; we write

$$d_H(A, B) = \max\{e(A, B), e(B, A)\} .$$

Regime switching functions. Defined on X is a two-parameter family of maps S_j^τ. Here j represents the regime. The space of available regimes will be denoted $\mathcal{J}$. We assume that $\mathcal{J}$ is a compact metric space with the metric $d_{\mathcal{J}}$. The superscript τ shows the dwell time (along which the regime j is engaged). We take τ from the interval $\mathcal{I} = [0, \tau_*]$, some $\tau_* > 0$. The metric on $\mathcal{I}$ is induced from $\mathbb{R}$.

A regime switching function is an infinite string of symbols from $\mathcal{J}$. Denote by $\Sigma_{\mathcal{J}}$ the space of all strings $w = w(0)w(1)w(2)\dots$ with $w(n) \in \mathcal{J}$. This is the full one-sided shift on the alphabet $\mathcal{J}$. Equipped with the metric

$$d_{\Sigma_{\mathcal{J}}}(u, v) = \sum_{i=0}^{\infty} \frac{1}{2^i} \; d_{\mathcal{J}}(u(i), v(i)) ,$$

$\Sigma_{\mathcal{J}}$ is a compact metric space. The shift operator σ, acting according to the rule

$$\sigma(w(0)w(1)w(2)w(3)\dots) = w(1)w(2)w(3)\dots ,$$

is continuous on $\Sigma_{\mathcal{J}}$. We use the notation $w[n] = w(0)w(1)w(2)\dots w(n-1)$.

Similarly, we define $\Sigma_{\mathcal{I}}$ with the metric $d_{\Sigma_{\mathcal{I}}}$. This space is a metric compact as well.

Along with infinite strings we need finite strings, or words. For a finite word, u, its length is the number of letters (symbols) in it; the length is denoted by $|u|$. Sometimes we put an asterisk after a word to indicate that the word is finite, e.g., $u*$. The space of all finite words over $\mathcal{J}$ will be denoted $\Sigma^*_{\mathcal{J}}$. If w is an infinite word, the notation $w[n]$ will be used for the finite word formed by the first n symbols of w, i.e., $w[n] = w(0) \dots w(n-1)$. Clearly, $|w[n]| = n$. For infinite words, $|w| = \infty$. The concatenation of u, a finite word, and v, a finite or infinite word, is denoted $u.v$.

For the words h in $\Sigma^*_{\mathcal{I}}$ and $\Sigma_{\mathcal{I}}$ we need to measure the total time span. The notation is this:

$$\|h\| = \sum_{n=0}^{|h|} h(n)\,.$$

Evolution maps. For every choice of $\tau \in \mathcal{I}$ and $j \in \mathcal{J}$, a continuous map $S^\tau_j : X \to X$ is given. Thanks to our assumptions on the state space X, each S^τ_j is a bounded map. The maps S^τ_j may not satisfy the semi-group property, i.e., $S^{\tau_1}_j \circ S^{\tau_2}_j$ is not assumed to be equal to $S^{\tau_1+\tau_2}_j$, but we assume that S^0_j is the identity map for any j. Starting with a point x and acting successively with different or the same maps S^τ_j, one obtains a switched trajectory.

Definition 1. *The switched trajectory corresponding to the regime switching $(h, w) \in \Sigma_{\mathcal{I}} \times \Sigma_{\mathcal{J}}$ and starting at $x \in X$ is the sequence of points defined recursively as*

$$y_0 = x, \quad y_{n+1} = S^{h(n)}_{w(n)}(y_n),\ n \geq 0\,.$$

As explained in our paper [7], a convenient way to study the switched trajectories is to use the scew-product construction and work in the product space $X \times \Sigma_{\mathcal{J}} \times \Sigma_{\mathcal{I}}$ with the metric of the product space:

$$d_{\mathfrak{X}}((x, w, h), (x', w', h')) \;=\; d_X(x, x') + d_{\Sigma_{\mathcal{J}}}(w, w') + d_{\Sigma_{\mathcal{I}}}(h, h')\,.$$

Here is the relevant definition.

Definition 2. *The variable time step dynamics with choice on X associated with the maps S^τ_j, $j \in \mathcal{J}$, $\tau \in \mathcal{I}$, is the discrete time dynamics generated on the product space $\mathfrak{X} = X \times \Sigma_{\mathcal{J}} \times \Sigma_{\mathcal{I}}$ by the iterations of the map*

$$\mathfrak{F} : (x, w, h) \mapsto (S^{h(0)}_{w(0)}(x), \sigma(w), \sigma(h))\,. \tag{7}$$

Remark 3. *This definition makes sense without any assumptions of metricity and compactness and such. But we are going to use this definition only when those assumptions are satisfied.*

Note that after n steps the point (x, w, h) is moved to the point

$$\mathfrak{F}^n(x, w, h) = (S_{w[n]}^{h[n]}(x), \sigma^n(w), \sigma^n(h)),$$

where we use a convenient shorthand

$$S_{w[n]}^{h[n]} = S_{w(n-1)}^{h(n-1)} \circ S_{w(n-2)}^{h(n-2)} \circ \cdots \circ S_{w(0)}^{h(0)}.$$

3 Continuous Limit Dynamics

In this section we define the continuous limit of a variable time step dynamics with choice. The resulting *continuous limit dynamics* (CLD for short) will be multivalued. We start by listing all the assumptions on the maps S_j^τ.

Assumptions on S_j^τ:

A0 Each $S_j^\tau : X \to X$ is a continuous map; S_j^0 is the identity map.

A1 For any closed, bounded set $B \subset X$, the maps S_j^τ, restricted to B, depend uniformly continuously on $j \in \mathcal{J}$ and $\tau \in \mathcal{I}$. More precisely, given a closed, bounded B, for every $\epsilon > 0$ there exist $\delta_{\mathcal{J}} > 0$ and $\delta_{\mathcal{I}} > 0$ such that

$$\sup_{x \in B} d_X(S_{j_1}^{\tau_1}(x), S_{j_2}^{\tau_2}(x)) \leq \epsilon$$

provided $d_{\mathcal{J}}(j_1, j_2) \leq \delta_{\mathcal{J}}$ and $|\tau_1 - \tau_2| = d_{\mathcal{I}}(\tau_1, \tau_2) \leq \delta_{\mathcal{I}}$.

A2 For every $x \in X$ and for every $T > 0$, there is a constant $C = C(x, T)$ such that, for any finite word $s* \in \Sigma_{\mathcal{I}}^*$ and for any finite word $w* \in \Sigma_{\mathcal{J}}^*$ of the same length as $s*$, if $\|s*\| \leq T$ then

$$d(x, S_{w*}^{s*}(x)) \leq C.$$

A3 For every $x \in X$ and for every $\epsilon > 0$ there exists $\delta = \delta(x, \epsilon) > 0$ such that, for any finite word $s* \in \Sigma_{\mathcal{I}}^*$ and for any finite word $w* \in \Sigma_{\mathcal{J}}^*$ of the same length as $s*$,

$$d(S_{w*}^{s*}(x), x) < \epsilon,$$

provided $\|s*\| \leq \delta$.

A4 Suppose $x_n \to x$. Suppose s_n is a sequence of finite words of symbols in $\mathcal{I}$ and u_n is a sequence of finite words of symbols in $\mathcal{J}$ of the same length as s_n. Suppose $\sup_n \|s_n\| \leq C$. Then for every $\epsilon > 0$ there exists $N = N(C, \epsilon)$ such that for all $n \geq N$,

$$d(S_{u_n}^{s_n}(x_n),\ S_{u_n}^{s_n}(x)) \leq \epsilon.$$

The first two assumptions refer to the properties of individual maps, while the last three assumptions refer to switched trajectories. In applications with differential or difference equations, verification of these assumptions relies on a priori estimates. Assumption **A2** says that every switched trajectory starting from (any point) x stays bounded over any finite time interval. Assumptions **A3** and **A4** say that over short time intervals the switched trajectories remain close to the initial position.

Lemma 4. *a) If $\tau_n \to \tau$ in $\mathcal{I}$, $j_n \to j$ in $\mathcal{J}$ and $x_n \to x$ in X, then $S_{j_n}^{\tau_n}(x_n) \to S_j^{\tau}(x)$.*

b) For every bounded set A, $S_{j_n}^{\tau_n}(A)$ converges to $S_j^{\tau}(A)$ in the Hausdorff metric when $\tau_n \to \tau$ and $j_n \to j$.

c) For every bounded A, for every integer $n > 0$, and for every $\epsilon > 0$, there exists $\delta = \delta(A, n, \epsilon) > 0$ such that

$$\sup_{h:\, \|h[n]\| \le \delta} \sup_{w} d_H(S_{w[n]}^{h[n]}(A), A) \le \epsilon\,.$$

Proof. To prove a) consider

$$d(S_{j_n}^{\tau_n}(x_n), S_j^{\tau}(x)) \le d(S_{j_n}^{\tau_n}(x_n), S_j^{\tau}(x_n)) + d(S_j^{\tau}(x_n), S_j^{\tau}(x))\,.$$

The first term on the right will become small due to **A1**, and the second term will become small due to continuity of S_j^{τ}.

For b), observe that

$$\sup_{x\in A} \inf_{y\in A} d(S_{j_n}^{\tau_n}(x), S_j^{\tau}(y)) \le \sup_{x\in A} d(S_{j_n}^{\tau_n}(x), S_j^{\tau}(x))$$

and the right side goes to zero due to **A1**. Swapping x and y completes the proof of b).

To prove c) use induction on n. For $n = 1$ we have $d_H(S_j^{\tau}(A), S_j^{0}(A)) \to 0$ as $\tau \to 0$ due to **A1**. Indeed,

$$\sup_{x\in A} \inf_{y\in A} d(S_j^{\tau}(x), y) \le \sup_{x\in A} d(S_j^{\tau}(x), x) \le \epsilon$$

and

$$\sup_{x\in A} \inf_{y\in A} d(S_j^{\tau}(y), x) \le \sup_{x\in A} d(S_j^{\tau}(x), x) \le \epsilon$$

by **A1**.

Suppose the statement is true for $n - 1$ and suppose that it is not true for n. Then there exist a bounded A, an $\epsilon > 0$, and the sequences h_k and w_k such that $\|h_k[n]\| < 1/k$, but

$$d_H(S_{w_k[n]}^{h_k[n]}(A), A) > \epsilon\,.$$

Each of $h_k(r)$ converges to zero, and we may assume that each of $w_k(r)$, $0 \le r \le n-1$, converges. We have

$$d_H(S_{w_k[n]}^{h_k[n]}(A), A) \le d_H(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(A), S_{w_k[n-1]}^{h_k[n-1]}(A)) + d_H(S_{w_k[n-1]}^{h_k[n-1]}(A), A)\,.$$

The second term on the right goes to 0 by the induction hypothesis, so we may assume that

$$d_H(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(A), S_{w_k[n-1]}^{h_k[n-1]}(A)) \ge \frac{\epsilon}{2}\,.$$

The Hausdorff distance can be estimated as follows.

$$\sup_{x\in A} \inf_{y\in A} d(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(x), S_{w_k[n-1]}^{h_k[n-1]}(y)) \le \sup_{x\in A} d(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(x), S_{w_k[n-1]}^{h_k[n-1]}(x))$$

and

$$\sup_{y\in A} \inf_{x\in A} d(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(x), S_{w_k[n-1]}^{h_k[n-1]}(y)) \le \sup_{y\in A} d(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(y), S_{w_k[n-1]}^{h_k[n-1]}(y))$$

Thus, for every sufficiently large k there is an $x_k \in A$ such that

$$d(S_{j_k}^{\tau_k} \circ S_{w_k[n-1]}^{h_k[n-1]}(x_k), S_{w_k[n-1]}^{h_k[n-1]}(x_k)) \ge \frac{\epsilon}{4}\,.$$

Because bounded sequences are relatively compact in X we may assume $x_k \to x_*$, and at the same time, because $\mathcal{J}$ and $\mathcal{I}$ are compact, we may assume that each of the sequences τ_k, j_k, $h_k[n-1]$ and $w_k[n-1]$ converges. We must have $\tau_k \to 0$ and $h_k[n-1] \to 0\ldots0$. By part a), the left side must go to 0, a contradiction. □

The following observation is a simple exercise.

Lemma 5. *The map $\mathfrak{F}$ is continuous and compact.*

3.1 Switched trajectories

Starting the switched trajectories from the initial point $x \in X$, we would like to know where we can end up at time $T > 0$ if we use all possible regime switching. This is the set reachable from x at time T,

$$R_T(x) = \{y \in X : y = S_{w[n]}^{h[n]}(x) \text{ for some } n \ge 0,\ (h, w) \in \Sigma_{\mathcal{I}} \times \Sigma_{\mathcal{J}} \text{ with } \|h[n]\| = T\} \quad (8)$$

Due to assumption **A2**, the set $R_T(x)$ is bounded. In fact, the set $\bigcup_{t\in[0,T]} R_T(x)$ is bounded. It makes sense to consider the closure $\overline{R_T(x)}$. This is the set of points that can be if not reached then approximated by switched trajectories. However, we propose to consider the following set instead.

The sets $\mathcal{F}_T(x)$. For every $x \in X$ and every $T > 0$ define $\mathcal{F}_T(x)$ as the set of all possible limits of sequences of the form $S^{h_k[n_k]}_{w_k[n_k]}(x)$, where $n_k \nearrow \infty$, $w_k \in \Sigma_{\mathcal{J}}$, and $h_k \in \Sigma_{[0,\epsilon_k]}$ with $\|h_k[n_k]\| = T$ and $\epsilon_k \searrow 0$. For $T = 0$ define $\mathcal{F}_T(x) = \{x\}$.

The difference between $\overline{R_T(x)}$ and $\mathcal{F}_T(x)$ is that $\mathcal{F}_T(x)$ may not contain some of the limits corresponding to the partitions where some intervals are not getting smaller in size.

Partitions. Consider the finite partitions of the interval $[0, T]$ into sufficiently small subintervals. If p is a partition, $|p|$ denotes the number of elements of p. Each partition into r subintervals is encoded by a finite word, which we can write as $h[r]$. Define the max-mesh size of the word $h[r]$ as $\mathit{max\text{-}mesh}\,(h[r]) = \max_{0\le k\le r-1} h(k)$, and define the min-mesh as $\mathit{min\text{-}mesh}\,(h[r]) = \min_{0\le k\le r-1} h(k)$ The strings $h_k[n_k]$ in the definition of $\mathcal{F}_T(x)$ have max-mesh going to 0.

Denote by $P(T; n, m)$ the set of all partitions p of $[0, T]$ with $\mathit{max\text{-}mesh}\,(p) \le 2^{-n}$ and $\mathit{min\text{-}mesh}\,(p) \ge 2^{-m}$. We identify $p \in P(T; n, m)$ with the corresponding word in $\mathcal{I}^{|p|}$.

Define the sets

$$A_N = \bigcup_{n\ge N} \bigcup_{m\ge n} \bigcup_{p\in P(T;n,m)} \bigcup_{w*\in \mathcal{J}^{\|p\|}} S^p_{w*}(x)\,.$$

Notice that $A_N \supset A_{N+1} \supset \dots$. Moreover,

$$\mathcal{F}_T(x) = \bigcap_{N\ge 0} \overline{A_N}\,. \tag{9}$$

Thus, $\mathcal{F}_T(x)$ is closed in addition to being bounded, and hence it is compact. This and other useful properties of $\mathcal{F}_T(x)$ are collected in Proposition 7 below. Here is a useful auxiliary result that exploits assumption **A4**.

Lemma 6. *Suppose s_n is a sequence of finite words of symbols in $\mathcal{I}$ and u_n is a sequence of finite words of symbols in $\mathcal{J}$ of the same length as s_n. Suppose $\|s_n\| \to 0$. Let x_n be a bounded sequence. Then*

$$\lim_{n\to\infty} d\left(S^{s_n}_{u_n}(x_n), x_n\right) \;=\; 0\,.$$

Proof. Suppose the assertion is wrong and there exist s_n, u_n, and x_n with the properties described in the statement but

$$d\left(S^{s_n}_{u_n}(x_n), x_n\right) > \epsilon$$

for some $\epsilon > 0$. In our settings, the sequence x_n has a convergent subsequence, and we call this subsequence x_n and work with it. Let x be the limit of x_n. Consider the inequality

$$d\left(S^{s_n}_{u_n}(x_n), x_n\right) \le d\left(S^{s_n}_{u_n}(x_n), S^{s_n}_{u_n}(x)\right) + d\left(S^{s_n}_{u_n}(x), x\right) + d(x, x_n)\,.$$

Each term on the right goes to 0 due to **A4** (the first term), to **A3** (the second term), and because $x_n \to x$. A contradiction. This proves the lemma. □

Proposition 7. *a) The set* $\mathcal{F}_T(x)$ *is compact.*

b) If $y_k = S^{h_k[n_k]}_{w_k[n_k]}(x) \to y$ *as* $k \to +\infty$ *with* $\|h_k[n_k]\| = T_k \to T$*,* max-mesh$\,(h_k[n_k]) \searrow 0$*, and* $n_k \nearrow +\infty$*, then* $y \in \mathcal{F}_T(x)$*.*

c) $\mathcal{F}_T(x)$ *is continuous with respect to* T*, i.e., if* $T_k \to T$*, then* $d_H(\mathcal{F}_{T_k}(x), \mathcal{F}_T(x)) \to 0$*.*

d) If $T_k \to T$ *and* $x_k \to x$*, then* $d_H(\mathcal{F}_{T_k}(x_k), \mathcal{F}_T(x)) \to 0$*.*

e) The set $\bigcup_{0 \le t \le T} \mathcal{F}_t(x)$ *is compact.*

Proof. We have already established compactness of $\mathcal{F}_T(x)$. To prove part b), first assume that there are infinitely many k for which $T_k \le T$ and work with those k only. Pick your favorite $j \in \mathcal{J}$ and let $z_k = S_j^{T-T_k}(y_k) = S_j^{T-T_k} \circ S^{h_k[n_k]}_{w_k[n_k]}(x)$. By Lemma 4.a), $z_k \to y$. On the other hand, the limit of z_k belongs to $\mathcal{F}_T(x)$. Hence, $y \in \mathcal{F}_T(x)$.

Next, assume that there are infinitely many k for which $T_k > T$ and work with those k. Denote by m_k the integer such that $\|h_k[m_k]\| \le T$, but $\|h_k[m_k+1]\| > T$. The sequence $z_k = S^{h_k[m_k]}_{w_k[m_k]}(x)$ is bounded, hence it has a convergent subsequence. Let z_* be its limit. Define $s_k = T - \|h_k[m_k]\|$ and consider the sequence $S^{s_k.h_k[m_k]}_{j.w_k[m_k]}(x)$. This sequence is bounded by Lemma 4.c), hence it has a convergent subsequence. The limit is in $\mathcal{F}_T(x)$ by construction, but this limit is z_* due to Lemma 6. Now we apply Lemma 6 to the sequence $S^{h_k[n_k]}_{w_k[n_k]}(x)$ to conclude that its limit is z_*, which then is equal to y and, hence, $y \in \mathcal{F}_T(x)$.

Now turn to the proof of continuity of $\mathcal{F}_T(x)$ with respect to T. We argue by contradiction. Suppose there is no continuity. Then there is a sequence $T_\ell \to T$ and there exists an $\epsilon > 0$ such that either

$$e(\mathcal{F}_{T_\ell}(x), \mathcal{F}_T(x)) > \epsilon \tag{10}$$

for all ℓ, or

$$e(\mathcal{F}_T(x), \mathcal{F}_{T_\ell}(x)) > \epsilon \tag{11}$$

for all ℓ, or both. First assume (10). Because each $\mathcal{F}_{T_\ell}(x)$ is compact, there exist $y^\ell \in \mathcal{F}_{T_\ell}(x)$ such that

$$d(y^\ell, \mathcal{F}_T(x)) = \sup_{y \in \mathcal{F}_{T_\ell}(x)} d(y, \mathcal{F}_T(x))$$

for every ℓ. Each y^ℓ is a limit of the form

$$y^\ell = \lim_k S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x)\,,$$

where $\|h^\ell_k[n^\ell_k]\| = T_\ell$ and max-mesh$\,(h^\ell_k[n^\ell_k]) \to 0$ as $k \to \infty$.

If there are infinitely many ℓ for which $T_\ell > T$, then use only this subsequence and do the following. There are integers $m_k^\ell < n_k^\ell$ such that $\|h_k^\ell[m_k^\ell]\| \le T$, but $\|h_k^\ell[m_k^\ell+1]\| > T$. The sequence $z_k^\ell = S_{w_k^\ell[m_k^\ell]}^{h_k^\ell[m_k^\ell]}(x)$ has a convergent subsequence as $k \to \infty$, and its limit, z_*^ℓ, belongs to $\mathcal{F}_T(x)$ by part b). Again, let s_k^ℓ and u_k^ℓ be such that $s_k^\ell.h_k^\ell[m_k^\ell] = h_k^\ell[n_k^\ell]$ and $u_k^\ell.w_k^\ell[m_k^\ell] = w_k^\ell[n_k^\ell]$. Consider the right hand side of the inequality

$$d(y^\ell, z_*^\ell) \le d(y^\ell, S_{u_k^\ell}^{s_k^\ell}(z_k^\ell)) + d(S_{u_k^\ell}^{s_k^\ell}(z_k^\ell), S_{u_k^\ell}^{s_k^\ell}(z_*^\ell)) + d(S_{u_k^\ell}^{s_k^\ell}(z_*^\ell), z_*^\ell)$$

written for $k \ge k_\ell$ large enough so that *max-mesh* $(h_k^\ell[n_k^\ell]) \le 2^{-\ell}$. Note that $\|s_k^\ell\| \le |T_\ell - T| + 2^{-\ell}$. By Lemma 6, $d(S_{u_k^\ell}^{s_k^\ell}(z_*^\ell), z_*^\ell)$ can be made arbitrarily small by choosing ℓ large enough. The remaining terms can be made small by choosing k large enough. Thus, $e(\mathcal{F}_{T_\ell}(x), \mathcal{F}_T(x)) \to 0$ as $\ell \to \infty$.

If there are at most finitely many ℓ for which $T_\ell > T$ but infinitely many ℓ for which $T_\ell < T$, then use only this subsequence and do the following. Fix your favorite $j \in \mathcal{J}$ and define $s_\ell = T - T_\ell$. The sequence y^ℓ is bounded as a subset of the bounded set

$$\bigcup_{0\le t\le 2T} \mathcal{F}_t(x) \subset \overline{\bigcup_{0\le t\le 2T} R_t(x)}\,. \tag{12}$$

By Lemma 6, we can find an ℓ such that

$$d(y^\ell, S_j^{s_\ell}(y^\ell)) \le \frac{\epsilon}{4}\,.$$

For this ℓ there exists a $k = k_\ell$ such that

$$d\left(S_j^{s_\ell} \circ S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x),\ S_j^{s_\ell}(y^\ell)\right) \le \frac{\epsilon}{4}\,.$$

The sequence $z^\ell = S_j^{s_\ell} \circ S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)$ (where $k = k_\ell$) has a subsequence converging to some point $z \in \mathcal{F}_T(x)$. Hence, $d(y^\ell, \mathcal{F}_T(x)) < \epsilon$ for large ℓ, a contradiction.

Now, assume (11). Again, for every ℓ there exist an $x^\ell \in \mathcal{F}_T(x)$ and a $y^\ell \in \mathcal{F}_{T_\ell}(x)$ such that

$$e(\mathcal{F}_T(x), \mathcal{F}_{T_\ell}(x)) = d(x^\ell, \mathcal{F}_{T_\ell}(x)) = d(x^\ell, y^\ell)\,.$$

Again, consider first the case $T_\ell \searrow T$. Every x^ℓ is a limit, $x^\ell = \lim_k x_k^\ell$,

$$x^\ell = \lim_k x_k^\ell\,, \quad x_k^\ell = S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)\,, \tag{13}$$

where $\|h_k^\ell[n_k^\ell]\| = T$. Denote by p_k^ℓ any partition of the interval $[T, T_\ell]$ into pieces of size not greater than *max-mesh* $(h_k^\ell[n_k^\ell])$. Pick your favorite j and denote by u_k^ℓ the word $j \dots j$ with as many j's as the length of p_k^ℓ. For every fixed ℓ, the sequence

$$\{S_{u_k^\ell}^{p_k^\ell}(x_k^\ell)\}_{k=1}^\infty$$

has a convergent subsequence, and its limit, call it z^ℓ, belongs to $\mathcal{F}_{T_\ell}(x)$. We have

$$d(z^\ell, x^\ell) \leq d\left(z^\ell, S_{u_k^\ell}^{p_k^\ell}(x_k^\ell)\right) + d\left(S_{u_k^\ell}^{p_k^\ell}(x_k^\ell), x_k^\ell\right) + d(x_k^\ell, x^\ell)\,. \tag{14}$$

Choose $k = k_\ell$ to make the first and the third terms on the right less than $\epsilon/4$. The second term then is small for large ℓ thanks to Lemma 6.

If $T_\ell \nearrow T$, then use the same approximation (13) to x^ℓ and split $h_k^\ell[n_k^\ell]$ into $h_k^\ell[n_k^\ell] = p_k^\ell.h_k^\ell[m_k^\ell]$, where $\|h_k^\ell[m_k^\ell]\| \leq T_\ell$ and $\|h_k^\ell[m_k^\ell + 1]\| > T_\ell$. Thus,

$$x_k^\ell = S_{u_k^\ell}^{p_k^\ell} \circ S_{w_k^\ell[m_k^\ell]}^{h_k^\ell[m_k^\ell]}(x)\,.$$

The sequence $z_k^\ell = S_{w_k^\ell[m_k^\ell]}^{h_k^\ell[m_k^\ell]}(x)$ has a convergent subsequence, and we work with this subsequence and let $z^\ell = \lim_k S_{w_k^\ell[m_k^\ell]}^{h_k^\ell[m_k^\ell]}(x)$. By part b), $z^\ell \in \mathcal{F}_{T_\ell}(x)$. Again, similar to (14), we have inequality

$$d(z^\ell, x^\ell) \leq d(z^\ell, z_k^\ell) + d\left(z_k^\ell, S_{u_k^\ell}^{p_k^\ell}(z_k^\ell)\right) + d(x_k^\ell, x^\ell)\,.$$

There exists k_ℓ such that $d(z^\ell, z_k^\ell) + d(x_k^\ell, x^\ell) \leq \epsilon/2$ for all $k \geq k_\ell$. With $k = k_\ell$, the middle term is small for large ℓ by Lemma 6. This contradicts (11) and completes the proof of part c).

Now assume $T_\ell \to T$ and $x^\ell \to x$. In order to prove part d) it suffices to prove that

$$\lim_\ell e(\mathcal{F}_{T_\ell}(x^\ell), \mathcal{F}_{T_\ell}(x)) = 0 \;\text{ and }\; \lim_\ell e(\mathcal{F}_{T_\ell}(x), \mathcal{F}_{T_\ell}(x^\ell)) = 0.$$

For the first limit, notice that every point $y^\ell \in \mathcal{F}_{T_\ell}(x^\ell)$ can be approximated by a sequence of the form $S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell)$ and that

$$d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell), \mathcal{F}_{T_\ell}(x)) \leq d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell), S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)) + d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x), \mathcal{F}_{T_\ell}(x))$$

Choose $k = k_\ell$ so that the distance between $S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell)$ and y^ℓ be less than $2^{-\ell}$. Then, for large ℓ, the first term on the right becomes small due to assumption **A4**, while the second becomes small because $S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)$ converges to a point in $\mathcal{F}_T(x)$ and $\mathcal{F}_{T_\ell}(x)$ converges to $\mathcal{F}_T(x)$ by part c). In a similar fashion, for the second limit, we represent any $y^\ell \in \mathcal{F}_{T_\ell}(x)$ as $y^\ell = \lim_k S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)$ and estimate

$$d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x), \mathcal{F}_{T_\ell}(x^\ell)) \leq d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x), S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell)) + d(S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x^\ell), \mathcal{F}_{T_\ell}(x^\ell))$$

Now, choose $k = k_\ell$ so that

$$d(y^\ell, S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell)) < \epsilon/3$$

and

$$d(S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell), \mathcal{F}_{T_\ell}(x^\ell)) < \epsilon/3$$

and then use assumption **A4** to choose ℓ large enough to have

$$d(S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x), S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell)) < \epsilon/3\,.$$

This show that for every $\epsilon > 0$, $e(\mathcal{F}_{T_\ell}(x), \mathcal{F}_{T_\ell}(x^\ell)) < \epsilon$ for all large enough ℓ. This proves d).

Next we show that the set $\bigcup_{0\le t\le T} \mathcal{F}_t(x)$ is closed. Suppose $x^\ell \in \mathcal{F}_{t_\ell}(x)$ and $x^\ell \to x_*$. We may assume that $t_\ell \to t_*$. Since $\mathcal{F}_{t_\ell}(x) \to \mathcal{F}_{t_*}(x)$ in the Hausdorff metric, $d(x^\ell, \mathcal{F}_{t_*}(x)) \to 0$. There exist $z^\ell \in \mathcal{F}_{t_*}(x)$ such that $d(x^\ell, \mathcal{F}_{t_*}(x)) = d(x^\ell, z^\ell)$. We see that z^ℓ must converge to x_*, and hence $x_* \in \mathcal{F}_{t_*}(x)$. This proves part e) and the proposition. □

The multi-valued map $\mathcal{F}_T : X \to 2^X$ can be extended to a multi-valued map on bounded sets as follows.

The sets $\mathcal{F}_T(B)$. For a bounded set $B \subset X$ define $\mathcal{F}_T(B)$ as the set of all possible limits of sequences of the form $S^{h_k[n_k]}_{w_k[n_k]}(x_k)$, where $x_k \in B$, $n_k \nearrow \infty$, $w_k \in \Sigma_{\mathcal{J}}$, and $h_k \in \Sigma_{\mathcal{I}}$ such that $\lim_k \textit{max-mesh}\,(h_k[n_k]) = 0$. For $T = 0$ define $\mathcal{F}_0(B) = \overline{B}$.

Here are a few basic properties of the map $B \mapsto \mathcal{F}_T(B)$.

Proposition 8. *a) If $A \subset B$, then $\mathcal{F}_T(A) \subset \mathcal{F}_T(B)$.*

b) $\mathcal{F}_T(\overline{B}) = \mathcal{F}_T(B)$.

c) If A is closed and bounded and $y \in \mathcal{F}_T(A)$, then there exists an $x \in A$ such that $y \in \mathcal{F}_T(x)$. As a corollary, $\mathcal{F}_T(A) = \cup_{x\in A}\mathcal{F}_T(x)$.

d) For every bounded B the set $\mathcal{F}_T(B)$ is compact.

e) For any T_1, $T_2 \ge 0$ we have $\mathcal{F}_{T_1}(\mathcal{F}_{T_2}(B)) = \mathcal{F}_{T_1+T_2}(B)$.

Proof. The monotonicity is obvious and so is the inclusion $\mathcal{F}_T(B) \subset \mathcal{F}_T(\overline{B})$. Suppose $x \in \mathcal{F}_T(\overline{B})$. This means

$$x = \lim_k S^{h_k[n_k]}_{w_k[n_k]}(x_k)$$

where $x_k \in \overline{B}$. If $x_k = \lim_\ell x_k^\ell$, where $x_k^\ell \in B$, then for every k there is an $\ell = \ell_k$ such that

$$d\left(S^{h_k[n_k]}_{w_k[n_k]}(x_k), S^{h_k[n_k]}_{w_k[n_k]}(x_k^\ell)\right) \le 2^{-k}\,.$$

This implies $x \in \mathcal{F}_T(B)$.

In the setting of part c), write y as the limit $y = \lim_k S^{h_k[n_k]}_{w_k[n_k]}(x_k)$, where $x_k \in A$. Sparse the sequence x_k to make it convergent, and let $x = \lim_k x_k$. Invoking assumption **A4** we see that $y = \lim_k S^{h_k[n_k]}_{w_k[n_k]}(x)$, and hence $y \in \mathcal{F}_T(x)$.

The set $\mathcal{F}_T(B)$ is bounded. To show this it suffices to to show that for every point $x \in B$ there is a ball $\mathcal{O}_\delta(x)$ around it such that its trajectory is bounded (then we assume that B is closed and cover B by a finite number of such balls). We claim that for every $x \in X$ and for every $T > 0$, there exists a $\delta > 0$ and there exists a constant $\tilde{C} = \tilde{C}(x, \delta, T)$ such that for any integer $n > 0$, any $h \in \Sigma_{\mathcal{I}}$ such that $\|h[n]\| \le T$, and any $w \in \Sigma_{\mathcal{J}}$,

$$d(y, S^{h[k]}_{w[k]}(y)) \le \tilde{C}\,, \quad \text{for} \quad 0 \le k \le n\,,$$

for any y in the δ-neighborhood of x. If the claim is wrong, then there exist x and T such that for every integer $m > 0$ one can find n_m, h_m with $\|h_m[n_m]\| \le T$, w_m, and $y_m \in \mathcal{O}_{1/m}(x)$ such that

$$d(y_m, S^{h_m[k_m]}_{w_m[k_m]}(y_m)) > m$$

for some $k_m \le n_m$. However, this is impossible because one has

$$d(y_m, S^{h_m[k_m]}_{w_m[k_m]}(y_m)) \le d(y_m, x) + d(x, S^{h_m[k_m]}_{w_m[k_m]}(x)) + d(S^{h_m[k_m]}_{w_m[k_m]}(x), S^{h_m[k_m]}_{w_m[k_m]}(y_m))\,,$$

where the first term goes to 0 as $m \to \infty$ because $y_m \to x$, the last term is small for large m due to **A4**, and the middle term is bounded due to **A2**.

The set $\mathcal{F}_T(B)$ is closed. This can be checked by constructing a representation similar to (9). This proves our assertion d).

Next we show that $\mathcal{F}_{T_2}(\mathcal{F}_{T_1}(B)) \subset \mathcal{F}_{T_1+T_2}(B)$. Let $y \in \mathcal{F}_{T_2}(\mathcal{F}_{T_1}(B))$. Then $y = \lim_\ell S^{s^\ell[m^\ell]}_{u^\ell[m^\ell]}(y^\ell)$ with $y^\ell \in \mathcal{F}_{T_1}(B)$, $\|s^\ell[m^\ell]\| = T_2$, and *max-mesh* $(s^\ell[m^\ell]) \to 0$. In turn, every y^ℓ is a limit of the form $y^\ell = \lim_k z_k^\ell$ where $z_k^\ell = S^{h_k^\ell[n_k^\ell]}_{w_k^\ell[n_k^\ell]}(y_k^\ell)$, $y_k^\ell \in B$ and $\|h_k^\ell[n_k^\ell]\| = T_1$. Given an $\epsilon > 0$, use the triangle inequality

$$d(y, S^{s^\ell[m^\ell]}_{u^\ell[m^\ell]}(z_k^\ell)) \le d(y, S^{s^\ell[m^\ell]}_{u^\ell[m^\ell]}(y^\ell)) + d\left(S^{s^\ell[m^\ell]}_{u^\ell[m^\ell]}(y^\ell), S^{s^\ell[m^\ell]}_{u^\ell[m^\ell]}(z_k^\ell)\right)\,,$$

choose ℓ to make the first term less than $\epsilon/2$, and then choose k to make the second term also less than $\epsilon/2$. This proves that $y \in \mathcal{F}_{T_1+T_2}(B)$.

To finish the proof of part e), pick $y \in \mathcal{F}_{T_1+T_2}(B)$. Writing y as a limit of the end-points of switched trajectories can be always done in the form

$$y = \lim_k S^{s_k[m_k]}_{u_k[m_k]} \circ S^{h_k[n_k]}_{w_k[n_k]}(x_k)\,,$$

where $x_k \in B$, $\|h_k[n_k]\| = T_{1,k}$, $\|s_k[m_k]\| = T_{2,k}$, $T_{1,k} + T_{2,k} = T_1 + T_2$, and $\lim_k T_{1,k} = T_1$, $\lim_k T_{2,k} = T_2$. Also, *max-mesh* $(h_k[n_k]) \to 0$ and *max-mesh* $(s_k[m_k]) \to 0$. Denote $z_k = S^{h_k[n_k]}_{w_k[n_k]}(x_k)$. We may assume B is closed and the sequence x_k converges to some $x_* \in B$. By assumption **A4**, $\lim_k d(z_k, S^{h_k[n_k]}_{w_k[n_k]}(x_*)) = 0$. In turn, we may pass to a subsequence and use Proposition 7 part b) to conclude that the (sub)sequence $S^{h_k[n_k]}_{w_k[n_k]}(x_*)$ converges to some point $z_* \in \mathcal{F}_{T_1}(x_*) \subset \mathcal{F}_{T_1}(B)$. Thus, (on this subsequence) $z_k \to z_* \in \mathcal{F}_{T_1}(x_*)$. Now,

$$d(y, S^{s_k[m_k]}_{u_k[m_k]}(z_*)) \le d(y, S^{s_k[m_k]}_{u_k[m_k]}(z_k)) + d(S^{s_k[m_k]}_{u_k[m_k]}(z_k), S^{s_k[m_k]}_{u_k[m_k]}(z_*))\,.$$

The first term on the right goes to 0 by definition of z_k, while the second term can be made arbitrarily small by **A4**. This completes the proof. □

3.2 Continuous Limit Dynamics

There are many ways one can think of continuous limit dynamics. For us, the CLD corresponding to the maps S^τ_j is generated by the multi-valued maps $\mathcal{F}_t$ on the space of closed bounded subsets of X.

Denote by $X^\sharp$ the space of compact (= closed and bounded, in our situation) subsets of X equipped with the Hausdorff metric. It is well know that $X^\sharp$ is a complete metric space, [8]. The maps $\mathcal{F}_t$ are defined on $X^\sharp$.

Lemma 9. *Each map $\mathcal{F}_T$ is continuous as a map from $X^\sharp$ to $X^\sharp$.*

Proof. Let A_n be a convergent sequence in $X^\sharp$, and let $A = \lim_n A_n$. We will show that $\mathcal{F}_T(A_n) \to \mathcal{F}_T(A)$. First show $e\,(\mathcal{F}_T(A_n), \mathcal{F}_T(A)) \to 0$. For this we show that any sequence $y^\ell \in \mathcal{F}_T(A_\ell)$ has a subsequence convergent to a point in $\mathcal{F}_T(A)$. That y^ℓ has a convergent subsequence follows from the fact that the set $B_1 = \bigcup_\ell A_\ell$ is bounded, and hence, $\mathcal{F}_T(B_1)$ is compact by Proposition 8.d), and $\mathcal{F}_T(A_\ell) \subset \mathcal{F}_T(B_1)$. Let y be the limit of this (sub)sequence y^ℓ. As we know from Proposition 8.c), for each y^ℓ there exists an $x^\ell \in A_\ell$ such that $y^\ell \in \mathcal{F}_T(x^\ell)$. Because $A_\ell \to A$, there is a subsequence of x^ℓ converging to some $x \in A$. Each y^ℓ can be represented as a limit,

$$y^\ell = \lim_k S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell)\,.$$

Choose $k = k_\ell$ so that $d(y^\ell, S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell)) \le 2^{-\ell}$. Then

$$S^{h^\ell_k[n^\ell_k]}_{w^\ell_k[n^\ell_k]}(x^\ell)\big|_{k=k_\ell} \underset{\ell\to\infty}{\to} y\,.$$

Because $x^\ell \to x$, we have (see **A4**) $S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)\big|_{k=k_\ell}$ arbitrarily close to y as well. This proves that $y \in \mathcal{F}_T(x) \subset \mathcal{F}_T(A)$.

Next, we show $e\,(\mathcal{F}_T(A), \mathcal{F}_T(A_n)) \to 0$. For this it will suffice to show that $d(z, \mathcal{F}_T(A_n))$ goes to 0 for any $z \in \mathcal{F}_T(A)$. There exists an $x \in A$ such that $z \in \mathcal{F}_T(x)$. Choose a sequence $x^\ell \in A_\ell$ converging to x. It will suffice to show that $\lim_\ell d(z, \mathcal{F}_T(x^\ell)) = 0$. Write z as a limit,

$$z = \lim_k S_{w_k[n_k]}^{h_k[n_k]}(x)\,.$$

Successively for $\ell = 1, 2, \ldots$ sparse the sequence of k's to obtain convergent sequences $S_{w_k[n_k]}^{h_k[n_k]}(x^\ell) \to z^\ell$. The limits, z^ℓ, belong to the corresponding sets $\mathcal{F}_T(x^\ell)$. Choosing the diagonal (sub)sequence and using the assumption **A4**, we see that $z^\ell \to z$, and we are done. □

Definition 10. *The CLD corresponding to the maps S_j^τ is the continuous semi-dynamical system generated by the maps $\mathcal{F}_T$ on the hyperspace $X^\sharp$.*

From a different angle, we may want to think of transition from variable time-step dynamics with choice to continuous limit dynamics as transition from switched trajectories to continuous "integral curves" where the switching is no longer so pronounced. We phrase this in the form of a generalization of the Filippov-Ważewski relaxation theorem.

Definition 11. *For every $B \in X^\sharp$, define $\mathcal{CLD}_{[0,T]}(B)$ as the set of all continuous maps (integral curves) $\gamma : [0,T] \to X$ such that $\gamma(t) \in \mathcal{F}_t(B)$ for all $t \in [0,T]$ and*

$$\gamma(t_1 + t_2) \in \mathcal{F}_{t_2}(\gamma(t_1)) \tag{15}$$

for all $t_1, t_2 \in [0,T]$ with $t_1 + t_2 \in [0,T]$. The set $\mathcal{CLD}_{[0,T]}(B)$ is a subset of $C([0,T], X)$.

For integral curves to make sense, we need additional assumptions on the space X. We require that any two points in X that are sufficiently close to each other could be connected by the unique "shortest curve". We do not have space to bring into discussion the metric length spaces and related notions. Here the "shortest curves" are (understood to be) assigned to pairs of points. If X is a (Montel) vector space, then the shortest curve is the line segment connecting the points. If X is a finite-dimensional Riemannian manifold, then the shortest curve is the geodesic (unique, because the points are close to each other). In any case, the metric, d, should satisfy the following condition.

AX For every compact set A the following is true: For every $\epsilon > 0$ there exist $\delta > 0$ and $\rho > 0$ such that whenever x_0 and x_1 lie in A together with the shortest curve $\gamma : [0,1] \to X$ connecting them and $d(x_0, x_1) \le \rho$, $d(\gamma(t), \gamma(s)) \le \epsilon$ if $|t - s| \le \delta$.

In applications we have in mind, when X is a finite-dimensional Riemannian manifold or X is a Fréchet-Montel space, condition **AX** is satisfied.

Theorem 12. *Assume the condition* **AX** *is satisfied. Then the set* $\mathcal{CLD}_{[0,T]}(B)$ *is nonempty. It is a compact in* $C([0,T],X)$*. At the same time, for every* $t \in [0,T]$*,*

$$\{x \in X : \, x = \gamma(t) \text{ for some } \gamma \in \mathcal{CLD}_{[0,T]}(B)\} = \mathcal{F}_t(B)\,. \tag{16}$$

Proof. The integral curves in $\mathcal{CLD}_{[0,T]}(B)$ will be constructed as limits of "broken trajectories" or, as we call them, approximation curves. Consider a sequence p^N of partitions of the interval $[0,T]$ with the property that $\mathit{max\text{-}mesh}(p^N) \searrow 0$. Pick an $x_0 \in B$, define $x_0^N = x_0$ and then successively pick the points $x_{n+1}^N \in \mathcal{F}_{t_{n+1}^N - t_n^N}(x_n^N)$. Connect x_0^N with x_1^N, x_1^N with x_2^N, etc., by the shortest curves. This yields a (continuous) curve $\gamma^N(t)$. We call the points x_n^N the nodes and the corresponding γ^N the approximation curve. We claim that the sequence of the approximation curves γ^N has a convergent subsequence in the space of X-valued continuous functions on $[0,T]$ and that its limit belongs to $\mathcal{CLD}_{[0,T]}(B)$.

The sequence $\{\gamma^N\}_N$ is uniformly bounded in the sense that there is a constant $C > 0$ such that

$$\sup_{0 \le t \le T} d(x_0, \gamma^N(t)) \le C \tag{17}$$

for all N. This is because the set $\bigcup_t \mathcal{F}_t(x)$ is compact by Proposition 7.d), and all points $\gamma^N(t)$ lie in the convex hull of this set, which is bounded.

Next, we show that the family $\{\gamma^N\}_N$ is equicontinuous. Denote temporarily by A the closure of the convex hull (in the case X is a vector space) of $\bigcup_t \mathcal{F}_t(x_0)$,

$$A = \overline{co}\left(\bigcup_t \mathcal{F}_t(x_0)\right)\,.$$

(If X is a manifold, and if r_0 is the number such that any two points of A_0 at distance $\le r_0$ are connected by a unique geodesic, then A denotes the closure of the union of the set $A_0 = \bigcup_t \mathcal{F}_t(x_0)$ with the the shortest geodesics connecting points of A_0 which are at most at distance r_0 from each other.) This is a compact set. Our assumptions **A3** and **A4** imply the following: for every $x \in X$ and for every $\epsilon > 0$ there exist $\delta > 0$ and $\rho > 0$ such that if $d(y,x) < \rho$ and $\|s*\| < \delta$, then $d(S_{w*}^{s*}(y), y) < \epsilon$. Fix an $\epsilon > 0$. Cover the set A by balls $\mathcal{O}_\rho(x)$ of radius $\rho = \rho(x,\epsilon)$. Because A is compact, there is a finite subcover, and we choose the smallest δ of the corresponding $\delta(x,\epsilon)$. For all large enough N, the consecutive nodes of the approximation curves γ^N will be at distance $\le \rho$ from each other, and we can use condition **AX** to assure that the curves γ^N are uniformly continuous. The set of values of the maps γ^N being a subset of A is totally bounded. Hence we can apply the Arzelà-Ascoli theorem, see, e.g., [13, §16.4] and obtain a subsequence γ^{N_k} converging in the space of continuous maps $C([0,T],X)$ to some γ. To show that $\gamma \in \mathcal{CLD}_{[0,T]}(B)$, we have to check (15).

For every t, $\gamma(t) \in \mathcal{F}_t(x_0)$. Indeed, consider the approximation curves γ^{N_k} (from the convergent subsequence) and let t_{j_k} be a point in the partition of $[0,T]$ such that $t_{j_k} \le t < t_{j_k+1}$. Both sequences, $\gamma^{N_k}(t_{j_k})$ and $\gamma^{N_k}(t_{j_k+1})$ converge to $\gamma(t)$. By the construction of approximation curves, and thanks to Proposition 8.e), $\gamma^{N_k}(t_{j_k}) \in \mathcal{F}_{t_{j_k}}(x)$. The points $\gamma^{N_k}(t_{j_k})$ can be approximated by sequences of the form $y_k^\ell = S_{w_k^\ell[n_k^\ell]}^{h_k^\ell[n_k^\ell]}(x)$ so that $\gamma^{N_k}(t_{j_k}) = \lim_\ell y_k^\ell$. Using diagonalization and Proposition 7.b), we see that $\gamma(t) \in \mathcal{F}_t(x_0)$. Now consider $\gamma(t+\tau)$. This point is a limit of the points of the form $\gamma^{N_k}(t_{j_k}+\tau_{j_k})$. By construction, $\gamma^{N_k}(t_{j_k}+\tau_{j_k}) \in \mathcal{F}_{\tau_{j_k}}(\gamma^{N_k}(t_{j_k}))$. We know that $\gamma^{N_k}(t_{j_k}+\tau_{j_k}) \to \gamma(t+\tau)$ and, by Proposition 7.d), $\mathcal{F}_{\tau_{j_k}}(\gamma^{N_k}(t_{j_k})) \to \mathcal{F}_{t+\tau}(x)$. This proves (15).

Next we show that every map γ in $\mathcal{CLD}_{[0,T]}(B)$ is a limit of approximation curves. If $\gamma \in \mathcal{CLD}_{[0,T]}(B)$ and p^N is a sequence of partitions of $[0,T]$ with decreasing max-mesh, we build the approximation curves γ^N by choosing the nodes $x_n^N = \gamma(t_n^N)$ and connecting them by the shortest curves. As before, $\gamma^N \to \gamma$ in $C([0,T],X)$.

The fact that $\mathcal{CLD}_{[0,T]}(B)$ is compact in $C([0,T],X)$ follows, because if γ_n is a sequence in $\mathcal{CLD}_{[0,T]}(B)$ we can build approximation curves γ_n^N so that $\sup_t d(\gamma_n^{N_n}(t), \gamma_n(t)) \to 0$, and then show as before that the Arzelà-Ascoli theorem is applicable to the family $\gamma_n^{N_n}$.

It remains to show that through every point $z \in \mathcal{F}_t(B)$ passes a whole curve $\gamma \in \mathcal{CLD}_{[0,T]}(B)$. Indeed, if $z \in \mathcal{F}_t(B)$, there exists an $x_0 \in B$ such that $z \in \mathcal{F}_t(x_0)$. As a result, z can be represented as a limit,

$$z = \lim_k S_{w_k[n_k]}^{h_k[n_k]}(x_0)\,,$$

where we can assume that $\|h_k[N_k]\| = T$ for some N_k, and, in addition, *max-mesh* $(h_k[N_k]) \to 0$ as $k \to \infty$. With k fixed, define the nodes $x_m^k = S_{w_k[m-1]}^{h_k[m-1]}(x_0)$ for $m = 1, \dots, N_k$. Connect the nodes with the shortest curves. This will give us the approximation curve $\gamma_k(\cdot)$. As before, we can show that the Arzelà-Ascoli theorem is applicable to this family, and obtain a limiting integral curve γ. Since $\gamma_k(t) = z_k \to z$ by construction, we have $\gamma(t) = z$. The theorem is proved. □

4 Differential Inclusions and CLD

In this section we present an example of a Continuous Limit Dynamics for a switched system and identify it with the solution set of the relaxed differential inclusion.

4.1 Switched solutions

Throughout this section $f_1, \dots, f_N : \mathbb{R}^d \to \mathbb{R}^d$ will be an arbitrary but fixed set of Lipschitz continuous vector-fields. Let L be the maximum of their Lipschitz constants, i.e., for all $x, y \in \mathbb{R}^d$,

$$|f_j(x) - f_j(y)| \le L\,|x-y|\,, \quad j = 1, \dots, N.$$

With each f_j we associate the system of ordinary differential equations

$$\dot{x}(t) = f_j(x(t)) \tag{18}$$

whose solutions define a flow on $\mathbb{R}^d$. We denote by S_j^τ the corresponding evolution map from $\mathbb{R}^d$ to $\mathbb{R}^d$, where $S_j^\tau(\xi) = x(\tau)$ is the value at time τ of the solution $x(t)$ of system (18) with the initial condition $x(0) = \xi$. The Lipschitz continuity of f_j guarantees that solutions of (18) are global in time and unique.

Now, let $\mathcal{J}$ be the finite set $\{1, 2, \dots, N\}$ with discrete metric. This is the set of possible regimes. The basic set of possible dwell-times will be $\mathcal{I} = [0, 1]$. As before, we denote by $\Sigma_{\mathcal{J}}$ the space of regime switching functions, and denote by $\Sigma_{\mathcal{I}}$ the space of dwell time switching functions. For every choice of $w \in \Sigma_{\mathcal{J}}$ and $h \in \Sigma_{\mathcal{I}}$, and for every choice of initial condition ξ, define $x_{(w,h)}(t, \xi)$ as follows. Set $t_0 = 0$, $t_1 = h(0)$, $t_2 = t_1 + h(1)$, $\dots$, $t_{n+1} = t_n + h(n)$, $\dots$. Then, starting with $x_{(w,h)}(0, \xi) = \xi$ define successively

$$x_{(w,h)}(t, \xi) = S_{w(n)}^{t-t_n}(x_{(w,h)}(t_n))\,, \quad \text{if} \quad t_n \le t \le t_{n+1}\,.$$

Definition 13. *A switched solution of a family of equations* (18), $j = 1, \dots, N$, *(or, a switched trajectory of the flows* $\{S_j\}$*) is any vector-function of the form* $x_{(w,h)}(t, \xi)$.

An equivalent definition: $x_{(w,h)}(t, \xi)$ is the unique continuous solution of the initial value problem

$$\dot{x}(t) = \sum_{j=1}^{N} \varphi_j(w, h; t)\, f_j(x(t))\,, \quad x(0) = \xi\,, \tag{19}$$

where

$$\varphi_j(w, h; t) = \begin{cases} 1\,, & \text{if } j = w(n) \text{ and } t_n \le t < t_{n+1} \\ 0\,, & \text{otherwise} \end{cases} \tag{20}$$

The solution of (19) is understood as a continuous function $x(t)$ satisfying

$$x(t) = \xi + \int_0^t \sum_{j=1}^{N} \varphi_j(w, h; s)\, f_j(x(s))\, ds\,. \tag{21}$$

To assure that the solution is defined for all t we need to assume that

$$\|h\| = \sum_{n=0}^{\infty} h(n) = +\infty\,. \tag{22}$$

This representation is convenient for obtaining estimates on the switched solutions. In particular, since

$$x(t) - \xi = \int_0^t \sum_{j=1}^N \varphi_j(w,h;s)\ (f_j(x(s)) - f_j(\xi))\ ds + \int_0^t \sum_{j=1}^N \varphi_j(w,h;s)\, f_j(\xi)\, ds\,,$$

we obtain

$$|x(t) - \xi| \le t\, f(\xi) + L \int_0^t |x(s) - \xi|\, ds\,,$$

where

$$|f(\xi)| = \max_j |f_j(\xi)|\,. \tag{23}$$

This leads to the estimate

$$\int_0^t |x(s) - \xi|\, ds \le |f(\xi)|\ \frac{e^{Lt} - 1 - Lt}{L^2}$$

which implies

$$|x_{(w,h)}(t,\xi) - \xi| \le |f(\xi)|\ \frac{e^{Lt} - 1}{L} \tag{24}$$

and, hence,

$$|x_{(w,h)}(t,\xi)| \le |\xi| + |f(\xi)|\,\frac{e^{Lt} - 1}{L}\,. \tag{25}$$

If $x_1(t)$ and $x_2(t)$ are solutions corresponding to different initial conditions, ξ_1 and ξ_2, then

$$x_1(t) - x_2(t) = \xi_1 - \xi_2 + \int_0^t \sum_{j=1}^N \varphi_j(w,h;s)\ (f_j(x_1(s)) - f_j(x_2(s)))\ ds$$

which implies

$$|x_1(t) - x_2(t)| \le |\xi_1 - \xi_2| + L \int_0^t |x_1(s) - x_2(s)|\, ds$$

and we obtain

$$|x_{(w,h)}(t,\xi_1) - x_{(w,h)}(t,\xi_2)| \le |\xi_1 - \xi_2|\,\frac{e^{Lt} - 1}{L}\,. \tag{26}$$

Notation:

- $\mathcal{SW}$ – the set of all switched solutions (associated with the maps $f_1, \ldots, f_N$),
- $\mathcal{SW}(\xi)$ – the set of all switched solutions initiating at the point ξ,
- $\mathcal{SW}_{[0,T]}(\xi)$ – restrictions to the time interval $[0,T]$ of all switched solutions initiating at the point ξ,
- $\mathcal{SW}(\xi; s)$ – the set of values at $t = s$ of all switched solutions $x_{(w,h)}(t,\xi)$.

4.2 Continuous Limit Dynamics (CLD)

Starting with the maps S_j^τ we define the corresponding CLD following the recipe of the previous section. The maps S_j^τ obviously satisfy assumption **A0**. Assumptions **A1**, **A2**, and **A3** are satisfied thanks to estimate (24). Assumption **A1** is satisfied because each $S_j^\tau(x)$ is continuous in τ and x, and the continuity is uniform when $\tau \in [0,1]$ and x varies on a compact set. Assumptions **A2** and **A3** are satisfied because the right-hand side of (24) depends on ξ and t only. Assumption **A4** is a consequence of (26), and assumption **AX** is valid since $X = \mathbb{R}^d$. Thus, Theorem 12 is applicable in our situation. Recall Definition 9 of the multi-valued map $\mathcal{F}_t$. It is instructive to compare the set $\mathcal{F}_t(\xi)$ with the set $\mathcal{SW}(\xi;t)$. Note that $\mathcal{SW}(\xi;t) \subset \mathcal{F}_t(\xi)$ because one can chop any time interval $h(n)$ into arbitrarily small subintervals and use the semi-group property of the maps S_j^τ. In other words, $S_{w(n)}^{h(n)}(x) = S_{w(n)}^{h(n)/N} \circ S_{w(n)}^{h(n)/N} \circ \cdots \circ S_{w(n)}^{h(n)/N}(x)$, so we can replace w by a string in which every symbol of w is repeated N times, and replace h by a string where each $h(n)$ is replaced by N repetitions of $h(n)/N$. The end-point, $x_{(w,h)}(t,\xi)$ will not change. Now it is clear that the closure of $\mathcal{SW}(\xi;t)$ in $\mathbb{R}^d$ is equal to $\mathcal{F}_t(\xi)$,

$$\overline{\mathcal{SW}(\xi;t)} = \mathcal{F}_t(\xi)\,. \tag{27}$$

In fact, the closure of the set $\mathcal{SW}_{[0,T]}(\xi)$ in the space $C([0,T],\mathbb{R}^d)$ yields exactly all integral curves in $\mathcal{CLD}_{[0,T]}(\xi)$,

$$\overline{\mathcal{SW}_{[0,T]}(\xi)} = \mathcal{CLD}_{[0,T]}(\xi)\,. \tag{28}$$

This shows the connection between the switched solutions and the CLD as defined in the previous section.

4.3 Differential inclusions

Now we turn to differential inclusion. Define the multi-valued map

$$F(x) = \{f_1(x), \ldots, f_N(x)\} \tag{29}$$

and consider the differential inclusion

$$\dot{x}(t) \in F(x(t))\,. \tag{30}$$

Recall that a solution of (30) with the initial condition $x(0) = \xi$ is an absolutely continuous (a.s. for short) function $x(t)$ such that $x(0) = \xi$ and $x(t) \in F(x(t))$ for a.e. t. See [1] for information on differential inclusions. For this particular map $F(x)$, the existence of solutions is not a problem because any solution of (18) is a solution of (30). We will be interested in the totality of solutions of (30), so, by analogy with the switched solutions, we define the sets $\mathcal{DI}$, $\mathcal{DI}(\xi)$, etc. In particular, $\mathcal{DI}_{[0,T]}(\xi)$ is the set of restrictions to the

the interval $[0,T]$ of all solutions to (30) with the initial condition $x(0) = \xi$. Along with (30), it is important to consider its relaxed version,

$$\dot{x}(t) \in \overline{co}\, F(x(t)) \tag{31}$$

where

$$\overline{co}\, F(x) = \overline{co}\, \{f_1(x), \ldots, f_N(x)\}\,, \tag{32}$$

and $\overline{co}$ stands for the closure of the convex hull. Again, any solution to (18) is a solution of (31). The totality of solutions of (31) is known to have much better properties than that of (30), see [1]. In particular, the set $\mathcal{RDI}_{[0,T]}(\xi)$ of restrictions to the the interval $[0,T]$ of all solutions to (31) with the initial condition $x(0) = \xi$ is known to be compact in the space of continuous functions $C([0,T];\mathbb{R}^d)$, [1, Theorem 2.2.1]. All solutions in $\mathcal{DI}$ and in $\mathcal{RDI}$ are bounded,

$$|x(t,\xi)| \le |\xi| + |f(\xi)|\, \frac{e^{LT}-1}{L}. \tag{33}$$

What is more important to us is that the set $\mathcal{DI}_{[0,T]}(\xi)$ is dense in $\mathcal{RDI}_{[0,T]}(\xi)$ by the Filippov-Ważewski theorem, [3].

It is clear that $\mathcal{SW}_{[0,T]}(\xi) \subset \mathcal{DI}_{[0,T]}(\xi)$. Next we show that the closure of $\mathcal{SW}_{[0,T]}(\xi)$ in the space of continuous functions $C([0,T],\mathbb{R}^d)$ is equal to $\mathcal{RDI}_{[0,T]}(\xi)$.

Theorem 14.

$$\mathcal{CLD}_{[0,T]}(\xi) = \mathcal{RDI}_{[0,T]}(\xi)\,. \tag{34}$$

Proof. The plan is to show that every solution $\gamma(t)$ of the relaxed inclusion (31) with $\gamma(0) = \xi$ can be approximated uniformly on the finite interval $[0,T]$ by switched solutions. Theorem 12 then will complete the proof of (34).

Let $\gamma(t)$ be a solution of (31) with the initial condition $\gamma(0) = \xi$. Fix the interval $[0,T]$. It is known, see [1, 1.14 Corollary 1] that there exist bounded measurable functions α_j on $[0,T]$ such that

$$0 \le \alpha_j(t) \le 1\,, \quad j = 1,\ldots,N, \quad \text{and} \quad \sum_{j=1}^{N} \alpha_j(t) = 1 \quad \text{a.e. on } [0,T], \tag{35}$$

and

$$\gamma(t) = \xi + \int_0^t \sum_{j=1}^{N} \alpha_j(s)\, f_j(\gamma(s))\, ds\,, \text{ for all } t \in [0,T]\,. \tag{36}$$

Approximate each α_i by simple functions as follows. For each positive integer n define the averages

$$\beta_{i,n}^k = \frac{n}{T}\int_{k\frac{T}{n}}^{(k+1)\frac{T}{n}} \alpha_i(s)\,ds\,,\ k = 0,\ldots,n-1\,,$$

and define the simple functions

$$\beta_{i,n}(t) = \beta_{i,n}^k \quad \text{for} \quad \frac{kT}{n} \leq t < \frac{(k+1)T}{n}. \tag{37}$$

Then $\sum_{i=1}^N \beta_{i,n}(t) = 1$ for all $t \in [0,T)$, and all n. Moreover, the functions $\beta_{i,n}$ converge to α_i in $L^1([0,T])$, i.e.,

$$\lim_{n\to\infty}\int_0^T |\beta_{i,n}(s) - \alpha_i(s)|\,ds = 0\,.$$

For each n denote by γ_n the unique solution of the corresponding equation

$$\gamma_n(t) = \xi + \int_0^t \sum_{j=1}^N \beta_{j,n}(s)\, f_j(\gamma_n(s))\,ds\,.$$

It is clear that $\gamma_n \in \mathcal{RDI}_{[0,T]}(\xi)$ and it is easy to show that $\gamma_n \to \gamma$ in $C([0,T],\mathbb{R}^d)$.

From now on we drop subscript n and work with $\gamma(t,\xi)$, the solution of the equation

$$\dot\gamma(t) = \sum_{i=1}^N \beta_i(t) f_i(\gamma(t))\,,\ x(0) = \xi\,, \tag{38}$$

where each β_i is a constant between 0 and 1 on the intervals $[kT/n,(k+1)T/n]$ and $\sum_i \beta_i(t) = 1$. We will approximate γ by switched solutions. To do this, we use Proposition 15 below which is motivated by the classical Lie product formula:

$$e^{t(A+B)} = \lim_{m\to\infty}\left(e^{\frac{t}{m}A}\, e^{\frac{t}{m}B}\right)^m\,, \tag{39}$$

where A and B are square matrices. This formula shows how to approximate the evolution function of the equation $\dot x = (A+B)x$ by the evolution functions of the equations $\dot x = Ax$ and $\dot x = Bx$. Proposition 15 is a similar in spirit result. We consider the equation

$$\dot x(t) = \sum_{i=1}^N b_i f_i(x)\,, \tag{40}$$

where $b_1, b_2, \ldots, b_N$ are positive real numbers such that $\sum_{i=1}^N b_i = 1$. This equation has the same form as equation (38) on each of the intervals $[kT/n,(k+1)T/n]$. To model this when dealing with (40), we consider a fixed interval $[0, \tau]$. Denote by $S^t_{\sum_{i=1}^N b_i f_i}$ the solution map for (40) and denote by S^t_i the solution maps of the individual equations $\dot{x} = f_i(x)$. The analogue of (39) is the following formula

$$S^t_{\sum_{i=1}^N b_i f_i} = \lim_{m\to\infty} \left(S_N^{b_N \frac{t}{m}} \circ S_{N-1}^{b_{N-1} \frac{t}{m}} \circ \cdots \circ S_1^{b_1 \frac{t}{m}} \right)^m , \tag{41}$$

where the equality is understood in the strong sense, i.e., on every $\xi \in \mathbb{R}^d$. This will be a corollary of a more general statement.

Given an integer $m > 0$, define $h_m \in \Sigma_{\mathcal{I}_1}$ as follows:

$$h_m(0) = b_1 \frac{\tau}{m} , \quad h_m(1) = b_2 \frac{\tau}{m} , \quad \dots , h_m(N-1) = b_N \frac{\tau}{m} , \tag{42}$$

and, for $k \geq N$,

$$h_m(k) = h_m(k(N)) , \tag{43}$$

where, to shorten the notation,

$$k(N) = k \mod N .$$

Also, define $w_m \in \Sigma_{\mathcal{J}}$ as follows:

$$w_m(k) = 1 + k(N) .$$

Simplify the notation for the switched solution $x_{(w_m, h_m)}(t, \xi)$ to $x_m(t, \xi)$.

Proposition 15. *For any choice of $\xi, \eta \in \mathbb{R}^d$, there exists a constant $C(\xi, \eta)$ such that*

$$|\gamma(t, \xi) - x_m(t, \eta)| \leq e^{L\tau} |\xi - \eta| + \frac{C(\xi, \eta)}{m} \tag{44}$$

for all $t \in [0, \tau]$. In particular, the functions $x_m(\cdot, \xi)$ converge to $\gamma(\cdot, \xi)$ in $C([0, \tau], \mathbb{R}^d)$ as $m \to \infty$.

Proof. Pick ξ and η in $\mathbb{R}^d$. We will compare $\gamma(t, \xi)$ with $x_m(t, \eta)$. We use the representations

$$\gamma(t, \xi) = \xi + \int_0^t \sum_{i=1}^N b_i f_i(\gamma(s, \xi))\, ds \tag{45}$$

and

$$x_m(t, \eta) = \eta + \int_0^t \sum_{i=1}^N \varphi_i(w_m, h_m, s)\, f_i(x_m(s, \eta))\, ds . \tag{46}$$

The bounds on these solutions are given in (25) and (33), and on the finite interval $[0, \tau]$ and with the fixed ξ, we get the bounds:

$$|\gamma(t,\xi)|, |x_m(t,\xi)| \le |\xi| + |f(\xi)| \frac{e^{Lt}-1}{L} \le |\xi| + |f(\xi)| \frac{e^{L\tau}-1}{L}, \quad 0 \le t \le \tau . \tag{47}$$

Also, using (24) and the Lipschitz property of f_j's, we obtain

$$\max_j |f_j(\gamma(t,\xi))|, |f_j(x_m(t,\xi))| \le |f(\xi)|\, e^{Lt} \le |f(\xi)|\, e^{L\tau}, \quad 0 \le t \le \tau . \tag{48}$$

We next estimate the difference $\gamma(t,\xi) - x_m(t,\eta)$. To handle the integral in (46), we use (42), (43) to partition the interval $[0, \tau]$. Denote $n_m = \lfloor mt/\tau \rfloor$. Next, denote

$$\lambda_0 = 0, \quad \lambda_j = \sum_{i=1}^{j} b_i, \quad j = 1, \ldots, N .$$

Define

$$t_0 = 0, \quad t_{k+1} = t_k + h_m(k).$$

Using λ_j's we can re-write this as

$$t_k = \left(\lfloor \frac{k}{N} \rfloor + \lambda_{k(N)} \right) \frac{\tau}{m}, \quad k = 0, 1, \ldots . \tag{49}$$

The interval $[0, t]$ with $t \le \tau$ will contain $N\, n_m$ of whole subintervals $[t_k, t_{k+1})$ of the partition plus possibly a leftover interval, $[t_{L_m}, t]$. Re-write the integral in (46) using this partition and the fact that

$$\varphi_i(w_m, h_m, s) = \delta_{i,k(N)} \quad \text{for} \quad t_{k-1} \le s < t_k .$$

We have

$$\int_0^t \sum_{i=1}^N \varphi_i(w_m, h_m, s)\, f_i(x_m(s,\eta))\, ds = \sum_{k=0}^{N n_m - 1} \int_{t_k}^{t_{k+1}} \sum_{i=1}^N \varphi_i(w_m, h_m, s)\, f_i(x_m(s,\eta))\, ds$$

$$+ \int_{n_m\tau/m}^t \sum_{i=1}^N \varphi_i(w_m, h_m, s)\, f_i(x_m(s,\eta))\, ds$$

We use (48) and the fact that $t - n_m\tau/m \le \tau/m$ to estimate the last integral:

$$\left| \int_{n_m\tau/m}^t \sum_{i=1}^N \varphi_i(w_m, h_m, s)\, f_i(x_m(s,\eta))\, ds \right| \le |f(\eta)| e^{L\tau}\, (\tau/m) . \tag{50}$$

Invoking (49) and change of variables, we obtain

$$\sum_{k=0}^{Nn_m-1}\int_{t_k}^{t_{k+1}}\sum_{i=1}^{N}\varphi_i(w_m,h_m,s)\,f_i(x_m(s))\,ds=\sum_{k=0}^{Nn_m-1}\int_{t_k}^{t_{k+1}}f_{k(N)}(x_m(s))\,ds$$

$$=\sum_{k=0}^{n_m-1}\sum_{i=1}^{N}\int_{(k+\lambda_{i-1})\tau/m}^{(k+\lambda_i)\tau/m}f_i(x_m(s))\,ds$$

$$=\sum_{k=0}^{n_m-1}\sum_{i=1}^{N}b_i\ \int_0^{\tau/m}f_i(x_m((k+\lambda_{i-1})\tau/m+b_is,\eta)\,ds\,.$$

Now turn to the integral in representation (45). Here we partition $[0,t]$ into the intervals of size τ/m plus what is left.

$$\int_0^t\sum_{i=1}^{N}b_if_i(\gamma(s))\,ds=$$

$$\sum_{k=0}^{n_m-1}\sum_{i=1}^{N}b_i\ \int_0^{\tau/m}f_i(\gamma(k\tau/m+s))\,ds+\int_{n_m\tau/m}^{t}\sum_{i=1}^{N}b_i\ f_i(\gamma(s))\,ds\,.$$

Again,

$$|\int_{n_m\tau/m}^{t}\sum_{i=1}^{N}b_i\,f_i(\gamma(s,\xi))\,ds|\le|f(\xi)|e^{L\tau}\,(\tau/m)\,. \tag{51}$$

Thus,

$$|\gamma(t,\xi)-x_m(t,\eta)|\le|\gamma(n_m\tau/m,\xi)-x_m(n_m\tau/m,\eta)|+(|f(\xi)|+|f(\eta)|)\,e^{L\tau}\,(\tau/m)\,.$$

Now we estimate the difference $\gamma(n\tau/m,\xi)-x_m(n\tau/m,\eta)$, where n is any integer between 0 and m. Write the difference as follows

$$\sum_{k=0}^{n-1}\sum_{i=1}^{N}b_i\ \int_0^{\tau/m}f_i(\gamma(k\tau/m+s,\xi))\,ds-\sum_{k=0}^{Nn-1}\int_{t_k}^{t_{k+1}}\sum_{i=1}^{N}\varphi_i(w_m,h_m,s)\,f_i(x_m(s,\eta))\,ds$$

which can be re-written as

$$\sum_{k=0}^{n-1}\sum_{i=1}^{N}b_i\ \int_0^{\tau/m}\left(f_i(\gamma(k\tau/m+s,\xi))-f_i(x_m((k+\lambda_{i-1})\tau/m+b_is,\eta))\right)\,ds\,.$$

Its absolute value is not greater than

$$L\ \sum_{k=0}^{n-1}\sum_{i=1}^{N}b_i\ \int_0^{\tau/m}|\gamma(k\tau/m+s,\xi)-x_m((k+\lambda_{i-1})\tau/m+b_is,\eta)|\,ds\,. \tag{52}$$

We use the triangle inequality

$$|\gamma(k\tau/m+s)-x_m((k+\lambda_{i-1})\tau/m+b_is)|\le|\gamma(k\tau/m,\xi)-x_m(k\tau/m,\eta)|+$$

$$|\gamma(k\tau/m+s)-\gamma(k\tau/m)|+|x_m((k+\lambda_{i-1})\tau/m+b_is)-x_m((k+\lambda_{i-1})\tau/m)|+$$

$$|x_m((k+\lambda_{i-1})\tau/m)-x_m(k\tau/m)|\,.$$

The contribution of each of the last three terms on the right when integrated over the interval $[0, \tau/m]$ can be estimated using a simple Lipschitz continuity of the solutions valid simultaneously for $\gamma(t,\xi)$ and $x_m(t,\eta)$:

$$|\gamma(t_2,\xi) - \gamma(t_1,\xi)|,\ |x_m(t_2,\xi) - x_m(t_1,\xi)| \leq |f(\xi)|\, e^{L\tau}\, (t_2 - t_1)\,,\ 0 \leq t_1 \leq t_2 \leq \tau\,. \quad (53)$$

Thus, we obtain

$$\int_0^{\tau/m} |\gamma(k\tau/m + s) - \gamma(k\tau/m)|\, ds \leq \frac{1}{2}\, |f(\xi)|\, e^{L\tau}\, (\tau/m)^2\,,$$

$$\int_0^{\tau/m} |x_m((k+\lambda_{i-1})\tau/m + b_i s) - x_m((k+\lambda_{i-1})\tau/m)|\, ds \leq \frac{1}{2}\, |f(\eta)|\, e^{L\tau}\, (\tau/m)^2\,,$$

and

$$\int_0^{\tau/m} |x_m((k+\lambda_{i-1})\tau/m) - x_m(k\tau/m)|\, ds \leq |f(\eta)|\, e^{L\tau}\, \lambda_{i-1}\, (\tau/m)^2 \leq |f(\eta)|\, e^{L\tau}\, (\tau/m)^2\,.$$

Collecting these inequalities to bound (52) we find that $|\gamma(n\tau/m,\xi) - x_m(n\tau/m,\eta)|$ is bounded by $|\xi - \eta|$ plus

$$L \sum_{k=0}^{n-1} \left(|\gamma(k\tau/m,\xi) - x_m(k\tau/m,\eta)|\, (\tau/m) + (\frac{1}{2}\, |f(\xi)| + \frac{3}{2}\, |f(\eta)|)\, e^{L\tau}\, (\tau/m)^2 \right)\,.$$

Denote temporarily $z_k = |\gamma(k\,\tau/m,\xi) - x_m(k\,\tau/m,\eta)|$. For any $n = 1,\dots,m$ we have the following inequalities:

$$z_n \leq z_0 + L\,(\tau/m) \sum_{k=0}^{n-1} z_k + \frac{3}{2}\, L\,(|f(\xi)| + |f(\eta)|)\, e^{L\tau}\, \tau^2/m\,.$$

Solving the discrete inequality

$$z_n \leq z_0 + L\,(\tau/m) \sum_{k=0}^{n-1} z_k + C/m$$

we obtain

$$z_n \leq (1 + L\,\tau/m)^{n+1}\, z_0 + (1 + L\,\tau/m)^n\, C/m$$

We conclude that

$$z_n \leq (1 + L\,\tau/m)\, e^{L\tau}\, z_0 + \frac{3}{2}\, L\,(|f(\xi)| + |f(\eta)|)\, e^{2\,L\tau}\, \tau^2/m\,.$$

This completes the proof of the proposition. □

Proof of Theorem 14 continued. Now it is clear how to approximate the solution of (38) by switched solutions. There is a finite number of intervals of length $\tau = T/n$. We approximate γ by x_m choosing appropriate m on consecutive intervals. If $|\gamma(t,\xi) - x_{m_1}(t,\xi)| \leq \epsilon$ on the first interval, then we use estimate (44) and choose m_2 so that $|\gamma(t,\xi) - x_{m_2}(t,\xi)| \leq e^{L\tau}\epsilon + \epsilon = (e^{L\tau}+1)\epsilon$ on the second interval. For the third interval we choose m_3 to have $|\gamma(t,\xi) - x_{m_3}(t,\xi)| \leq (e^{L\tau}+1)^2\epsilon$ there. This way we construct a switched solution which is at most $(e^{L\tau}+1)^n\epsilon$ away from $\gamma(t)$ on the whole interval $[0,T]$. The theorem is proved. □

In the assumptions of the preceding theorem, consider the variable time-step dynamics with choice associated with the maps $S_j^\tau(x) = x + \tau\, f_j(x)$ (S_j^τ is the evolution map of the Euler method for the equation (18)). The corresponding switched trajectories are given by $x_{n+1} = S_{j_n}^{\tau_n}(x_n)$. Using the fact that the Euler method approximates the solutions of the constituent differential equations, and using Theorem 14, we conclude that the CLD for this variable time-step dynamics with choice coincides with the dynamics of the relaxed differential inclusion (31) and that equality (34) holds.

References

[1] Aubin, J. P.; Cellina, A. Differential inclusions. Set-valued maps and viability theory. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 264. Springer-Verlag, Berlin, 1984.

[2] Branicky, M. "Introduction to hybrid systems", in D. Hristu-Varsakelis and W.S. Levine (eds.), Handbook of Networked and Embedded Control Systems, pp. 91-116. Birkhauser, 2005.

[3] Filippov, A. F. Differential equations with discontinuous righthand sides. Translated from the Russian. Mathematics and its Applications (Soviet Series), 18. Kluwer Academic Publishers Group, Dordrecht, 1988.

[4] Goebel, Rafal; Sanfelice, Ricardo G.; Teel, Andrew R. (2009), "Hybrid dynamical systems", IEEE Control Systems Magazine 29 (2): 2893

[5] Kapitanski, L.; Zivanovic, S.: *Dynamics with choice*. Nonlinearity 22 (2009) 163-186.

[6] Kapitanski, L.; Zivanovic, S.: *Dynamics with a range of choice*. Reliable Computing 15, (2011) 290-299.

[7] Kapitanski, L.; Živanović Gonzalez, S.; Variable time-step dynamics with choice, submitted.

[8] K. Kuratowski, Topology, Volume I. New York - London - Warszawa, Academic Press: Polish Scientific Publishers, 1966.

[9] D. Liberzon, Switching in Systems and Control, Birkhäuser, Boston, 2003.

[10] A. van der Schaft, H. Schumacher. An Introduction to Hybrid Dynamical Systems. Lecture Notes in Control and Information Sciences 251, Springer-Verlag, 2000.

[11] Shorten, Robert; Wirth, Fabian; Mason, Oliver; Wulff, Kai; King, Christopher Stability criteria for switched and hybrid systems. SIAM Rev. 49 (2007), no. 4, 545592.

[12] Zivanovic, S.; *Attractors in Dynamics with Choice*. PhD Thesis, University of Miami, Coral Gables, Fl, USA, 2009. Open Access Dissertations. Paper 210.
$http://scholarlyrepository.miami.edu/oa\_dissertations/210$

[13] Zorich, Vladimir A. Mathematical analysis. II. Translated from the 2002 fourth Russian edition by Roger Cooke. Universitext. Springer-Verlag, Berlin, 2004