Shadowing and Numerical Analysis of Set-Valued Dynamical Systems

Dissertation zur Erlangung
des Doktorgrades der Fakultät für Mathematik
der Universität Bielefeld

vorgelegt von
Janosch Rieger

Betreuender Hochschullehrer / Supervisor:
Prof. Dr. Wolf-Jürgen Beyn, Universität Bielefeld

Gutachter / Reviewers:
Prof. Dr. Wolf-Jürgen Beyn, Universität Bielefeld
Prof. Dr. Frank Lempio, Universität Bayreuth

Datum des Kolloquiums / Date of thesis defense and acceptance:
21. Juli 2009 / 21st of Juli, 2009

Gedruckt auf alterungsbeständigem Papier ISO 9706

Contents

List of symbols

$\langle \cdot, \cdot \rangle$	Euclidean inner product
$\vert \cdot \vert$	Euclidean norm
$\Vert \cdot \Vert_\infty$	supremum norm
$B_a(x)$	ball of radius a with center x
id	identity mapping
$L^1(\cdot)$	Lebesgue integrable functions
$C(\cdot)$	continuous functions
$AC(\cdot)$	absolutely continuous functions
$C^1(\cdot)$	continuously differentiable functions
$C^\infty(\cdot)$	infinitely many times differentiable functions
co	convex hull
$\overline{\text{co}}$	closure of the convex hull
$\mathcal{A}(\mathbb{R}^m)$	collection of closed subsets of $\mathbb{R}^m$
$\mathcal{C}(\mathbb{R}^m)$	collection of compact subsets of $\mathbb{R}^m$
$\mathcal{CC}(\mathbb{R}^m)$	collection of convex and compact subsets of $\mathbb{R}^m$
dist	one-sided Hausdorff distance
dist_H	symmetric Hausdorff distance
$\sigma(p, A)$	Hamilton function (also called support function)
$\text{Proj}(x, A)$	projection of the vector x to the set A
$\text{Dev}(x, A)$	set of shortest vectors joining x and the set A
T_xM	tangent space to a manifold M at x
$T_K(x)$	contingent cone at x to K
$\mathcal{R}(T, t, x)$	reachable set (of a differential inclusion) from (t, x) at time T
$\text{Viab}_F(K)$	viability kernel of K under F
$E^s(x)$, $E^u(x)$	stable/unstable subspace at x
$W^s(x)$, $W^u(x)$	stable/unstable manifold of x
$W^{s,\varepsilon}(x)$	local unstable manifold of x
$Q(x)$, $P(x)$	projection to the stable/unstable subspace

Introduction

Discrete and continuous-time set valued dynamical systems arise whenever the behaviour of a system is not uniquely determined. There is an abundance of applications featuring non-unique trajectories such as control systems, economical models, and the deterministic treatment of uncertainty (cp. e.g. [38], [1], and [19]).

Though there is a great number of results about particular aspects of multivalued dynamics in the literature, there exists no closed theory that could be named a theory of set-valued dynamical systems. In set-valued numerical analysis, the situation is similar. There exist many singular contributions to the topic, but a transparent general concept still has to be developed.

Shadowing theory examines the impact of uniformly small perturbations on the behaviour of dynamical systems on infinite time intervals. For classical dynamical systems, shadowing theory is an established field of research (cf. [30] and [31]) which is intimately related to structural stability, ergodic theory, and the notion of hyperbolicity.

Moreover, shadowing theory can be interpreted as a theory oriented branch of numerical analysis. Since a numerical method is a perturbation of the original system, shadowing theorems provide estimates for the accuracy of numerical methods on infinite time intervals when they are applied to the time-t flow of a differential equation or inclusion.

In the context of set-valued dynamical systems, only few attempts have been made to establish shadowing results (see [36] and [21]).

The content of this thesis is organized as follows.

In Chapter 1, the vocabulary and basic elements of set-valued analysis are introduced, while the most important facts related to differential inclusions, i.e. generalized ordinary differential equations of the form

$$\dot{x}(t) \in F(x(t)) \text{ for almost all } t \in [0,T],$$

are presented in Chapter 2. The density theorems proved in Section 2.4 are applied in Section 2.5 where the time-t flow of a differential inclusion with relaxed one-sided Lipschitz right hand side is shown to be a set-valued contraction.

Chapter 3 briefly summarizes classical shadowing theory for diffeomorphisms in order to display those concepts and ideas to which the content of the following chapters is linked.

In Chapter 4, first shadowing results for discrete-time set-valued dynamical systems of the form

$$p_{k+1} \in F(p_k) \text{ for all } k \in \mathbb{Z}$$

are given. Section 4.1 provides adaptations of the notions of pseudotrajectories, of the shadowing property, and of the inverse shadowing property to the set-valued environment.

The relatively simple class of contracting mappings analyzed in Section 4.2 deserves attention, because it contains the time-t flows of differential inclusions with relaxed one-sided Lipschitz right hand sides with negative Lipschitz constants.

In Section 4.3 a first definition of hyperbolicity for set-valued mappings is proposed, and it is shown that it implies the shadowing and the inverse shadowing property. The essence of the coresponding results is further refined to a selection-based and less restrictive notion of hyperbolicity in Section 4.4.

In Chapter 5, shadowing theory is applied to the Viability Kernel Algorithm, which is one of the most important numerical schemes in the set-valued context. This algorithm computes the largest subset of a given domain that is weakly invariant under the flow induced by a differential inclusion. It is natural to use shadowing theorems in order to derive error estimates for the accuracy of this algorithm, because the behaviour of exact and numerical trajectories on the unbounded time interval $[0,\infty)$ must be controlled.

Eventually, explicit error bounds and linear convergence of the Viability Kernel Algorithm are proved for the class of one-sided Lipschitz right hand sides.

Future prospects

According to my opinion, it is incomprehensible that there is no systematic approach to set-valued differentiation, to set-valued dynamical systems, and to set-valued numerical analysis. I believe that we do not need a *ménagerie* of tangent cones and corresponding set-valued differentials, cp. [3], but one solid differential calculus in order to understand dynamics and numerical analysis. Furthermore, I believe that such a differential calculus will be the right language for formulating a powerful hyperbolicity condition which will be easily verifiable in concrete applications.

As far as convex-valued maps and convex reachable sets are concerned, it seems relatively easy to establish such a theory, because the Hörmander embedding (see Chapter 1 or [24]) into the Banach space of continuous real-valued functions on the sphere provides a framework in which sets become computable objects. In the significantly more important non-convex case it is unclear what can be achieved.

Statement of originality

The results displayed in Sections 4.2.1 and 4.3 have partly been developed in the framework of a cooperation with Prof. Dr. Sergei Pilyugin (St. Petersburg). We worked out the set-valued notions of pseudotrajectories, the shadowing property, and the inverse shadowing property as well as Theorem 67, the original version of Theorem 70, Example 72, Definition 75, and Theorem 77. The remaining parts of the corresponding Sections have been established by me alone.

The original statement and the proof of Theorem 70 given in Section 4 of [33] contain a serious inaccuracy. The version displayed in the present text is based on the same fixed point argument, but the space of sequences in which the fixed point theorem is eventually applied must be fixed *after* the pseudotrajectory has been specified; in addition, the diameters of the images of the defining mapping must be uniformly bounded.

Please note that Theorem 70 is of considerable importance, because it is presently almost indispensable for the treatment of the time-t flow of differential inclusions. In [32], I proved that contractive set-valued mappings with not necessarily convex images but sufficiently large 'continuous convex kernels' still have the shadowing property. On the basis of this theorem, it is possible to prove a shadowing result for one-sided Lipschitz differential inclusions, but Theorem 70 is by far more elegant and natural.

While I profited from Sergei Pilyugin's remarkable knowledge about shadowing technique, I could contribute my intuition for set-valued concepts. In the hyperbolic setup, I made the key observation that, in contrast to the single-valued case, the operator must be defined before the projections are applied. In Section 4.4, I refined this idea on my own and developed a fully selection-based notion of hyperbolicity.

I worked out the entire content of Chapter 5 by myself.

The search for a class of differential inclusions with contractive or hyperbolic time-T flow turned out to be a veritable odyssey through the literature.

As I was not aware of the work of Tzanko Donchev and Elza Farkhi, see e.g. [13], I reinvented the relaxed one-sided Lipschitz condition and proved contractivity of the time-t flow of differential inclusions with one-sided Lipschitz right hand sides.

In order to use classical techniques, I needed the graph of the mapping (2.35) to be closed. Thus I proved Density Theorem 40 (which was also well-known, but originally proved in a very different way, see Section 2.4) in order to approximate arbitrary solutions by smooth ones.

When I learned about the existence theorem for the Caratheodory case and the Inverse Intersection Lemma, I realized that with these new tools, it was fairly easy to prove contractivity (see Theorem 42).

In summary, I have developed and proved all results mentioned above independently, but only the C^∞ Density Theorem (Theorem 39) is a truly new result.

Chapter 1

Some set-valued analysis

Set-valued analysis provides tools for the study of differential inclusions, optimization problems, and less popular topics such as the study of inverses of single-valued mappings which are not one-to-one. In this text, only basic definitions and results which are necessary for a self-contained presentation will be given. For a complete overview over the topic, the reader is referred to [3]. An introduction to the matter with a focus on convex optimization is presented in [34].

1.1 Basic notation

Let $\mathbb{R}^m$ be equipped with the Euclidean norm $|\cdot|$, and let $B_r(x)$ and $B(x,r)$ denote the ball of radius $r \geq 0$ around the point $x \in \mathbb{R}^m$. Set $B := B_1(0)$, and define

$$B_r(A) := B(A,r) := \cup_{a\in A} B_r(a)$$

for any $A \subset \mathbb{R}^m$. The collection of all subsets of $\mathbb{R}^m$ will be denoted by $\mathcal{P}(\mathbb{R}^m)$, and for any $A \subset \mathbb{R}^m$, the convex hull and the closure of the convex hull of A will be denoted by $\operatorname{co} A$ and $\overline{\operatorname{co}} A$.

The symbols $\mathcal{A}(\mathbb{R}^m)$, $\mathcal{C}(\mathbb{R}^m)$, and $\mathcal{CC}(\mathbb{R}^m)$ will denote the collections of the nonempty closed, the nonempty compact, and the nonempty convex and compact subsets of $\mathbb{R}^m$, respectively.

For any $A, B \subset \mathbb{R}^m$, the Minkowski sum is defined by

$$A + B := \{a + b : a \in A, b \in B\}, \tag{1.1}$$

and similarly $\mu A := \{\mu a : a \in A\}$.

The projection $\mathrm{Proj} : \mathbb{R}^m \times \mathcal{P}(\mathbb{R}^m) \to \mathcal{P}(\mathbb{R}^m)$ of a vector to a set is defined by

$$\mathrm{Proj}(x, A) := \{a \in A : |x - a| \leq |x - a'| \ \forall a' \in A\}, \tag{1.2}$$

while the deviation $\mathrm{Dev} : \mathbb{R}^m \times \mathcal{P}(\mathbb{R}^m) \to \mathcal{P}(\mathbb{R}^m)$ is given by

$$\mathrm{Dev}(x, A) := \mathrm{Proj}(x, A) - x \tag{1.3}$$

and can be interpreted as the set of shortest vectors joining x and A. If Proj is single-valued and continuous, then so is Dev.

For compact sets $A, B \in \mathcal{C}(\mathbb{R}^m)$, the one-sided and the symmetric Hausdorff distance are defined by

$$\mathrm{dist}(A, B) := \sup_{a \in A} \inf_{b \in B} |a - b| \tag{1.4}$$

and

$$\mathrm{dist}_H(A, B) := \max\{\mathrm{dist}(A, B), \mathrm{dist}(B, A)\}, \tag{1.5}$$

respectively. While dist_H is a distance, i.e. it satisfies

- $\mathrm{dist}_H(A, B) = 0 \Leftrightarrow A = B$,
- $\mathrm{dist}_H(A, B) = \mathrm{dist}_H(B, A)$, and
- $\mathrm{dist}_H(A, C) \leq \mathrm{dist}_H(A, B) + \mathrm{dist}_H(B, C)$

for all $A, B, C \in \mathcal{C}(\mathbb{R}^m)$, the semidistance dist is neither definite nor symmetric. A weak substitute for these properties is the equivalence

$$\mathrm{dist}(A, B) = 0 \Leftrightarrow A \subset B \tag{1.6}$$

for $A, B \in \mathcal{C}(\mathbb{R}^m)$.

The maximal norm of the elements of a set $A \subset \mathbb{R}^m$ is denoted by

$$\|A\| := \sup_{a \in A} |a|. \tag{1.7}$$

This notation is standard, but slightly misleading, because the expression is not a norm at all.

The nonempty convex and compact sets $\mathcal{CC}(\mathbb{R}^m)$ can be embedded into the linear vector space $C(S^{m-1})$ of continuous functions from the sphere into the real numbers by setting

$$\sigma(p, A) := \max_{a \in A} \langle p, a \rangle \tag{1.8}$$

for any $A \in \mathcal{CC}(\mathbb{R}^m)$ and $p \in S^{m-1}$, see [24]. The function $\sigma(\cdot, A) : S^{m-1} \to \mathbb{R}$ is called the Hamilton function of the set A. The properties of these functions are discussed e.g. in [3] and [34].

1.2 Set-valued mappings, notions of continuity

A set-valued mapping $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ is a function $F : \mathbb{R}^{m_1} \to \mathcal{P}(\mathbb{R}^{m_2})$. Throughout this text it will be assumed that the images of set-valued mappings are nonempty.

Definition 1. *Let $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ be a set-valued mapping. The image $F(A)$ of a subset $A \subset \mathbb{R}^{m_1}$ is the union*

$$F(A) := \cup_{x \in A} F(x), \tag{1.9}$$

and the inverse image $F^{-1}(B)$ of a set $B \subset \mathbb{R}^{m_2}$ is defined by

$$F^{-1}(B) := \{x \in \mathbb{R}^{m_1} : F(x) \cap B \neq \emptyset\}. \tag{1.10}$$

Definition 2. *Let $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ be a set-valued mapping. Any single-valued function $f : \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$ such that $f(x) \in F(x)$ for all $x \in \mathbb{R}^{m_1}$ is called a selection of F.*

Definition 3. *A set-valued mapping $F : \mathbb{R}^{m_1} \to \mathcal{C}(\mathbb{R}^{m_2})$ is called upper semicontinuous (usc) at $x \in \mathbb{R}^{m_1}$ if*

$$\operatorname{dist}(F(x'), F(x)) \to 0 \text{ as } x' \to x. \tag{1.11}$$

It is called lower semicontinuous (lsc) at $x \in \mathbb{R}^{m_1}$ if

$$\operatorname{dist}(F(x), F(x')) \to 0 \text{ as } x' \to x, \tag{1.12}$$

and it is called continuous at $x \in \mathbb{R}^{m_1}$ whenever

$$\mathrm{dist}_H(F(x), F(x')) \to 0 \text{ as } x' \to x. \tag{1.13}$$

As usual, F is called usc, lsc, or continuous whenever it has this property at every point $x \in \mathbb{R}^{m_1}$.

In the case $m = 1$, the relationship between single-valued and set-valued upper and lower semicontinuity can be visualized in an aesthetic way: Let a single-valued function $f : \mathbb{R} \to \mathbb{R}$ be given, and define a set-valued map $F : \mathbb{R} \rightrightarrows \mathbb{R}$ by $F(x) := (-\infty, f(x)]$. If f is usc (resp. lsc), then so is F.

A more elaborate statement about this relationship has been given in [3], Corollary 1.4.17:

Proposition 4. *If a set-valued mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is lsc (resp. usc with compact values), then the distance function $(x, y) \mapsto \mathrm{dist}(y, F(x))$ is usc (resp. lsc).*

It is in general false that the intersection of two continuous set-valued mappings is continuous. For usc mappings, however, there is the following result, see Theorem 1.1.1 in [2].

Theorem 5. *Let $F, G : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ be set-valued mappings such that $F(x) \cap G(x) \neq \emptyset$ for all $x \in \mathbb{R}^{m_1}$. If F is usc at x_0, $F(x_0)$ is compact, and the graph of G is closed, then the set-valued mapping $x \mapsto F(x) \cap G(x)$ is usc at x_0.*

Definition 6. *A set-valued mapping $F : \mathbb{R}^m \to \mathcal{C}(\mathbb{R}^m)$ is called one-sided Lipschitz (OSL), if there exists a constant $\mu \in \mathbb{R}$ such that for every $x, x' \in \mathbb{R}^m$, $y \in F(x)$, and $y' \in F(x')$*

$$\langle y - y', x - x' \rangle \leq \mu |x - x'|^2 \tag{1.14}$$

holds. It is called relaxed one-sided Lipschitz (ROSL) if for every $x, x' \in \mathbb{R}^m$ and $y \in F(x)$, there exists some $y' \in F(x')$ such that (1.14) holds. In both cases, μ is called the one-sided Lipschitz constant of F.

If $F : \mathbb{R}^{m_1} \to \mathcal{C}(\mathbb{R}^{m_2})$ satisties

$$\mathrm{dist}_H(F(x), F(x')) \leq L|x - x'| \ \forall x, x' \in \mathbb{R}^m, \tag{1.15}$$

with some fixed $L \geq 0$ then F is called Lipschitz continuous with Lipschitz constant L.

Lipschitz continuity implies continuity, which in turn implies upper and lower semicontinuity. Furthermore, relaxed one-sided Lipschitz continuity generalizes both, one-sided Lipschitz continuity and Lipschitz continuity.

The concept of one-sided Lipschitz continuity is quite rigid. It is closely related to the notion of monotone multifunctions which are single-valued almost everywhere. A simple application of the Gronwall Lemma shows that differential inclusions with a OSL right hand side have at most one solution, while ROSL right hand sides generically allow the existence of many solutions.

In spite of their names, one-sided Lipschitz continuity and relaxed one-sided Lipschitz continuity are monotonicity rather than continuity concepts. The single-valued real function $x \mapsto -\operatorname{sign}(x)$ is OSL and ROSL, but discontinuous at zero.

1.3 Tangent cones and set-valued derivatives

There exists a whole *ménagerie* (see [3], Chapter 4.5.4) of tangent cones which were defined for a variety of purposes. In this text, only the contingent cone and the corresponding derivative will be presented.

Definition 7. *Let $K \subset \mathbb{R}^m$ be an arbitrary subset, and let $x \in \overline{K}$. Then the contingent cone at x to K is defined by*

$$T_K(x) := \{v \in \mathbb{R}^m : \liminf_{h \to 0^+} h^{-1} \operatorname{dist}(x + hv, K) = 0\}. \tag{1.16}$$

The contingent cone is closely related to the classical subtangent condition, and its use will be discussed in the context of viability theory in Section 2.6.

Definition 8. *Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be a set-valued mapping, and let $(x, y) \in$ graph(F). Then the contingent differential of F at (x, y) is defined by*

$$\operatorname{graph}(DF(x, y)) := T_{\operatorname{graph}(F)}(x, y). \tag{1.17}$$

The set-valued differential generalizes the single-valued one from a geometric point of view: The single-valued derivative $Df(x)$ is a linear mapping which describes the tangent space of the graph of the original function, regarded as a manifold, at $(x, f(x))$. Following this concept, the set-valued derivative at $(x, y) \in \text{graph}(F)$ is a cone containing all vectors which are tangent to $\text{graph}(F)$ at (x, y).

The disadvantage of this notion is that it does not generalize the single-valued differential in the sense that it provides an $o(h)$ approximation of the original function. As it is a function of the independent *and* the dependent variable, it is impossible to formulate an analog of the fundamental theorem of calculus.

1.4 Fixed point theorems

In the single-valued case, an element $x \in \mathbb{R}^m$ is called a fixed point of a mapping $f : \mathbb{R}^m \to \mathbb{R}^m$ if $f(x) = x$. This notion is too restrictive in the multivalued context.

Definition 9. *Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. An element $x \in \mathbb{R}^m$ such that $x \in F(x)$ is called a fixed point of F.*

This more general setting preserves the dynamical features of the single-valued case: If $x \in F(x)$ is a fixed-point, then the constant sequence $\{x\}_{k\in\mathbb{Z}}$ is an orbit of the discrete time dynamical system induced by F. Furthermore, $x \in F(x)$ is equivalent to $0 \in (\text{id} - F)(x)$, which means that the existence of a fixed point is still equivalent to the existence of a zero for set-valued mappings.

Please note that the definitions (1.4) and (1.5) of the Hausdorff distances can be extended to arbitrary subsets of metric spaces with the drawback that distances may become infinite. The notions of continuity are generalized accordingly. The symbols $\mathcal{A}(M)$, $\mathcal{C}(M)$, and $\mathcal{CC}(M)$ denote the collections of the nonempty closed, the nonempty compact, and the nonempty convex and compact subsets of a given subset M of a topological vector space X.

The Kakutani Fixed Point Theorem is probably the most popular fixed point theorem for set-valued mappings. A detailed proof for the finite dimensional case is given in [3].

Theorem 10 (Kakutani)**.** *Let X be a locally convex topological vector space and $M \in \mathcal{CC}(X)$ be nonempty. Then any usc set-valued mapping $F : M \to \mathcal{CC}(M)$ has a fixed point.*

The Tikhonov-Schauder Fixed Point Theorem is the single-valued version of Kakutani's theorem.

Theorem 11 (Tikhonov-Schauder)**.** *Let X be a locally convex topological vector space and $M \in \mathcal{CC}(X)$ be nonempty. Then any continuous function $f : M \to M$ has a fixed point.*

The set-valued analog of the contraction mapping principle is Nadler's Theorem, cp. [45].

Theorem 12 (Nadler)**.** *Let (X, d) be a complete metric space, $M \in \mathcal{A}(X)$ be nonempty, and $F : M \to \mathcal{A}(M)$ be a set-valued mapping such that*

$$\mathrm{dist}_H(F(x), F(x')) \leq \lambda d(x, x') \tag{1.18}$$

for all $x, x' \in M$ and a fixed $\lambda \in [0, 1)$. Then F has a fixed point in M.

The Frigon-Granas Fixed Point Theorem, which is a strengthened version of Nadler's Theorem, has been given in [18].

Theorem 13 (Frigon-Granas)**.** *Let (X, d) be a complete metric space, let $x \in X$, $r > 0$, and $\lambda \in [0, 1)$. If $F : B_r(x) \rightrightarrows X$ is a set-valued mapping with closed and bounded values such that*

$$\mathrm{dist}_H(F(x'), F(x'')) \leq \lambda d(x', x'') \tag{1.19}$$

for all $x', x'' \in B_r(x)$ and

$$\mathrm{dist}(x, F(x)) \leq (1 - \lambda)r, \tag{1.20}$$

then F has a fixed point in $B_r(x)$.

1.5 Measurability and integration

Definition 14. *A mapping $F : \mathbb{R}^{m_1} \to \mathcal{A}(\mathbb{R}^{m_2})$ is called measurable if the inverse image $F^{-1}(B)$ of every open subset $B \subset \mathbb{R}^{m_2}$ is Borel measurable.*

Please note that the definition of the inverse image does not preserve complements: For a single-valued $f : \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$ and any $B \subset \mathbb{R}^{m_2}$,

$$f^{-1}(B^c) = (f^{-1}(B))^c$$

holds, whereas for set-valued $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$,

$$F^{-1}(B^c) = (F^{-1}(B))^c$$

is in general false. For this reason, set-valued measurability is not connatural to the conventional notion in the single-valued case.

A highlight from the calculus of measurable maps is the Inverse Intersection Lemma (cf. [3], Theorem 8.2.9).

Lemma 15 (Inverse Intersection)**.** *Let $F : \mathbb{R}^{m_0} \to \mathcal{A}(\mathbb{R}^{m_1})$ and $G : \mathbb{R}^{m_0} \to \mathcal{A}(\mathbb{R}^{m_2})$ be measurable mappings, and let $f : \mathbb{R}^{m_0} \times \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$ be a Caratheodory map. Then the set-valued mapping H defined by*

$$H(x) := \{y \in F(x) : f(x, y) \in G(x)\} \tag{1.21}$$

is measurable, and there exists a measurable selection of H whenever $H(x) \neq \emptyset$ for all $x \in \mathbb{R}^{m_0}$.

If $m_1 = m_2$ and $f(x, y) = y$ for all $x, y \in \mathbb{R}^{m_0}$, then $H(x) = F(x) \cap G(x)$. Thus the Inverse Intersection Lemma implies that the intersection of two measurable mappings is measurable.

Definition 16. *A set-valued map $F : \mathbb{R}^{m_1} \to \mathcal{A}(\mathbb{R}^{m_2})$ is integrably bounded if there exists a nonnegative function $k \in L^1(\mathbb{R}^{m_1})$ such that $F(x) \subset B_{k(x)}(0)$ for almost every $x \in \mathbb{R}^{m_1}$.*

If a set-valued mapping is integrably bounded then every measurable selection is integrable by Lebesgue's Theorem.

Definition 17. *The Aumann integral of an integrably bounded set-valued map $F : \mathbb{R}^{m_1} \to \mathcal{A}(\mathbb{R}^{m_2})$ is defined as the set of integrals*

$$\int_{\mathbb{R}^{m_1}} F(x)dx := \left\{ \int_{\mathbb{R}^{m_1}} f(x)dx : f \text{ is a measurable selection of } F \right\}. \tag{1.22}$$

Some important features of the Aumann integral are listed in the following theorem. The corresponding proofs are given in Chapter 8.6 of [3].

Theorem 18. *Let $F : \mathbb{R}^{m_1} \to \mathcal{A}(\mathbb{R}^{m_2})$ be a measurable and integrably bounded mapping. Then*

1. $\int_{\mathbb{R}^{m_1}} F(x)dx \in \mathcal{CC}(\mathbb{R}^{m_2})$.

2. $\forall p \in \mathbb{R}^{m_2}$, $\sigma(p, \int_{\mathbb{R}^{m_1}} F(x)dx) = \int_{\mathbb{R}^{m_1}} \sigma(p, F(x))dx$.

3. *If for some* $y \in \int_{\mathbb{R}^{m_1}} F(x)dx$ *and* $p \in \mathbb{R}^{m_2}$ *with* $|p| = 1$

$$\langle p, y \rangle = \sigma(p, \int_{\mathbb{R}^{m_1}} F(x)dx)$$

   holds, then every measurable selection f *of* F *with* $y = \int_{\mathbb{R}^{m_1}} f(x)dx$ *satisfies*

$$\langle p, f(x) \rangle = \sigma(p, F(x)) \text{ almost everywhere.}$$

Chapter 2

Differential inclusions

An ordinary differential inclusion (ODI) is a set-valued generalization of an ordinary differential equation (ODE). It is usually given by an inclusion of the form

$$\dot{x}(t) \in F(t, x(t)) \text{ almost everywhere,} \tag{2.1}$$

where F is a set valued mapping and $x(\cdot)$ is required to be absolutely continuous.

The concept of ODIs allows a rigorous treatment of ODEs with discontinuous right hand sides, which can be embedded into the class of upper semicontinuous set-valued mappings in a natural way, cf. Chapter 2.1 in [2] or [26].

The right hand side of a continuous-time control system

$$\dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U \tag{2.2}$$

is usually smooth in the second and third argument (cf. [38, 43]), and it can be interpreted as an ODI of type (2.1) by the embedding

$$F(t, x) := \{f(t, x, u) : u \in U\}, \tag{2.3}$$

see Chapter 10 in [3] for an overview. The respective set-valued right hand sides are generically locally Lipschitz continuous and almost everywhere multivalued.

Fundamental results have been published in [2], whereas advanced existence theory for the Caratheodory case and infinite dimensional state spaces can be found in [10]. The monograph [37] is a comprehensible text, which avoids technical difficulties and provides an overview of modern concepts such as optimality and stabilization. A whole book has been dedicated to the important aspect of viability theory, see [1], and two more volumes dealing with applications will follow.

2.1 Absolutely continuous functions

A short, but readable introduction to the matter can be found in [42], Chapter 9.22. The proofs are based on a careful study of the analytical features of absolutely continuous functions.

Definition 19. *Let $J \in \mathbb{R}$ be an interval. A function $f : J \to \mathbb{R}$ is called absolutely continuous if for every $\varepsilon > 0$ there exists some $\delta > 0$ such that for any finite system $((\alpha_i, \beta_i))_{i=1,\dots,p}$ of pairwise disjoint open subintervals of J with $\sum_{i=1}^{p}(\beta_i - \alpha_i) < \delta$ inequality*

$$\sum_{i=1}^{p} |f(\beta_i) - f(\alpha_i)| < \varepsilon \tag{2.4}$$

holds.

Lipschitz continuity implies absolute continuity, which in turn implies uniform continuity.

For arbitrary J, the collection $AC(J)$ of all absolutely continuous functions on J is a vector space. If J is compact, the product of two functions $f, g \in AC(J)$ is again absolutely continuous.

Definition 20. *A function $f : J \to \mathbb{R}^m$ is called absolutely continuous if every component of f is absolutely continuous.*

The absolutely continuous functions are the most general class of functions for which the Fundamental Theorem of Calculus holds true:

Theorem 21 (Fundamental Theorem)**.** *Let $J := [a, b]$ be an interval. A function $f \in AC(J)$ is differentiable almost everywhere in J with derivative $f' \in L^1(J)$, and*

$$f(b) - f(a) = \int_a^b f'(t)dt \tag{2.5}$$

holds.

Conversely, if $\varphi \in L^1(J)$, then $f(t) := \int_a^t \varphi(s)ds$ is absolutely continuous and satisfies $f' = \varphi$ almost everywhere in J.

Thus all techniques for real-valued differentiable functions which are derived from the Fundamental Theorem such as integration by parts and transformation can be applied to absolutely continuous functions.

Another important consequence of the Fundamental Theorem is the Gronwall Lemma. It is difficult to find an explicit proof for the AC case in the literature, cf. [2], [37], or [41].

Theorem 22 (Gronwall Lemma)**.** *Let $\alpha(\cdot), \beta(\cdot) \in L^1([0,T])$.*

(i) If $\varphi(\cdot) \in AC([0,T])$ satisfies

$$\dot{\varphi}(t) \leq \alpha(t) + \beta(t)\varphi(t) \ \text{a.e.},$$

then

$$\varphi(t) \leq \varphi(0)e^{\int_0^t \beta(\tau)d\tau} + \int_0^t \alpha(s)e^{\int_s^t \beta(\tau)d\tau}ds$$

for all $t \in [0,T]$.

(ii) If $\varphi(\cdot) \in L^1([0,T])$ satisfies

$$\varphi(t) \leq \alpha(t) + \int_0^t \beta(s)\varphi(s)ds \ \text{a.e.},$$

where in addition $\beta(\cdot) \in L^\infty([0,T])$ with $\beta(\cdot) \geq 0$, then

$$\varphi(t) \leq \alpha(t) + \int_0^t \alpha(s)\beta(s)e^{\int_s^t \beta(\tau)d\tau}ds$$

for almost all $t \in [0,T]$.

Proof. (i) The function $t \mapsto \int_0^t \beta(\tau)d\tau$ is absolutely continuous. Hence, $\gamma(t) := e^{-\int_0^t \beta(\tau)d\tau}$ satisfies

$$\dot{\gamma}(t) = -e^{-\int_0^t \beta(\tau)d\tau}\left(\frac{d}{dt}\int_0^t \beta(\tau)d\tau\right) = -e^{-\int_0^t \beta(\tau)d\tau}\beta(t) \text{ a.e.}$$

as a consequence of the Fundamental Theorem. Thus

$$(\gamma\varphi)' = -\beta\gamma\varphi + \gamma\dot{\varphi} \leq -\beta\gamma\varphi + \gamma\beta\varphi + \gamma\alpha = \gamma\alpha$$

holds almost everywhere. By the Fundamental Theorem and since $\gamma(0) = 1$,

$$\gamma(t)\varphi(t) \leq \varphi(0) + \int_0^t \alpha(s)\gamma(s)gs$$

for all $t \in [0, T]$. Multiplication of both sides with $\gamma(t)^{-1}$ yields the result.

(ii) The function $v(t) := \int_0^t \beta(s)\varphi(s)ds$ is absolutely continuous. By the Fundamental Theorem and by assumption,

$$\dot{v}(t) = \beta(t)\varphi(t) \leq \beta(t)\alpha(t) + \beta(t)v(t) \text{ a.e..}$$

By part (i),

$$v(t) \leq \int_0^t \alpha(s)\beta(s)e^{\int_s^t \beta(\tau)d\tau}ds,$$

which implies the desired result.

□

2.2 Existence and properties of solutions

Consider the differential inclusion

$$\dot{x}(t) \in F(x(t)), \quad t \in [0, T], \tag{2.6}$$

where $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is some set-valued mapping. A solution of (2.6) with initial state $x_0 \in \mathbb{R}^m$ is a function $x(\cdot) \in AC([0, T], \mathbb{R}^m)$ such that $x(0) = x_0$ and (2.6) is satisfied almost everywhere in $[0, T]$. The set of all solutions of this initial value problem will be denoted by $\mathcal{S}(T, 0, x_0)$. The reachable set at time T is the set

$$\mathcal{R}(T, 0, x_0) := \{x(T) : x(\cdot) \in \mathcal{S}(T, 0, x_0)\}, \tag{2.7}$$

i.e. the set of all states which are attained by solutions at time T.

In this text, existence results will only be sketched for ODIs with convex-valued right-hand-sides. Proofs of the following theorems and a coverage of the non-convex case can be found in [2].

Existence theorems are usually based on selections with suitable smoothness properties. If a right hand side F admits a continuous selection f, then every solution $x(\cdot)$ of the ODE

$$\dot{x}(t) = f(x(t)), \quad t \in [0, T] \tag{2.8}$$

solves (2.6).

Please note that an arbitrary solution of (2.6) need not be induced by a selection. Solutions of ODIs with Lipschitz continuous right hand sides can intersect themselves or become constant in finite time and exhibit a much more complicated behaviour than solutions of ODEs.

If a mapping is lsc, i.e. its values cannot collapse instantaneously, it is relatively easy to prove the existence of a continuous selection:

Theorem 23 (Michael's Selection Theorem)**.** *Let $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ be lsc with closed convex values. Then there exists a continuous selection f of F.*

The continuous selection gives rise to a solution of (2.6):

Theorem 24. *Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be lsc with closed convex values, and let any initial value $x_0 \in \mathbb{R}^m$ be given. Then there exist an interval $J := (t_-, t_+)$ with $t_- < 0 < t_+$ and a continuously differentiable solution $x(\cdot) : J \to \mathbb{R}^m$ of (2.6) satisfying $x(0) = x_0$. Moreover, either $t_+ = \infty$ or $\lim_{t \to t_+} x(t) = \infty$, and analogously for t_-.*

If a convex-valued set-valued mapping is continuous, the projection of zero to F is single-valued and continuous:

Definition 25. *Let $F : \mathbb{R}^{m_1} \to \mathcal{CC}(\mathbb{R}^{m_2})$ be a set-valued mapping. Then the selection $m_F(\cdot)$ given by $m_F(x) := \mathrm{Proj}(0, F(x))$ is called the minimal selection.*

Theorem 26 (Minimal Selection)**.** *Let $F : \mathbb{R}^{m_1} \rightrightarrows \mathbb{R}^{m_2}$ be continuous with closed convex values. Then the minimal selection $x \mapsto m_F(x)$ is single-valued and continuous.*

An analog of Theorem 24 holds for solutions induced by the minimal selection:

Theorem 27. *Let* $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ *be continuous with closed convex values, and let any initial value* $x_0 \in \mathbb{R}^m$ *be given. Then there exist an interval* $J := (t_-, t_+)$ *with* $t_- < 0 < t_+$ *and a continuously differentiable solution* $x(\cdot) : J \to \mathbb{R}^m$ *of (2.6) satisfying* $x(0) = x_0$ *and* $\dot{x}(t) = m(F(x(t)))$*. Moreover, either* $t_+ = \infty$ *or* $\lim_{t\to t_+} x(t) = \infty$*, and analogously for* t_-*.*

Existence results for upper semicontinuous right hand sides are more difficult to prove, because usc set-valued mappings need not possess any continuous selections. Thus most proofs are based on approximate selections and their corresponding solutions. The following theorem from [10] is weaker than the corresponding result from [2], but its assumptions can be verified more easily.

Theorem 28. *Let* $F : [0,T] \times \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ *be usc with closed convex values such that*

$$\|F(t,x)\| \leq k(t)(1 + |x|) \tag{2.9}$$

holds for all $t \in [0,T]$ *and* $x \in \mathbb{R}^m$*, where* $k(\cdot) \in L^1([0,T])$*. Then there exists an absolutely continuous solution of the initial value problem*

$$\dot{x}(t) \in F(t, x(t)), \quad x(0) = x_0$$

for any $x_0 \in \mathbb{R}^m$*.*

The Caratheodory case is covered in [10]:

Theorem 29. *Let* $F : [0,T] \times \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ *be a set-valued mapping with closed and convex images such that* $F(t,\cdot)$ *is usc,* $F(\cdot,x)$ *is measurable, and the growth condition*

$$\|F(t,x)\| \leq k(t)(1 + |x|) \ \forall t \in [0,T], x \in \mathbb{R}^m \tag{2.10}$$

is satisfied for some $k(\cdot) \in L^1([0,T])$*. Then there exists an absolutely continuous solution of the initial value problem*

$$\dot{x}(t) \in F(t, x(t)), \quad x(0) = x_0$$

for any $x_0 \in \mathbb{R}^m$*.*

Clearly, Theorem 28 is a consequence of Theorem 29. It is displayed nevertheless, because it can be proved by relatively simple techniques, while the treatment of the Caratheodory case requires some deeper results about measurable functions.

The following important statement about the behaviour of solutions is taken from [10], Theorem 7.1. In [2], Theorem 2.2.1, upper semicontinuity of the mapping $x \mapsto \mathcal{S}(T,0,x)$ is proved under a so-called boundedness assumption.

Theorem 30. *Under the assumptions of Theorem 29, the mapping* $x \mapsto \mathcal{R}(T,0,x)$ *is upper semicontinuous with compact values.*

In contrast to the autonomous case, where lower semicontinuity is a favourable property, it is impossible to prove a general existence result for lsc right hand sides in the Caratheodory situation, see Example 6.2 in [10].

2.3 Filippov Theorem and Relaxation Theorem

The Filippov Theorem is a central result for differential inclusions, because it is an existence and a stability theorem at the same time. It has important consequences for the error analysis of numerical schemes for ODIs. For a proof consider [2] or the original publication [16].

Theorem 31 (Filippov)**.** *Let* $y \in AC([0,T],\mathbb{R}^m)$ *and a constant* $\beta > 0$ *be given and denote* $Q := \{(t,x) \in \mathbb{R} \times \mathbb{R}^m : |x - y(t)| \leq \beta\}$*. Let* $F : Q \to \mathcal{A}(\mathbb{R}^m)$ *be continuous and such that*

$$\operatorname{dist}_H(F(t,x), F(t,x')) \leq k(t)|x - x'| \tag{2.11}$$

for some $k(\cdot) \in L^1([0,T])$*. Assume moreover that*

$$\delta := |y(0) - x_0| \leq \beta \text{ and } \operatorname{dist}(\dot{y}(t), F(t,y(t))) \leq p(t) \text{ a.e.}$$

for some $p \in L^1([0,T])$*. Define*

$$\xi(t) := \delta e^{\int_0^t k(\tau)d\tau} + \int_0^t e^{\int_s^t k(\tau)d\tau} p(s)ds$$

and let $t_+ > 0$ *be such that* $\xi(t_+) \leq \beta$*. Then there exists a solution* $x(\cdot) : [0,t_+] \to \mathbb{R}^m$ *of the ODI*

$$\dot{x}(t) \in F(t,x(t)), \quad x(0) = x_0 \tag{2.12}$$

such that

$$|x(t) - y(t)| \le \xi(t) \ \forall t \in [0, t_+] \tag{2.13}$$

and

$$|\dot{x}(t) - \dot{y}(t)| \le k(t)\xi(t) + p(t) \ \ a.e. \ in \ [0, t_+]. \tag{2.14}$$

For autonomous ODIs, the Filippov Theorem can be reduced to the following rule of thumb: If F is Lipschitz continuous with Lipschitz constant L, then the initial error δ is propagated with a factor e^{LT}, whereas the defect $p(\cdot)$ causes an additional error of size $\int_0^T e^{L(t-s)} p(s) ds$.

The Filippov Theorem has been generalized to a setting, where the Lipschitz-like condition (2.11) is replaced by relaxed one-sided Lipschitz continuity, see [13].

The Relaxation Theorem states that the solutions of an ODI with a Lipschitz continuous right hand side are dense in the set of solutions of the convexified problem. As a consequence, it is usually sufficient to consider convex-valued multifunctions which are easier to handle.

Theorem 32 (Relaxation Theorem)**.** *Let $F : \mathbb{R}^m \to \mathcal{C}(\mathbb{R}^m)$ be Lipschitz continuous and let a solution $x : [-T, T] \to \mathbb{R}^m$ of the ODI*

$$\dot{x}(t) \in \overline{\text{co}} F(x(t)), \quad x(0) = x_0 \tag{2.15}$$

and $\varepsilon > 0$ be given. Then there exists a solution $y : [-T, T] \to \mathbb{R}^m$ of the ODI

$$\dot{y}(t) \in F(x(t)), \quad y(0) = x_0 \tag{2.16}$$

such that $|y(t) - x(t)| \le \varepsilon$ for all $t \in [-T, T]$.

Inclusion (2.15) is called the relaxed version of the original problem (2.16). For a proof, see [2].

2.4 Density theorems

The first density theorem is given in the amazing paper [16] by Filippov:

Definition 33. *A subset $A \subset \mathbb{R}^m$ is called uniformly locally connected if there exists a function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\lim_{s \to 0} \eta(s) = 0$ and such that any two points $y, y' \in A$ with $|y - y'| \le s$ can be joined by a connected set $B \subset A$ with $\text{diam}(B) \le \eta(s)$.*

Theorem 34. *Let $F : [0,T] \times \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ be continuous with uniformly locally connected images such that*

$$\mathrm{dist}_H(F(t,x), F(t,x')) \leq L|x - x'|$$

for some $L \geq 0$ and all $t \in [0,T]$ and $x, x' \in \mathbb{R}^m$. If $x(\cdot)$ is a solution of

$$\dot{x}(t) \in F(t, x(t)), \quad x(0) = x_0, \tag{2.17}$$

and $\varepsilon > 0$ is given, then there exists a continuously differentiable solution $\bar{x}(\cdot)$ of (2.17) with $|x(t) - \bar{x}(t)| < \varepsilon$ for all $t \in [0,T]$ which in addition satisfies $\dot{\bar{x}}(0) = v_0$, where $v_0 \in F(0, x_0)$ is arbitrary.

The proof is based on a skillfully performed construction of a sequence of absolutely continuous solutions such that a suitable measure of discontinuity of the derivatives converges to zero along the sequence.

A weaker density theorem which can be proved by standard techniques is due to Wolenski, see [44]:

Theorem 35. *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be locally Lipschitz continuous, and let $y(\cdot) \in C^1([0,T])$, $K \in \mathcal{C}(\mathbb{R}^m)$ and $\delta > 0$ be such that $B_\delta(y([0,T])) \subset K$. If $\int_0^T \mathrm{dist}(\dot{y}(t), F(y(t)))dt < \delta e^{-LT}$ for some Lipschitz constant L of F on K, then there exists a continuously differentiable solution $\bar{x}(\cdot)$ of the ODI*

$$\dot{x}(t) \in F(x(t)), \quad x(0) = y(0)$$

satisfying

$$|y(t) - \bar{x}(t)| \leq e^{LT} \int_0^T \mathrm{dist}(\dot{y}(s), F(y(s)))ds$$

for all $t \in [0,T]$.

In contrast to the approach pursued by Filippov, this proof uses a sequence of continuously differentiable approximations obtained by a modified Picard-Lindelöf iteration which converges to a solution.

The density of the C^1 solutions follows from Lusin's Theorem together with Theorem 35.

Following an alternative concept, it is possible to show under strengthened assumptions that also the infinitely many times differentiable solutions are dense:

Definition 36. *A set-valued mapping* $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ *is called* (L, δ_0)*-stout with constants* $\delta_0 > 0$ *and* $L > 0$ *if there exists a Lipschitz continuous mapping* $F_{\delta_0} : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ *with Lipschitz constant* L *such that*

$$F(x) = B(F_{\delta_0}(x), \delta_0) \ \forall x \in \mathbb{R}^m. \tag{2.18}$$

Remark 37. *Since* F_{δ_0} *is Lipschitz continuous with Lipschitz constant* L*, the mappings defined by* $x \mapsto B(F_{\delta_0}(x), \delta)$ *with* $\delta > 0$*, and in particular* F*, are Lipschitz continuous with the same constant. Obviously, the images* $B(F_{\delta_0}(x), \delta)$ *are compact and convex. Thus an* (L, δ_0)*-stout mapping is* (L, δ)*-stout for every* $\delta \in (0, \delta_0]$*, where*

$$F_\delta(x) := B(F_{\delta_0}(x), \delta_0 - \delta). \tag{2.19}$$

Please note that stoutness is closely related to smoothness properties of set-valued mappings: If a map $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ is (L, δ_0)-stout, its images cannot have any 'edges'.

The following lemma formalizes a simple geometric principle.

Lemma 38. *Let* $F_i \subset \mathbb{R}^m$*,* $i \in I$ *and* $G \subset \mathbb{R}^m$ *be closed and convex, where* I *is some index set. Then*

$$\operatorname{dist}(\operatorname{co}(\cup_{i \in I} F_i), G) \leq \sup_{i \in I} \operatorname{dist}(F_i, G). \tag{2.20}$$

Proof. Let $f \in \operatorname{co}(\cup_{i \in I} F_i)$. There exist $\lambda_0, \ldots, \lambda_m \in [0, 1]$ and $f_0, \ldots, f_m \in \cup_{i \in I} F_i$, $f_j \in F_{i_j}$ such that $f = \sum_{j=0}^m \lambda_j f_j$ and $\sum_{j=0}^m \lambda_j = 1$. Let $g_j \in G$ be such that

$$|f_j - g_j| = \operatorname{dist}(f_j, G) \leq \operatorname{dist}(F_{i_j}, G).$$

Then $g := \sum_{j=0}^m \lambda_j g_j \in G$, and

$$|f - g| \ \leq \ \sum_{j=0}^m \lambda_j |f_j - g_j| \leq \sum_{j=0}^m \lambda_j \operatorname{dist}(F_{i_j}, G) \leq \sup_{i \in I} \operatorname{dist}(F_i, G).$$

□

The proof of the following C^∞ Density Theorem follows the tradition of proving existence results by considering selections. The problem is that an arbitrary solution need not be induced by any selection. This difficulty is overcome by a suitable non-autonomous reformulation which admits a smooth selection close to the derivative of the original absolutely continuous solution.

Theorem 39 (C^∞ Density Theorem)**.** *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be (L, δ_0)-stout with $\delta_0 \in (0,1]$. Then the infinitely many times differentiable solutions of the initial value problem*

$$\dot{x}(t) \in F(x(t)) \text{ } a.e. \text{ } in \text{ } [0,T], \quad x(0) = x_0 \tag{2.21}$$

are dense in the set of all solutions with respect to the maximum norm.

Proof. Let $x(\cdot)$ be a solution of (2.21). We will construct smooth solutions $a_\delta(\cdot)$ arbitrarily close to $x(\cdot)$.

Step 1: A-priori bounds. The solution x is à-priori bounded: Let $z(s) \in F(x(0))$ such that $|\dot{x}(s) - z(s)| = \text{dist}(\dot{x}(s), F(x(0)))$. Then

$$\begin{aligned}
|x(t) - x(0)| &\leq \int_0^t |\dot{x}(s)| ds \\
&\leq \int_0^t |\dot{x}(s) - z(s)| + |z(s)| ds \\
&\leq \int_0^t \text{dist}_H(F(x(s)), F(x(0))) + ||F(x(0))|| ds \\
&\leq t||F(x(0))|| + \int_0^t L|x(s) - x(0)| ds,
\end{aligned}$$

and by the Gronwall lemma,

$$\begin{aligned}
|x(t) - x(0)| &\leq t||F(x(0))|| + \int_0^t s||F(x(0))|| L e^{L(t-s)} ds \\
&= t||F(x(0))|| + \frac{1}{L}||F(x(0))||(e^{Lt} - Lt - 1) \\
&= \frac{1}{L}||F(x(0))||(e^{Lt} - 1).
\end{aligned}$$

In particular,

$$\begin{aligned} |x(t+\eta)-x(t)| &\le \frac{1}{L}||F(x(t))||(e^{L\eta}-1) \\ &\le \frac{1}{L}\big[||F(x(0))||+\mathrm{dist}_H(F(x(t)),F(x(0)))\big](e^{L\eta}-1) \\ &\le \frac{1}{L}\big[||F(x(0))||+L|x(t)-x(0)|\big](e^{L\eta}-1) \\ &\le \frac{1}{L}\big[||F(x(0))||+||F(x(0))||(e^{Lt}-1)\big](e^{L\eta}-1) \\ &\le \underbrace{\frac{1}{L}||F(x(0))||e^{LT}}_{=:C_1}(e^{L\eta}-1). \qquad (2.22) \end{aligned}$$

Step 2: Regular approximation. Now we construct a regular approximation x_δ of x. Without loss of generality we can assume that

$$\dot{x}(t) \in F(x(t)) \;\forall t \in [0,T]$$

as a function. We formally continue it as a function $\dot{x} \in L^1_{loc}(\mathbb{R},\mathbb{R}^m)$ by setting

$$\dot{x}(t) := \begin{cases} \dot{x}(T), & T < t \\ \dot{x}(t), & 0 < t \le T \\ \dot{x}(0), & t \le 0. \end{cases} \qquad (2.23)$$

For given $\delta \in (0,\delta_0]$, there exists a function $\varphi_\delta \in C_0^\infty(\mathbb{R},\mathbb{R}_+)$ satisfying $\mathrm{supp}(\varphi_\delta) \subset [-\delta,\delta]$ and $\int_{\mathbb{R}} \varphi_\delta(\tau)d\tau = 1$ such that

$$y_\delta(s) := \int_{\mathbb{R}} \varphi_\delta(\tau)\dot{x}(s-\tau)d\tau$$

is a function $y_\delta \in C^\infty(\mathbb{R},\mathbb{R}^m)$ (see Theorem 2.16 in [28]) with

$$\int_0^T |y_\delta(s)-\dot{x}(s)|ds \le \delta.$$

Hence $x_\delta \in C^\infty(\mathbb{R},\mathbb{R}^m)$ given by

$$x_\delta(t) := x(0) + \int_0^t y_\delta(s)ds$$

satisfies

$$|x_\delta(t) - x(t)| \le \delta \ \forall t \in [0,T].$$

Note that y_δ is Lipschitz continuous in $[-1, T+1]$ with Lipschitz constant $K_\delta > 0$.

Step 3: Construction of a regular selection. Consider the time dependent mappings

$$\tilde{F} : \mathbb{R} \times \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m), \tilde{F}(t,x) := F(x) - y_\delta(t) \tag{2.24}$$

and

$$\tilde{F}_\delta : \mathbb{R} \times \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m), \tilde{F}_\delta(t,x) := F_\delta(x) - y_\delta(t), \tag{2.25}$$

where the F_δ is the δ-retract of F defined in (2.19). Since $\tilde{F}_\delta(t,x)$ is Lipschitz continuous in t and x, its minimal selection

$$(t,x) \mapsto m(\tilde{F}_\delta(t,x))$$

is continuous by theorem 1.7.1 in [2]. Take a $\psi_\delta \in C_0^\infty(\mathbb{R} \times \mathbb{R}^m, \mathbb{R}_+)$ with $\mathrm{supp}(\psi_\delta) \subset B(0, \frac{\delta}{2K_\delta}) \times B(0, \frac{\delta}{2L})$ and $\int_{\mathbb{R}\times\mathbb{R}^m} \psi_\delta(t,x) d(t,x) = 1$, so that the function

$$\tilde{m}(t,x) := \int_{\mathbb{R}\times\mathbb{R}^m} \psi_\delta(\theta,\xi)\, m(\tilde{F}_\delta(t-\theta, x-\xi))\, d(\theta,\xi) \tag{2.26}$$

is an element of $C^\infty(\mathbb{R} \times \mathbb{R}^m, \mathbb{R}^m)$. According to Theorem 1.6.13 in [43], it satisfies

$$\begin{aligned}
\tilde{m}(t,x) &\in \overline{\mathrm{co}}\{m(\tilde{F}_\delta(t-\theta, x-\xi)) : (\theta,\xi) \in \mathrm{supp}(\psi_\delta)\} \\
&\subset \overline{\mathrm{co}}(\tilde{F}_\delta(B(t, \frac{\delta}{2K_\delta}), B(x, \frac{\delta}{2L}))) \\
&\subset \overline{\mathrm{co}}(B(\tilde{F}_\delta(B(t, \frac{\delta}{2K_\delta}, x), \frac{\delta}{2}))) \\
&\subset \overline{\mathrm{co}}(B(\tilde{F}_\delta(t,x), \delta)) = \tilde{F}(t,x)
\end{aligned}$$

for $t \in [0,T]$ and $x \in \mathbb{R}^m$, which implies

$$y_\delta(t) + \tilde{m}(t,x) \in F(x). \tag{2.27}$$

On the other hand,

$$
\begin{aligned}
|\tilde{m}(t,x)| &\leq \|\overline{\text{co}}\{m(\tilde{F}_\delta(t-\theta,x-\xi)) : (\theta,\xi) \in \text{supp}(\psi_\delta)\}\| \\
&= \|\{m(\tilde{F}_\delta(t-\theta,x-\xi)) : (\theta,\xi) \in \text{supp}(\psi_\delta)\}\| \\
&\leq \sup\{\text{dist}(0,\tilde{F}_\delta(\theta,\xi)) : \theta \in B(t,\frac{\delta}{2K_\delta}), \xi \in B(x,\frac{\delta}{2L})\} \\
&\leq \sup\{\text{dist}(0,\tilde{F}(\theta,\xi)) : \theta \in B(t,\frac{\delta}{2K_\delta}), \xi \in B(x,\frac{\delta}{2L})\} + \delta \\
&\leq \sup\{\text{dist}(0,\tilde{F}(t,x)) + \text{dist}(\tilde{F}(t,x),\tilde{F}(\theta,x)) \\
&+ \text{dist}(\tilde{F}(\theta,x),\tilde{F}(\theta,\xi)) : \theta \in B(t,\frac{\delta}{2K_\delta}), \xi \in B(x,\frac{\delta}{2L})\} + \delta \\
&\leq \text{dist}(0,\tilde{F}(t,x)) + 2\delta \quad = \quad \text{dist}(y_\delta(t),F(x)) + 2\delta.
\end{aligned}
$$

Step 4: Corresponding solution. By the Cauchy-Peano theorem, the initial value problem

$$\dot{a}_\delta(t) = y_\delta(t) + \tilde{m}(t,a_\delta(t)), \quad a_\delta(0) = x(0) \tag{2.28}$$

admits a solution $a_\delta(\cdot)$ on a maximal subinterval $J \subset [0,T]$ with $0 \in J$. It is an element of $C^\infty(J,\mathbb{R}^m)$, and because of (2.27) it is also a solution of the original differential inclusion (2.21). For $t \in J$ one obtains

$$
\begin{aligned}
|x_\delta(t) - a_\delta(t)| &\leq \int_0^t |y_\delta(s) - (y_\delta(s) + \tilde{m}(s,a_\delta(s)))| ds \\
&= \int_0^t |\tilde{m}(s,a_\delta(s))| ds \\
&= \int_0^t \text{dist}(y_\delta(s),F(a_\delta(s))) + 2\delta \, ds.
\end{aligned}
$$

By Theorem 1.6.13 in [43],

$$y_\delta(s) \in \overline{co}\{\dot{x}(\tau) : \tau \in s - \text{supp}(\varphi_\delta)\}. \tag{2.29}$$

Hence

$$\begin{aligned}
|x_\delta(t)-a_\delta(t)| &\leq \int_0^t \operatorname{dist}(\overline{co}\{\cup_{s-\operatorname{supp}(\varphi_\delta)}F(x(\tau))\}, F(a_\delta(s))) + 2\delta\, ds \\
&\overset{(2.20)}{\leq} \int_0^t \sup_{\tau\in[s-\delta,s+\delta]} \operatorname{dist}(F(x(\tau)), F(a_\delta(s))) + 2\delta\, ds \\
&\leq \int_0^t \sup_{\tau\in[s-\delta,s+\delta]} L|x(\tau)-a_\delta(s)| + 2\delta\, ds \\
&\overset{(2.22)}{\leq} \int_0^t L(|x(s)-a_\delta(s)| + C_1(e^{L\delta}-1)) + 2\delta\, ds \\
&\leq \int_0^t L|x_\delta(s)-a_\delta(s)| + \underbrace{L(\delta + C_1(e^{L\delta}-1)) + 2\delta}_{=:C_2(\delta)}\, ds.
\end{aligned}$$

The Gronwall lemma yields

$$\begin{aligned}
|x_\delta(t)-a_\delta(t)| &\leq C_2(\delta)t + \int_0^t C_2(\delta)sLe^{L(t-s)}ds \\
&= C_2(\delta)t + \frac{1}{L}C_2(\delta)(e^{Lt}-Lt-1) \\
&= \frac{1}{L}C_2(\delta)t(e^{Lt}-1),
\end{aligned}$$

and thus

$$\begin{aligned}
|x(t)-a_\delta(t)| &\leq \delta + \frac{1}{L}C_2(\delta)t(e^{Lt}-1) \\
&\leq \delta + \frac{1}{L}C_2(\delta)T(e^{LT}-1). \qquad (2.30)
\end{aligned}$$

In particular, a_δ is bounded on J. Hence $J=[0,T]$, and

$$\begin{aligned}
||x-a_\delta||_\infty &\leq \delta + \frac{1}{L}C_2(\delta)T(e^{LT}-1) \qquad (2.31)\\
&\longrightarrow 0 \ as\ \delta\to 0.
\end{aligned}$$

□

The C^∞ Density Theorem is linked to the numerical analysis of differential inclusions. In order to design and prove convergence of higher order methods, it is necessary to specify an object with sufficient smoothness properties, e.g. a suitable subset $\tilde{\mathcal{S}}(T,0,x_0)$ of smooth elements of the set of all solutions $\mathcal{S}(T,0,x_0)$ which can be approximated efficiently.

There is a characteristic tradeoff between two errors of different nature: If the Hausdorff distance between $\mathcal{S}(T,0,x_0)$ and $\tilde{\mathcal{S}}(T,0,x_0)$ is kept small the norm of the derivatives of the approximating C^∞ solutions may become large, resulting in larger errors of the numerical scheme.

This tradeoff seems to be the fundamental dilemma in the study of higher order methods. In the well-known paper [40] of Vladimir Veliov, the impact of non-smoothness and the resulting numerical errors are unfortunately hidden in the constants. It is an absolute necessity to study this phenomenon very carefully in the future.

It is possible to give an alternative proof for the classical result of Filippov and Wolenski using the above techniques. If only C^1 solutions are constructed the minimal selection does not need to be smoothened, and thus it is not necessary to assume that the set-valued mapping F is stout.

Theorem 40. *Let $F:\mathbb{R}^m\to\mathcal{CC}(\mathbb{R}^m)$ be Lipschitz continuous. Then the continuously differentiable solutions of the initial value problem*

$$\dot{x}(t)\in F(x(t)) \text{ a.e. in } [0,T],\quad x(0)=x_0 \tag{2.32}$$

are dense in the set of all solutions with respect to the maximum norm.

Proof. The à-priori estimate and the regular approximation $x_\delta(\cdot)$ can be obtained exactly as in the previous proof. For the construction of a regular selection, we can consider the time dependent mapping

$$\tilde{F}:\mathbb{R}\times\mathbb{R}^m\to\mathcal{CC}(\mathbb{R}^m), \tilde{F}(t,x):=F(x)-y_\delta(t).$$

Since y_δ is continuous, $\tilde{F}$ is continuous w.r.t the Hausdorff metric and consequently, the minimal selection $(t,x)\mapsto m(t,x)$ of $\tilde{F}$ is also continuous. Obviously

$$|m(t,x)|=\operatorname{dist}(y_\delta(t),F(x)),$$

and

$$y_\delta(t)+m(t,x)\in F(x)\ \forall t\in[0,T],\ \forall x\in\mathbb{R}^m.$$

By the Cauchy-Peano theorem, the initial value problem

$$\dot{a}_\delta(t)=y_\delta(t)+m(t,a_\delta(t)),\quad a_\delta(0)=x(0) \tag{2.33}$$

admits a solution $a_\delta(\cdot)$ on a maximal subinterval $J \subset [0,T]$ with $0 \in J$. The following estimates are merely a simplified version of the previous calculations. Of course, the solution $a_\delta(\cdot)$ is defined on the whole interval $[0,T]$ and it is continuously differentiable, because the right hand side of (2.33) is a continuous function. □

2.5 Relaxed one-sided Lipschitz continuity

In the context of differential equations and inclusions, Lipschitz continuity can sometimes be replaced by the notion of relaxed one-sided Lipschitz continuity (ROSL), see Definition 6. Multivalued mappings with this property and the corresponding differential inclusions have been thoroughly analyzed by Tzanko Donchev, see e.g. [12], [13], and [14].

The ROSL condition is the most accurate stability concept for differential inclusions, because it imposes conditions on the right hand side only in those directions which matter. Furthermore, it is possible to define ROSL continuity with negative Lipschitz constant implying contractivity of the solution sets, see Theorems 41 and 42. Thus it generalizes the principle of eigenvalues, which is the most important stability criterion in the single-valued case.

Thanks to the Density Theorem 40, it is possible to prove a stability theorem for differential inclusions without using sophisticated results about measurable and Caratheodory mappings.

Theorem 41. *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be Lipschitz continuous with Lipschitz constant $L > 0$ and ROSL with constant $\mu \in \mathbb{R}$. Then the mapping $x_0 \mapsto \mathcal{R}(T, 0, x_0)$ from the initial states to the reachable sets of the differential inclusions*

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0 \tag{2.34}$$

for some fixed time T is Lipschitz continuous with Lipschitz constant $e^{\mu T}$.

In particular, $x_0 \mapsto \mathcal{R}(T, 0, x_0)$ is a contraction whenever $\mu < 0$.

Proof. Let $x(\cdot)$ be any solution of (2.34). Without loss of generality we can assume that $x(0) = 0$. Let $v \in \mathbb{R}^m$ be given. By Theorem 40, for every $\varepsilon > 0$ there exists a solution $x_\varepsilon \in C^1([0,T], \mathbb{R}^m)$ of (2.34) such that $x_\varepsilon(0) = 0$ and $||x_\varepsilon - x||_\infty < \varepsilon$.

The graph of the mapping

$$G(t,x) := \{y \in \mathbb{R}^m : \langle y - \dot{x}_\varepsilon(t), x - x_\varepsilon(t)\rangle \leq \mu|x - x_\varepsilon(t)|^2\} \tag{2.35}$$

is closed, because x_ε, $\dot{x}_\varepsilon$, and the inner product are continuous. As F is ROSL, the right hand side of the differential inclusion

$$\dot{y}(t) \in F(y(t)) \cap G(t, y(t)) \tag{2.36}$$

is nonempty, and it is obviously convex and compact. By Theorem 5, it is simultaneously upper semicontinuous in both arguments. Hence there exists a solution $x_{v,\varepsilon}$ of (2.36) with $x_{v,\varepsilon}(0) = v$ according to Theorem 28 (growth condition (2.9) follows from Lipschitz continuity of F).

Now

$$\begin{aligned} \frac{d}{dt}|x_\varepsilon(t) - x_{v,\varepsilon}(t)|^2 &= 2\langle \dot{x}_\varepsilon(t) - \dot{x}_{v,\varepsilon}(t), x_\varepsilon(t) - x_{v,\varepsilon}(t)\rangle \\ &\leq 2\mu|x_\varepsilon(t) - x_{v,\varepsilon}(t)|^2 \end{aligned}$$

implies that

$$|x_\varepsilon(T) - x_{v,\varepsilon}(T)| \leq |x_\varepsilon(0) - x_{v,\varepsilon}(0)|e^{\mu T} = e^{\mu T}|v|, \tag{2.37}$$

and

$$|x(T) - x_{v,\varepsilon}(T)| \leq \varepsilon + e^{\mu T}|v|. \tag{2.38}$$

Since this estimate holds for every $\varepsilon > 0$, $e^{\mu T}$ is a Lipschitz constant for the T-flow w.r.t. the Hausdorff distance. □

It is possible to replace Lipschitz continuity by upper semicontinuity, the ROSL property, and a linear growth condition and an existence theorem for the Caratheodory case. This theorem is stronger than the previous one, but it requires advanced results such as the Inverse Intersection Lemma. A similar result is presented in [13].

Theorem 42. *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be an upper semicontinuous set-valued mapping which satisfies*

$$\|F(x)\| \leq c(1 + |x|) \ \forall x \in \mathbb{R}^m, \tag{2.39}$$

for a constant $c > 0$ and the ROSL condition with constant $\mu \in \mathbb{R}$. Then the mapping $x_0 \mapsto \mathcal{R}(T, 0, x_0)$ from the initial states to the reachable sets of the differential inclusions

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0 \tag{2.40}$$

for some fixed time T is Lipschitz continuous with Lipschitz constant $e^{\mu T}$.

In particular, $x_0 \mapsto \mathcal{R}(T, 0, x_0)$ is a contraction whenever $\mu < 0$.

Proof. Let $x(\cdot)$ be any solution of (2.40). Without loss of generality we can assume that $x(0) = 0$. Let $v \in \mathbb{R}^m$ be given.

Consider the mapping

$$H(t,x) := \{y \in F(x) : \langle y - \dot{x}(t), x - x(t)\rangle \leq \mu |x - x(t)|^2\}. \tag{2.41}$$

It inherits the linear growth condition from F. The images of H are obviously convex, compact, and because of the ROSL property also non-empty. Since H is an intersection $H(t,x) = F(x) \cap G(t,x)$, where

$$G(t,x) := \{y \in \mathbb{R}^m : \langle y - \dot{x}(t), x - x(t)\rangle \leq \mu |x - x(t)|^2\}$$

and $\operatorname{graph} G(t,\cdot)$ is closed because of the continuity of the inner product, $H(t,\cdot)$ is usc by Theorem 5.

As H can also be represented as

$$H(t,x) = \{y \in F(x) : f(t,x,y) \in \tilde{G}(t,x)\},$$

where

$$f(t,x,y) := \langle y - \dot{x}(t), x - x(t)\rangle$$

is measurable in (t,x) and continuous in y and

$$\tilde{G}(t,x) := (-\infty, \mu |x - x(t)|^2]$$

is measurable in (t,x) and has closed values, the Inverse Intersection Lemma 15 guarantees that $H(\cdot,x)$ is measurable.

Thus H satisfies the assumptions of Theorem 29 and there exists a solution x_v of the initial value problem

$$\dot{x}_v(t) \in G(t, x_v(t)), \quad x_v(0) = v \tag{2.42}$$

on $[0,T]$. Now

$$\begin{aligned} \frac{d}{dt}|x(t) - x_v(t)|^2 &= 2\langle \dot{x}(t) - \dot{x}_v(t), x(t) - x_v(t)\rangle \\ &\leq 2\mu |x(t) - x_v(t)|^2 \end{aligned}$$

implies that

$$|x(T) - x_v(T)| \leq |x(0) - x_v(0)| e^{\mu T} = e^{\mu T}|v|, \tag{2.43}$$

and $e^{\mu T}$ is a Lipschitz constant for the T-flow w.r.t. the Hausdorff distance. □

Remark 43. *Please note that the ROSL condition characterizes stability in the following sense: Suppose that the assumptions of Theorem 42 hold, but that there is an open set* $U \subset \mathbb{R}^m$ *such that for any* $x, x' \in U$ *there exists some* $y \in F(x)$ *with*

$$\langle y - y', x - x' \rangle \geq \mu |x - x'|^2 \tag{2.44}$$

for any $y' \in F(x')$.

Fix $x_1, x_2 \in U$. *Because of the growth condition, there exists some* $t_+ \in (0, T]$ *such that any solutions* $x_1(\cdot)$ *and* $x_2(\cdot)$ *with initial values* $x_1(0) = x_1$ *and* $x_2(0) = x_2$ *are contained in* U *for all* $t \in [0, t_+]$. *But then the computation of the previous proof can be repeated with '*$\geq$*' instead of '*$\leq$*', implying*

$$|x_1(t) - x_2(t)| \geq e^{\mu t} |x_1 - x_2| \ \forall t \in [0, t_+] \tag{2.45}$$

so that

$$\mathrm{dist}_H(\mathcal{R}(t, 0, x_1), \mathcal{R}(t, 0, x_2)) \geq e^{\mu t} |x_1 - x_2| \ \forall t \in [0, t_+].$$

2.6 Viability theory

Consider solutions $x(\cdot)$ of the ODI

$$\dot{x}(t) \in F(x(t)) \text{ a.e.}, \tag{2.46}$$

and orbits $\{x_k\}_{k \in \mathbb{N}}$ of the discrete-time set-valued dynamical system

$$x_{k+1} \in G(x_k) \ \forall k \in \mathbb{Z}, \tag{2.47}$$

where $F, G : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ are set-valued mappings.

Definition 44. *A subset* $D \subset \mathbb{R}^m$ *is a viability domain of* (2.46)*, if for any* $x_0 \in D$ *there exists a solution* $x : [0, \infty) \to D$ *such that* $x(0) = x_0$. *The viability kernel* $\mathrm{Viab}_F(K)$ *is the largest closed viability domain of* (2.46) *contained in some set* $K \subset \mathbb{R}^m$.

A subset $D \subset \mathbb{R}^m$ *is a viability domain of* (2.47)*, if for any* $x_0 \in D$ *there exists a solution of* (2.47) *starting at* x_0 *and remaining in* D *for all time. The viability kernel* $\mathrm{Viab}_G(K)$ *is the largest closed discrete viability domain of* (2.47) *contained in some set* $K \subset \mathbb{R}^m$.

Under mild assumptions on F and G, both types of viability kernels are well-defined, compare e.g. [1].

The Viability Theorem from [3] is a set-valued version of the well-known subtangent principle:

Theorem 45 (Viability Theorem)**.** *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be usc with linear growth, and let $D \in \mathcal{A}(\mathbb{R}^m)$ be such that*

$$F(x) \cap T_D(x) \neq \emptyset \ \forall x \in D. \tag{2.48}$$

Then D is a viability domain of (2.46)*.*

If $G : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is an arbitrary mapping and

$$G(x) \cap D \neq \emptyset \ \forall x \in D, \tag{2.49}$$

then D is a viability domain of (2.47)*.*

For a detailed coverage of this topic, the reader is referred to [1].

Chapter 3

Shadowing in dynamical systems

The main aim of Shadowing Theory is the characterization of dynamical systems which are robust under uniformly small perturbations. Most shadowing theorems provide sufficient conditions for the existence of an orbit of the dynamical system close to a given faulty trajectory on the bi-infinite time interval.

Various types of shadowing properties have been stated and investigated in the literature. The monograph [31] provides an overview over this research area and its connections to the delicate subject of structural stability, while [30] presents an elaborate analysis of the classical Shadowing Lemma and the intricate geometric features of dynamical systems with hyperbolic structure.

In this text, only few concepts and facts will be displayed in order to show where the results of Section 4 are located in the landscape of Shadowing Theory. Most statements and the corresponding proofs can be found in [30].

3.1 Hyperbolic fixed points

Hyperbolicity is an essential concept in shadowing theory. Its key feature is the assumption that there exists a decomposition of the tangent space into a direct sum of two subspaces such that the linearized dynamical system contracts in forward time on one of the subspaces and in backward time on the other.

The simplest example of a hyperbolic set for a C^1 diffeomorphism, i.e.

for a bijective mapping $f : U \to f(U)$ such that $f, f^{-1} \in C^1$, is a hyperbolic fixed point.

Definition 46. *Let $U \subset \mathbb{R}^m$ be open. A point $x_0 \in U$ is said to be a hyperbolic fixed point of a C^1 diffeomorphism $f : U \to \mathbb{R}^m$ if $f(x_0) = x_0$ and the eigenvalues of $Df(x_0)$ lie off the unit circle.*

The sum of the generalized eigenspaces corresponding to the eigenvalues inside (outside) the unit circle is called the stable (unstable) subspace and is denoted by E^s (E^u).

Note that the stable and unstable subspaces are invariant under $Df(x_0)$. It is well-known that for any $\lambda_1, \lambda_2 > 0$ such that $|\lambda| < \lambda_1 < 1$ for all eigenvalues λ of $Df(x_0)$ with $|\lambda| < 1$ and $1 < \lambda_2^{-1} < |\lambda|$ for all eigenvalues λ with $|\lambda| > 1$, there exist $K_1, K_2 > 0$ such that for all $k \geq 0$

$$|[Df(x_0)]^k \xi| \leq K_1 \lambda_1^k |\xi| \text{ for } \xi \in E^s \tag{3.1}$$

and

$$|[Df(x_0)]^{-k} \xi| \leq K_2 \lambda_2^k |\xi| \text{ for } \xi \in E^u. \tag{3.2}$$

Thus the behaviour of the linearized system is characterized by

$$[Df(x_0)]^k \xi \to 0 \text{ as } k \to \infty \text{ if and only if } \xi \in E^s$$

and

$$[Df(x_0)]^{-k} \xi \to 0 \text{ as } k \to \infty \text{ if and only if } \xi \in E^u.$$

This motivates the following definition for the original nonlinear system:

Definition 47. *Let x_0 be a hyperbolic fixed point of the C^1 diffeomorphism $f : U \to \mathbb{R}^m$. Then the sets*

$$W^s(x_0) := \{x \in U : f^k(x) \to x_0 \text{ as } k \to \infty\}$$

and

$$W^u(x_0) := \{x \in U : f^{-k}(x) \to x_0 \text{ as } k \to \infty\}$$

are called the stable and the unstable manifold of x_0, respectively.

Despite its name, the stable manifold may not be a submanifold of $\mathbb{R}^m$, but it can be described in terms of the local stable manifold which is a smooth submanifold of $\mathbb{R}^m$.

Definition 48. *Let x_0 be a hyperbolic fixed point of the C^1 diffeomorphism $f : U \to \mathbb{R}^m$. For given $\varepsilon > 0$, the local stable manifold is defined by*

$$W^{s,\varepsilon}(x_0) := \{x \in U : f^k(x) \to x_0 \text{ as } k \to \infty \text{ and } |f^k(x) - x_0| < \varepsilon \text{ for } k \geq 0\}.$$

Please note that for any $\varepsilon > 0$,

$$W^s(x_0) = \bigcup_{k \geq 0} f^{-k}(W^{s,\varepsilon}(x_0)), \tag{3.3}$$

and that the invariance properties

$$f(W^s(x_0)) = W^s(x_0) \text{ and } f(W^{s,\varepsilon}(x_0)) \subset W^{s,\varepsilon}(x_0) \tag{3.4}$$

hold.

The proof of the smoothness property is based on the contraction mapping principle.

Theorem 49. *Let $U \subset \mathbb{R}^m$ be open and let $f : U \to \mathbb{R}^m$ be a C^r diffeomorphism with hyperbolic fixed point x_0 and associated stable subspace E^s. Then for $\varepsilon > 0$ sufficiently small, $W^{s,\varepsilon}(x_0)$ is a C^r submanifold of R^m containing x_0, and $T_{x_0}W^{s,\varepsilon}(x_0) = E^s$.*

The following statement is often referred to as the saddle-point property, cf. [30].

Theorem 50. *Let $U \subset \mathbb{R}^m$ be open, and let $x_0 \in U$ be a hyperbolic fixed point of the C^1 diffeomorphism $f : U \to \mathbb{R}^m$. Then there exists some $\Delta > 0$ such that $f^k(x) \to x_0$ as $k \to \infty$ whenever $|f^k(x) - x_0| \leq \Delta$ for all $k \geq 0$.*

Hence $W^{s,\varepsilon}(x_0)$ is already characterized by

$$W^{s,\varepsilon}(x_0) := \{x \in U : |f^k(x) - x_0| < \varepsilon \text{ for } k \geq 0\}. \tag{3.5}$$

if ε is small enough.

3.2 Hyperbolic sets

A hyperbolic periodic orbit of length n is a set $\{x_0, f(x_0), \ldots, f^{n-1}(x_0)\}$ of distinct points, where $f^n(x_0) = x_0$ and the eigenvalues of $Df^n(x_0)$ lie off the unit circle.

As x_0 is a hyperbolic fixed point of the mapping f^n, the state space $\mathbb{R}^m$ can be represented as the direct sum of the stable and the unstable subspaces $E^s(x_0)$ and $E^u(x_0)$. It is easy to show that for any iterate $x = f^k(x_0)$, the splitting $\mathbb{R}^m = E^s(x) \oplus E^u(x)$ defined by

$$E^s(x) := Df^k(x_0)[E^s(x_0)], \quad E^u(x) := Df^k(x_0)[E^u(x_0)]$$

has the invariance property

$$Df(x)[E^s(x)] = E^s(f(x)), \quad Df(x)[E^u(x)] = E^u(f(x)),$$

and that there are suitable constants $0 < \lambda_1, \lambda_2 < 1$ and $K_1, K_2 > 0$ such that the linearized dynamical system satisfies

$$|Df^k(x)\xi| \leq K_1 \lambda_1^{k/n} |\xi| \text{ for } \xi \in E^s(x)$$

and

$$|Df^{-k}(x)\xi| \leq K_2 \lambda_1^{k/n} |\xi| \text{ for } \xi \in E^u(x)$$

for all $k \geq 0$. These facts motivate the following general definition.

Definition 51. *A compact set $S \subset U$ is said to be hyperbolic if*

(i) S is invariant, i.e. $f(S) = S$;

(ii) there is a continuous splitting $\mathbb{R}^m = E^s(x) \oplus E^u(x)$, $x \in S$

such that the subspaces $E^s(x)$ and $E^u(x)$ have constant dimensions. Moreover, the invariance properties

$$Df(x)[E^s(x)] = E^s(f(x)), \quad Df(x)[E^u(x)] = E^u(f(x)) \tag{3.6}$$

hold and there are constants $0 < \lambda_1, \lambda_2 < 1$ and $K_1, K_2 > 0$ such that for all $k \geq 0$ and $x \in S$

$$|Df^k(x)\xi| \leq K_1 \lambda_1^k |\xi| \text{ for } \xi \in E^s(x) \tag{3.7}$$

and

$$|Df^{-k}(x)\xi| \leq K_2 \lambda_2^k |\xi| \text{ for } \xi \in E^u(x). \tag{3.8}$$

The numbers K_1 and K_2 are called the constants and λ_1 and λ_2 are called the rates for the hyperbolic set S.

Remark 52. *A splitting* $\mathbb{R}^m = E^s(x) \oplus E^u(x)$, $x \in S$ *is called continuous if the projection* $x \mapsto Q(x)$ *with* $\operatorname{Im} Q(x) = E^s(x)$ *is continuous. Please note that the continuity of the splitting is already implied by the other assumptions in Definition 51 (cf. [27]).*

For a general hyperbolic set it is still possible to define the stable and unstable manifolds.

Definition 53. *Let* $f : U \to \mathbb{R}^m$ *be a* C^1 *diffeomorphism and let* S *be a compact hyperbolic set. Its stable and unstable manifolds are defined by*

$$W^s(S) := \{x \in U : \operatorname{dist}(f^k(x), S) \to 0 \text{ as } k \to \infty\} \tag{3.9}$$

and

$$W^u(S) := \{x \in U : \operatorname{dist}(f^{-k}(x), S) \to 0 \text{ as } k \to \infty\}. \tag{3.10}$$

Furthermore, the stable and unstable manifolds of a point $x \in S$ *are given by*

$$W^s(x) := \{y \in U : |f^k(y) - f^k(x)| \to 0 \text{ as } k \to \infty\} \tag{3.11}$$

and

$$W^u(x) := \{y \in U : |f^{-k}(y) - f^{-k}(x)| \to 0 \text{ as } k \to \infty\}. \tag{3.12}$$

Whenever a hyperbolic set is isolated, the stable and unstable manifolds of the whole set coincide with the unions of the stable and unstable manifolds of its elements, respectively.

Definition 54. *Let* $f : U \to \mathbb{R}^m$ *be a* C^1 *diffeomorphism. An invariant subset* S *of* U *is said to be isolated if there is a neighbourhood* W *of* S *such that* S *is the maximal invariant set contained in* W*.*

Though the following statement is very natural, its proof is not trivial at all.

Theorem 55. *Let* S *be an isolated compact hyperbolic set for the* C^1 *diffeomorphism* $f : U \to \mathbb{R}^m$*. Then*

$$W^s(S) = \bigcup_{x \in S} W^s(x), \quad W^u(S) = \bigcup_{x \in S} W^u(x). \tag{3.13}$$

Hyperbolicity is intimately related to the notion of exponential dichotomies.

Definition 56. *Let $J \subset \mathbb{Z}$ be an interval, and let A_k, $k \in J$, be invertible $m \times m$ matrices. Denote*

$$\Phi(k,n) := \begin{cases} A_{k-1} \dots A_n & \text{for } k > n \\ \mathrm{id} & \text{for } k = n \\ \Phi(n,k)^{-1} & \text{for } k < n. \end{cases}$$

The difference equation

$$u_{k+1} = A_k u_k$$

is said to have an exponential dichotomy on J if there are projections Q_k and constants $0 < \lambda_1, \lambda_2 < 1$ and $K_1, K_2 > 0$ such that for all k and n in J the projections satisfy the invariance conditions

$$\Phi(k,n)Q_n = Q_k \Phi(k,n),$$

and the inequalities

$$|\Phi(k,n)Q_n| \leq K_1 \lambda_1^{k-n}, \quad k \geq m$$

and

$$|\Phi(k,n)(\mathrm{id} - Q_n)| \leq K_2 \lambda_2^{n-k}, \quad k \leq m$$

hold.

It is clear that the linearized dynamical system has an exponential dichotomy on hyperbolic sets.

Theorem 57. *If S is a compact invariant set for the diffeomorphism $f : U \to \mathbb{R}^m$, then S is hyperbolic if and only if for all $x \in S$ the difference equation*

$$u_{k+1} = Df(f^k(x))u_k \tag{3.14}$$

has an exponential dichotomy on $(-\infty, \infty)$ with constants, exponents, and rank of projection independent of x.

Exponential dichotomies are more flexible than the hyperbolicity condition, because the time interval J need not be infinite. This feature makes them an attractive tool in the field of numerical analysis, whenever the restriction of a problem on $(-\infty, \infty)$ to a finite time interval J is studied (see e.g. the recent paper [25]).

A diffeomorphism f has the so-called expansivity property on any hyperbolic set. The expansivity property is a natural generalization of the saddle-point property of hyperbolic fixed points (see Theorem 50).

Theorem 58. *Let S be a compact hyperbolic set for the C^1 diffeomorphism $f : U \to \mathbb{R}^m$. Then f is expansive on S, i.e. there exists some $\delta > 0$ such that for any $x, y \in S$ the inequalities*

$$|f^k(x) - f^k(y)| \leq \delta \quad \forall k \in \mathbb{Z} \tag{3.15}$$

imply $x = y$.

It is clear by intuition that the hyperbolicity property is robust under C^1 small perturbations which do not shift the eigenvalues of the linearization too much. The following result which is often referred to as the Roughness Theorem states this more precisely.

Theorem 59. *Let S be a compact hyperbolic set for the C^1 diffeomorphism $f : U \to \mathbb{R}^m$, and let U be convex. Choose $\beta_1, \beta_2 > 0$ such that*

$$\lambda_1 < \beta_1 < 1 \text{ and } \lambda_2 < \beta_2 < 1.$$

Then there exist positive numbers σ_0 and d_0 depending only on f, S, β_1, and β_2 such that if O is an open neighbourhood of S with

$$\max_{x \in \overline{O}} \operatorname{dist}(x, S) \leq d_0$$

and $g : U \to \mathbb{R}^m$ is a C^1 diffeomorphism satisfying

$$\sigma := \sup_{x \in U} |g(x) - f(x)| + \sup_{x \in U} |Dg(x) - Df(x)| \leq \sigma_0,$$

the set

$$S_O := \{x \in \overline{O} : g^k(x) \in \overline{O} \text{ for all } k \in \mathbb{Z}\}$$

is a compact hyperbolic set for g with exponents β_1 and β_2 and the dimension of the stable bundle is the same as for f and S.

Furthermore, there exists a constant $M > 0$ such that for every g with $\sigma < \sigma_0$ there is a homeomorphism $h : S \to S_O$ satisfying

$$|h(x) - x| \leq M\sigma$$

and $h \circ f = g \circ h$.

3.3 The Shadowing Lemma

Various types of shadowing properties have been stated and analyzed, but it is still partly unclear how these properties are related to each other. It is also an open question whether there are necessary conditions for shadowing. In this text, only the most common notion of shadowing will be discussed.

Definition 60. *A sequence $\{x_k\}_{k\in\mathbb{Z}}$ of points in U is said to be a d-pseudo-trajectory of a mapping $f : U \to \mathbb{R}^m$ if*

$$|x_{k+1} - f(x_k)| \leq d \text{ for } k \in \mathbb{Z}. \tag{3.16}$$

Pseudotrajectories can be regarded as approximate or perturbed orbits with a uniformly bounded perturbation.

Definition 61. *Let $f : U \to \mathbb{R}^m$ be a C^1 diffeomorphism, and let $d, \varepsilon > 0$ be positive numbers. A d-pseudotrajectory $\{x_k\}_{k\in\mathbb{Z}}$ in U is said to be ε-shadowed by a real orbit, i.e. a sequence $\{p_k\}_{k\in\mathbb{Z}}$ with $p_k = f^k(p_0)$, if $|x_k - p_k| \leq \varepsilon$ for all $k \in \mathbb{Z}$.*

The diffeomorphism f is said to have the (d, ε)-shadowing property on U if any d-pseudotrajectory $\{x_k\}_{k\in\mathbb{Z}}$ in U is ε-shadowed by a real orbit $\{p_k\}_{k\in\mathbb{Z}}$.

Thus the long-term behaviour of the induced dynamical system is robust w.r.t. uniformly small perturbations whenever f has the shadowing property.

The following result is often called the Shadowing Lemma.

Theorem 62 (Shadowing Lemma)**.** *Let S be a compact hyperbolic set for a C^1 diffeomorphism $f : U \to \mathbb{R}^m$. Then there exist positive constants d_0, σ_0, and M depending only on f and S such that for any C^1 diffeomorphism $g : U \to \mathbb{R}^m$ satisfying*

$$|f(x) - g(x)| + |Df(x) - Dg(x)| \leq \sigma \text{ for } x \in U \tag{3.17}$$

with $\sigma \leq \sigma_0$, any d-pseudotrajectory of f in S with $d \leq d_0$ is ε-shadowed by a unique true orbit of g with $\varepsilon = M(d + \sigma)$.

The proof is based on a fixed point argument on the space of sequences. Every true orbit of g is a fixed point of the operator T which maps a sequence $\{x_k\}_{k\in\mathbb{Z}}$ to a sequence $\{y_k\}_{k\in\mathbb{Z}}$ given by $y_{k+1} := g(x_k)$, and it must be shown that there exists a fixed point of T close to any given d-pseudotrajectory.

The technical details vary according to the fixed point theorem which is eventually applied.

Of course the Shadowing Lemma is of great interest for numerical computations, because it guarantees that the errors caused by a numerical scheme will not explode on arbitrarily long time intervals whenever it is applied to hyperbolic systems.

The conjugacy statement in the Shadowing Lemma is closely related to the behaviour of the numerical method regarded as a discrete-time dynamical system. In the pioneering work [4] it has been shown that the phase portrait of a dynamical system near a stationary hyperbolic point is reproduced correctly by numerical one-step methods. Many similar results for systems with special or hyperbolic structure followed, see e.g. [39] and [20].

Chapter 4

Shadowing for set-valued dynamical systems

Though Shadowing Theory is an established field of research, very little is known about the shadowing property in set-valued dynamical systems.

In [19], spatial discretization effects were investigated by means of a multivalued extension, and first attempts to generalize the concept of unstable manifolds were made. These ideas were developed further in [7], where the properties of multivalued stable and unstable manifolds were investigated systematically.

The papers [21] and [22] analyze set-valued dynamical systems induced by contractive set-valued mappings. Unfortunately, the proof of the shadowing theorem stated therein contains a critical error.

In [29] and [36], hyperbolicity was defined for smooth relations. This hyperbolicity condition is designed to study features of classical dynamical systems such as stable and unstable manifolds in the framework of non-invertible maps. Due to the nature of the analyzed objects, this hyperbolicity condition does not allow the graph of a relation to have nonempty interior, which is generically the case in the set-up discussed in the present text.

4.1 Basic notation

A set-valued dynamical system on $\mathbb{R}^m$ is determined by a set-valued mapping $F : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ and its iterates. Throughout this text, the mapping F and the corresponding dynamical system are identified. An orbit of the set-valued system is defined as follows.

Definition 63. *A sequence $\eta = \{p_k\}$ is a trajectory of the system F if*

$$p_{k+1} \in F(p_k) \text{ for any } k \in \mathbb{Z}. \tag{4.1}$$

As in the single-valued case, d-pseudotrajectories are sequences which satisfy the defining relations (4.1) up to a small error.

Definition 64. *A sequence $\xi = \{x_k\}$ is called a d-pseudotrajectory of F if an error of size $d > 0$ is allowed in every step, i.e., if*

$$\operatorname{dist}(x_{k+1}, F(x_k)) \leq d \quad \text{for any } k \in \mathbb{Z}. \tag{4.2}$$

The shadowing property is defined analogously to the classicial case as well.

Definition 65. *Let d and ε be positive numbers. The system F has the (d, ε)-shadowing property on a subset $K \subset \mathbb{R}^m$, if for any d-pseudotrajectory $\xi = \{x_k\} \subset K$ of F there exists a trajectory $\eta = \{p_k\}$ with*

$$\operatorname{dist}(x_k, p_k) \leq \varepsilon \text{ for any } k \in \mathbb{Z}. \tag{4.3}$$

The inverse shadowing property has been discussed in the context of single-valued dynamical systems, see e.g. [4], [8], and [15]. The definition of the inverse shadowing property for set-valued mappings introduced in [32] and [33] is local in contrast to the case of shadowing.

Definition 66. *Let $\eta = \{p_k\}_{k \in \mathbb{Z}}$ be an orbit of the system F. The system F has the local inverse (a, d, ε)-shadowing property at η if for every sequence of mappings*

$$\Phi = \{\Phi_k : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)\} \tag{4.4}$$

such that each Φ_k is continuous w.r.t. dist_H and

$$\operatorname{dist}_H(F(p_k + v), \Phi_k(p_k + v)) \leq d \text{ for } k \in \mathbb{Z} \text{ and } |v| \leq a \tag{4.5}$$

there exists a trajectory $\xi = \{x_k\}_{k \in \mathbb{Z}}$ of Φ, i.e. a sequence of points satisfying $x_{k+1} \in \Phi_k(x_k)$, such that

$$|x_k - p_k| \leq \varepsilon \text{ for all } k \in \mathbb{Z}. \tag{4.6}$$

Of course it is possible to state a global inverse shadowing property by imposing the same conditions on the sequence Φ on the whole phase space, and global inverse shadowing theorems can be derived from local ones.

Please note that inverse shadowing is qualitatively different from shadowing. In the case of shadowing, one is interested in finding an orbit of a system F with strong properties such as hyperbolicity close to a given pseudotrajectory. In the inverse setup, an orbit of the system F is given and an orbit of a system, which is only known to be continuous and close to F in the sense of inequalities (4.5) is desired.

4.2 Contractive systems

Contractive systems are a subclass of hyperbolic systems which will be treated in Section 4.3. Nevertheless, there are good reasons for discussing them separately. The underlying principles can be displayed in the absence of technical difficulties, and contractivity allows the use of the Frigon-Granas Theorem (see Section 1.4) which does not require the defining mapping to be convex-valued. Moreover, it is possible to apply shadowing theorems to the time-T flow of differential inclusions with relaxed one-sided Lipschitz right hand sides which induce contractive dynamics (cf. Section 2.5), while it is still unclear whether there is a nontrivial class of right hand sides with a hyperbolic time-T flow.

4.2.1 Shadowing theorems

The following shadowing theorem is based on the Tikhonov-Schauder Fixed Point Theorem, cf. Theorem 11.

Theorem 67. *Let $K \subset \mathbb{R}^m$ be any subset, let $a > 0$ and $\lambda \in [0, 1)$, and let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be a set-valued mapping which satisfies*

$$\mathrm{dist}_H(F(x), F(x')) \leq \lambda |x - x'| \tag{4.7}$$

for any $x \in K$ and $x' \in \mathbb{R}^m$ with $|x - x'| \leq a$. Then F has the $(d, \frac{d}{1-\lambda})$-shadowing property on K whenever $d \leq (1 - \lambda)a$.

Proof. Define the sets $H_d := \{v \in \mathbb{R}^m : |v| \leq \frac{d}{1-\lambda}\}$ and $H_d^\infty := (H_d)^{\mathbb{Z}}$. Then $H_d \subset \mathbb{R}^m$ is compact w.r.t the Euclidean topology and $H_d^\infty \subset (\mathbb{R}^m)^{\mathbb{Z}}$ is compact w.r.t. the Tikhonov topology.

Let $\{x_k\}_{k\in\mathbb{Z}}$ be a d-pseudotrajectory in K, let $V = \{v_k\} \in H_d^\infty$, and define a sequence $W = \{w_k\}$ by

$$w_{k+1} = \mathrm{Dev}(x_{k+1}, F(x_k + v_k)). \tag{4.8}$$

Such a sequence W is unique since the sets $F(x_k + v_k)$ are convex.

Condition (4.7) implies that the mapping F is continuous w.r.t. dist_H. Hence, the mapping $v_k \mapsto w_{k+1}$ is continuous by Theorem 26. Furthermore,

$$\begin{aligned} |w_{k+1}| &\leq \mathrm{dist}(x_{k+1}, F(x_k)) + \mathrm{dist}(F(x_k), F(x_k + v_k)) \\ &\leq d + \lambda|v_k| \;\leq\; d + \lambda\frac{d}{1-\lambda} \;=\; \frac{d}{1-\lambda}, \end{aligned}$$

and $W \in H_d^\infty$. Thus, the operator σ defined by $\sigma(V) = W$ maps the compact convex set H_d^∞ into itself.

Since the $(k+1)$th element of $\sigma(V)$ depends on the kth element of V only, the operator σ is continuous w.r.t. the Tikhonov topology. By the Tikhonov-Schauder Fixed Point Theorem, there is a sequence $V = \{v_k\} \in H_d^\infty$ such that $\sigma(V) = V$. Thus,

$$x_{k+1} + v_{k+1} = x_{k+1} + \mathrm{Dev}(x_{k+1}, F(x_k + v_k)) \in F(x_k + v_k), \tag{4.9}$$

and the trajectory $\eta = \{p_k\} \in (\mathbb{R}^m)^{\mathbb{Z}}$ given by $p_k := x_k + v_k$ is a solution of (4.1) with

$$\|\eta - \xi\|_\infty = \|V\|_\infty \leq \frac{d}{1-\lambda}. \tag{4.10}$$

□

Remark 68. *Note that this type of proof can easily be adapted to sequences $\{x_k\}_{k\in\mathbb{N}}$ and $\{p_k\}_{k\in\mathbb{N}}$ by the trivial continuation $x_k = p_k = v_k = w_k = 0$ for $k < 0$. With the setting $w_0 := v_0$ the operator σ remains continuous w.r.t. the product topology and possesses the desired fixed point.*

It is possible to formulate an inverse shadowing theorem and prove it in a very similar way. Note that for inverse shadowing, only the images of the approximations Φ_k and not the images of the mapping F are required to be convex.

Theorem 69. *Let $\eta = \{p_k\} \in (\mathbb{R}^m)^{\mathbb{Z}}$ be a trajectory of a set-valued dynamical system generated by a mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ with compact values. Assume that there exist numbers $a > 0$ and $\lambda \in [0, 1)$ such that F satisfies*

$$\mathrm{dist}_H(F(p_k), F(p_k + v)) \leq \lambda|v| \quad \text{for any } k \in \mathbb{Z} \quad \text{and} \quad |v| \leq a. \tag{4.11}$$

Then F has the inverse $(a, d, \frac{d}{1-\lambda})$-shadowing property at η whenever $d \leq (1-\lambda)a$.

Proof. Let H_d and H_d^∞ be as in the proof of Theorem 67, and let a family Φ of mappings be given according to Definition 66. Take some $V = \{v_k\} \in H_d^\infty$ and define a sequence $W = \{w_k\}$ by

$$w_{k+1} = \mathrm{Dev}(p_{k+1}, \Phi_k(p_k + v_k)). \tag{4.12}$$

By the properties of the mappings Φ_k, the operator defined by $\sigma(V) := W$ is continuous w.r.t. the Tikhonov topology. Furthermore,

$$\begin{aligned} |w_{k+1}| &\leq \mathrm{dist}(p_{k+1}, F(p_k)) + \mathrm{dist}_H(F(p_k), F(p_k + v_k)) \\ &+ \mathrm{dist}_H(F(p_k + v_k), \Phi_k(p_k + v_k)) \\ &\leq \lambda|v_k| + d \leq \frac{d}{1-\lambda}, \end{aligned}$$

and $W \in H_d^\infty$.

By the Tikhonov-Schauder theorem, there exists a sequence $V = \{v_k\} \in H_d^\infty$ such that $\sigma(V) = V$. Thus,

$$p_{k+1} + v_{k+1} = p_{k+1} + \mathrm{Dev}(p_{k+1}, \Phi_k(p_k + v_k)) \in \Phi_k(p_k + v_k),$$

and $\xi = \{x_k\} \in (\mathbb{R}^m)^{\mathbb{Z}}$ given by $x_k := p_k + v_k$ is an orbit of Φ such that

$$||\eta - \xi||_\infty = ||V||_\infty \leq \frac{d}{1-\lambda}.$$

$\square$

The following shadowing theorem uses the Frigon-Granas Fixed Point Theorem to prove the existence of a shadowing trajectory without assuming convexity of the values of the defining mapping F.

Theorem 70. *Let $K \subset \mathbb{R}^m$ be a subset, and let $F : \mathbb{R}^m \to \mathcal{C}(\mathbb{R}^m)$ be a set-valued mapping. If there exist constants $a > 0$, $M > 0$, and $\lambda \in [0, 1)$ such that*

$$\mathrm{dist}_H(F(x), F(x')) \leq \lambda|x - x'| \text{ for all } x, x' \in B_a(K) \tag{4.13}$$

and

$$\mathrm{diam}\, F(x) \leq M \text{ for all } x \in B_a(K), \tag{4.14}$$

then F has the $(d, \frac{d}{1-\lambda})$-shadowing property on K whenever

$$d \leq d_0 := a(1-\lambda).$$

Proof. Let $\xi^0 = \{x_k^0\}_{k\in\mathbb{Z}} \subset K$ be a d-pseudotrajectory of F with $d \le d_0$. The set

$$X := \left\{ \{x_k\}_{k\in\mathbb{Z}} \subset \mathbb{R}^m : \sup_{k\in\mathbb{Z}} |x_k - x_k^0| < \infty \right\} \tag{4.15}$$

equipped with the distance

$$\mathrm{Dist}(\{x_k\}_{k\in\mathbb{Z}}, \{x_k'\}_{k\in\mathbb{Z}}) := \sup_{k\in\mathbb{Z}} |x_k - x_k'| \tag{4.16}$$

is a complete metric space. Let $\mathrm{Dist}(A,B)$ and $\mathrm{Dist}_H(A,B)$ denote the non-symmetric and the symmetric Hausdorff distance of two subsets $A, B \subset X$, respectively.

Consider the mapping $\mathcal{F} : B_a(\xi_0) \rightrightarrows X$ given by

$$\mathcal{F}(\xi = \{x_k\}_{k\in\mathbb{Z}}) = \{\eta = \{p_k\}_{k\in\mathbb{Z}} \subset \mathbb{R}^m : p_{k+1} \in F(x_k)\ \forall k \in \mathbb{Z}\}. \tag{4.17}$$

For any $\xi = \{x_k\}_{k\in\mathbb{Z}} \in B_a(\xi_0)$, set $d_1 := \mathrm{Dist}(\xi, \xi^0) < \infty$. If $\eta = \{p_k\}_{k\in\mathbb{Z}} \in \mathcal{F}(\xi)$,

$$\begin{aligned} |x_{k+1}^0 - p_{k+1}| &\le \mathrm{dist}(x_{k+1}^0, F(x_k)) + \mathrm{diam}\, F(x_k) \\ &\le \mathrm{dist}(x_{k+1}^0, F(x_k^0)) + \mathrm{dist}(F(x_k^0), F(x_k)) + M \\ &\le d + \lambda d_1 + M < \infty. \end{aligned} \tag{4.18}$$

Thus every element of $\mathcal{F}(\xi)$ is an element of X, and the mapping $\mathcal{F}$ is well-defined. Furthermore, estimate (4.18) implies that the values of $\mathcal{F}$ are bounded. Since F has closed values, so does $\mathcal{F}$.

Set $r := \frac{d}{1-\lambda}$ and let $\xi', \xi'' \in B_r(\xi^0)$. For any $\eta' = \{p_k'\}_{k\in\mathbb{Z}} \in \mathcal{F}(\xi')$,

$$\mathrm{Dist}(\eta', \mathcal{F}(\xi'')) \le \sup_{k\in\mathbb{Z}} \mathrm{dist}(p_{k+1}', F(x_k'')) \le \sup_{k\in\mathbb{Z}} \mathrm{dist}_H(F(x_k'), F(x_k'')) \tag{4.19}$$

$$\le \sup_{k\in\mathbb{Z}} \lambda |x_k' - x_k''| = \lambda\, \mathrm{Dist}(\xi', \xi''), \tag{4.20}$$

and hence

$$\mathrm{Dist}_H(\mathcal{F}(\xi'), \mathcal{F}(\xi'')) \le \lambda\, \mathrm{Dist}(\xi', \xi'') \text{ for all } \xi', \xi'' \in B_r(\xi^0). \tag{4.21}$$

In addition,

$$\mathrm{Dist}(\xi^0, \mathcal{F}(\xi^0)) \le \sup_{k\in\mathbb{Z}} \mathrm{dist}(x_{k+1}^0, F(x_k^0)) \le d \le (1-\lambda) r. \tag{4.22}$$

By the Frigon-Granas Fixed Point Theorem, there is a fixed point η^0 of $\mathcal{F}$ such that

$$\|\eta^0 - \xi^0\|_\infty \leq r = \frac{d}{1-\lambda}. \tag{4.23}$$

It remains to note that the definition of a fixed point η^0 of $\mathcal{F}$ implies that η^0 is an orbit of F. □

Remark 71. *It is unclear if the above proof can be adapted to the case of forward sequences $\{x_k\}_{k\in\mathbb{N}}$ and $\{p_k\}_{k\in\mathbb{N}}$, because it seems that it is impossible to choose component p_0 of η in* (4.17) *without prior knowledge about the fixed point and without violating the conditions of the Frigon-Granas Fixed Point Theorem.*

The Frigon-Granas Theorem belongs to the group of contraction mapping principles, and thus it can only be applied to strictly contracting maps. Thus it is impossible to prove an inverse shadowing theorem based on the Frigon-Granas Theorem in the above manner: If F is contractive and Φ is a family of approximating mappings, the Φ_k need not be contractions.

Using the Kakutani Fixed Point Theorem, it is possible to give alternative proofs of Shadowing Theorem 67 and the Inverse Shadowing Theorem 69 which are based on truly set-valued methods and deserve to be displayed.

Alternative proof of Theorem 67. If $\xi = \{x_k\}_{k\in\mathbb{Z}}$ is any d-pseudotrajectory in K,

$$\begin{aligned} \operatorname{dist}(x_{k+1}, F(x_k+v)) &\leq \operatorname{dist}(x_{k+1}, F(x_k)) + \operatorname{dist}(F(x_k), F(x_k+v)) \\ &\leq d + \lambda|v| \leq d + \lambda\frac{d}{1-\lambda} \leq \frac{d}{1-\lambda} \end{aligned}$$

for all $v \in B_{\frac{d}{1-\lambda}}(0)$. Thus the mappings $\mathcal{F}_k : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ defined by

$$\mathcal{F}_{k+1}(v = \{v_k\}_{k\in\mathbb{Z}}) := B_{\frac{d}{1-\lambda}}(0) \cap (F(x_k+v) - x_{k+1}) \tag{4.24}$$

have nonempty convex and compact values on $B_{\frac{d}{1-\lambda}}(0)$, and they are upper semicontinuous by Propositions 1.4.8 and 1.4.9 in [3].

The ball $B^* := \{x \in (\mathbb{R}^m)^{\mathbb{Z}} : \|x\|_\infty \leq \frac{d}{1-\lambda}\}$ is compact in the product space $(\mathbb{R}^m)^{\mathbb{Z}}$ equipped with the product topology. Consider the mapping $\mathcal{F} : B^* \rightrightarrows B^*$ given by

$$\mathcal{F}(v) := \prod_{k\in\mathbb{Z}} \mathcal{F}_k(v_{k-1}). \tag{4.25}$$

It is upper semicontinuous and its values are nonempty, convex, and compact. Thus the Kakutani Fixed Point Theorem implies the existence of a fixed point $\bar{v} \in B^*$ of $\mathcal{F}$, which means that $\xi + \bar{v}$ is a shadowing trajectory close to ξ. $\square$

Alternative proof of Theorem 69. If Φ is a family of approximating mappings,

$$\begin{aligned}\mathrm{dist}(p_{k+1}, \Phi_k(p_k+v_k)) &\leq \mathrm{dist}(p_{k+1}, F(p_k+v_k)) + \mathrm{dist}(F(p_k+v_k), \Phi_k(p_k+v_k)) \\ &\leq \lambda|v_k| + d \leq \frac{d}{1-\lambda}\end{aligned}$$

for all $v_k \in B_{\frac{d}{1-\lambda}}(0)$. Thus the mappings $\mathcal{F}_k : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ defined by

$$\mathcal{F}_{k+1}(v_k) := B_{\frac{d}{1-\lambda}}(0) \cap (\Phi(p_k+v_k) - p_{k+1}) \tag{4.26}$$

have nonempty convex and compact values on $B_{\frac{d}{1-\lambda}}(0)$, and they are upper semicontinuous by Propositions 1.4.8 and 1.4.9 in [3].

The Kakutani Theorem applied to the product mapping

$$\mathcal{F}(v = \{v_k\}_{k\in\mathbb{Z}}) := \prod_{k\in\mathbb{Z}} \mathcal{F}_k(v_{k-1}). \tag{4.27}$$

yields a fixed point $\bar{v} \in B^*$ of $\mathcal{F}$ as in the previous proof, and $\eta + \bar{v}$ is an orbit of the approximating family Φ. $\square$

The following remark illustrates that shadowing trajectories are not necessarily unique.

Example 72. *Consider the Euclidean plane $\mathbb{R}^2$ with coordinates $z = (z_1, z_2)$ and the segment*

$$I = \{z : z_1 = 0,\ 0 \leq z_2 \leq 1\}. \tag{4.28}$$

Define a set-valued dynamical system generated by the constant mapping $F(z) = I$, $z \in \mathbb{R}^2$. A sequence $\eta = \{p_k\}$ is a trajectory of F if and only if $p_k \in I$ for all $k \in \mathbb{Z}$.

The mapping F satisfies the conditions of Theorems 67 and 70 for any $a > 0$ and $\lambda \in [0, 1)$. Fix $d > 0$ and consider the sequence $\xi = \{x_k\}$ given by

$$x_k = (d(1 - 2^{-|k|}), 0). \tag{4.29}$$

Clearly, ξ is a d-pseudotrajectory of F, but not a δ-pseudotrajectory for any $\delta < d$. There is no exact orbit η of F such that $\|\xi-\eta\|_\infty < d$, but $\|\xi-\eta\|_\infty = d$ holds for any sequence $\eta = \{p_k\}$, where

$$p_k \in I \cap B_{d\sqrt{1-(1-2^{-|k|})^2}}(0). \tag{4.30}$$

4.2.2 Application to differential inclusions

Consider the differential inclusion

$$\dot{x}(t) \in F(x(t)) \text{ a.e.} \tag{4.31}$$

Mild smoothness and growth assumptions guarantee compactness of its reachable set $\mathcal{R}(T,0,x)$ for any $x \in \mathbb{R}^m$, see Theorem 30. If $\operatorname{diam} F(x) \le M$ for all $x \in \mathbb{R}^m$, then any two solutions $y(\cdot)$ and $z(\cdot)$ of (4.31) with identical initial value $y(0) = z(0) = x_0$ satisfy

$$\begin{aligned} |y(T) - z(T)| &\le \int_0^T |\dot{y}(s) - \dot{z}(s)| ds \le \int_0^T \operatorname{diam} F(y(s)) + L|y(s) - z(s)| ds \\ &\le MT + \int_0^T L|y(s) - z(s)| ds, \end{aligned}$$

and the Gronwall Lemma yields

$$|y(T) - z(T)| \le MT + \int_0^T LMse^{L(T-s)} ds = \frac{M}{L}(e^{LT} - 1). \tag{4.32}$$

It follows that $\operatorname{diam} \mathcal{R}(T,0,x_0) \le \frac{M}{L}(e^{LT} - 1)$ for all $x_0 \in \mathbb{R}^m$.

If F satisfies the relaxed one-sided Lipschitz property with constant $\mu < 0$ (see Sections 1.2 and 2.5) the time-T flow $x \mapsto \mathcal{R}(T,0,x)$ is a set-valued contraction according to Theorem 42.

Hence the following global statement holds.

Theorem 73. *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be a Lipschitz continuous set-valued mapping which also satisfies the relaxed one-sided Lipschitz property with negative constant $\mu < 0$. If the diameter of the images of F is uniformly bounded, then the time-t flow of* (4.31) *satisfies the assumptions of the Shadowing Theorem 70 and the Inverse Shadowing Theorem 69 with arbitrary $a > 0$, $K = \mathbb{R}^m$, and $\lambda := e^{\mu T}$.*

In view of Remark 43, the relaxed one-sided Lipschitz property with negative constant is an 'almost necessary' condition for contractivity of the time-T flow of an ODI.

The lengthy discussion of this topic in [32] turns out to be unnatural and unnecessary. At that time, we were not aware of the Frigon-Granas Fixed

Point Theorem and had to prove shadowing via approximating selections of the time-T flow. Please note that the inverse shadowing theorems do not require the defining mapping, but the approximating family to be convex-valued. This assumption does not cause any problems, because the images of the Euler scheme and the Viability Kernel Algorihm (cf. Chapter 5) are convex by construction.

Example 74. *The right hand side $f : \mathbb{R} \to \mathbb{R}$, $f(x) = -x$ of the differential equation*

$$\dot{x}(t) = -x(t) \tag{4.33}$$

is one-sided Lipschitz with constant $\mu = -1$. Thus the right hand side $F : \mathbb{R} \to \mathcal{CC}(\mathbb{R})$, $F(x) = [-x - \varepsilon, -x + \varepsilon]$ of the ODI

$$\dot{x}(t) \in [-x(t) - \varepsilon, -x(t) + \varepsilon] \tag{4.34}$$

is relaxed one-sided Lipschitz with the same constant. The time-t flow satisfies

$$G_{\log 2}(x) := \mathcal{R}(\log 2, 0, x) = [\frac{x}{2} - \frac{\varepsilon}{2}, \frac{x}{2} + \frac{\varepsilon}{2}], \tag{4.35}$$

i.e. it is a set-valued contraction with constant $\lambda = \frac{1}{2}$, and in particular,

$$G_{\log 2}([-\varepsilon, \varepsilon]) = [-\varepsilon, \varepsilon]. \tag{4.36}$$

Note that the set $[-\varepsilon, \varepsilon]$ attracts every compact subset of $\mathbb{R}$, that it contains a dense trajectory of $G_{\log 2}$, and that its diameter depends on the diameter of the values of F.

Since $G^{-1}_{\log 2}(x) = [2x - \varepsilon, 2x + \varepsilon]$, every orbit $\{p_k\}_{k\in\mathbb{Z}}$ such that $p_0 \notin [-\varepsilon, \varepsilon]$ satisfies $p_k \to \infty$ as $k \to -\infty$, but any sequence $\{x_k\}_{k\in\mathbb{Z}}$ such that $x_0 \in [-\varepsilon - 2d, \varepsilon + 2d]$ and $\mathrm{dist}(x_{k+1}, G_{\log 2}(x_k)) \le d$ for all $k \ge 0$ can be extended to a d-pseudotrajectory such that $x_k \in [-\varepsilon - 2d, \varepsilon + 2d]$ for all $k < 0$.

Hence $G_{\log 2}$ satisfies the conditions of Theorem 70 for any subset K, any $a > 0$ and $M := \varepsilon$. Furthermore, the statement of Theorem 70 still holds for d-pseudotrajectories ξ_+ which are only defined for nonnegative times whenever $K = [-\delta, \delta]$ with $0 < \delta \le \varepsilon + 2d$, any $a > 0$ and $M := \varepsilon$, because it is possible to extend ξ_+ artificially to a d-pseudotrajectory ξ on $\mathbb{Z}$ which does not leave K.

4.3 Hyperbolic systems

The aim of this section is to propose a hyperbolicity condition for set-valued mappings on $\mathbb{R}^m$ and to show that this condition implies the shadowing and the inverse shadowing property. One can give a similar definition of hyperbolicity on a compact subset without changing the essence of Theorem 77 below.

Definition 75. *Let a set-valued mapping $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ of the form*

$$F(x) = L(x) + M(x) \tag{4.37}$$

be given, where $L : \mathbb{R}^m \to \mathbb{R}^m$ is a continuous single-valued mapping and $M : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ is a set-valued mapping with compact and convex images. The mapping (4.37) *is said to be hyperbolic in $\mathbb{R}^m$ if there exist constants $N \geq 1$, $a, \kappa, l > 0$, and $\lambda \in (0, 1)$ such that the following conditions hold.*

(P1) For any point $x \in \mathbb{R}^m$ there exist linear subspaces $E^s(x)$ and $E^u(x)$ of $\mathbb{R}^m$ such that

$$E^s(x) \oplus E^u(x) = \mathbb{R}^m, \tag{4.38}$$

and

$$\|Q(x)\|, \|P(x)\| \leq N, \tag{4.39}$$

where $Q(x)$ and $P(x)$ are the corresponding complementary projections from $\mathbb{R}^m$ to $E^s(x)$ and $E^u(x)$, which are called the stable and unstable subspaces, respectively.

(P2) If $x, y, v \in \mathbb{R}^m$ satisfy $|v| \leq a$ and $\operatorname{dist}(y, F(x)) \leq a$, then $L(x + v)$ can be represented as

$$L(x + v) = L(x) + A(x)v + B(x, v), \tag{4.40}$$

where $A(x) : \mathbb{R}^m \to \mathbb{R}^m$ is a linear mapping that is continuous with respect to x and such that

$$|Q(y)A(x)v| \leq \lambda|v| \text{ for } v \in E^s(x), \tag{4.41}$$
$$|Q(y)A(x)v| \leq \kappa|v| \text{ for } v \in E^u(x), \tag{4.42}$$

and

$$|P(y)A(x)v| \leq \kappa|v| \text{ for } v \in E^s(x). \tag{4.43}$$

Finally, the restriction $P(y)A(x)|_{E^u(x)} : E^u(x) \to E^u(y)$ *is assumed to be a linear isomorphism satisfying*

$$|P(y)A(x)v| \geq \frac{1}{\lambda}|v| \text{ for } v \in E^u(x). \tag{4.44}$$

(P3) For $v \in \mathbb{R}^m$ *with* $|v| \leq a$,

$$|B(x,v)| \leq l|v| \tag{4.45}$$

and

$$\mathrm{dist}_H(M(x), M(x+v)) \leq l|v| \text{ for } x \in \mathbb{R}^m. \tag{4.46}$$

Since $L(x)$ and $A(x)$ are assumed to be continuous, $B(x,v)$ is continuous for any x and v with $|v| \leq a$. Moreover, condition (4.46) implies the continuity of M w.r.t. the Hausdorff distance.

Example 76. *Let* A *be an* $m \times m$ *matrix such that all eigenvalues lie off the unit circle, and let* $M \in \mathcal{CC}(\mathbb{R}^m)$ *be a fixed set. Then the mapping* $x \mapsto Ax + M$ *is hyperbolic in the sense of the above definition, where the stable and unstable subspaces and projections are those associated with the hyperbolic matrix* A.

In [7], the dynamics of multifunctions similar to those in Example 76 are analyzed. It is assumed that the single-valued component has a saddle point and that the set M is a small ball. The analysis is given from a completely different point of view. Every trajectory of the multivalued system is considered as a trajectory of the single-valued system which is perturbed by some sequence with values in M. Conjugacy-type results are obtained, analogs of the stable and unstable manifolds are proposed, and their smoothness properties are discussed.

The following shadowing theorem is based on the Tikhonov-Schauder Fixed Point Theorem. In the single-valued case it is possible to decompose the dynamics into their stable and unstable components, to prove the existence of a fixed point, and to piece the components together again. It is not possible to transfer this technique directly to the set-valued case, mainly because

$$x = Px + Qx \tag{4.47}$$

for any point $x \in \mathbb{R}^m$, but

$$M \subsetneqq PM + QM \tag{4.48}$$

for almost all subsets $M \subset \mathbb{R}^m$, which means that a fixed point constructed by direct projection might not be an element of the set, but only an element of the sum of its projections.

Theorem 77. *Let F be a set-valued hyperbolic mapping as described above. If*

$$\lambda + \kappa + 4lN < 1, \tag{4.49}$$

then F has the $(d, \mathcal{L}d)$-shadowing property whenever $d \leq a/\mathcal{L}$, where

$$\mathcal{L}^{-1} = \frac{1}{2N}\left(1 - \lambda - \kappa - 4lN\right).$$

Remark 78. *Condition* (4.49) *implies the inequality*

$$\lambda(1 + \kappa + 4lN) < 1. \tag{4.50}$$

In addition,

$$\mathcal{L}^{-1} \leq \frac{1}{2N}\left(\frac{1}{\lambda} - 1 - \kappa - 4lN\right). \tag{4.51}$$

Proof of Theorem 77. Set $d_0 = a/\mathcal{L}$ and consider a d-pseudotrajectory $\{x_k\}_{k\in\mathbb{Z}}$ of F with $d \leq d_0$. The aim is to find a sequence $V = \{v_k \in \mathbb{R}^m : k \in \mathbb{Z}\}$ such that

$$x_{k+1} + v_{k+1} \in F(x_k + v_k) \tag{4.52}$$

and

$$\|V\|_\infty \leq \mathcal{L}d; \tag{4.53}$$

in this case, $\{p_k = x_k + v_k\}$ is the desired trajectory of F.

By (4.37), relations (4.52) take the form

$$x_{k+1} + v_{k+1} \in L(x_k + v_k) + M(x_k + v_k). \tag{4.54}$$

If $|v_k| \leq a$, it follows from property (P2) that

$$L(x_k + v_k) = L(x_k) + A(x_k)v_k + B(x_k, v_k). \tag{4.55}$$

Thus, relation (4.54) can be rewritten as

$$x_{k+1} + v_{k+1} \in L(x_k) + A(x_k)v_k + B(x_k, v_k) + M(x_k + v_k),$$

or

$$v_{k+1} \in L(x_k) + A(x_k)v_k + B(x_k, v_k) + M(x_k + v_k) - x_{k+1}. \tag{4.56}$$

Consider the vector

$$\sigma_k = \mathrm{Dev}(x_{k+1}, L(x_k) + B(x_k, v_k) + M(x_k + v_k)).$$

The compact and convex set $L(x_k) + B(x_k, v_k) + M(x_k + v_k)$ depends continuously on v_k w.r.t. the Hausdorff distance for $|v_k| \le a$ (see the definition of F and properties (P2) and (P3)). By (4.45),

$$|B(x_k, v_k)| \le l|v_k|;$$

by (4.46),

$$\mathrm{dist}_H(M(x_k), M(x_k + v_k)) \le l|v_k|.$$

Since

$$\mathrm{dist}(x_{k+1}, L(x_k) + M(x_k)) < d,$$

estimate

$$|\sigma_k| = \mathrm{dist}(x_{k+1}, L(x_k) + B(x_k, v_k) + M(x_k + v_k)) \le d + 2l|v_k| \tag{4.57}$$

holds. If

$$\sigma_k = v_{k+1} - A(x_k)v_k, \tag{4.58}$$

then the inclusion

$$x_{k+1} + \sigma_k \in L(x_k) + B(x_k, v_k) + M(x_k + v_k)$$

implies that

$$x_{k+1} + v_{k+1} - A(x_k)v_k \in L(x_k) + B(x_k, v_k) + M(x_k + v_k),$$

which is equivalent to the desired inclusion (4.52). Thus, a solution $V = \{v_k\}$ of (4.58) yields a shadowing trajectory.

Consider the projections

$$Q(x_{k+1})v_{k+1} = Q(x_{k+1})A(x_k)v_k + Q(x_{k+1})\sigma_k, \quad (4.59)$$
$$P(x_{k+1})v_{k+1} = P(x_{k+1})A(x_k)v_k + P(x_{k+1})\sigma_k \quad (4.60)$$

of equality (4.58) to $S(x_{k+1})$ and $U(x_{k+1})$, respectively. Denote $b := d\mathcal{L}/2$ and let

$$H_k = \{v_k \in \mathbb{R}^m : |P(x_k)v_k|, |Q(x_k)v_k| \leq b\}$$

and

$$H = \prod_{k\in\mathbb{Z}} H_k. \quad (4.61)$$

Each H_k is compact and convex; hence, H is convex and compact w.r.t. the Tikhonov product topology.

If $V = \{v_k\} \in H$, then

$$|v_k| \leq |P(x_k)v_k| + |Q(x_k)v_k| \leq 2b = \mathcal{L}d \leq a; \quad (4.62)$$

hence, all the terms in (4.59) and (4.60) are defined. Thus an operator T that maps a sequence $V = \{v_k \in \mathbb{R}^m\}$ to a sequence $W = \{w_k \in \mathbb{R}^m\}$ can be defined as follows:

The stable components of w_k are defined by

$$Q(x_{k+1})w_{k+1} = Q(x_{k+1})A(x_k)v_k + Q(x_{k+1})\sigma_k. \quad (4.63)$$

To obtain the unstable components, equation (4.60) must be transformed. Consider the mapping

$$G(w) = P(x_{k+1})A(x_k)w, \ w \in U(x_k). \quad (4.64)$$

Clearly, $G(0) = 0$. It follows from (4.44) that

$$|G(w) - G(w')| \geq \frac{1}{\lambda}|w - w'|, \ w, w' \in U(x_k). \quad (4.65)$$

Since the restriction of $P(x_{k+1})A(x_k)$ to $U(x_k)$ is assumed to be a linear isomorphism,

$$G(D(b, x_k)) \supset D(b', x_{k+1}), \quad (4.66)$$

where $b' = b/\lambda$,

$D(b, x_k) = \{z \in U(x_k) : |w| \leq b\}$, and $D(b', x_{k+1}) = \{z \in U(x_{k+1}) : |z| \leq b'\}$.

By (4.65) and (4.66), the inverse Γ of G is defined on $D(b', x_{k+1})$. By (4.65),

$$|\Gamma(z) - \Gamma(z')| \leq \lambda |z - z'|, \; z, z' \in D(b', x_{k+1}). \tag{4.67}$$

Now the unstable components of w_k can be defined by

$$P(x_k)w_k = \Gamma\{P(x_{k+1})[v_{k+1} - \sigma_k - A(x_k)Q(x_k)v_k]\}. \tag{4.68}$$

Lemma 79. *The operator T maps H to itself.*

Proof of Lemma 79. For the stable components, estimate

$$\begin{aligned} & |Q(x_{k+1})w_{k+1}| \\ \leq \;& |Q(x_{k+1})A(x_k)v_k| + |Q(x_{k+1})\sigma_k| \\ \leq \;& |Q(x_{k+1})A(x_k)P(x_k)v_k| + |Q(x_{k+1})A(x_k)Q(x_k)v_k| + |Q(x_{k+1})\sigma_k| \\ \leq \;& \kappa|P(x_k)v_k| + \lambda|Q(x_k)v_k| + N(d + 2l|v_k|) \end{aligned}$$

holds (see (4.39), (4.41), (4.42), and (4.57)). Since $|v_k| \leq 2b$ (see (4.62)),

$$|Q(x_{k+1})w_{k+1}| \leq (\lambda + \kappa + 4lN)b + Nd = \left(\lambda + \kappa + 4lN + \frac{2N}{\mathcal{L}}\right) b \leq b \tag{4.69}$$

by the definition of b and $\mathcal{L}$.

It must be checked that the argument in the right hand side of (4.68) is contained in the domain of Γ. Since $v_{k+1} \in H_{k+1}$,

$$|P(x_{k+1})v_{k+1}| \leq b. \tag{4.70}$$

By (4.39) and (4.57),

$$|P(x_{k+1})\sigma_k| \leq N(d + 2l|v_k|) \leq N(d + 4lb). \tag{4.71}$$

By (4.43),

$$|P(x_{k+1})A(x_k)Q(x_k)v_k| \leq \kappa|Q(x_k)v_k| \leq \kappa b. \tag{4.72}$$

By inequalities (4.70)–(4.72) and (4.51), the argument of Γ satisfies

$$\begin{aligned} |P(x_{k+1})[v_{k+1} - \sigma_k - A(x_k)Q(x_k)v_k]| &\leq (1 + \kappa + 4lN)b + Nd \\ &= \left(1 + \kappa + 4lN + \tfrac{2N}{\mathcal{L}}\right) b \leq \tfrac{b}{\lambda} = b'. \end{aligned} \tag{4.73}$$

Thus, $\Gamma\{\ldots\}$ is defined, and it follows from (4.73) and (4.67) that

$$|P(x_k)w_k| \leq b. \tag{4.74}$$

Inequalities (4.69) and (4.74) show that if $V \in H$ and $W = T(V)$, then $W \in H$. □

Since σ_k depends on v_k only, formulas (4.63) and (4.68) show that $(T(V))_k$ depends on v_{k-1}, v_k, v_{k+1}. Hence, the operator is continuous w.r.t. the Tikhonov topology on H.

The Tikhonov-Schauder fixed point theorem implies that T has a fixed point in H. To complete the proof of Theorem 77, it remains to show that if $T(V) = V$, then V solves equation (4.58). By (4.63),

$$Q(x_{k+1})v_{k+1} = Q(x_{k+1})A(x_k)v_k + Q(x_{k+1})\sigma_k \tag{4.75}$$

if $T(V) = V$. Apply G to the equality

$$P(x_k)v_k = \Gamma\{P(x_{k+1})[v_{k+1} - \sigma_k - A(x_k)Q(x_k)v_k]\}$$

in order to show that

$$P(x_{k+1})A(x_k)P(x_k)v_k = G(P(x_k)v_k) = P(x_{k+1})[v_{k+1} - \sigma_k - A(x_k)Q(x_k)v_k].$$

Hence,

$$\begin{aligned} & P(x_{k+1})v_{k+1} \\ = \; & P(x_{k+1})\sigma_k + P(x_{k+1})A(x_k)Q(x_k)v_k + P(x_{k+1})A(x_k)P(x_k)v_k \\ = \; & P(x_{k+1})[\sigma_k + A(x_k)v_k]. \end{aligned} \tag{4.76}$$

Statements (4.75) and (4.76) imply that

$$v_{k+1} = \sigma_k + A(x_k)v_k,$$

i.e., V solves equation (4.58). Since

$$\|V\|_\infty \le \mathcal{L}d$$

by (4.62), the proof is complete. □

For the local notion of inverse shadowing, it is necessary to modify the hyperbolicity condition given above. The mapping F is said to be hyperbolic at the trajectory $\eta = \{p_k\}_{k\in\mathbb{Z}}$ if there exist constants $N \ge 1$, $a, \kappa, l > 0$, and $\lambda \in (0, 1)$ such that condition (P1) holds for points $x = p_k$, condition (P2) holds for points $x = p_k, y = p_{k+1}$, and vectors v with $|v| \le a$, and, finally, condition (P3) holds for points $x = p_k$ and vectors v with $|v| \le a$.

Theorem 80. *Assume that F is hyperbolic at a trajectory $\eta = \{p_k\}$ in the above sense. If*

$$\lambda + \kappa + 4lN < 1, \tag{4.77}$$

then F has the inverse $(a, d, \mathcal{L}d)$-shadowing property whenever $d \leq a/\mathcal{L}$, where

$$\mathcal{L}^{-1} = \frac{1}{2N}(1 - \lambda - \kappa - 4lN).$$

Proof. The line of argument is very similar to the proof of Theorem 77. Here, a trajectory $\{x_k\}$ of Φ is constructed by proving the existence of a sequence $\{v_k\}$ such that

$$p_{k+1} + v_{k+1} \in \Phi_k(p_k + v_k)$$

and

$$\|V\|_\infty \leq \mathcal{L}d. \tag{4.78}$$

The mappings Φ_k can be represented as

$$\Phi_k(p_k + v) = L(p_k) + A(p_k)v + B(p_k, v) + \tilde{M}_k(p_k + v) \tag{4.79}$$

for small v, where each $\tilde{M}_k : \mathbb{R}^m \to CC(\mathbb{R}^m)$ is a continuous mapping w.r.t. dist_H such that

$$\text{dist}_H(M(p_k + v), \tilde{M}_k(p_k + v)) \leq d. \tag{4.80}$$

Indeed,

$$F(p_k + v) = L(p_k) + A(p_k)v + B(p_k, v) + M(p_k + v),$$

inequalities (4.5) hold, and the Hausdorff distance between the sets $\Phi_k(p_k+v)$ and $F(p_k + v)$ is preserved when these sets are shifted by the same vector $-(L(p_k) + A(p_k)v + B(p_k, v))$.

The aim is to prove the existence of a sequence $\{v_k\}$ such that

$$p_{k+1} + v_{k+1} \in L(p_k) + A(p_k)v_k + B(p_k, v_k) + \tilde{M}_k(p_k + v_k), \quad k \in \mathbb{Z}. \tag{4.81}$$

As in the previous proof, it is enough to define

$$\tilde{\sigma}_k = \text{Dev}(p_{k+1}, L(p_k) + B(p_k, v_k) + \tilde{M}(p_k + v_k)).$$

and to show that there exists a sequence of vectors $V = \{v_k\}$ such that

$$\tilde{\sigma}_k = v_{k+1} - A(p_k)v_k \tag{4.82}$$

and inequality (4.78) holds.

Indeed, it follows from (4.82) that

$$p_{k+1} + v_{k+1} = p_{k+1} + \tilde{\sigma}_k + A(p_k)v_k$$

$$\in L(p_k) + B(p_k, v_k) + A(p_k)v_k + \tilde{M}(p_k + v_k) = \Phi_k(p_k + v_k).$$

For $|v_k| \le a$,

$$|\tilde{\sigma}_k| = \operatorname{dist}(p_{k+1}, L(p_k) + B(p_k, v_k) + \tilde{M}(p_k + v_k))$$

$$\le \operatorname{dist}(p_{k+1}, F(p_k)) + \operatorname{dist}_H(L(p_k) + M(p_k), L(p_k) + B(p_k, v_k) + \tilde{M}(p_k + v_k))$$

$$= \operatorname{dist}_H(M(p_k), B(p_k + v_k) + \tilde{M}(p_k + v_k))$$

$$\le |B(p_k + v_k)| + \operatorname{dist}_H(M(p_k), M(p_k + v_k)) + \operatorname{dist}_H(M(p_k + v_k), \tilde{M}_k(p_k + v_k))$$

$$\le d + 2l|v_k|,$$

which corresponds to estimate (4.57).

Now the operator $\tilde{T} : H \to H$ is defined by

$$Q(p_{k+1})w_{k+1} = Q(p_{k+1})A(p_k)v_k + Q(p_{k+1})\tilde{\sigma}_k \tag{4.83}$$

and

$$P(p_k)w_k = \Gamma(P(p_{k+1})[v_{k+1} - \tilde{\sigma}_k - A(p_k)Q(p_k)v_k]) \tag{4.84}$$

with H and Γ defined in (4.61) and via (4.64). (Of course, $P(x_k)$ must be replaced by $P(p_k)$ etc. in these definitions.) Since the estimates of $|\sigma_k|$ and $|\tilde{\sigma}_k|$ are the same and the operators $A(p_k)$ have the same properties as the operators $A(x_k)$ in Theorem 77, the rest of the proof is identical with that of Theorem 77, and all constants remain unchanged. □

Remark 81. *The contractive set-valued mappings F discussed in Section 4.2 are hyperbolic according to the definition proposed above. For any $\lambda \in (0, 1)$, take $E^s(x) = \mathbb{R}^m, E^u(x) = \{0\}$, and $L(x) = 0$ (thus, $A = 0$) for all $x \in \mathbb{R}^m$. Then conditions (P1), (P2), and (4.45) hold with $N = 1$ and any $l, \kappa > 0$, while inequalities (4.46) are a reformulation of the contractivity condition* (4.7).

4.4 Selection-based hyperbolicity

In the previous section, a hyperbolicity concept for set-valued mappings was proposed which generalizes both, the classical single-valued hyperbolicity condition as well as the set-valued contractive case discussed in Section 4.2. The drawback of this approach is that it imposes rigid restrictions on the behaviour of the set-valued mappings. On the other hand, the assumptions imposed on the mappings are relatively easy to check.

It is possible to state a more general selection-based hyperbolicity condition for set-valued dynamical systems and to prove that every such system has the shadowing and inverse shadowing property. In a way, this selection-based approach is the first satisfactory notion of hyperbolicity from the point of view of set-valued analysis. Unfortunately, it turns out to be very difficult to specify sufficient conditions which can be verified easily, because it is impossible to characterize dynamics by means of the set-valued differentials proposed in the literature (see Section 1.3).

4.4.1 Shadowing theorems

Throughout this section, the following notion of hyperbolicity will be considered.

Definition 82. *A set-valued mapping $F : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ is said to be hyperbolic if it is locally parametrized by a family of hyperbolic selections:*

(P1) For every $x \in \mathbb{R}^m$ there exist linear subspaces $E^s(x), E^u(x) \subset \mathbb{R}^m$ called the stable and unstable subspaces, respectively, such that

$$E^s(x) \oplus E^u(x) = \mathbb{R}^m. \tag{4.85}$$

If $Q(x)$ and $P(x)$ are the corresponding complementary projections from $\mathbb{R}^m$ to $E^s(x)$ and $E^u(x)$, then there exists an $N \geq 1$ such that

$$|Q(x)|, |P(x)| \leq N \tag{4.86}$$

for all $x \in \mathbb{R}^m$.

(P2) There exist constants $\lambda \in (0, 1)$, $\kappa > 0$, $l > 0$, and $a > 0$ such that for every point $(x, z) \in \mathrm{graph}(F)$ there exists a local selection f_z of F, which is a single-valued function $f_z : B_a(x) \to \mathbb{R}^m$ with $f_z(x) = z$,

$f_z(x') \in F(x')$ for all $x' \in B_a(x)$, and such that the following property holds: For any $y, v \in \mathbb{R}^m$ with $|v| \le a$ and $|z - y| \le a$,

$$f_z(x + v) = z + A_z(x)v + b_z(x, v), \tag{4.87}$$

where the $A_z(x) : \mathbb{R}^m \to \mathbb{R}^m$ is a linear map, the restriction

$$P(y)A_z(x)_{|_{E^u(x)}} : E^u(x) \to E^u(y) \tag{4.88}$$

is an isomorphism such that

$$|P(y)A_z(x)P(x)v| \ge \lambda^{-1}|P(x)v|, \tag{4.89}$$

$$|P(y)A_z(x)Q(x)v| \le \kappa|Q(x)v|, \tag{4.90}$$

$$|Q(y)A_z(x)P(x)v| \le \kappa|P(x)v|, \tag{4.91}$$

$$|Q(y)A_z(x)Q(x)v| \le \lambda|Q(x)v|, \tag{4.92}$$

and $b_z(x, \cdot)$ is a small perturbation continuous in v and bounded by

$$|b_z(x, v)| \le l|v|. \tag{4.93}$$

Formula (4.87) and the above condition on b_z imply that f_z is continuous for $|v| \le a$.

Remark 83. *The definition of hyperbolicity stated in Section 4.3 is a special case of the general definition given above.*

In Section 4.3, set-valued mappings of the form

$$F(x) = L(x) + M(x) \tag{4.94}$$

were considered, where $L : \mathbb{R}^m \to \mathbb{R}^m$ is a continuous single-valued mapping, and $M : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ is a set-valued mapping with compact and convex images. It was assumed that there exist constants $N \ge 1$, $\lambda \in (0, 1)$, $\kappa > 0$, $l > 0$, and $a > 0$ such that

- *condition (P1) above is satisfied;*
- *if $x, y, v \in \mathbb{R}^m$ satisfy the inequalities $|v| \le a$ and $\operatorname{dist}(y, F(x)) \le a$, then $L(x + v)$ can be represented as*

  $$L(x + v) = L(x) + A(x)v + b(x, v),$$

  where $A(x) : \mathbb{R}^m \to \mathbb{R}^m$ is a linear map that is continuous in x and such that (after the replacement of $A_z(x)$ by $A(x)$ in (P2)) the restriction (4.88) is an isomorphism that satisfies estimate (4.89), and inequalities (4.90)-(4.92) hold;

- $|b(x,v)| \le l|v|$;
- $\operatorname{dist}_H(M(x+v), M(x)) \le l|v|$.

Take a mapping F be of the form (4.94) *that satisfies the above conditions, let $z \in F(x)$, and define the corresponding local selection f_z by*

$$f_z(x+v) = z + A(x)v + \operatorname{Dev}(z + A(x)v, F(x+v)),$$

i.e. take $A(x)$ as $A_z(x)$ and set

$$b_z(x,v) = \operatorname{Dev}(z + A(x)v, F(x+v)).$$

Clearly, $f_z(x) = z$, $f_z(x+v) \in F(x+v)$, and $A_z(x) = A(x)$ satisfies the corresponding properties formulated in (P2) above. Since F is convex and continuous, b_z is continuous.

The inclusion $z \in F(x)$ implies that

$$\operatorname{dist}(z + A(x)v, F(x+v)) \le \operatorname{dist}(F(x) + A(x)v, F(x+v))$$
$$= \operatorname{dist}(L(x) + A(x)v + M(x), L(x) + A(x)v + b(x,v) + M(x+v))$$
$$= \operatorname{dist}(M(x), b(x,v) + M(x+v))$$
$$\le |b(x,v)| + \operatorname{dist}_H(M(x), M(x+v)) \le 2l|v|,$$

and inequality (4.93) is verified (with l replaced by $2l$).

The following shadowing theorem reduces the shadowing problem to a discussion of selections.

Theorem 84. *Let $F : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ be a hyperbolic set-valued mapping such that*

$$\lambda + \kappa + 2lN < 1.$$

Then F has the (d, Ld)-shadowing property whenever $d < a/L$, where

$$L^{-1} = \frac{1}{2N}(1 - \lambda - \kappa - 2lN).$$

Proof. For every $k \in \mathbb{Z}$, fix $y = x_{k+1}$ and a point $z \in \text{Proj}(x_{k+1}, F(x_k))$. Then $|y - z| < a$, and there exists a continuous hyperbolic selection f_z of F (denoted below f_k) such that

$$f_k(x_k + v) = f_k(x_k) + A_k(x_k)v + b_k(x_k, v) \in F(x_k + v), \quad |v| \leq a,$$

according to (P2). Thus any sequence v_k with $|v_k| \leq Ld$ such that

$$x_{k+1} + v_{k+1} = f_k(x_k + v_k)$$

yields a shadowing trajectory.

Take $b = d/(2L)$ and define

$$H_k := \{v \in \mathbb{R}^m : |P(x_k)v|, |Q(x_k)v| \leq b\}$$

and $H := \prod_{k\in\mathbb{Z}} H_k$. Note that if $v \in H_k$, then $|v| \leq 2b = Ld$. Since each H_k is compact and convex, so is H w.r.t. the Tikhonov topology.

The mapping $G_k : U(x_k) \to U(x_{k+1})$ given by

$$G_k(w) := -P(x_{k+1})A_k(x_k)w \tag{4.95}$$

satisfies $G_k(0) = 0$,

$$|G_k(w)| \geq \lambda^{-1}|w|, \quad w \in U(x_k), \tag{4.96}$$

and $G_k(B^b_{x_k}) \supset B^{b/\lambda}_{x_{k+1}}$, where

$$B^c_x := \{z \in E^u(x) : |z| \leq c\}, \tag{4.97}$$

because of property (P2). Thus the inverse G_k^{-1} of G_k is defined on $B^{b/\lambda}_{x_{k+1}}$, and

$$|G_k^{-1}(z) - G_k^{-1}(z')| \leq \lambda|z - z'|, \quad z, z' \in B^{b/\lambda}_{x_{k+1}}. \tag{4.98}$$

The operator $T : H \to H$ which is given by

$$Q(x_{k+1})T_{k+1}(V) := Q(x_{k+1})\left(f_k(x_k + v_k) - x_{k+1}\right) \tag{4.99}$$

and

$$\begin{aligned} P(x_k)T_k(V) \;\; := \;\; & G_k^{-1}\left(P(x_{k+1})\left\{b_k(x_k, v_k)\right.\right. \\ & \left.\left. + A_k(x_k)Q(x_k)v_k + f_k(x_k) - x_{k+1} - v_{k+1}\right\}\right) \end{aligned} \tag{4.100}$$

for $V = \{v_k\}_{k\in\mathbb{Z}} \in H$, is well-defined. The argument in (4.100) satisfies

$$
\begin{aligned}
&|P(x_{k+1})\left\{b_k(x_k, v_k) + A_k(x_k)Q(x_k)v_k + f_k(x_k) - x_{k+1} - v_{k+1}\right\}| \\
&\le Nl|v_k| + \kappa|Q(x_k)v_k| + Nd + b \le 2lNb + \kappa b + Nd + b \\
&\le (2lN + \kappa + \frac{Nd}{b} + 1)b \le \lambda^{-1}b
\end{aligned}
$$

for $V = \{v_k\}_{k\in\mathbb{Z}} \in H$, because

$$
b^{-1} = \frac{1}{Nd}(1 - \lambda - \kappa - 2lN) \le \frac{1}{Nd}(\lambda^{-1} - 1 - \kappa - 2lN),
$$

so that the argument in (4.100) is an element of $B^{b/\lambda}_{x_{k+1}}$. Furthermore,

$$
\begin{aligned}
&|Q(x_{k+1})T_{k+1}(V)| \\
&\le |Q(x_{k+1})A_k(x_k)P(x_k)v_k| + |Q(x_{k+1})A_k(x_k)Q(x_k)v_k| \\
&\quad + |Q(x_{k+1})b_k(x_k, v_k)| + |Q(x_{k+1})(f_k(x_k) - x_{k+1})| \\
&\le \kappa|P(x_k)v_k| + \lambda|Q(x_k)v_k| + lN|v_k| + Nd \le \kappa b + \lambda b + 2lNb + \frac{2N}{L}b = b,
\end{aligned}
$$

and $T(V) \in H$. The operator T is continuous w.r.t. the Tikhonov topology, because every component T_k depends on v_{k-1}, v_k, v_{k+1} only. Hence T has a fixed point $V \in H$, which implies that

$$
Q(x_{k+1})v_{k+1} = Q(x_{k+1})\left(f_k(x_k + v_k) - x_{k+1}\right) \tag{4.101}
$$

and

$$
\begin{aligned}
&-P(x_{k+1})A_k(x_k)P(x_k)v_k = G_k(P(x_k)v_k) \\
&= P(x_{k+1})\left\{b_k(x_k, v_k) + A_k(x_k)Q(x_k)v_k + f_k(x_k) - x_{k+1} - v_{k+1}\right\}
\end{aligned}
$$

or

$$
P(x_{k+1})v_{k+1} = P(x_{k+1})(f_k(x_k + v_k) - x_{k+1}). \tag{4.102}
$$

By (4.101) and (4.102), the sequence $\eta := \{p_k\}_{k\in\mathbb{Z}}$ with $p_k = x_k + v_k$ is the desired shadowing trajectory. $\square$

As the line of argument for inverse shadowing is very similar to the previous discussion, only the elements of the setup which have to be modified will be highlighted.

A mapping $F : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ is said to be hyperbolic at a given trajectory $\eta = \{p_k\}$ if there exist constants $N \geq 1$, $a, \kappa, l > 0$, and $\lambda \in (0, 1)$ such that condition (P1) holds for points $x = p_k$, and condition (P2) holds for points $x = p_k$, $y = z = p_{k+1}$, and vectors v with $|v| \leq a$.

It is possible to consider two classes of sequences of mappings that approximate the set-valued mapping F. Fix a number $d > 0$.

Class 1.

Consider a sequence of mappings

$$\Phi = \{\Phi_k : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)\}$$

such that each Φ_k is continuous w.r.t. dist_H and

$$\text{dist}(F(p_k + v), \Phi_k(p_k + v)) \leq d \text{ for } k \in \mathbb{Z} \text{ and } |v| \leq a. \tag{4.103}$$

Class 2.

Let

$$CS(\Psi, x, a) = \{\psi \in C(B_a(x), \mathbb{R}^m) \ : \ \psi(y) \in \Psi(y), \quad y \in B_a(x)\}$$

be the set of all continuous local selections of a set-valued mapping Ψ and let $C(B_a(x), \mathbb{R}^m)$ be equipped with the supremum norm.

Consider a sequence of mappings

$$\Phi = \{\Phi_k : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)\}$$

such that

$$\text{dist}(CS(F, p_k, a), CS(\Phi_k, p_k, a)) \leq d, \quad k \in \mathbb{Z}. \tag{4.104}$$

For both classes, a sequence of points $x_k \in \mathbb{R}^m$ is a trajectory of the sequence Φ if $x_{k+1} \in \Phi_k(x_k)$.

Remark 85. *According to my opinion, the following problem has received far too little attention:*

Which conditions can be imposed on two set-valued mappings F and G from some set K to $\mathbb{R}^m$ such that the sets $CS(F)$ and $CS(G)$ of their continuous selections are close to each other w.r.t. the Hausdorff distance in the space of continuous functions, and is it possible to relate this distance $\mathrm{dist}_H(CS(F), CS(G))$ to the Hausdorff distance $\mathrm{dist}_H(F(\cdot), G(\cdot))$ between the images of F and G?

It is well-known that convexity and compactness of the images together with continuity imply that

$$\mathrm{dist}_H(CS(F), CS(G)) \leq \sup_{x \in K} \mathrm{dist}_H(F(x), G(x)), \tag{4.105}$$

but these conditions are by no means necessary: Consider a continuous mapping F with arbitrary images and define $G(\cdot) := v + F(\cdot)$, where v is any vector. Then the Hausdorff distance between the sets of continuous selections of both mappings is $|v| = \mathrm{dist}_H(F(x), G(x))$.

As most fixed point theorems can only be applied to continuous functions, this question has a considerable impact on existence results.

Theorem 86. *Assume that a set-valued mapping $F : \mathbb{R}^m \to \mathcal{A}(\mathbb{R}^m)$ is hyperbolic at a trajectory $\eta = \{p_k\}$ in the above sense. If*

$$\lambda + \kappa + 2lN < 1, \tag{4.106}$$

then F has the inverse (a, d, Ld)-shadowing property: Whenever a family Φ of mappings is in one of the above classes with $d < a/L$, there exists a trajectory $\xi = \{x_k\}$ of Φ such that

$$\|\xi - \eta\|_\infty \leq Ld,$$

where

$$L^{-1} = \frac{1}{2N}(1 - \lambda - \kappa - 2lN).$$

Proof. By assumption, there exist hyperbolic selections f_k of F such that $f_k(p_k) = p_{k+1}$,

$$f_k(p_k + v) = f_k(p_k) + A_k(p_k)v + b_k(p_k, v),$$

$$|P(p_{k+1})A_k(p_k)P(p_k)v| \geq \lambda^{-1}|P(p_k)v|,$$

and so on.

Case 1. Because of (4.103), $\varphi_k(p_k+v) := \mathrm{Proj}(f_k(p_k+v), \Phi_k(p_k+v))$ is a selection of Φ_k such that $|f_k(p_k+v) - \varphi_k(p_k+v)| \le d$ for all $|v| \le a$. Since Φ_k is continuous w.r.t. the Hausdorff distance and has convex values, the φ_k are also continuous according to Theorem 26.

Case 2. Assumption (4.104) implies the existence of continuous selections φ_k of Φ_k such that

$$|f_k(p_k+v) - \varphi_k(p_k+v)| \le d, \quad |v| \le a.$$

In both cases, the aim is to find a sequence v_k with $|v_k| \le Ld$ such that

$$p_{k+1} + v_{k+1} = \varphi_k(p_k+v_k) \in \Phi_k(p_k+v_k).$$

As before, $b = d/(2L)$, $H_k := \{v \in \mathbb{R}^m : |P(p_k)v|, |Q(p_k)v| \le b\}$, and $H := \prod_{k\in\mathbb{Z}} H_k$. Here,

$$G_k(w) := -P(p_{k+1})A_k(p_k)w, \tag{4.107}$$

and the operator $T : H \to H$ is defined by

$$\begin{aligned} Q(p_{k+1})T_{k+1}(V) &:= Q(p_{k+1})\left(\varphi_k(p_k+v_k) - p_{k+1}\right), &(4.108)\\ P(p_k)T_k(V) &:= G_k^{-1}\left(P(p_{k+1})\{b_k(p_k,v_k) + A_k(p_k)Q(p_k)v_k\right. &(4.109)\\ &\quad \left. -(f_k-\varphi_k)(p_k+v_k) + f_k(p_k) - p_{k+1} - v_{k+1}\}\right). \end{aligned}$$

The estimates are essentially unchanged, merely the error Nd is now caused by the term $P(\cdot)(\varphi_k - f_k)(\cdot)$ instead of $P(\cdot)(f_k(x_k) - x_{k+1})$ as before. □

4.4.2 Application to polytope-valued mappings

Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be a polytope-valued mapping, i.e. a set-valued mapping which is characterized by its vertices $s_1, \ldots, s_n : \mathbb{R}^m \to \mathbb{R}^m$ via

$$F(x) = co\{s_1(x), \ldots, s_n(x)\} \text{ for all } x \in \mathbb{R}^m. \tag{4.110}$$

Assume that there exist $N \ge 1$, $a, \kappa, l > 0$, and $\lambda \in [0,1]$ such that

(P1′) condition (P1) of Section 4.4.1 holds, and the dimensions of the spaces $E^u(x)$ are the same for all $x \in \mathbb{R}^m$.

(P2′) For any $x, y, v \in \mathbb{R}^m$ with $|v| \leq a$ and $|s_i(x) - y| \leq a$, the vertices can be represented as

$$s_i(x + v) = s_i(x) + A_i(x)v + b_i(x, v) \tag{4.111}$$

for $1 \leq i \leq n$, where any $A_i(x) : \mathbb{R}^m \to \mathbb{R}^m$ is a linear map such that for each v there exists a direction of expansion $p(x, v) \in \mathbb{R}^m$ with $|p(x, v)| = 1$ and

$$\langle p(x, v), P(y)A_i(x)P(x)v \rangle \geq \lambda^{-1}|P(x)v|, \tag{4.112}$$

analogs of conditions (4.90)-(4.92) hold (with $A_z(x)$ replaced by $A_i(x)$), and $b_i(x, \cdot)$ are small continuous perturbations for which analog of condition (4.93) is valid.

Remark 87. *From the geometric point of view, inequality (4.112) ensures that the unstable perturbations $P(y)A_i(x)P(x)v$ drive all vertices in the same direction, so that their movements cannot cancel each other when combined.*

In the case of polytope-valued mappings, the general notion of hyperbolicity introduced in Section 4.4.1 is implied by conditions on the behavior of a finite set of points. Please note that there are polytope-valued mappings which satisfy the conditions of Theorem 88, but not the restrictive setup of Section 4.3, because inequality (4.112) bounds the expansion of each single vertex from below but not from above.

Theorem 88. *Let $F : \mathbb{R}^m \to \mathcal{CC}(\mathbb{R}^m)$ be a polytope-valued mapping such that its vertices satisfy conditions (P1′) and (P2′). Assume that the projections P and Q are Lipschitz continuous with Lipschitz constant $K \geq 0$ such that*

$$K \operatorname{diam}(F(x)) \max_{1 \leq i \leq n} \|A_i(x)\| < \lambda^{-1}, \quad x \in \mathbb{R}^m. \tag{4.113}$$

If

$$\lambda_0 := \sup_{x \in \mathbb{R}^m} \max(\lambda_1(x), \lambda_2(x)) < 1,$$

where

$$\lambda_1(x) := (\lambda^{-1} - K \operatorname{diam}(F(x)) \max_{1 \leq i \leq n} \|A_i(x)\|)^{-1}$$

and

$$\lambda_2(x) := \lambda + K \operatorname{diam}(F(x)) \max_{1 \leq i \leq n} \|A_i(x)\|,$$

then F is a hyperbolic set-valued mapping with constants N, λ_0,

$$\kappa_0 := \kappa + \sup_{x \in \mathbb{R}^m} K \operatorname{diam}(F(x)) \max_{1 \le i \le n} \|A_i(x)\|,$$

l, and a.

Proof. Let any point $(x, z) \in \operatorname{graph}(F)$ be given. Because of (4.110), there exist $\theta_1, \dots, \theta_n \in [0, 1]$ with $\sum_{i=1}^n \theta_i = 1$ and

$$z = \sum_{i=1}^{n} \theta_i s_i(x).$$

Define the selection $f_z : \mathbb{R}^m \to \mathbb{R}^m$ as the convex combination

$$f_z(x') := \sum_{i=1}^{n} \theta_i s_i(x'). \tag{4.114}$$

of the vertices of F with the above coefficients. Then

$$\begin{aligned} f_z(x + v) &= \sum_{i=1}^{n} \theta_i s_i(x + v) = \sum_{i=1}^{n} \theta_i \left(s_i(x) + A_i(x)v + b_i(x, v)\right) \\ &= s(x) + \sum_{i=1}^{n} \theta_i A_i(x)v + \sum_{i=1}^{n} \theta_i b_i(x, v) =: z + A(x)v + b(x, v). \end{aligned}$$

In order to check condition (P2), take y with $|y - z| \le a$ and define $y_i = y - z + s_i(x)$, so that $|y_i - s_i(x)| \le a$. Since the projections P are

Lipschitz continuous with Lipschitz constant K,

$$\begin{aligned}
& |P(y)A(x)P(x)v| \\
\geq\ & \langle p(x,v), P(y)A(x)P(x)v\rangle \\
=\ & \langle p(x,v), \sum_{i=1}^{n} \theta_i P(y)A_i(x)P(x)v\rangle \\
=\ & \langle p(x,v), \sum_{i=1}^{n} \theta_i P(y_i)A_i(x)P(x)v\rangle \\
& + \langle p(x,v), \sum_{i=1}^{n} \theta_i \left(P(y) - P(y_i)\right) A_i(x)P(x)v\rangle \\
\geq\ & \lambda^{-1}|P(x)v| - K\,\mathrm{diam}(F(x)) \max_{1\leq i\leq n} \|A_i(x)\||P(x)v| \\
=\ & \left(\lambda^{-1} - K\,\mathrm{diam}(F(x)) \max_{1\leq i\leq n} \|A_i(x)\|\right)|P(x)v| \\
=\ & \lambda_1^{-1}(x)|P(x)v|
\end{aligned}$$

by estimates (4.112), which implies that the restriction

$$P(y)A(x)_{|_{E^u(x)}} : E^u(x) \to E^u(y)$$

is an isomorphism (the dimensions of $E^u(x)$ and $E^u(y)$ coincide). The same estimate proves inequality (4.89).

To prove inequalities (4.90)-(4.92), note that

$$\begin{aligned}
& |P(y)A(x)Q(x)v| \\
=\ & |P(y)\sum_{i=1}^{n} \theta_i A_i(x)Q(x)v| \\
\leq\ & |\sum_{i=1}^{n} \theta_i P(y_i)A_i(x)Q(x)v| + |\sum_{i=1}^{n} \theta_i \left(P(y) - P(y_i)\right) A_i(x)Q(x)v| \\
\leq\ & \kappa|Q(x)v| + K\,\mathrm{diam}(F(x)) \max_{1\leq i\leq n} \|A_i(x)\||Q(x)v| \\
=\ & \left(\kappa + K\,\mathrm{diam}(F(x)) \max_{1\leq i\leq n} \|A_i(x)\|\right)|Q(x)v| \\
\leq\ & \kappa_0|Q(x)v|,
\end{aligned}$$

$$\begin{aligned}
& |Q(y)A(x)P(x)v| \\
= \;& |Q(y)\sum_{i=1}^{n}\theta_i A_i(x)P(x)v| \\
\le \;& |\sum_{i=1}^{n}\theta_i Q(y_i)A_i(x)P(x)v| + |\sum_{i=1}^{n}\theta_i \left(Q(y)-Q(y_i)\right)A_i(x)P(x)v| \\
\le \;& \kappa|P(x)v| + K\,\mathrm{diam}(F(x)) \max_{1\le i\le n}\|A_i(x)\||P(x)v| \\
= \;& \left(\kappa + K\,\mathrm{diam}(F(x)) \max_{1\le i\le n}\|A_i(x)\|\right)|P(x)v| \\
\le \;& \kappa_0|P(x)v|,
\end{aligned}$$

and

$$\begin{aligned}
& |Q(y)A(x)Q(x)v| \\
= \;& |Q(y)\sum_{i=1}^{n}\theta_i A_i(x)Q(x)v| \\
\le \;& |\sum_{i=1}^{n}\theta_i Q(y_i)A_i(x)Q(x)v| + |\sum_{i=1}^{n}\theta_i \left(Q(y)-Q(y_i)\right)A_i(x)Q(x)v| \\
\le \;& \lambda|Q(x)v| + K\,\mathrm{diam}(F(x)) \max_{1\le i\le n}\|A_i(x)\||Q(x)v| \\
= \;& \left(\lambda + K\,\mathrm{diam}(F(x)) \max_{1\le i\le n}\|A_i(x)\|\right)|Q(x)v| \\
= \;& \lambda_2(x)|Q(x)v|.
\end{aligned}$$

Finally,

$$|b(x,v)| \le \sum_{i=1}^{n}\theta_i|b_i(x,v)| \le l|v|,$$

which proves estimate (4.93). □

Corollary 89. *If the assumptions of Theorem 88 hold and*

$$\lambda_0 + \kappa_0 + 2lN < 1,$$

then F has the shadowing property due to Theorem 84 and the inverse shadowing property according to Theorem 86.

Chapter 5

An application: The Viability Kernel Algorithm

Viability kernels (cf. Section 2.6) of differential inclusions are of considerable interest, because many theoretical and practical problems can be reformulated as viability problems.

The viability approach enjoys an increasing popularity in a variety of applications where constrained dynamics are analyzed. It has been used e.g. in [6] in oder to derive conditions under which ecosystems are viable in the sense that no species dies out.

Most problems related to the prevention of collisions arising from traffic control or robotics can be reformulated as viability problems in a natural way: The set K of desirable states is defined as the union of all states in which the distance between the vehicles or between robot and obstacles, respectively, is bigger than some given safety distance. In this setup, the viability kernel is the set of states from which collisions can successfully be prevented.

Unfortunately, it is very difficult to calculate viability kernels analytically and thus reliable numerical methods are required. In [17], Frankowska and Quincampoix proposed a first algorithm for the computation of viability kernels, and Saint-Pierre succeeded to prove the convergence of a fully discretized and hence implementable algorithm in [35]. This Viability Kernel Algorithm was later generalized to impulsive differential inclusions in [9]. However, it is still an open question how fast this algorithm converges, and until now no error estimates have been available.

As viability theory describes the behaviour of trajectories on the unbounded time interval $[0, \infty)$, it seems natural to use shadowing results as tools in this context. The aim of this chapter is to derive the first rigorous estimates for the accuracy of the fully discretized viability algorithm.

5.1 Algorithm and general estimates

Consider the autonomous differential inclusion

$$\dot{x}(t) \in F(x(t)) \text{ for almost every } t \geq 0, \quad x(0) = x_0 \in \mathbb{R}^m \tag{5.1}$$

and its time-h-flow

$$G_h : \mathbb{R}^m \rightrightarrows \mathbb{R}^m, \quad x \mapsto \mathcal{R}(h, 0, x), \tag{5.2}$$

where $\mathcal{R}(h, 0, x)$ denotes the reachable set of (5.1) at time h. For any $\rho > 0$ let $X_\rho \subset \mathbb{R}^m$ be a countable subset with

$$\forall x \in \mathbb{R}^m, \ \exists x_\rho \in X_\rho \text{ with } |x - x_\rho| \leq \rho. \tag{5.3}$$

In most applications, X_ρ is simply a subgrid of $\mathbb{R}^m$. Given any subset $A \subset \mathbb{R}^m$ and $\varepsilon > 0$, define

$$A^\varepsilon := A + \varepsilon B \text{ and } A_\rho^\varepsilon := A^\varepsilon \cap X_\rho. \tag{5.4}$$

Consider the semi-discretized Euler scheme

$$\Gamma_h : \mathbb{R}^m \rightrightarrows \mathbb{R}^m, \quad x \mapsto x + hF(x) \tag{5.5}$$

and the fully discretized scheme

$$\Gamma_{h,\rho} : X_\rho \rightrightarrows X_\rho, \quad x_\rho \mapsto \big(x_\rho + hF(x_\rho) + (2 + Lh)\rho B\big) \cap X_\rho, \tag{5.6}$$

where $L > 0$ will be a Lipschitz constant of a restriction of F (see assumption (iii) below).

The Viability Kernel Algorithm for the computation of the viability kernel $\text{Viab}_F(K)$ of a compact set K is straight forward:

1. Set $V_0 := K \cap X_\rho$ and $Z_0 := \emptyset$.

2. For each $x_\rho \in V_0$: If $\Gamma_{h,\rho}(x_\rho) \cap V_0 = \emptyset$, set $Z_0 := Z_0 \cup \{x_\rho\}$.
3. Set $V_1 := V_0 \backslash Z_0$ and $Z_1 := \emptyset$.
4. For each $x_\rho \in V_1$, ...

The sequence $V_0 \supset V_1 \supset V_2 \supset \ldots$ will eventually be constant. The computed set $V_\infty := \cap_{n=0}^{\infty} V_n$ is the largest weakly positively invariant set, i.e. the discrete viability kernel, under the inflated Euler method.

It should be mentioned that the Viability Kernel Algorithm is closely related to the well-known subdivision method for the approximation of attractors and unstable manifolds of ODEs proposed by Dellnitz and Hohmann, see [11]. While the subdivision method computes in every step a covering of the desired object which is backward invariant w.r.t. a suitable space discretization, the Viability Kernel Algorithm computes a forward invariant subset of a grid. Whenever the time-h flow is locally Lipschitz continuous, these concepts coincide up to a reversal of time, so that the subdivision method could be regarded as a particularly important special case of the Viability Kernel Algorithm.

Throughout this chapter, the following assumptions will be supposed:

(i) The viability kernel $\mathrm{Viab}_F(K)$ is stable in the sense that there exist an $\varepsilon_0 > 0$ and a Lipschitz constant $L_V > 0$ such that for all $0 \le \varepsilon \le \varepsilon_0$, $\mathrm{Viab}_F(K) \subset \mathrm{Viab}_F(K^\varepsilon)$ and

$$\mathrm{dist}(\mathrm{Viab}_F(K^\varepsilon), \mathrm{Viab}_F(K)) \le L_V \varepsilon. \tag{5.7}$$

(ii) There exist a $d_0^{(s)} > 0$ and a $d_0^{(is)} > 0$, possibly dependent on h, such that the h-flow G_h has

(iia) the $(d, \varphi(d))$-shadowing property in K^{ε_0} for $d \in (0, d_0^{(s)}]$ and

(iib) the (global) inverse $(d, \psi(d))$-shadowing property in K for $d \in (0, d_0^{(is)}]$,

where $\varphi, \psi : \mathbb{R}_+ \to \mathbb{R}_+$ are increasing functions with $\lim_{d\to 0} \varphi(d) = 0$ and $\lim_{d\to 0} \psi(d) = 0$, which can also depend on h.

(iii) The mapping F is Lipschitz-continuous in K^{ε_0} with Lipschitz constant $L > 0$ and has compact and convex values.

The following observations will be used frequently throughout this Chapter.

Observation 90. *For any compact set $A \subset \mathbb{R}^m$, the viability kernels can be characterized by*

$$\mathrm{Viab}_G(A) = \{x_0 \in A : \exists (p_n)_{n\in\mathbb{N}} \subset A \text{ with } p_0 = x_0 \text{ and } p_{n+1} \in G(p_n)\ \forall n \in \mathbb{N}\}$$

and

$$\mathrm{Viab}_F(A) = \{x_0 \in A : \exists \text{ a solution } x : [0,\infty) \to A \text{ of } (5.1) \text{ with } x(0) = x_0\}.$$

It is obvious that the right hand sides are the largest viability domains contained in A. Under mild assumptions on F and G they are closed, compare e.g. Theorem 3.5.3 in [1].

Observation 91. *Because of assumption (iii), F is bounded on K^{ε_0} by $\|F\|_\infty = M < \infty$. Thus any solution x of (5.1) remaining in K^{ε_0} satisfies*

$$|x(t) - x(0)| \le \int_0^t |\dot{x}(s)| ds \le Mt. \tag{5.8}$$

If $x(0) \in K^\varepsilon$ with $0 < \varepsilon < \varepsilon_0$, it follows that $x(t) \in K^{\varepsilon_0}$ for all $t \in [0,h]$ with $0 < h < \frac{\varepsilon_0 - \varepsilon}{M}$. Otherwise $0 < t_0 := \inf\{t \in [0,h] : x(t) \notin K^{\varepsilon_0}\} < h$ and (5.8) holds for all $0 \le t \le t_0$. But then

$$|x(t_0) - x(0)| \le Mt_0 < Mh \le \varepsilon_0 - \varepsilon$$

implies that $x(t_0)$ is in the interior of K^{ε_0}, which is a contradiction. Thus (5.8) holds for all $t \in [0,h]$. If $x(0) \in K^\varepsilon$ and $x(h) \in K^\varepsilon$, (5.8) can be applied forwards in time from $x(0)$ and backwards in time from $x(h)$ in order to obtain

$$\mathrm{dist}(x(s), K^\varepsilon) \le \frac{1}{2} Mh\ \forall s \in [0,h]. \tag{5.9}$$

Lemma 92. *The relations $\mathrm{Viab}_F(K) \subset \mathrm{Viab}_{G_h}(K^\varepsilon)$ and*

$$\mathrm{dist}(\mathrm{Viab}_{G_h}(K^\varepsilon), \mathrm{Viab}_F(K)) \le L_V(\varepsilon + \frac{1}{2}Mh)$$

hold whenever $0 \le \varepsilon < \varepsilon_0$ and $0 < Mh < \varepsilon_0 - \varepsilon$.

Proof. Obviously $\mathrm{Viab}_F(K) \subset \mathrm{Viab}_{G_h}(K^\varepsilon)$. But

$$\begin{aligned} x_0 \in \mathrm{Viab}_{G_h}(K^\varepsilon) &\Rightarrow \exists (p_n)_{n\in\mathbb{N}} \subset K^\varepsilon : \; p_0 = x_0, \; p_{n+1} \in G_h(p_n) \; \forall n \in \mathbb{N} \\ &\Rightarrow \exists \text{ a solution } x : [0,\infty) \to \mathbb{R}^m \text{ of } (5.1): \; x(nh) = p_n \in K^\varepsilon \\ &\Rightarrow \mathrm{dist}(x(t), K^\varepsilon) \le \frac{1}{2}Mh \; \forall t \ge 0 \\ &\Rightarrow x_0 \in \mathrm{Viab}_F(K^{\varepsilon+\frac{1}{2}Mh}) \end{aligned}$$

by Observation 91, and thus

$$\begin{aligned} & \mathrm{dist}(\mathrm{Viab}_{G_h}(K^\varepsilon), \mathrm{Viab}_F(K)) \\ \le\; & \mathrm{dist}(\mathrm{Viab}_{G_h}(K^\varepsilon), \mathrm{Viab}_F(K^{\varepsilon+\frac{1}{2}Mh})) + \mathrm{dist}(\mathrm{Viab}_F(K^{\varepsilon+\frac{1}{2}Mh}), \mathrm{Viab}_F(K)) \\ \le\; & L_V(\varepsilon + \frac{1}{2}Mh) \end{aligned}$$

by assumption (i). □

The following Lemma is contained implicitly in many works, because it estimates the local error of the semi-discretized Euler-scheme.

Lemma 93. *The error of approximation between G_h and Γ_h is*

$$\mathrm{dist}_H(G_h(x_0), \Gamma_h(x_0)) \le Mh(e^{Lh} - 1)$$

for all $0 < \varepsilon < \varepsilon_0$, $x_0 \in K^\varepsilon$, and $h > 0$ such that $Mhe^{Lh} \le \varepsilon_0 - \varepsilon$.

Proof. Let a solution $x : [0,h] \to \mathbb{R}^m$ of (5.1) with $x(0) = x_0$ be given. Because of (5.8), $x(s) \in K^{\varepsilon_0}$ for all $s \in [0,h]$. As F has convex values, the image of the Euler-step $\Gamma_h(x_0)$ is identical with the reachable set of the constant differential inclusion

$$\dot{e}(t) \in F(x_0), \quad e(0) = x_0 \tag{5.10}$$

at time h. Since

$$\mathrm{dist}(\dot{x}(t), F(x(0))) \le \mathrm{dist}(F(x(t)), F(x(0))) \le L|x(t) - x(0)| \le LMt,$$

the Filippov Theorem (cf. Theorem 31) guarantees the existence of a solution $e : [0,h] \to \mathbb{R}^m$ of (5.10) satisfying

$$|x(t) - e(t)| \le \int_0^t e^{L(t-s)} LMt ds = Mt(e^{Lt} - 1)$$

for all $t \in [0, h]$, and in particular

$$\operatorname{dist}(x(h), \Gamma_h(x_0)) \le |x(h) - e(h)| \le Mh(e^{Lh} - 1).$$

Conversely, let $\eta \in F(x_0)$ be given and consider the corresponding linear trajectory $e(t) := x_0 + t\eta$ for $t \in [0, h]$ of Γ_h. As

$$\operatorname{dist}(\dot{e}(t), F(e(t))) \le \operatorname{dist}(F(x_0), F(e(t))) \le LMt,$$

the Filippov theorem yields a solution $x : [0, h] \to \mathbb{R}^m$ of (5.1) with $x(0) = x_0$ and

$$|x(t) - e(t)| \le \int_0^t e^{L(t-s)} LMt ds = Mt(e^{Lt} - 1)$$

for all $t \in [0, h]$, and in particular

$$\operatorname{dist}(e(h), G_h(x_0)) \le |e(h) - x(h)| \le Mh(e^{Lh} - 1).$$

□

5.2 Estimates using the shadowing and the inverse shadowing property

Lemma 94. *If $Mh(e^{Lh}-1) \le d_0^{(is)}$, $\varepsilon_1 := \psi(Mh(e^{Lh}-1)) \le \varepsilon_0$ and $Mhe^{Lh} \le \varepsilon_0$ then*

$$\operatorname{dist}(\operatorname{Viab}_F(K), \operatorname{Viab}_{\Gamma_h}(K^{\varepsilon_1})) \le \varepsilon_1.$$

Proof. Let $p_0 \in \operatorname{Viab}_F(K) \subset \operatorname{Viab}_{G_h}(K)$ be given. Then there exists an orbit $(p_n)_{n\in\mathbb{N}}$ of G_h such that $p_n \in K$ for all $n \in \mathbb{N}$. As Γ_h is continuous with compact and convex values, Lemma 93 ensures that Γ_h is an approximation of G_h in the sense of assumption (iib), which in turn yields the existence of an orbit $(x_n)_{n\in\mathbb{N}}$ of Γ_h such that $|p_n - x_n| \le \varepsilon_1$. Thus $x_n \in K^{\varepsilon_1}$ for all $n \in \mathbb{N}$, and $x_0 \in \operatorname{Viab}_{\Gamma_h}(K^{\varepsilon_1})$ by Observation 90. □

The following lemma uses a simple fact: For any $A \subset X$, the estimate $\operatorname{dist}(A, A_\rho^\rho) \le \rho$ holds, because for every $a \in A$ there is an $x_\rho \in X_\rho$ such that $|a - x_\rho| \le \rho$, and then $x_\rho \in A^\rho \cap X_\rho$.

Lemma 95. *If $\varepsilon_1 + \rho \le \varepsilon_0$, then*

$$\operatorname{dist}(\operatorname{Viab}_{\Gamma_h}(K^{\varepsilon_1}), \operatorname{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho})) \le \rho.$$

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a viable orbit of Γ_h in K^{ε_1}. By definition, there exists a $\xi_0 \in K_\rho^{\varepsilon_1+\rho}$ such that $|x_0 - \xi_0| \le \rho$. Since

$$\mathrm{dist}(x_0 + hF(x_0), \xi_0 + hF(\xi_0)) \le (1 + Lh)\rho, \tag{5.11}$$

it follows that

$$\mathrm{dist}(x_0 + hF(x_0), \xi_0 + hF(\xi_0) + (1 + Lh)\rho B) = 0, \tag{5.12}$$

and thus

$$\begin{aligned} & \mathrm{dist}(\Gamma_h(x_0), \Gamma_{h,\rho}(\xi_0)) \\ = \; & \mathrm{dist}(x_0 + hF(x_0), (\xi_0 + hF(\xi_0) + (2 + Lh)\rho B) \cap X_\rho) \\ \le \; & \rho. \end{aligned}$$

Thus there exists a $\xi_1 \in \Gamma_{h,\rho}(\xi_0)$ such that $|x_1 - \xi_1| \le \rho$, and by induction there exists a whole orbit $(\xi_n)_{n\in\mathbb{N}}$ of $\Gamma_{h,\rho}$ with $|x_n - \xi_n| \le \rho$ for all $n \in \mathbb{N}$. Consequently $\xi_n \in K_\rho^{\varepsilon_1+\rho}$ for all $n \in \mathbb{N}$, and $\xi_0 \in \mathrm{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho})$. □

Lemma 96. *Let $\varepsilon_2 := \varphi((2+Lh)\rho + Mh(e^{Lh} - 1))$. If $Mhe^{Lh} \le \varepsilon_0 - \varepsilon_1 - \rho$, $Mh \le \varepsilon_0 - \varepsilon_1 - \rho - \varepsilon_2$, and $(2 + Lh)\rho + Mh(e^{Lh} - 1) \le d_0^{(s)}$, then*

$$\mathrm{dist}(\mathrm{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho}), \mathrm{Viab}_F(K)) \le \varepsilon_2 + L_V(\varepsilon_1 + \rho + \varepsilon_2 + \frac{1}{2}Mh).$$

Proof. By Lemma 93,

$$\begin{aligned} & \mathrm{dist}(\Gamma_{h,\rho})(x_\rho), G_h(x_\rho)) \\ \le \; & \mathrm{dist}(x_\rho + hF(x_\rho) + (2 + Lh)\rho B, x_h + hF(x_\rho)) + \mathrm{dist}(x_\rho + hF(x_\rho), G_h(x_\rho)) \\ \le \; & (2 + Lh)\rho + Mh(e^{Lh} - 1) =: d \end{aligned}$$

for every $x_\rho \in K_\rho^{\varepsilon_1+\rho}$. Thus any trajectory $(\xi_n)_{n\in\mathbb{N}}$ of $\Gamma_{h,\rho}$ which is viable in $K_\rho^{\varepsilon_1+\rho}$ is a d-pseudotrajectory of G_h, and assumption (iia) implies the existence of an orbit $(p_n)_{n\in\mathbb{N}}$ of G_h such that $|p_n - \xi_n| \le \varepsilon_2$ for all $n \in \mathbb{N}$. Hence $p_0 \in \mathrm{Viab}_{G_h}(K^{\varepsilon_1+\rho+\varepsilon_2})$ by Observation 1, which means that

$$\begin{aligned} & \mathrm{dist}(\mathrm{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho}), \mathrm{Viab}_F(K)) \\ \le \; & \mathrm{dist}(\mathrm{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho}), \mathrm{Viab}_{G_h}(K_\rho^{\varepsilon_1+\rho+\varepsilon_2})) \\ & + \mathrm{dist}(\mathrm{Viab}_{G_h}(K_\rho^{\varepsilon_1+\rho+\varepsilon_2}), \mathrm{Viab}_F(K)) \\ \le \; & \varepsilon_2 + L_V(\varepsilon_1 + \rho + \varepsilon_2 + \frac{1}{2}Mh) \end{aligned}$$

by Lemma 92. □

Altogether, an estimate for the accuracy of the Viability Kernel Algorithm is obtained:

Theorem 97. *If*

$$\begin{aligned} (2+Lh)\rho + Mh(e^{Lh}-1) &\le d_0^{(s)}, && (5.13)\\ Mh(e^{Lh}-1) &\le d_0^{(is)}, && (5.14)\\ Mhe^{Lh} + \varepsilon_1 + \rho &\le \varepsilon_0, && (5.15)\\ Mh + \varepsilon_1 + \varepsilon_2 + \rho &\le \varepsilon_0, && (5.16) \end{aligned}$$

and if assumptions (i) to (iii) are satisfied, then

$$\begin{aligned} &\mathrm{dist}_H(\mathrm{Viab}_F(K), \mathrm{Viab}_{\Gamma_{h,\rho}}(K_\rho^{\varepsilon_1+\rho})) \\ \le\ & \max\{\varepsilon_1+\rho, \varepsilon_2 + L_V(\varepsilon_1+\rho+\varepsilon_2+\frac{1}{2}Mh)\}. \qquad (5.17) \end{aligned}$$

The conditions of Theorem 97 do not look very appealing. Please note that they can be verified easily in a quite practical sense: In concrete applications, it is usually reasonable to express ρ in terms of h, e.g. $\rho := h^2$. Under moderate assumptions, the left hand sides of (5.13) and (5.14) converge considerably faster to zero in ρ than the corresponding shadowing constants $d_0^{(s)}$ and $d_0^{(is)}$ (compare Section 5.4). If a desired accuracy $\delta > 0$ of the approximation of the viability kernel and concrete monotone functions φ and ψ are given, the inequalities (5.13) to (5.16) together with

$$\max\{\varepsilon_1 + h^2, \varepsilon_2 + L_V(\varepsilon_1 + h^2 + \varepsilon_2 + \frac{1}{2}Mh)\} \le \delta \qquad (5.18)$$

can be regarded as scalar constraints which are monotone w.r.t. h. Thus one can determine the maximal h which respects all constraints using a simple interval subdivision algorithm. The same method can be used in the context of Theorem 100.

5.3 Estimates using the shadowing property only

It is possible to dispense with the inverse shadowing property by inflating the right hand sides of the numerical schemes so much that the numerical errors are 'swallowed' by the inflation. To this end, define

$$\tilde{\Gamma}_h : \mathbb{R}^m \rightrightarrows \mathbb{R}^m, \quad x \mapsto x + hF(x) + Mh(e^{Lh}-1)B \qquad (5.19)$$

and

$$\tilde{\Gamma}_{h,\rho} : X_\rho \rightrightarrows X_\rho, \qquad (5.20)$$
$$x_\rho \mapsto \big(x_\rho + hF(x_\rho) + Mh(e^{Lh} - 1)B + (2 + Lh)\rho B\big) \cap X_\rho.$$

For these schemes the following estimates hold.

Lemma 98. *If $Mhe^{Lh} \le \varepsilon_0$ and $\rho \le \varepsilon_0$,*

$$\mathrm{dist}(\mathrm{Viab}_F(K), \mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho)) \le \rho.$$

Proof. According to Lemma 93,

$$\mathrm{dist}(G_h(x_0), \Gamma_h(x_0)) \le Mh(e^{Lh} - 1)$$

for all $x_0 \in K$, and thus

$$\mathrm{dist}(G_h(x_0), \tilde{\Gamma}_h(x_0)) = 0$$

and

$$\mathrm{dist}(\mathrm{Viab}_{G_h}(K), \mathrm{Viab}_{\tilde{\Gamma}_h}(K)) = 0.$$

Adapting the proof of Lemma 95 to $\tilde{\Gamma}_h$ and $\tilde{\Gamma}_{h,\rho}$ yields the desired result. □

Lemma 99. *If $\varepsilon_3 := \varphi(2Mh(e^{Lh}-1)+(2+Lh)\rho) \le \varepsilon_0 - \rho$, $Mhe^{Lh} \le \varepsilon_0 - \rho$, and $2Mh(e^{Lh} - 1) + (2 + Lh)\rho \le d_0^{(s)}$, then*

$$\mathrm{dist}(\mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho), \mathrm{Viab}_F(K)) \le \varepsilon_3 + L_V(\rho + \varepsilon_3 + \frac{1}{2}Mh).$$

Proof. By Lemma 93,

$$\begin{aligned} & \mathrm{dist}(\tilde{\Gamma}_{h,\rho}(x_\rho), G_h(x_\rho)) \\ \le\; & \mathrm{dist}(x_\rho + hF(x_\rho) + Mh(e^{Lh} - 1)B + (2 + Lh)\rho B, x_\rho + hF(x_\rho)) \\ & + \mathrm{dist}(x_\rho + hF(x_\rho), G_h(x_\rho)) \\ \le\; & 2Mh(e^{Lh} - 1) + (2 + Lh)\rho =: \tilde{d} \end{aligned}$$

for any $x_\rho \in K_\rho^\rho$.

Thus any trajectory $(\xi_n)_{n\in\mathbb{N}}$ of $\tilde{\Gamma}_{h,\rho}$ which is viable in K_ρ^ρ is a $\tilde{d}$-pseudotrajectory of G_ρ, and assumption (iia) implies that there exists an orbit $(p_n)_{n\in\mathbb{N}}$

of G_h such that $|p_n - \xi_n| \leq \varepsilon_3$ for all $n \in \mathbb{N}$. Hence $p_0 \in \mathrm{Viab}_{G_h}(K^{\rho+\varepsilon_3})$, and therefore

$$\begin{aligned} & \mathrm{dist}(\mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho), \mathrm{Viab}_F(K)) \\ \leq \; & \mathrm{dist}(\mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho), \mathrm{Viab}_{G_h}(K_\rho^{\rho+\varepsilon_3})) \\ & + \mathrm{dist}(\mathrm{Viab}_{G_h}(K_\rho^{\rho+\varepsilon_3}), \mathrm{Viab}_F(K)) \\ \leq \; & \varepsilon_3 + L_V(\rho + \varepsilon_3 + \frac{1}{2}Mh) \end{aligned}$$

by Lemma 92. □

Summarizing the following estimate for the accuracy of the Viability Kernel Algorithm can be obtained for systems which have the shadowing but not the inverse shadowing property:

Theorem 100. *If*

$$Mhe^{Lh} + \rho \leq \varepsilon_0, \tag{5.21}$$

$$2Mh(e^{Lh} - 1) + (2 + Lh)\rho \leq d_0^{(s)}, \tag{5.22}$$

$$\varepsilon_3 + \rho \leq \varepsilon_0, \tag{5.23}$$

and if assumptions (i), (iia), and (iii) are satisfied, then

$$\begin{aligned} & \mathrm{dist}_H(\mathrm{Viab}_F(K), \mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho)) \\ \leq \; & \max\{\rho, \varepsilon_3 + L_V(\rho + \varepsilon_3 + \frac{1}{2}Mh)\}. \end{aligned} \tag{5.24}$$

5.4 One-sided Lipschitz right hand sides

In this section, the above results are applied to differential inclusions with relaxed one-sided Lipschitz right hand sides. This discussion should serve as a kind of template which shows how one can derive statements about the accuracy of the viability kernel algorithm from shadowing theorems. The behaviour, the stability properties, and the shadowing properties of these systems are well understood (see [13, 14], and Section 2.5).

Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be a set-valued mapping which is Lipschitz continuous with Lipschitz constant $L > 0$ and satisfies the relaxed one-sided Lipschitz condition with one-sided Lipschitz constant $\mu \in \mathbb{R}$. Then Theorem 41 states that F defines a differential inclusion such that the reachable sets at time $h > 0$ satisfy

$$\mathrm{dist}_H(G_h(x), G_h(x')) = \mathrm{dist}_H(\mathcal{R}(h,0,x), \mathcal{R}(h,0,x')) \leq e^{\mu h}|x - x'| \quad (5.25)$$

for all $x, x' \in \mathbb{R}^m$. If $\mu < 0$, the time-h flow G_h is a contraction with contraction constant $\lambda := e^{\mu h} < 1$. It is a well-known fact that the reachable sets of a differential inclusion with Lipschitz continuous right hand side in a finite dimensional vector space are nonempty and compact, cf. Theorem 30.

As F is Lipschitz continuous on K^{ε_0}, the reachable sets $\mathcal{R}(h,0,x)$ are uniformly bounded for $x \in K^{\varepsilon}$ with $0 < \varepsilon < \varepsilon_0$ and $h > 0$ small enough. Thus $G_h(x) = \mathcal{R}(h,0,x)$ satisfies the assumptions of Theorem 70 and has the $(d, \frac{d}{1-e^{\mu h}})$-shadowing property whenever $d < d_0^{(s)} := \frac{1-e^{\mu h}}{2}(\varepsilon_0 - \varepsilon)$. Please note that $\varphi(d) = \frac{d}{1-e^{\mu h}}$ and that $d_0^{(s)}$ and φ indeed depend on the time-step h.

Repeating the line of argument of section 5.3 in this setup, one obtains

Theorem 101. *Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ be a set-valued mapping with convex and compact values which is Lipschitz continuous with Lipschitz constant $L > 0$ and satisfies the relaxed one-sided Lipschitz condition with one-sided Lipschitz constant $\mu < 0$. Then*

$$Mhe^{Lh} + \rho \leq \varepsilon_0 \quad (5.26)$$

and

$$4Mh\frac{e^{Lh}-1}{1-e^{\mu h}} + (4 + 2Lh)\frac{\rho}{1-e^{\mu h}} + \rho \leq \varepsilon_0 \quad (5.27)$$

imply

$$\begin{aligned} &\mathrm{dist}_H(\mathrm{Viab}_F(K), \mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_\rho^\rho)) \\ &\leq (1 + L_V)\left(2Mh\tfrac{e^{Lh}-1}{1-e^{\mu h}} + (2 + Lh)\tfrac{\rho}{1-e^{\mu h}}\right) + L_V\rho + \tfrac{1}{2}ML_V h. \quad (5.28) \end{aligned}$$

In classical numerical analysis, the order of convergence of a scheme is regarded as one of the most important indicators for its quality. It is doubtful if this way of thinking is appropriate here, but it is possible to obtain a notion of convergence by setting $\rho := h^2$. In this case,

$$Mhe^{Lh} + h^2 \leq \varepsilon_0 \tag{5.29}$$

and

$$4Mh\frac{e^{Lh}-1}{1-e^{\mu h}} + (4+2Lh)\frac{h^2}{1-e^{\mu h}} + h^2 \leq \varepsilon_0 \tag{5.30}$$

imply

$$\begin{aligned} &\mathrm{dist}_H(\mathrm{Viab}_F(K), \mathrm{Viab}_{\tilde{\Gamma}_{h,\rho}}(K_h^{h^2})) \\ &\leq (1+L_V)\left(2Mh\tfrac{e^{Lh}-1}{1-e^{\mu h}} + (2+Lh)\tfrac{h^2}{1-e^{\mu h}}\right) + L_V h^2 + \tfrac{1}{2}ML_V h. \end{aligned} \tag{5.31}$$

Since $\frac{e^{Lh}-1}{1-e^{\mu h}} \to 0$ and $\frac{h}{1-e^{\mu h}} \to \frac{1}{|\mu|}$ as $h \to 0$, the algorithm converges linearly in h. These findings are in tune with the behaviour of set-valued numerical methods for initial value problems with spatial discretization, see [5] and [23].

Please note that every shadowing theorem for the time-h flow of a differential inclusion can be used to derive a concrete error estimate for the Viability Kernel Algorithm in the way sketched above. As the shadowing theory for set-valued systems is still under development, there is hope that the reasoning of this chapter can soon be applied to more general classes of right hand sides.

Acknowledgement

I would like to express my gratitude to Prof. Dr. Wolf-Jürgen Beyn (Bielefeld), who has been teaching me literally from the first minute.

Furthermore, I would like to thank Prof. Dr. Sergei Pilyugin (St. Petersburg), who taught me his view on shadowing theory by exploring set-valued dynamical systems together with me.

Finally, I would like to express my respect for the profound knowledge as well as the mathematical craftsmanship of both my teachers.

Bibliography

[1] J.-P. Aubin. *Viability Theory*. Birkhäuser, Boston, 1991.

[2] J. P. Aubin and A. Cellina. *Differential Inclusions*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1984.

[3] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

[4] W.-J. Beyn. On the numerical approximation of phase portraits near stationary points. *SIAM J. Numer. Anal.*, 24:1095–1113, 1987.

[5] W.-J. Beyn and J. Rieger. Numerical fixed grid methods for differential inclusions. *Computing*, 81(1):91–106, 2007.

[6] N. Bonneuil. Making ecosystem models viable. *Bulletin of Mathematical Biology*, 65:1081–1094, 2003.

[7] G. Colombo, M. Fečkan, and B. M. Garay. Multivalued perturbations of a saddle dynamics. *Differential Equations and Dynamical Systems*, to appear.

[8] R. M. Corless and S. Y. Pilyugin. Approximate and real trajectories for generic dynamical systems. *J. Math. Anal. Appl.*, 189:409–423, 1995.

[9] E. Crück and P. Saint-Pierre. Nonlinear impulse target problems under state constraint: A numerical analysis based on viability theory. *Set-Valued Analysis*, 12:383–416, 2004.

[10] K. Deimling. *Multivalued Differential Equations*. De Gruyter, Berlin, 1992.

[11] M. Dellnitz and A. Hohmann. A subdivision algorithm for the computation of unstable manifolds and global attractors. *Numerische Mathematik*, 75:293–317, 1997.

[12] T. Donchev. Properties of one sided Lipschitz multivalued maps. *Nonlinear Analysis*, 49:13–20, 2002.

[13] T. Donchev and E. Farkhi. Stability and Euler approximation of one-sided Lipschitz differential inclusions. *SIAM J. Control Optim.*, 36(2):780–796, 1998.

[14] T. Donchev, E. Farkhi, and S. Reich. Fixed set iterations for relaxed Lipschitz multimaps. *Nonlinear Analysis*, 53:997–1015, 2003.

[15] A. Al-Nayef et al. Bi-shadowing and delay equations. *Dynamics and Stability of Systems*, 11:121–134, 1996.

[16] A. F. Filippov. Classical solutions of differential equations with multivalued right-hand side. *SIAM J. Control*, 5(4):609–621, 1967.

[17] H. Frankowska and M. Quincampoix. Viability kernels of differential inclusions with constraints: Algorithm and applications. *J. Math. Systems Estimation Control*, 1(3):371–388, 1991.

[18] M. Frigon and A. Granas. Résultats du type de Leray-Schauder pour des contractions multivoques. *Topol. Meth. Nonl. Anal.*, 4(1):197–208, 1994.

[19] B. M. Garay and P. E. Kloeden. Discretization near compact invariant sets. *Random Comput. Dynam.*, 5(2-3):93–123, 1997.

[20] B. M. Garay and L. Lóczi. Discretizing the fold bifurcation – a conjugacy result. *Period. Math. Hungar.*, 56(1):37–53, 2008.

[21] V. Glavan and V. Gutu. On the dynamics of contracting relations. *Analysis and optimization of differential systems, Kluwer Acad. Publ., Boston, MA*, pages 179–188, 2003.

[22] V. Glavan and V. Gutu. Attractors and fixed points of weakly contracting relations. *Fixed Point Theory*, 5(2):265–284, 2004.

[23] G. Grammel. Towards fully discretized differential inclusions. *Set-Valued Analysis*, 11:1–8, 2003.

[24] L. Hörmander. Sur la fonction d'appui des ensembles convexes dans un espace localement convexe. *Arkiv för Matematik*, 3:181–186, 1955.

[25] T. Hüls. Homoclinic trajectories of non-autonomous maps. *Journal of Difference Equations and Applications*, to appear.

[26] M. Kunze. *Non-smooth Dynamical Systems*. Springer, Berlin, 2000.

[27] O. E. Lanford. Introduction to hyperbolic sets. *NATO ASI Series B, Physics*, 118:73–102, 1985.

[28] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. AMS, 2001.

[29] R. McGehee and E. Sander. A new proof of the stable manifold theorem. *Z. angew. Math. Phys.*, 47:497–513, 1996.

[30] K. Palmer. *Shadowing in Dynamical Systems. Theory and Applications.* Kluwer, Dordrecht, 2000.

[31] S. Y. Pilyugin. *Shadowing in Dynamical Systems*, volume 1706 of *Lect. Notes Math.* Springer, Berlin, 1999.

[32] S. Y. Pilyugin and J. Rieger. Shadowing and inverse shadowing in set-valued dynamical systems. Contractive case. *Top. Meth. Nonl. Anal.*, 32(1):139–150, 2008.

[33] S. Y. Pilyugin and J. Rieger. Shadowing and inverse shadowing in set-valued dynamical systems. Hyperbolic case. *Top. Meth. Nonl. Anal.*, 32(1):151–164, 2008.

[34] R. T. Rockafellar. *Convex Analysis*, volume 28 of *Princeton mathematical series*. Princeton Univ. Press, 1970.

[35] P. Saint-Pierre. Approximation of the viability kernel. *Appl. Math. Optim.*, 29:187–209, 1994.

[36] E. Sander. Hyperbolic sets for noninvertible maps and relations. *Disc. Cont. Dyn. Syst.*, 8(2):339–357, 1999.

[37] G. Smirnov. *Introduction to the Theory of Differential Inclusions*, volume 41 of *Graduate Studies in Mathematics*. AMS, 2002.

[38] E. Sontag. *Mathematical Control Theory*, volume 6 of *Texts in Applied Mathematics*. Springer, New York, 1998.

[39] A. M. Stuart and A. R. Humphries. *Dynamical systems and numerical analysis*, volume 2 of *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, 1996.

[40] V. Veliov. Second order discrete approximations to strongly convex differential inclusions. *Systems and Control Letters*, 13:263–269, 1989.

[41] W. Walter. *Differential and Integral Inequalities*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer, Berlin, 1970.

[42] W. Walter. *Analysis 1,2*. Springer, Berlin, 1997.

[43] J. Warga. *Optimal Control of Differential and Functional Equations*. Academic Press, New York, 1972.

[44] P. R. Wolenski. The exponential formula for the reachable set of a Lipschitz differential inclusion. *SIAM J. Control and Optimization*, 28:1148–1161, 1990.

[45] E. Zeidler. *Nonlinear Functional Analysis and its Aplications*, volume 1. Springer, Heidelberg, 1985.

Index