Set-Valued and Convex Analysis in Dynamics and Control

Rafal Goebel *†

November 12, 2013

Abstract

Many developments in convex, nonsmooth, and set-valued analysis were motivated by optimization and optimal control theory. Real-life problems frequently yield convex objective functions that need to be minimized; sets of solutions to optimization problems depending on a parameter naturally lead to set-valued mappings; and optimal value functions in optimal control problems, when parameterized by the initial condition, are often nonsmooth. Many of the theoretical developments then found applications to optimization and optimal control, pushing these fields to places not accessible using classical analysis. Several excellent books outline these developments and their applications to optimization and optimal control.

The goal of the lectures is to present several basic elements of convex and set-valued analysis, and motivate them by and apply them to questions not in optimal control but rather in dynamical systems and general control theory. Topics from control theory include uncertain dynamics, which involve switching between several linear dynamical systems; uniformity and robustness to perturbations of asymptotic stability in a dynamical system, especially when the dynamics are discontinuous or uncertain; invariance concepts for hybrid dynamical systems, that is, dynamical systems which combine continuous-time and discrete-time characteristics; and necessary and sufficient conditions for convergence to a consensus, motivated by problems where several agents or vehicles are to be steered to a common position. These topics call for the introduction of elements of convex analysis, including the construction of the conjugate function and the convex subdifferential; and of set-valued analysis, for example convergence of sets, continuity properties of set-valued mappings, etc. The lectures are not meant to be a comprehensive treatment of either control theory or set-valued analysis but, rather, are meant to motivate the audience for further study.

*Department of Mathematics and Statistics, Loyola University Chicago, 1032 W. Sheridan Road, Chicago, IL 60660 *rgoebel1@luc.edu*. This work was partially supported by the National Science Foundation grant 1008602.

†Rafal Goebel received his M.Sc. degree in mathematics in 1994 from the University of Maria Curie Sklodowska in Lublin, Poland, and his Ph.D. degree in mathematics in 2000 from University of Washington, Seattle. He held a postdoctoral position at the Departments of Mathematics at the University of British Columbia and Simon Fraser University in Vancouver, Canada, 2000 – 2002; a postdoctoral and research positions at the Electrical and Computer Engineering Department at the University of California, Santa Barbara, 2002 – 2004; and a part-time teaching position at the Department of Mathematics at the University of Washington. In 2008, he joined the Department of Mathematics and Statistics at Loyola University Chicago. His background includes optimization, optimal control and control theory in general, and convex and nonsmooth analysis. His current research interests include applications of set-valued analysis and of convex duality theory to control problems; generalized concepts of convexity; and hybrid dynamical systems. Rafal Goebel is the recipient of the 2009 SIAM Activity Group on Control and Systems Theory Prize and a co-recipient of the 2010 IEEE Control Systems Society award for the Control Systems Magazine Outstanding Paper.

1 Introduction

Variational analysis is now a mature and broad discipline of mathematics, which, as the name suggests, grew out of calculus of variations and subsumes convex, set-valued, and non-smooth analysis. Convex analysis, the foundations of which can be found in the books *Convex analysis* by Rockafellar [29], has seen motivation from and important applications in optimization, as seen in books *Convex optimization* by Boyd and Vandenberghe [8] or *Conjugate duality and optimization* by Rockafellar [30], where convex duality is applied to broadly understood optimization problems beyond finite-dimensional spaces, in finance, control, etc. Nonsmooth analysis, studying nondifferentiable functions beyond the convex case, has seen motivation from and applications to optimal control, as evidenced by the books *Optimization and nonsmooth analysis* by Clarke [9], *Optimal control* by Vinter [36], and control theory in general, as seen in *Nonsmooth analysis and control theory* by Clarke, Ledyaev, Stern, and Wolenski [10]. Set-valued analysis, on its own, has seen treatment in the book *Set-valued analysis* by Aubin and Frankowsa [4], and applications to dynamical systems described by differential inclusions, as seen in the books *Differential Inclusions* by Aubin and Cellina [3] and *Viability theory* by Aubin [2].

Variational analysis as a whole has now seen extensive treatment in volumes such as *Variational analysis* by Rockafellar and Wets [31], where finite-dimensional objects are studied, and *Variational analysis and generalized differentiation, I and II* by Mordukhovich [27], [28], which deals with the infinite-dimensional case. Graduate-level textbooks which can serve as a great introduction to the area include *Convex analysis and nonlinear optimization* by Borwein and Lewis [7], *Introduction to the theory of differential inclusions* by Smirnov [33] and *Optimal control via nonsmooth analysis* by Loewen [23]. New developments in variational analysis abound, motivated by problems where constraints, nonsmoothness, discontinuities, impulsive behavior, and uncertainty are present. Variational analysis can also provide novel insights into classical topics, as seen in the book *Implicit functions and solution mappings - A view from variational analysis* by Dontchev and Rockafellar [12].

This set of notes presents several basic topics in variational analysis, mostly from convex and set-valued analysis, and highlights their applications to dynamical and control systems. While the presentation of the topics is necessarily brief, some topics are connected to recent research, by the author and by others.

2 Linear switching systems and convex Lyapunov functions

For a linear differential equation $\dot{x} = Ax$, asymptotic stability can be equivalently characterized by the existence of a quadratic Lyapunov function, i.e., a positive definite function which decreases along every solution to the differential equation. Furthermore, this asymptotic stability is equivalent to that same property for the adjoint system $\dot{y} = A^T y$. Now consider two — or more — linear differential equations $\dot{x} = A_1 x$, $\dot{x} = A_2 x$, each of which displays asymptotic stability, and look at behaviors resulting from switching between these two differential equations, as often as one wants. What happens to asymptotic stability? If asymptotic stability persists, is there a Lyapunov function verifying it? Is there a quadratic one? A convex one? What about the adjoint switching system, resulting from switching between $\dot{y} = A_1^T y$ and $y \dot{=} A_2^T y$? How are Lyapunov functions for the switching system and its adjoint related? These questions are addressed in this section.

2.1 Background: linear differential equations and quadratic Lyapunov functions

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function with $f(0) = 0$. The equilibrium $x = 0$ (or the equilibrium solution $x(t) \equiv 0$, $t \in [0, \infty)$) of the differential equation

$$\dot{x} = f(x) \tag{1}$$

is called *stable*, if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every solution ϕ to (1), $\phi(0) \in \delta I\!B$ implies $\operatorname{rge} \phi \in \varepsilon I\!B$; *locally attractive*, if there exists $\delta > 0$ such that for every solution ϕ to (1) on $[0, \infty)$, $\phi(0) \in \delta I\!B$ implies $\phi(t) \to 0$ as $t \to \infty$; *globally attractive*, if every solution ϕ to (1) on $[0, \infty)$ satisfies $\phi(t) \to 0$ as $t \to \infty$; and *locally*, respectively, *globally asymptotically stable* if it is both stable and locally, respectively, globally attractive.

Particularly good description of asymptotic stability is possible when the differential equation (1) is linear:

$$\dot{x} = Ax. \tag{2}$$

Fact 2.1 *Let A be a $n \times n$ matrix. The following are equivalent:*

(a) The origin is locally attractive for (2)*.*

(b) The origin is globally asymptotically stable for (2)*.*

(c) Eigenvalues of the matrix A have negative real parts.

(d) There exists a symmetric and positive definite $n \times n$ matrix P such that the quadratic function

$$V(x) = \frac{1}{2} x \cdot Px \tag{3}$$

has the following property: there exists $\gamma > 0$ such that, for every solution ϕ to (2)*,*

$$\frac{d}{dt} V(\phi(t)) \leq -\gamma V(\phi(t)). \tag{4}$$

(e) There exists a symmetric and positive definite $n \times n$ matrix P and $\gamma > 0$ such that

$$PA + A^T P \leq -\gamma P, \tag{5}$$

in the sense that $PA + A^T P + \gamma P$ be negative definite.

(f) The origin is globally asymptotically stable for $\dot{y} = A^T y$.

Functions satisfying (4) were first introduced in Lyapunov's dissertation [24] at the end of 19th century and are called *Lyapunov functions*. Proving that the Lyapunov condition (4) implies asymptotic stability is straightforward. In fact, with $a, b > 0$ such that $a\|x\|^2 \leq V(x) \leq b\|x\|^2$ for every $x \in \mathbb{R}^n$, (4) implies that for every solution ϕ to (2),

$$a\|\phi(t)\|^2 \leq V(\phi(t)) \leq e^{-\gamma t} V(\phi(0)) \leq b e^{-\gamma t} \|\phi(0)\|^2,$$

and so $\|\phi(t)\| \leq K e^{-\gamma t} \|\phi(0)\|$ for some $K > 0$, which is referred to as *exponential stability*. The converse implication, from (b) to (d), is far more involved, but can be illustrated in the 2×2 case.

Example 2.2 Let A be a 2×2 matrix. Then $A = MJM^{-1}$ for a nonsingular 2×2 matrix M and a $2\times$ matrix J (the real Jordan form of A) of the following form:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \tag{6}$$

and these forms occur, respectively, when A has two distinct real eigenvalues λ_1, λ_2; when A has a single (repeated) real eigenvalue λ and two distinct eigenvectors; when A has a single (repeated) real eigenvalue λ and one eigenvector; and when A has complex eigenvalues $\alpha + i\beta$, $\alpha - i\beta$. Consider the change of coordinates

$$z = M^{-1}x, \qquad x = Mz.$$

which turns $\dot{x} = Ax$ to

$$\dot{z} = Jz. \tag{7}$$

Suppose that A is such that 0 is asymptotically stable for (2). Then J is such that 0 is asymptotically stable for (7). This can be immediately seen by looking at the eigenvalues of A and J, which are the same, but can also be argued directly from the definition of asymptotic stability and the relation $z = M^{-1}x$. For the first two cases of J in (6), with $\lambda_1 \leq \lambda_2 < 0$, and for the last case, with $\alpha < 0$, the function $W(z) = \frac{1}{2}\|z\|^2$ is a Lyapunov function: the reader should check that for every solution ψ to (7) one obtains $\frac{d}{dt}\|\psi(t)\|^2 \leq 2\lambda_2\|\psi(t)\|^2$ for the first two cases, while for the last case, $\frac{d}{dt}\|\psi(t)\|^2 \leq \alpha\|\psi(t)\|^2$. For the case of a repeated real eigenvalue $\lambda < 0$ and one eigenvector, the norm squared fails to be a Lyapunov function, but the reader can verify that $W(z) = \frac{1}{2}z_1^2 + \frac{c}{2}z_2^2$ is a Lyapunov function for $\dot{z} = Jz$ as long as $c > 1/4\lambda^2$. Overall, when J is such that 0 is asymptotically stable for (7), there exists a symmetric and positive definite, in fact diagonal, 2×2 matrix Q such that

$$W(z) = \frac{1}{2} z \cdot Qz$$

is a Lyapunov function for (7): one has $QJ + J^TQ \leq -\gamma Q$ for some $\gamma > 0$. Then, for the symmetric and positive definite matrix $P = M^{-T}QM^{-1}$,

$$\begin{aligned} PA + A^TP &= M^{-T}QM^{-1}MJM^{-1} + M^{-T}J^TM^TM^{-T}QM^{-1} = M^{-T}(QJ + J^TQ)M^{-1} \\ &\leq M^{-T}(-\gamma Q)M^{-1} = -\gamma P \end{aligned}$$

and so the function $V(x) = \frac{1}{2}x \cdot Px$ is a Lyapunov function for (2). △

2.2 Switching systems

Let $A_1, A_2, \ldots, A_m$ be $n \times n$ matrices. A solution to the *switching system*

$$\dot{x} = A_\sigma x \tag{8}$$

consists of a piecewise continuous function $\sigma : [0, \infty) \to \{1, 2, \ldots, m\}$ and a piecewise differentiable function $\phi : [0, \infty) \to \mathbb{R}^n$ such that

$$\dot{\phi}(t) = A_{\sigma(t)}x(t) \quad \text{for almost all } t \in [0, \infty).$$

Piecewise continuity/differentiability is understood here as finitely many discontinuities/points of nondifferentiability on each compact subinterval of $[0, \infty)$. A general reference for switching systems, which includes control engineering motivation, is the book [22].

Example 2.3 With $a > 0$, consider two 2×2 matrices given by

$$A_1 = \begin{pmatrix} 0 & -1 \\ a^2 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & a^2 \\ -1 & 0 \end{pmatrix}.$$

Solutions to $\dot{x} = A_1 x$ rotate counterclockwise along the ellipses $x_1^2 + \frac{x_2^2}{a^2} = \text{const}$; solutions to $\dot{x} = A_2 x$ rotate clockwise along the ellipses $\frac{x_1^2}{a^2} + x_2^2 = \text{const}$. For both differential equations, 0 is stable but not asymptotically stable. Note also that for both A_1 and A_2, eigenvalues are $\lambda = \pm ai$. Appropriately picking the switching signal σ produces solutions to (8) that converge to 0 and solutions that diverge. Furthermore, one can pick σ so that the solution to $\dot{x} \in \{A_1 x, A_2 x\}$ converges to 0 but an arbitrarily slow rate. Considering small $\varepsilon > 0$ and matrices

$$A_1^\varepsilon = \begin{pmatrix} 0 & -1 \\ a^2 & -\varepsilon \end{pmatrix}, \quad A_2^\varepsilon = \begin{pmatrix} 0 & a^2 \\ -1 & -\varepsilon \end{pmatrix}$$

with eigenvalues $\lambda = -\varepsilon/2 \pm \sqrt{a^2 - \varepsilon^2/4}i$ which have negative real parts yields two linear differential equations for which 0 is asymptotically stable. Still, picking the switching signal σ appropriately produces solutions to (8) given by A_1^ε, $A_2\varepsilon$ that that diverge, and σ can be such that the solution converges to 0 arbitrarily slow. Indeed, consider the matrices A_1^ε, A_2^ε with $a = 1$ and the function on $\mathbb{R}^2$ given by $V(x) = \frac{1}{2}\|x\|^2$. Then, for every solution to (8),

$$\frac{d}{dt} V(\phi(t)) = -\varepsilon \phi_2(t)^2,$$

for almost all t, and hence the distance, to 0, of every solution is strictly decreasing. However, the decrease can be slow enough to prevent solutions from converging to 0. △

Asymptotic stability of 0 for a linear differential inclusion is understood similarly to what is done for differential equations, with the requirements placed by stability and attractivity holding for all solutions to (8). One conclusion of the example above is that a linear differential inclusion needs not have 0 asymptotically stable even if for each $i = 1, 2, \ldots, m$, 0 is asymptotically stable for $\dot{x} = A_i x$.

If there exists a Lyapunov function for (8), that is, a differentiable function $V : \mathbb{R}^n \to [0, \infty)$ such that $a\|x\|^2 \leq V(x) \leq b\|x\|^2$ for some $a, b > 0$ and all $x \in \mathbb{R}^n$ and such that, for some $\gamma > 0$ and for every solution ϕ to (8),$\frac{d}{dt} V(\phi(t)) \leq -\gamma V(\phi(t))$ for almost all t, then the origin is globally asymptotically stable for (8); in fact, it is exponentially stable, in the sense that there exist $M, \gamma > 0$ such that every solution ϕ to (8) satisfies

$$\|\phi(t)\| \leq K e^{-\gamma t} \|\phi(0)\| \tag{9}$$

The questions that will be answered next, with the help of convex analysis, are:

- If every solution ϕ to (8) satisfies (9), does there exist a Lyapunov function verifying this?
- If every solution ϕ to (8) satisfies (9), what can be said about asymptotic stability for the dual switching system $\dot{y} = A_\sigma^T y$?

Remark 2.4 A Lyapunov function for (8), equivalently, a differentiable function V with bounds as above and such that

$$\nabla V(x) \cdot A_i x \leq -\gamma V(x)$$

for every $i = 1, 2, \ldots, m$, every $x \in \mathbb{R}^n$ is callaed a *common Lyapunov function.* The existence of a common quadratic Lyapunov function for a family of matrices $A_1, A_2, \ldots, A_m$, when (8) is asymptotically stable, is an active area of research; see [32] for an overview. In general, asymptotic stability of (8) is not sufficient for the existence of a common quadratic Lyapunov function; see . This topic is not pursued further here. △

2.3 Convex conjugate functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for every $x, y \in \mathbb{R}^n$, every $\lambda \in [0, 1]$,

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$

The standard reference for convex analysis is [29]; Fact 2.6 can be found in Chapter 10. In what follows, the class of convex, positive definite , and positively homogeneous of degree 2 functions $f : \mathbb{R}^n \to [0, \infty)$ will play an important role; the class is denoted by $\mathcal{C}$.

Example 2.5 Examples of functions $f \in \mathcal{C}$ include

(a) $f(x) = \frac{1}{2} x \cdot Px$ for a symmetric and positive definite $n \times n$ matrix P.

(b) $f(x) = g(Ax)$ for $g \in \mathcal{C}$ and any $n \times n$ matrix A. It is worth noting that, in general, a composition of a convex function g with an affine mapping $x \mapsto Ax + b$ is a convex function; that is $x \mapsto g(Ax + b)$ is a convex function.

(c) $f(x) = \sup_{\gamma \in \Gamma} g_\gamma(x)$, where $\{g_\gamma \mid \gamma \in \Gamma\}$ is a family of functions $g_\gamma \in \mathcal{C}$, under the condition that $f(x) < \infty$ for every $x \in \mathbb{R}^n$. This condition is automatically true if the family Γ is finite. It is worth noting that, in general, a pointwise supremum of convex functions is a convex function; even when infinite values are allowed.

(d) $f(x) = \operatorname{con} \inf_{\gamma \in \Gamma} g_\gamma(x)$, where $\{g_\gamma \mid \gamma \in \Gamma\}$ is a family of functions $g_\gamma \in \mathcal{C}$, $\operatorname{con} \inf_{\gamma \in \Gamma} g_\gamma(x)$ denotes the *convex hull* of $\inf_{\gamma \in \Gamma} g_\gamma(x)$ which is the greatest convex function bounded above by $\inf_{\gamma \in \Gamma} g_\gamma(x)$, under the condition that f is positive definite. This condition is automatically true if the family Γ is finite.

△

Fact 2.6 *Every convex function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, in fact locally Lipschitz continuous.*

An important operation on convex functions is that of convex conjugacy. For a function $f : \mathbb{R}^n \to \mathbb{R}$, its *convex conjugate* function is defined, at each $y \in \mathbb{R}^n$, by

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{y \cdot x - f(x)\}. \tag{10}$$

In general, f^* can attain infinite values. In an appropriate sense, f^* is always convex and lower semicontinuous, because it is a supremum of affine functions. One can also consider $(f^*)^*$, denoted, for simplicity, f^{**}.

Example 2.7 If $f(x) = \frac{1}{2} x \cdot Px$ for a symmetric and positive definite $n \times n$ matrix P then $f^*(y) = \frac{1}{2} y \cdot P^{-1} y$. Indeed,

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left\{ y \cdot x - \frac{1}{2} x \cdot Px \right\}$$

and because $x \mapsto y \cdot x - \frac{1}{2} x \cdot Px$ is a strictly concave function which diverges to $-\infty$ when $\|x\| \to \infty$, its unique supremum is attained when the gradient is 0, so when $y = Px$. Plugging in $x = P^{-1} y$ yields the formula for f^*. △

For general lower semicontinuous convex functions, the convex conjugacy gives a one to one correspondence between a function and its conjugate. Here, only a special case is considered:

Theorem 2.8 *Let $f \in \mathcal{C}$. Then $f^* \in \mathcal{C}$ and $f^{**} = f$.*

Proof. Suppose that $f \in \mathcal{C}$. Let $\mu = 2\min\{f(x) \,|\, \|x\| = 1\}$. Because f is continuous and positive definite, $\mu > 0$. Because f is positively homogeneous of degree 2, for $x \neq 0$,

$$f(x) = \|x\|^2 f\left(\frac{x}{\|x\|}\right) \geq \frac{\mu}{2}\|x\|^2.$$

Then $f^*(y)$ is bounded above by the conjugate of the quadratic function $x \mapsto \frac{\mu}{2}\|x\|^2$, which, thanks to Example (2.7), is $y \mapsto \frac{1}{2\mu}\|y\|^2$. Hence f^* is finite-valued and $f^*(0) \leq 0$. A symmetric argument shows that f^* is bounded below by a multiple of the norm squared. Thus, f^* is positive definite. Now consider $\lambda > 0$ and note that, for every $y \in \mathbb{R}^n$,

$$\begin{aligned} f^*(\lambda y) &= \sup_{x\in\mathbb{R}^n} \{(\lambda y)\cdot x - f(x)\} = \lambda^2 \sup_{x\in\mathbb{R}^n} \left\{ y \cdot \frac{x}{\lambda} - \frac{1}{\lambda^2} f(x) \right\} = \lambda^2 \sup_{x\in\mathbb{R}^n} \left\{ y \cdot \frac{x}{\lambda} - f\left(\frac{x}{\lambda}\right) \right\} \\ &= \lambda^2 \sup_{z\in\mathbb{R}^n} \{y\cdot z - f(z)\} = \lambda^2 f^*(y). \end{aligned}$$

Consequently, $f^* \in \mathcal{C}$. To prove that $f^{**} = f$, note that by the definition of f^*, for every $x, y \in \mathbb{R}^n$, $f^*(y) \geq y \cdot x - f(x)$, hence $f(x) \geq x \cdot y - f^*(y)$. Then for every $x \in \mathbb{R}^n$, $f(x) \geq \sup_{y\in\mathbb{R}^n} \{x \cdot y - f^*(y)\}$, and so $f(x) \geq f^{**}(x)$. Suppose then that for some $\overline{x} \in \mathbb{R}^n$, $f(\overline{x}) > (f^*)^*(\overline{x})$. Then there exists $\overline{y} \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ such that

$$f(x) > x \cdot \overline{y} + \beta \quad \forall x \in \mathbb{R}^n \quad \text{and} \quad \overline{x} \cdot \overline{y} + \beta > f^{**}(\overline{x}).$$

The first inequality implies that

$$f^*(\overline{y}) = \sup_{x\in\mathbb{R}^n} \{\overline{y} \cdot x - f(x)\} \leq \sup_{x\in\mathbb{R}^n} \{\overline{y} \cdot x - (\overline{y} \cdot x + \beta)\} = -\beta$$

and

$$f^{**}(\overline{x}) = \sup_{y\in\mathbb{R}^n} \{\overline{x} \cdot y - f^*(y)\} \geq \overline{x} \cdot \overline{y} - f^*(\overline{y}) \geq \overline{x} \cdot \overline{y} + \beta,$$

which contradicts the second claimed inequality. □

2.4 Switching systems, Lyapunov functions, and duality

Consider the discrete-time version of a switching linear system first. For a recent work with extensive bibliography on the topic see [26]. Given $n \times n$ matrices $A_1, A_2, \ldots, A_m$, a solution to the *switching system*

$$x^+ = A_\sigma x \tag{11}$$

consists of a function $\sigma : \mathbb{N}_0 \to \{1, 2, \ldots, m\}$ and a function $\phi : \mathbb{N}_0 \to \mathbb{R}^n$ such that

$$\phi(j+1) = A_{\sigma(j)} x(j) \quad \text{for every } j \in \mathbb{N}_0.$$

If there exist $M, \gamma > 0$ such that every solution ϕ to (11) satisfies $\|\phi(j)\| \leq K e^{-\gamma j}\|\phi(0)\|$, then 0 is *exponentially stable* for (11). The result below goes back to [25]. The continuous-time version includes smoothness of V, appeared in [11], and will be further mentioned in Section 2.5.

Proposition 2.9 *The origin is exponentially stable for* (11) *if and only if there exists* $V \in \mathcal{C}$ *such that, for some* $\delta \in (0, 1)$,

$$V(A_i x) \leq \delta V(x) \quad \textit{for every } i = 1, 2, \ldots, m, \ x \in \mathbb{R}^n.$$

Proof. Suppose that 0 is exponentially stable for (11). Consider $V : \mathbb{R}^n \to [0, \infty)$ defined by

$$V(x) = \sup\left\{\|\phi(j)\|^2 e^{2\gamma j} \mid \phi \text{ is a solution to (11)}, \ \phi(0) = x, \ j \in \mathbb{N}_0\right\}. \tag{12}$$

For each fixed $j \in \mathbb{N}_0$ and each fixed switching signal σ on $\{0, 1, \ldots, j-1\}$, the function $x \mapsto \phi(j)$ is linear, and so $x \mapsto \|\phi(j)\|^2$ is convex. Then $x \mapsto \|\phi(j)\|^2 e^{2\gamma j}$ is convex, and V is the supremum of such functions over all $j \in \mathbb{N}_0$ and all switching signals. Hence V is convex. It is positive definite because $V(x) \geq \|x\|^2$, which is seen by considering $j = 0$, and because the unique solution from 0 is $\phi \equiv 0$. Now, pick $i \in \{1, 2, \ldots m\}$ and note that

$$\begin{aligned}
V(A_i x) &= \sup\left\{\|\phi(j)\|^2 e^{2\gamma j} \mid \phi \text{ is a solution to (11)}, \ \phi(0) = A_i x, \ j \in \mathbb{N}_0\right\} \\
&= \sup\left\{\|\psi(j)\|^2 e^{2\gamma(j-1)} \mid \psi \text{ is a solution to (11)}, \ \psi(0) = x, \ \psi(1) = A_i x, \ j \in \mathbb{N}\right\} \\
&\leq \sup\left\{\|\psi(j)\|^2 e^{2\gamma(j-1)} \mid \psi \text{ is a solution to (11)}, \ \psi(0) = x, \ j \in \mathbb{N}_0\right\} \\
&= e^{-2\gamma} V(x)
\end{aligned}$$

The result follows, with $\delta = e^{-2\gamma}$. The reverse implication follows from the fact that if V is continuous (as it is when it is convex and finite-valued), positive definite, and positively homogeneous of degree 2, then for some $a, b > 0$ and all $x \in \mathbb{R}^n$, $a\|x\|^2 \leq V(x) \leq b\|x\|^2$. □

The next observation appeared in [13] and will lead to equivalent characterization of exponential stability for (11) through the dual system

$$x^+ = A_\sigma^T x. \tag{13}$$

in Theorem 2.11, obtained first in [5], [6] without the use of convex Lyapunov functions.

Lemma 2.10 *Let $V \in \mathcal{C}$, A be a $n \times n$ matrix, and $\delta > 0$. The following are equivalent:*

- *(a) $V(Ax) \leq \delta V(x)$ for every $x \in \mathbb{R}^n$;*
- *(b) $V^*(A^T y) \leq \delta V^*(y)$ for every $y \in \mathbb{R}^n$;*

Proof. Suppose that (a) holds. Then

$$\begin{aligned}
V^*(A^T y) &= \sup_{x \in \mathbb{R}^n} \left\{\left(A^T y\right) \cdot x - V(x)\right\} \leq \sup_{x \in \mathbb{R}^n} \left\{y \cdot Ax - \frac{1}{\delta} V(Ax)\right\} = \sup_{z \in \operatorname{rge} A} \left\{y \cdot z - \frac{1}{\delta} V(z)\right\} \\
&\leq \sup_{x \in \mathbb{R}^n} \{y \cdot x - V(x)\} = \frac{1}{\delta} \sup_{x \in \mathbb{R}^n} \left\{(\delta y) \cdot x - \frac{1}{\delta} V(x)\right\} = \frac{1}{\delta} V^*(\delta y) \\
&= \delta V^*(y).
\end{aligned}$$

Hence (a) implies (b). The implication from (b) to (a) follows from considering A^T, V^*, and $(V^*)^* = V$ in place of A, V, and V^* in the implication from (a) to (b). □

Theorem 2.11 *Consider $n \times n$ matrices $A_1, A_2, \ldots, A_m$. Then* 0 *is exponentially stable for* (11) *if and only if* 0 *is exponentially stable for* (13).

Proof. Combine Proposition 2.9 and Lemma 2.10. □

2.5 More convex analysis, continuous-time switching systems, and duality

The concept of a Lyapunov function for a discrete-time switching system, as described in Proposition 2.9, does not require the Lyapunov function to be differentiable. For continuous-time Lyapunov functions, their decrease is described using gradients. Hence, to provide continuous-time results similar to those in Lemma 2.10 or Theorem 2.11, one needs some differentiability properties of convex functions and their conjugates. The fact below follows from [29, Theorem 26.3].

Fact 2.12 *A convex function $f \in \mathcal{C}$ is differentiable if and only if its conjugate function $f^* \in \mathcal{C}$ is strictly convex.*

Let $\mathcal{CS}$ be the class of function $f \in \mathcal{C}$ that are differentiable and strictly convex. The fact above and Theorem 2.8 imply:

Corollary 2.13 *$f \in \mathcal{CS}$ if and only if $f^* \in \mathcal{CS}$.*

The lemma below is a special case of [29, Theorem 23.5] which is applicable to nondifferentiable convex functions.

Lemma 2.14 *Assume that $f \in \mathcal{CS}$, equivalently, that $f^* \in \mathcal{CS}$. The following are equivalent:*

(a) $y = \nabla f(x)$;

(b) $x = \nabla f^*(y)$;

(c) $f(x) + f^*(y) = x \cdot y$.

Proof. By definition, for every $x, y \in \mathbb{R}^n$, $f^*(y) \geq x \cdot y - f(x)$. If $y = \nabla f(x)$ then x maximizes $x \cdot y - f(x)$ and (c) holds. Thus (a) implies (c). On the other hand, (c) means that x maximizes $x \cdot y - f(x)$, so (a) must hold. The equivalence of (b) and (c) follows by symmetry. □

Lemma 2.14 implies the equivalence between $V \in \mathcal{CS}$ decreasing along solutions to (8) and $V^* \in \mathcal{CS}$ decreasing along solutions to

$$\dot{x} = A_\sigma^T x. \tag{14}$$

Indeed, for any matrix A, suppose that $\nabla V(x) \cdot Ax < 0$ for every $x \neq 0$. Pick any $y \neq 0$, let $x = \nabla V^*(x)$ which is equivalent to $y = \nabla V(x)$. Then $\nabla V^*(y) \cdot A^T y = x \cdot A^T y = y \cdot Ax = \nabla V(x) \cdot Ax < 0$. Hence $\nabla V(x) \cdot Ax < 0$ for every $x \neq 0$ implies $\nabla V^*(y) \cdot A^T y$ for every $y \neq 0$, and a reverse implication follows by duality. A more useful equivalence, taken from [17], is in fact true:

Lemma 2.15 *Let $V \in \mathcal{CS}$, A be a $n \times n$ matrix, and $\delta > 0$. The following are equivalent:*

(a) $\nabla V(x) \cdot Ax \leq -\delta V(x)$ *for every* $x \in \mathbb{R}^n$;

(b) $\nabla V^*(y) \cdot A^T y \leq -\delta V^*(y)$ *for every* $y \in \mathbb{R}^n$.

Proof. Assume (a) and note that $\nabla V(x) \cdot Ax \leq -\delta V(x)$ for every $x \in \mathbb{R}^n$ is equivalent to $\nabla V(x) \cdot Ax \leq -\delta$ for every x with $V(x) = 1$. This comes from homogeneity of V, ∇V, and the fact that for $x \neq 0$, $V\left(x/\sqrt{V(x)}\right) = 1$. For every $y \in \mathbb{R}^n$ there is $x \in \mathbb{R}^n$ such that $y = \nabla V(x)$; indeed, one takes $x = \nabla V^*(y)$. If, furthermore, y is such that $V^*(y) = 1$ then $V(x) = 1$. Indeed,

$$V(x) \quad = \quad x \cdot y - V^*(y) = \max_{z \in \mathbb{R}^n} \{x \cdot z - V^*(z)\} = \max_{\lambda \in \mathbb{R}} \{x \cdot \lambda y - V^*(\lambda y)\}$$

$$= \max_{\lambda\in\mathbb{R}} \left\{\lambda\left(V(x)+V^*(y)\right)-\lambda^2 V^*(y)\right\} = \max_{\lambda\in\mathbb{R}}\left\{\lambda\left(V(x)+1\right)-\lambda^2\right\}$$

and setting the derivative of the last expression above, in λ, equal to 0 when $\lambda = 1$, gives $V(x) = 1$. Hence, $\nabla V(x)\cdot Ax \le -\delta$ for every x with $V(x) = 1$ implies $\nabla V^*(y)\cdot A^T y \le -\delta$ for every y with $V^*(y) = 1$, and this implies (b). The opposite implication holds by symmetry. □

With this in hand, and with the result on existence of smooth Lyapunov functions [11], one can now show the equivalence of exponential stability for (8) and (14).

3 Difference and differential inclusions

Besides the pursuit of generalization, there are other natural reasons to consider differential and difference inclusions rather than equations. Section 3.1 below illustrates this. Basic properties of set-valued mappings are needed to study differential and difference equations, they are collected in Section 3.2. Semicontinuity and boundedness properties of the set-valued mapping defining a difference inclusion have implications on uniformity of asymptotic stability, its robustness, etc., this is presented in Section 3.3. In particular, some descriptions of asymptotic stability for a switching system and its dual are improved there. A simple example illustrates in Section 3.4 how control engineering problems. naturally lead to differential inclusions.

3.1 Some motivation

In Proposition 2.9, it was shown that exponential stability of 0 for a discrete time switching system (11) is equivalent to the existence of a Lyapunov function $V \in \mathcal{C}$, which satisfies

$$V(A_i x) \le \delta V(x) \quad \text{for every } i = 1, 2, \ldots, m,\ x \in \mathbb{R}^n.$$

This condition in fact guarantees that V is also a Lyapunov function for the *linear difference inclusion*

$$x^+ \in \mathcal{A}(x), \tag{15}$$

where $\mathcal{A}(x)$ is the convex hull of $\{A_1x, A_2x, \ldots, A_mx\}$, that is, the smallest convex set containing $\{A_1x, A_2x, \ldots, A_mx\}$ A solution to (15) is a function $\phi : \mathbb{N}_0 \to \mathbb{R}^n$ such that, for every $j \in \mathbb{N}_0$, $\phi(j+1) \in \mathcal{A}(\phi(j))$. Take any $w \in \mathcal{A}(x)$, which can be equivalently characterized as the set of all convex combinations

$$\mathcal{A}(x) = \left\{\sum_{i=1}^m \lambda_i A_i x \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \ge 0 \text{ for } i = 1, 2, \ldots, m\right\}.$$

Then, thanks to convexity of V,

$$V(w) = V\left(\sum_{i=1}^m \lambda_i A_i x\right) \le \sum_{i=1}^m \lambda_i V(A_i x) \le \sum_{i=1}^m \lambda_i \delta V(x) = \delta V(x).$$

Hence, exponential stability of 0 for the switching system (11) is equivalent to the existence of a Lyapunov function for the linear difference inclusion (15), which in turn implies asymptotic of 0 for (15).

The same idea applied to continuous-time systems. If a differentiable $V \in \mathcal{C}$ is a Lyapunov function for (8), that is, if

$$\nabla V(x)\cdot A_i x \le -\gamma V(x) \quad \text{for every } i = 1, 2, \ldots, m,\ x \in \mathbb{R}^n,$$

then V is a Lyapunov function for the *linear differential inclusion*

$$x^+ \in \mathcal{A}(x), \tag{16}$$

that is, $\nabla V \cdot v \leq -\gamma V(x)$ for every $v \in \mathcal{A}(x)$, $x \in \mathbb{R}^n$. It was illustrated before that just asymptotic stability is not sufficient for the existence of a Lyapunov function for a the switching system (8). Through the analysis of the set-valued mapping $\mathcal{A}$, it will be shown below that for the linear differential inclusion (16), asymptotic stability is equivalent to exponential stability, which in turn implies the existence of a convex Lyapunov function.

Another motivation to study difference inclusions and the set-valued mappings defining them comes from the effect of perturbations on difference equations given by discontinuous functions. For $x \in \mathbb{R}$ let $\lfloor x \rfloor$ be the greatest integer less than x; so $\lfloor 3 \rfloor = 2$, $\lfloor \pi \rfloor = 3$, etc. Define $f : \mathbb{R} \to \mathbb{R}$ to be an odd function such that, $f(0) = 0$, and for $x > 0$, $f(x) = \lfloor x \rfloor$. Then the difference equation

$$x^+ = f(x)$$

has 0 globally asymptotically stable. In fact, convergence to 0 occurs in finitely many steps from any initial condition. For example, the solution ϕ from initial condition π is $\phi(0) = \pi$, $\phi(1) = 3$, $\phi(2) = 2$, $\phi(3) = 1$, $\phi(j) = 0$ for all $j \geq 4$. Consider arbitrarily small $\varepsilon > 0$ and the difference equation

$$x^+ = f(x + \varepsilon).$$

Local asymptotic stability of 0 is preserved, as $f(x+\varepsilon) = 0$ for every $x \in [-1-\varepsilon, 1-\varepsilon]$. However, convergence to 0 from large initial conditions fails: for example, $f(1+\varepsilon) = 1$, and so the solution from 1 is $\phi \equiv 1$. Such dramatic failure of convergence to 0 under small perturbations is due to discontinuity of f. This behavior can be predicted, without considering perturbations, by looking at a set-valued "closure" of f, which is the set-valued mapping F given by $F(0) = 0$, $F(x) = f(x)$ for $x \notin \mathbb{Z}$, $F(x) = \{x - 1, x\}$ for $x \in \mathbb{N}$. Then, $x^+ \in F(x)$ has constant solutions from every integer initial point.

More precisely, one has the following result. Its continuous-time version was predicted by [19], proved by [18], and appears here as Fact 3.13. The discrete-time version is not hard.

Proposition 3.1 *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a function, $I = \{0, 1, 2, \ldots, J\}$, and $\phi_i : I \to \mathbb{R}^n$ and $\phi : I \to \mathbb{R}^n$ be functions. The following are equivalent*

(a) There exist sequences of functions $\phi_i, e_i, p_i : I \to \mathbb{R}^n$ such that, for every $j \in I$, $\phi_i(j) \to \phi(j)$, $e_i(j) \to 0$, $p_i(j) \to 0$ as $i \to \infty$ and, for every $j \in \{0, 1, \ldots, J-1\}$,

$$\phi_i(j+1) = f(\phi_i(j) + e_i(j)) + p_i(j). \tag{17}$$

- *For every $j \in \{0, 1, \ldots, J-1\}$,*

  $$\phi(j+1) \in F(\phi(j)),$$

  where, for every $x \in \mathbb{R}^n$,

  $$F(x) = \bigcap_{\delta > 0} \overline{f(x + \delta \mathbb{B})}. \tag{18}$$

Proof. Assume (a) and pick $j \in \{0, 1, \ldots, J-1\}$. Then, for every $\delta > 0$ and all large enough i,

$$\phi_i(j+1) \in f(\phi_i(j) + \delta \mathbb{B}) + \delta \mathbb{B} \subset \overline{f(\phi(j) + 2\delta \mathbb{B})}.$$

Passing to the limit yields $\phi(j+1) \in F(\phi(j))$. Now assume (b) and pick $j \in \{0, 1, \ldots, J-1\}$. Considering $\delta = 1/i$ in (18) shows that $\phi(j+1) \in F(\phi(j))$ implies that for every $i \in \mathbb{N}$ there are $e_i(j), p_i(j) \in i^{-1}\mathbb{B}$ such that (17) holds. ☐

Through the analysis of set-valued mappings F and difference inclusions $x^+ \in F(x)$, robustness of asymptotic stability to perturbations can be studied.

3.2 Set-valued mappings

The definitions below follow [31]. A *set-valued mapping* from $\mathbb{R}^m$ to $\mathbb{R}^n$ assigns, to every $x \in \mathbb{R}^m$, a subset of $\mathbb{R}^n$. To distinguish such a mapping F from a function, the notation $F : \mathbb{R} \rightrightarrows \mathbb{R}^n$ will be used. The *effective domain* of $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, denoted $\operatorname{dom} F$, is the set $\{x \in \mathbb{R}^m \mid F(x) \neq \emptyset\}$. The mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is called

- *locally bounded* if every point $x \in \mathbb{R}^m$ has a neighborhood U such that $F(U)$ is a bounded set, where $F(U) = \bigcup_{u \in U} F(u)$.
- *outer semicontinuous* if for every convergent sequence $\{x_i\}$ of points in $\mathbb{R}^m$ and every convergent sequence $\{y_i\}$ of points $y_i \in F(x_i)$, one has $\lim_{n\to\infty} y_i \in F\left(\lim_{n\to\infty} x_i\right)$.
- *inner semicontinuous* if for every convergent sequence $\{x_i\}$ of points in $\mathbb{R}^m$ and every $y \in F\left(\lim_{n\to\infty} x_i\right)$ there exists a sequence $\{y_i\}$ of points $y_i \in F(x_i)$ such that $\lim_{n\to\infty} y_i = y$.
- *continuous* if it is both outer and inner semicontinuous.

The *graph* of a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the set

$$\operatorname{gph} F = \left\{(x, y) \in \mathbb{R}^{2n} \mid y \in F(x)\right\}.$$

The graph provides a convenient characterization of outer semicontinuity, given below. The proof is straightforward.

Proposition 3.2 *A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous if and only if $\operatorname{gph} F$ is closed.*

Example 3.3 Examples of set-valued mappings and their continuity properties include:

(a) *Reachable sets:* Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous and $K \subset \mathbb{R}^n$ be compact. Define $R : \mathbb{R} \rightrightarrows \mathbb{R}^n$ for $\tau \geq 0$ by

$$R(\tau) = \{\phi(\tau) \mid \phi : [0, \tau] \to \mathbb{R}^n \text{ solves } \dot{x} = f(x),\ \phi(0) \in K\},$$

and extend the definition to $\tau < 0$ by considering solutions on $[\tau, 0]$ and their initial points $\phi(\tau)$. Then R is inner semicontinuous.

(b) *Optimal solution mappings:* Let $f : \mathbb{R}^{m+n}$ be a function, and define $M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ by

$$M(x) = \left\{ y \in \mathbb{R}^n \mid f(x, y) = \inf_{y \in \mathbb{R}^n} f(x, y) \right\}.$$

(c) *Inflated functions:* Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a function and pick $\delta > 0$. Define $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$F(x) = f\left(x + \delta \mathbb{B}\right) + \delta \mathbb{B}.$$

If f is locally bounded then so is F. Similarly, if f is continuous then so is F. Note also that $y \in F(x)$ iff there exist x', y' such that $y \in f(x)$, $x' \in x + \delta \mathbb{B}$, $y \in y' + \delta \mathbb{B}$. Hence,

$$\operatorname{gph} F = \operatorname{gph} f + (\delta \mathbb{B} \times \delta \mathbb{B}).$$

This, Proposition 3.2, and the fact that a continuous function has a closed graph can be used to verify that a continuous f leads to an upper semicontinuous F.

(d) For any function $f : \mathbb{R}^n \to \mathbb{R}^n$ the set-valued mapping F defined in (18) is outer semicontinuous. Indeed, pick $x_i \to x$, $y_i \in F(x_i)$ such that $y_i \to y$. Because $y_i \in F(x_i)$, considering $\delta = 1/i$ in (18) implies that there exists $u_i \in x + i^{-1}\mathbb{B}$ and $w_i \in f(u_i)$ such that $\|y_i - w_i\| < 1/i$. Then $u_i \to x$ and $w_i \to y$. Because $u_i \to x$, for every $\delta > 0$ and every large enough i, $u_i \in x + \delta\mathbb{B}$ and hence $w_i \in f(x + \delta\mathbb{B})$. Because $w_i \to y$, it must be that $y \in \overline{f(x + \delta\mathbb{B})}$. Since this holds for every $\delta > 0$, $y \in F(x)$, which verifies outer semicontinuity. Furthermore, if f is locally bounded then so is F.

(e) *Tangent cones:* Let $C \subset \mathbb{R}^n$. The tangent cone $T_C(x)$ to the set C at x is defined as

$$T_C(x) = \left\{ v \in \mathbb{R}^n \mid \exists x_i \in C,\ x_i \to x,\ \text{and } \lambda_i \searrow 0 \text{ so that } \frac{x_i - x}{\lambda_i} \to v \right\}. \tag{19}$$

Tangent cones play a significant role in Section 4.4.

△

Proposition 3.4 *Consider set-valued mappings $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and $G : \mathbb{R}^k \rightrightarrows \mathbb{R}^m$ and define $M : \mathbb{R}^k \to \mathbb{R}^n$ by $M(x) = F(G(x)) = \bigcup_{y \in G(x)} F(y)$.*

(a) If F and G are locally bounded then M is locally bounded.

(b) If F and G are outer semicontinuous and G is locally bounded then M is outer semicontinuous.

Proof. Local boundedness of F can equivalently be characterized as: for every bounded $K \subset \mathbb{R}^m$, $F(K)$ is bounded. Indeed, the closure of a bounded K is compact and relying on the finite open covering property makes this characterization equivalent to the definition. Using this and a similar characterization for G proves (a). For (b), pick $x_i \to x$ and $z_i \to z$ with $z_i \in F(G(x_i))$. There exist $y_i \in G(x_i)$ such that $z_i \in F(y_i)$. Because G is locally bounded, some subsequence of y_i converges to a point y, and because of outer semicontinuity of G, $y \in G(x)$. Outer semicontinuity of F and $z_{i_k} \in F(y_{i_k})$ then gives $z \in F(y)$, and hence $z \in F(G(x))$. □

To see how local boundedness is necessary in (b), consider, for example, a function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 1/x$ for $x \neq 0$ and $f(0) = 7$. This function happens to be an outer semicontinuous mapping which is not locally bounded. Then, consider f composed with itself.

Proposition 3.5 *A locally bounded set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous if and only if it has closed values and for every $x \in \mathbb{R}$ and $\varepsilon > 0$ there exists $\delta > 0$ such that, for every $x' \in x + \delta\mathbb{B}$,*

$$F(x') \subset F(x) + \varepsilon\mathbb{B}.$$

3.3 Difference inclusions and asymptotic stability

A solution to a difference inclusion

$$x^+ \in F(x), \tag{20}$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping with $\operatorname{dom} F = \mathbb{R}^n$, is a function $\phi : \mathbb{N}_0 \to \mathbb{R}^n$ such that $\phi(j+1) \in F(\phi(j))$ for every $j \in \mathbb{N}_0$. A sequence of solutions $\{\phi_i\}$ is *locally bounded* if for every $J \in \mathbb{N}$ there exists $M > 0$ such that, for every $i \in \mathbb{N}$, $j \leq J$, one has $\phi_i(j) \in M\mathbb{B}$.

Theorem 3.6 *Suppose that the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous. Then, from every locally bounded sequence $\{\phi_i\}$ of solutions to* (20) *one can extract a subsequence that converges pointwise to a solution to* (20).

Proof. Local boundedness ensures that $\{\phi_i(0)\}$ is a bounded sequence and a convergent subsequence can be extracted. From that subsequence, one can further extract a subsequence such that $\{\phi_i(1)\}$ converge, etc. Cantor's diagonalization argument produces a subsequence of $\{\phi_i\}$, which we do not relabel, such that $\{\phi_i(j)\}$ converges for every $j \in \mathbb{N}_0$ to some point $\phi(j)$. Now, for each $j \in \mathbb{N}_0$, $(\phi_i(j), \phi_i(j+1)) \in \operatorname{gph} F$, and because the graph of F is closed, $(\phi(j), \phi(j+1)) \in \operatorname{gph} F$ and hence the mapping $\phi : \mathbb{N}_0 \to \mathbb{R}^n$ is a solution to (20). □

Corollary 3.7 *Suppose that the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded. Then, from every sequence $\{\phi_i\}$ of solutions to* (20) *such that the sequence of initial points $\{\phi_i(0)\}$ is bounded one can extract a subsequence that converges pointwise to a solution to* (20)*.*

Proof. Let K be a compact set such that $\phi_i(0) \in K$ for every $i \in \mathbb{N}$. Then, for every $i \in \mathbb{N}$ and $j \in \mathbb{N}$, $\phi_i(j) \in F^j(K)$. By 3.4, F^j is locally bounded, and so $F^j(K)$ is a bounded set. Noting that $i \in \mathbb{N}$, $j \leq J$, one has $\phi_i(j) \in K \cup F(K) \cup \cdots \cup F^J(K)$ concludes the proof. □

Lemma 3.8 *Suppose that the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded and let $K \subset \mathbb{R}^n$ be compact.*

(a) Let $O \subset \mathbb{R}^n$ be open. If, for every solution ϕ to (20) *with $\phi(0) \in k$ there exists $j \in \mathbb{N}_0$ such that $\phi(j) \in O$ then there exists J such that, for every solution ϕ to* (20) *with $\phi(0) \in k$ there exists $j \leq J$ such that $\phi(j) \in O$.*

(b) For every $J \in \mathbb{N}_0$, the reachable set

$$\mathcal{R}_{\leq J}(K) = \{\phi(j) \,|\, \phi \text{ is a solution to } (20),\ \phi(0) \in K,\ j \in \{0, 1, 2, \ldots, J\}\}$$

is compact.

Proof. Suppose that (a) fails. Then there exists a sequence of solutions $\{\phi_i\}$ with $\{\phi_i(0)\} \in K$ such that $\|\phi_i(j)\| \notin O$ for every $i \in \mathbb{N}$, every $j \leq J$ Because K is compact and the complement of O is closed, Corollary 3.7 yields a solution ϕ which satisfies $\|\phi(j)\| \notin O$ for every $j \in \mathbb{N}_0$. This is a contradiction. For (b), it is enough to note that

$$\mathcal{R}_{\leq J}(K) = K \cup F(K) \cup F^2(K) \cup \cdots \cup F^J(K),$$

that for every $j \in \mathbb{N}_0$ F^j is outer semicontinuous and locally bounded thanks to Proposition 3.4, and hence $F^j(K)$ is compact for every $j \in \mathbb{N}_0$. □

Proposition 3.9 *Suppose that the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded. If* 0 *is globally asymptotically stable for* (20) *then*

(a) 0 *is* uniformly (Lyapunov) stable*, that is, there exists a continuous and nondecreasing function $\alpha : [0, \infty) \to [0, \infty)$ such that $\alpha(0) = 0$ and, for every $\phi \in \mathcal{S}$ and every $j \in \mathbb{N}_0$, $\|\phi(j)\| \leq \alpha(\|\phi(0)\|)$.*

(b) 0 *is* uniformly attractive *for* (20)*, that is, for every $\Delta, \delta > 0$ there exists $J > 0$ such that, for every $\phi \in \mathcal{S}(\Delta\mathbb{B})$ and every $j > J$ one has $\phi(j) \in \delta\mathbb{B}$.*

Proof. For (a), consider a function defined at each $r \geq 0$ by

$$\gamma(r) = \sup\{\|\phi(j)\| \,|\, \phi \in \mathcal{S}(r\mathbb{B}),\ j \in \mathbb{N}_0\}.$$

Clearly, $\gamma(0) = 0$; by Lyapunov stability of 0 for (20), γ is continuous at 0; and, by construction, γ is nondecreasing on $[0, \infty)$. If γ has finite values for every r, then there exists a continuous α

with $\alpha(r) \geq \gamma(r)$ which has all of the needed properties. Suppose then, to the contrary, that for some $r > 0$ there exists a sequence of solutions $\phi_i \in \mathcal{S}(r\mathbb{B})$ to (20) and a sequence of $j_i \in \mathbb{N}_0$ such that $\phi_i(j_i) \to \infty$. Local boundedness of F ensures that $\{\phi_i\}$ is locally bounded. By Theorem 3.6, there exists a subsequence $\{\phi_{i_k}\}$ of $\{\phi_i\}$ which converges pointwise to a solution ϕ to (20). This solution converges to 0 because of the assumed attractivity. Invoke Lyapunov stability of 0 to obtain $\delta > 0$ that corresponds to $\varepsilon = 1$. Then, for some $J \in \mathbb{N}_0$, $\phi(J) \in 0.5\delta\mathbb{B}$, and so, for every large enough k, $\phi_{i_k}(J) \in \delta\mathbb{B}$. Then, by the choice of δ, $\phi_{i_k}(j) \in \mathbb{B}$ for every $j > J$. Hence, it must be that $j_{i_k} \leq J$. But this is impossible because of local boundedness of the sequence $\{\phi_i\}$. This proves that $\gamma(r) < \infty$ for every $r \geq 0$. The proof of (b) is similar and left to the reader. □

Building on the arguments in the proof of Propostion 3.9 one can show that there exists a function $\beta : [0, \infty)^2 \to [0, \infty)$ which is 0 at 0 and strictly increasing in the first argument and decreasing to 0 as the second argument goes to infinity, and such that, for every solution ϕ, one has $\|\phi(j)\| \leq \beta(\|\phi(0)\|, j)$. In fact, the following can be said, thanks to [34].

Fact 3.10 *If* 0 *is uniformly Lyapunov stable and globally uniformly attractive for* (20)*, then there exist two functions* $\alpha_1, \alpha_2 : [0, \infty) \to [0, \infty)$ *which are continuous, strictly increasing,* $\alpha_i(0) = 0$, *and* $\lim_{r\to\infty} \alpha_i(r) = \infty$*, and such that, for every* $\phi \in \mathcal{S}$,

$$\alpha_1(\|\phi(j)\|) \leq \alpha_2(\|\phi(0)\|)\, e^{-j}. \tag{21}$$

With the bound (21) one can construct a Lyapunov function for (20), following the idea of the proof of Proposition 2.9. For each $x \in \mathbb{R}^n$, set

$$V(x) = \sup\left\{\alpha_1(\|\phi(j)\|)\, e^j \mid \phi \in \mathcal{S}(x),\, j \in \mathbb{N}_0\right\}.$$

Then $\alpha_1(x) \leq V(x) \leq \alpha_2(x)$, so the function V is positive definite, finite-valued, and continuous at 0. Arguments as in the proof of Proposition 2.9 show that $V(y) \leq e^{-1}V(x)$ for every x, every $y \in F(x)$. Hence this V is a Lyapunov function for (20), but its regularity is questionable. It can be argued that V is upper semicontinuous, but an interesting question remains: can one build from this V a continuous Lyapunov function for (20), or obtain such a function using other methods?

Proposition 3.11 *Suppose that* 0 *is locally attractive for the linear difference inclusion* (15)*. Then* 0 *is globally asymptotically stable, in fact exponentially stable, for* (15).

Proof. First, note that the set-valued mapping $\mathcal{A}$ defining (15) is outer semicontinuous and locally bounded, and hence results like Theorem 3.6 and Corollary 3.7 apply. Second, note that local attractivity implies global attractivity for (15), thanks to homogeneity of $\mathcal{A}$. Indeed, let $\delta > 0$ come from the definition of local attractivity of 0 for (15). Because for every solution ϕ to (15) and every $\lambda \in \mathbb{R}$ the function $\lambda\phi$ is a solution to (15), and because for every $\phi(0)$ there exists $\lambda \neq 0$ such that $\lambda\phi(0) \in \delta\mathbb{B}$, every solution ϕ converges to 0 and hence 0 is globally attractive.

Now, for $k \in \mathbb{N}_0$ let $C_0 = \{x \in \mathbb{R}^n \mid 2^{-k-1} \leq \|x\| \leq 2^{-k}\}$. By Lemma 3.8 (a), there exists J such that for every solution ϕ to (15) with $\phi(0) \in C_0$ there exists $j \leq J$ with $\|\phi(j)\| < 1/2$. By Lemma 3.8 (b), there exists $\rho > 0$ such that $\mathcal{R}_{\leq J}(C_0) \subset \rho\mathbb{B}$. Homogeneity implies that for every $k \in \mathbb{N}$, for every solution ϕ to (15) with $\phi(0) \in C_k$ there exists $j \leq J$ with $\|\phi(j)\| < 2^{-k-1}$ and $\mathcal{R}_{\leq J}(C_k) \subset 2^{-k-1}\rho\mathbb{B}$.

Let ϕ be any solution to (15) with $\phi(0) \in \mathbb{B}$. Pick k_0 so that $\phi(0) \in C_{k_0}$. There exist $0 < j_1 \leq J$, $k_1 > k_0$ such that $\phi(j_1) \in C_{k_1}$, and for $j = 0, 1, \ldots, j_1$, $\phi(j) \in 2^{-k_0-1}\rho\mathbb{B}$. Similarly, there exist $j_1 < j_2 < j_1 + J$, $k_2 > k_1$ such that $\phi(j_2) \in C_{k_2}$, and for $j = j_1, j_1 + 2, \ldots, j_2$, $\phi(j) \in 2^{-k_1-1}\rho\mathbb{B}$. Repeating this argument shows that $\phi(j) \in 2^{-k_0-1}\rho\mathbb{B}$ for every $j \in \mathbb{N}_0$, and hence, any solution to (15) with $\phi(0) \in \mathbb{B}$ satisfies $\phi(j) \in \rho\mathbb{B}$ for every $j \in \mathbb{N}_0$. Homogeneity then suggests that given $\varepsilon > 0$ one can take $\delta = \varepsilon/\rho$ to satisfy the conditions for Lyapunov stability. This verifies global asymptotic stability of 0 for (15)

The arguments above in fact imply that every solution ϕ to (15) with $\phi(0) \in \mathbb{B}$ satisfies $\phi(j) \in \rho\mathbb{B}$ for every $j \in \{0, 1, \ldots, J\}$; $\phi(j) \in 2^{-1}\rho\mathbb{B}$ for every $j \in \{J, J+1, \ldots, 2J\}$; etc. Then the exponential bound (9) holds for every such solution, with $K = 2\rho$ and $\gamma = (\ln 2)/J$, and by homogeneity, this bound holds for every solution to (15). Hence, 0 is exponentially stable for (15) $\square$

Recall that the right-hand side of (15) is given in terms of $\mathcal{A}$, where $\mathcal{A}(x)$ is the convex hull of $\{A_1 x, A_2 x, \ldots, A_m x\}$. Define $\mathcal{A}^T$ by setting $\mathcal{A}^T(x)$ to be the convex hull of $\{A_1^T x, A_2^T x, \ldots, A_m^T x\}$ and consider the *dual linear difference inclusion*

$$x^+ \in \mathcal{A}^T(x). \tag{22}$$

The developments of Proposition 2.9, Lemma 2.10, Proposition 3.11 are summarized in the following theorem.

Theorem 3.12 *Consider $n \times n$ matrices $A_1, A_2, \ldots, A_m$. The following are equivalent:*

- 0 *is locally attractive for* (15).
- 0 *is exponentially stable for* (15).
- *There exists a convex Lyapunov function $V \in \mathcal{C}$ for* (15)*, in the sense that, for some $\delta \in (0, 1)$,*

$$V(x^+) \leq \delta V(x) \quad \text{for every } x^+ \in \mathcal{A}(x),\ x \in \mathbb{R}^n.$$

- *There exists a convex Lyapunov function $V \in \mathcal{C}$ for* (22).
- 0 *is exponentially stable for* (22).
- 0 *is locally attractive for* (22).

3.4 From discontinuous differential equations to differential inclusions

The following lovely example comes from [1]. In $\mathbb{R}^2$, consider the control system

$$\dot{x}_1 = (x_1^2 - x_2^2)u, \quad \dot{x}_2 = 2x_1 x_2 u, \tag{23}$$

The variable $x = (x_1, x_2)$ is the state of the system and u is the control variable or input. This control system has the following properties:

- For any choice of a function $u : [0, \infty) \to \mathbb{R}$ regular enough to not cause issues with existence of solutions to the non-autonomous differential equation

$$\dot{x} = (x_1^2 - x_2^2)u(t), \quad \dot{y} = 2x_1 x_2 u(t),$$

  the solution $x(t)$ is such that, for some $r \in \mathbb{R}$, $x_1^2(t) + (x_2(t) - r)^2 = r^2$ for every $t \geq 0$, except the case when $x_2(0) = 0$.

- For every initial condition $x(0)$ there exists $u : [0, \infty) \to \mathbb{R}$ such that the resulting solution converges to 0 as $t \to \infty$. In fact, one can accomplish this with using constant control values $u \equiv 1$ or $u \equiv -1$. Setting $u \equiv 1$ yields a differential equation

$$\dot{x}_1 = x_1^2 - x_2^2, \quad \dot{x}_2 = 2x_1 x_2.$$

  Solutions with $x_2(0) > 0$ move counterclockwise along circles $x_1^2 + (x_2 - r)^2 = r^2$ and converge to 0; solutions with $x_2(0) < 0$ move clockwise along circles $x_1^2 + (x_2 - r)^2 = r^2$ and converge to 0; and solutions with $x_2(0) = 0$, $x_1(0) < 0$ clearly converge to 0 as well. For $x_2(0) = 0$, $x_1(0) > 0$, one picks $u \equiv -1$ and obtains convergence to 0. This behavior does not lead to asymptotic stability of 0; Lyapunov stability is violated.

- There exists no continuous function $u : \mathbb{R}^2 \to \mathbb{R}$ such that the autonomous differential equation

$$\dot{x}_1 = (x_1^2 - x_2^2)u(x), \quad \dot{x}_2 = 2x_1x_2u(x), \tag{24}$$

has 0 asymptotically stable. To see this, note first that if $u(x) = 0$ for some $x \neq 0$, that x is an equilibrium and hence attractivity of 0 fails. Furthermore, Lyapunov stability dictates that, given $r > 0$ and x such that $x_1^2 + (x_2 - r)^2 = r^2$, one must have $u(x) < 0$ for every small enough $x_1 > 0$ and $u(x) > 0$ for every small enough $x_1 < 0$. Hence $u(x) = 0$ for some x on $x_1^2 + (x_2 - r)^2 = r^2$.

- There exists a discontinuous function $u : \mathbb{R}^2 \to \mathbb{R}$ such that (24) has 0 asymptotically stable. For example, one can consider

$$u(x) = \begin{cases} -1 & \text{if} \quad x_1 \geq 0 \\ 1 & \text{if} \quad x_1 < 0 \end{cases}. \tag{25}$$

Similarly to what was shown for discrete-time dynamics in Proposition 3.1, the effect of perturbations on a discontinuous differential equation is captured by passing to a differential inclusion. The set-valued mapping (27) is the *Krasovskii regularization* of f. The result below was envisioned by [19] and proved by [18]. Alternative approaches are in [9, Theorem 3.1.6] and [15, Section 4.5].

Fact 3.13 *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a locally bounded function, $T > 0$, and $\phi : [0, T] \to \mathbb{R}^n$ be absolutely continuous. The following are equivalent*

(a) There exist sequences of function $\phi_i, e_i, p_i : [0, T] \to \mathbb{R}^n$ with ϕ_i absolutely continuous and e_i, p_i measurable such that $\phi_i \to \phi$, $e_i(j) \to 0$, $p_i(j) \to 0$ uniformly on $[0, T]$ as $i \to \infty$ and, for almost every $t \in [0, T]$,

$$\dot{\phi}_i(t) = f(\phi_i(t) + e_i(t)) + p_i(t). \tag{26}$$

- *For almost every $t \in [0, T]$,*

$$\dot{\phi}(t) \in F(\phi(t)),$$

where, for every $x \in \mathbb{R}^n$,

$$F(x) = \bigcap_{\delta > 0} \operatorname{con} \overline{f(x + \delta\mathbb{B})}. \tag{27}$$

Consequently, the behavior of the control system (24) with the discontinuous feedback (25) and under small perturbations is reflected in the behavior resulting from

$$U(x) = \begin{cases} -1 & \text{if} \quad x_1 > 0 \\ [-1, 1] & \text{if} \quad x_1 = 0 \\ 1 & \text{if} \quad x_1 < 0 \end{cases}. \tag{28}$$

More precisely, the behavior of solutions to $\dot{x} = f(x, u(x))$ with $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined in (23) and with u as in (25) under perturbations is reflected in the behavior of solutions to $\dot{x} \in f(x, U(x))$, where U comes from (28). In particular, each point with $x_1 = 0$ is an equilibrium, in the sense that there exists a constant solution from that point.

This furthermore implies that, for the differential inclusion resulting from (24), (25), there exists no smooth Lyapunov function. Indeed, it can be argued, using the continuity of V and ∇V, that if $\nabla V(x) \cdot f(x, u(x)) \leq -\gamma V(x)$ for every $x \in \mathbb{R}^n$ then also $\nabla V(x) \cdot v \leq -\gamma/2V(x)$ for every $x \in \mathbb{R}^n$, and every $v \in f(x, U(x))$. That is, if V is a Lyapunov function for $\dot{x} = f(x, u(x))$ then it is also a Lyapunov function for $\dot{x} \in f(x, U(x))$. The latter is impossible because of equilibria other than 0.

To conclude this section, some general comments about differential inclusions are given. First, it is noted that if $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally bounded then the set-valued mapping $F : \mathbb{R} \rightrightarrows \mathbb{R}^n$ defined in (27) is locally bounded and outer semicontinuous. For differential inclusions given in terms of such mappings, asymptotic stability theory can be developed similarly to what was outlined in Section 3.3 for difference inclusions.

Below, a *solution* on $[0, T]$ to the differential inclusion

$$\dot{x} \in F(x) \tag{29}$$

is an absolutely continuous function $\phi : [0, T] \to \mathbb{R}^n$ such that, for almost every $t \in [0, T]$, $\dot{\phi}(t) \in F(\phi(t))$. The following result is the key to further results about differential inclusions, for example about asymptotic stability or reachable sets, just like Theorem 3.6 and Corollary 3.7 were the foundations for what was done for difference inclusions.

Theorem 3.14 *Suppose that the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded. Then, from every locally bounded sequence $\{\phi_i\}$ of solutions to* (20) *on* $[0, T]$ *one can extract a subsequence that converges uniformly to a solution to* (20) *on* $[0, T]$.

4 Invariance, set convergence, and hybrid dynamics

Invariance of a set under a differential equation or inclusion is the property that solutions from the set remain in the set. When solutions are not unique for some initial points, this concept splits into weak invariance, also called viability, which requires that some solution from each initial point remain in the set, and strong invariance, requiring that every solution remain in the set. Necessary and sufficient conditions for these properties, utilizing the concept of tangent cones, are stated in Section 4.4.

Invariance plays a role in predicting the asymptotic behavior of solutions to differential equations or inclusions. The classical invariance principle dates back to the work of Krasovskii [20] and LaSalle [21]. It lets one predict the asymptotic behavior of solutions to a differential equation if one has a nondecreasing function, rather than a strictly decreasing Lyapunov function. The classical result is recalled below as Theorem 4.6 in Section 4.2. Some ingredients of the classical result can be conveniently viewed through the concept of set convergence, summarized in Section 4.1 below, and this then lets one extend the result far beyond the classical setting. One such extension is presented, on an example, in Section 4.3.

4.1 Set convergence

Let $\{S_i\}$ be a sequence of sets in $\mathbb{R}^n$. The following define an "upper" limit, "lower" limit, and the limit of the sequence. Terminology follows that of [31].

- the *outer limit* of the sequence $\{S_i\}$, denoted $\limsup_{i\to\infty} S_i$, is the set of all $x \in \mathbb{R}^n$ such that, for some subsequence $\{S_{i_k}\}$ and some $x_k \in S_{i_k}$, $x = \lim_{k\to\infty} x_i$;
- the *inner limit* of the sequence $\{S_i\}$, denoted $\liminf_{i\to\infty} S_i$, is the set of all $x \in \mathbb{R}^n$ such that, for some $x_i \in S_i$, $x = \lim_{i\to\infty} x_i$;
- the *limit* of the sequence $\{S_i\}$, denoted $\lim_{i\to\infty} S_i$, is said to exist if the inner and the outer limits are the same, and then $\lim_{i\to\infty} S_i = \limsup_{i\to\infty} S_i = \liminf_{i\to\infty} S_i$

Example 4.1 The following illustrate set convergence:

- If $S_i = \{x_i\}$ for every $i \in \mathbb{N}$, then $\limsup S_i$ is the set of cluster points of the sequence $\{x_i\}$, $\liminf S_i$ exists only if $\lim x_i$ exists in the usual sense, and then $\limsup S_i = \lim S_i = \lim x_i$.
- If $S_i = S$ for every $i \in \mathbb{N}$, for some set $S \subset \mathbb{R}^n$, then $\lim S_i = \overline{S}$. In general, the outer and inner limits are always closed.
- If $S_i = \{x \in \mathbb{R}^n \,|\, x_2 = a_i x_1\}$ and $\lim a_i = a$ then $\lim S_i = \{x \in \mathbb{R}^n \,|\, x_2 = a x_1\}$. Note that, for every $\varepsilon > 0$, it is never the case that $S_i \subset S + \varepsilon \mathbb{B}$.

△

Concepts of set convergence can be used to describe continuity of set-valued mappings. That is, a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous if for every $x_i \to x$, $\limsup_{i\to\infty} F(x_i) \subset F(x)$; it is inner semicontinuous if for every $x_i \to x$, $\liminf_{i\to\infty} F(x_i) \supset F(x)$; and continuous if for every $x_i \to x$, $\lim_{i\to\infty} F(x_i)$ exists and equals $F(x)$.

For compact sets, one can characterize their convergence through the Hausdorff metric. With some care, this can be applied to sets that are not necessarily bounded. The following is a part of [31, Theorem 4.10].

Fact 4.2 *Let $S_i, S \subset \mathbb{R}^n$ be closed.*

- $\limsup_{i\to\infty} S_i \subset S$ *if and only if for every $\rho > 0$, $\varepsilon > 0$ there exists I such that for every $i > I$,*
$$S_i \cap \rho \mathbb{B} \subset S + \varepsilon \mathbb{B};$$
- $\liminf_{i\to\infty} S_i \supset C$ *if and only if for every $\rho > 0$, $\varepsilon > 0$ there exists I such that for every $i > I$,*
$$S \cap \rho \mathbb{B} \subset S_i + \varepsilon \mathbb{B};$$

The classical Bolzano-Weierstrass theorem says that a bounded sequence of points in $\mathbb{R}^n$ has a convergent subsequence. This generalizes to sequences of sets, as stated below, following [31, Theorem 4.18]. In particular, a bounded sequence of sets has a convergent subsequence.

Fact 4.3 *For every sequence $\{S_i\}$ of sets $S_i \subset \mathbb{R}^n$, either the sequence diverges, in the sense that for every $M > 0$ there exists I such that for every $i > I$, $S_i \cap M\mathbb{B} = \emptyset$; or the sequence has a subsequence which converges to a nonempty limit.*

4.2 The classical invariance principle

Knowing the invariant sets for a differential equation helps one predict the asymptotic behavior of bounded solutions. The reason for this is as follows. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz continuous. Let $\phi : [0, \infty) \to \mathbb{R}^n$ be a solution. The Ω*-limit* of ϕ, denoted $\omega(\phi)$, is the set defined by
$$\omega(\phi) = \left\{ x \in \mathbb{R}^n \,|\, \exists\, t_i \geq 0, \ \lim_{i\to\infty} t_i = \infty \ \text{ such that } \ \lim_{i\to\infty} \phi(t_i) = x \right\}.$$
Then

- $\omega(\phi)$ is closed and, unless $\lim_{t\to\infty} \|\phi(t)\| = \infty$, nonempty.

It is easy to show this directly. Alternatively, one can consider sets $S_i = \{\phi(t) \,|\, t \geq i\}$, note that $\omega(\phi) = \limsup S_i$, and recall that the outer limit of a sequence of sets is always closed and, unless the sets "escape to infinity", nonempty. Suppose now that ϕ is bounded. Then:

- $\omega(\phi)$ is compact and ϕ converges to $\omega(\phi)$ uniformly, in the sense that for every $\varepsilon > 0$ there exists T such that for every $t > T$, $\phi(t) \in \omega(\phi) + \varepsilon \mathbb{B}$.

The proof of convergence by contradiction is simple. Finally,

- $\omega(\phi)$ is *invariant*: for every $x \in \Omega(\phi)$ there exists a solution $\psi : \mathbb{R} \to \mathbb{R}^n$ such that $\psi(t) \in \Omega(\phi)$ for every $t \in \mathbb{R}$.

To prove this, pick $x \in \Omega(\phi)$, find $t_i \to \infty$ such that $\phi(t_i) \to x$, define $\psi_i : [0, \infty) \to \mathbb{R}^n$ by $\psi_i(t) = \phi(t - t_i)$, note that $\psi_i(0) \to x$, and because solutions to $\dot{x} = f(x)$ depend continuously on initial conditions, ψ_i converge locally uniformly to a solution ψ with $\psi(0) = x$. Because ϕ converges uniformly to $\omega(\phi)$, $\psi(t) \in \omega(\phi)$ for every $t \geq 0$. Hence, ψ verifies forward invariance (the domain of ψ is $[0, \infty)$), and this simpler proof was given for illustration purposes. To prove invariance, given $x \in \Omega(\phi)$, pick $t_i > 2i$ such that $\phi(t_i) \to x$ and consider $\psi_i : [-i, \infty) \to \mathbb{R}^n$ given by$\psi_i(t) = \phi(t - t_i)$. The rest of the argument is the same.

Remark 4.4 Ω-limits are invariant for differential and difference inclusions as well, if the dynamics are given by outer semicontinuous and locally bounded set-valued mappings. Indeed, the continuous dependence of solutions on initial conditions, as used in the discussion above, is not essential. Given a bounded sequence of solutions ψ_i with $\psi_i(0) \to x$ it is enough to know that a subsequence converges uniformly and in the limit yields a solution with initial point x. For differential inclusions, this follows from Theorem 3.14; for difference inclusions one invokes Theorem 3.6 and Corollary 3.7. Roughly, outer semicontinuous dependence of solutions on initial conditions is sufficient for the argument to work. △

Example 4.5 Consider the differential equation $\dot{x} = f(x)$ where $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x) = \begin{pmatrix} -x_2 + \left(1 - x_1^2 - x_2^2\right) x_1 \\ x_1 + \left(1 - x_1^2 - x_2^2\right) x_2 \end{pmatrix}.$$

One can see that 0 is an equilibrium, and one can verify that the set $\{x \in \mathbb{R}^2 \,|\, x_1^2 + x_2^2 = 1\}$ is invariant: on this set, the differential equation reduces to $\dot{x}_1 = -x_2$, $\dot{x}_2 = x_1$. To predict the behavior of solutions from other initial points, consider how the solution's distance from 0 changes. For every solution ϕ, one has

$$\frac{d}{dt} \left(\|\phi(t)\|\right)^2 = 2\left(1 - \|\phi(t)\|^2\right) \|\phi(t)\|^2.$$

This suggests that when $\|\phi(t)\| > 1$, $\|\phi(t)\|$ is decreasing and when $0 < \|\phi(t)\| < 1$, $\|\phi(t)\|$ is increasing. △

Theorem 4.6 *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz continuous. Suppose that there exists $V : \mathbb{R}^n \to \mathbb{R}$ such that, for every solution ψ to $\dot{x} = f(x)$ one has*

$$\frac{d}{dt} V(\psi(t)) \leq 0.$$

Let $\phi : [0, \infty) \to \mathbb{R}^n$ be a bounded solution to $\dot{x} = f(x)$. Then, for some $r \in \mathbb{R}$, ϕ converges to the union of all invariant subsets of $V^{-1}(r)$.

Proof. Let $\phi : [0, \infty) \to \mathbb{R}^n$ be a bounded solution to $\dot{x} = f(x)$. The discussion at the beginning of this section showed that $\omega(\phi)$ is nonempty, compact, invariant, and that ϕ converges to this set. Because $V(\phi(t))$ is nonincreasing on $[0, \infty)$ and bounded below, $V(\phi(t))$ converges. Let $r = \lim_{t \to \infty} V(\phi(t))$. By definition of $\omega(\phi)$, for every $x \in \omega(\phi)$ there exist $t_i \to \infty$ such that $\phi(t_i) \to x$. Then $V(x) = \lim_{i \to \infty} V(\phi(t_i)) = r$. This finishes the proof. □

Example 4.7 Consider the differential equation from Example 4.5. The analysis in that example suggests that solutions from nonzero initial conditions converge to the unit circle. To fully justify this, with the help of Theorem 4.6, consider the function $V(x) = (x^2 + y^2 - 1)^2$. Then, for every solution ψ, one has

$$\frac{d}{dt}V(\psi(t)) = -8(x_1^2 + x_2^2 - 1)^2(x_1^2 + x_2^2) = -8V(x)\|x\|^2.$$

Theorem 4.6 implies that every solution ϕ converges to the largest invariant subset of $V^{-1}(r)$, for some real r. The set $V^{-1}(r)$ does not have any invariant subsets unless $r = 0$ or $r = 1$. For $r = 1$, $V^{-1}(1)$ is the unit circle, which is invariant For $r = 1$, $V^{-1}(1)$ is the union of $\{0\}$ and the circle of radius $\sqrt{2}$. The only invariant subset of this union is $\{0\}$. Hence, every solution converges either to $\{0\}$ or to the unit circle. △

4.3 Invariance in a hybrid system

This section discusses a dynamical system in $\mathbb{R}^2$ which can be informally introduced as follows. Let C be the union of the square with vertices $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ and the triangle with vertices $(2,0)$, $(2,1)$, and $(3,1)$. Let D_1 be the segment between $(1,0)$ and $(1,1)$; let D_2 be the segment between $(2,0)$ and $(3,1)$. If $x \in C$ and when possible, the solution from x should satisfy the differential equation $\dot{x} = f(x)$, with $f(x) \equiv (1,0)$. When $x \in D_1$, the solution from x should satisfy the difference equation $x^+ = (2, x_2^2)$. When $x \in D_2$, the solution from x should satisfy $x^+ = (0, x_2)$.

For background on such, and similar, dynamical systems, which combine continuous-time and discrete-time dynamics, the reader can consult [35], [15], [14] and the references therein.

One can try to parameterize the behaviors of this roughly introduced dynamics by $t \in \mathbb{R}$. Then, for example, the "solution" ϕ from $(0,1)$ is $\phi(t) = (t,1)$ for $t \in [0,1)$, $\phi(t) = (t+1,1)$ for $t \in [1,2)$, and $\phi(t) = \phi(t-2)$ for $t \geq 2$. This is a discontinuous function of t, taken to be right-continuous just by choice, as one could use a left-continuous function instead. Another example of a "solution", from $(0,.5)$ is $\phi(t) = (t,.5)$ for $t \in [0,1)$, $\phi(t) = (t+1,.25)$ for $t \in [1,1.25)$, $\phi(t) = (t-1.25,.25)$ for $t \in [1.25,2.25)$, $\phi(t) = (t-.25,.0625)$ for $t \in [2.25,2.3125)$, etc. Notice that the discontinuities of the two presented "solutions" occur at different times. This can have unexpected consequences: given the "solution" ϕ from $(0,1)$ as above, and the "solution" ϕ_ε from $(0,1-\varepsilon)$, one has $\sup_{t\in[0,T]} \|\phi(t) - \phi_\varepsilon(t)\| \geq 1$ for any large enough T.

Questions that may arise after a closer look at the dynamics described above are:

- What is the "solution" from $(0,0)$? How to define a solution in general.
- If $[0,1] \times \{0\} \cup \{(2,0)\}$ is the Ω-limit of a solution from, say, $(0,.5)$, in what sense is it invariant? How to state and prove an appropriate invariance principle?
- The behaviors from $(0,1)$ and $(0,1-\varepsilon)$ appear similar — how should the distance between the corresponding solutions be measured to reflect this?

The remainder of this section outlines one particular approach to the issues above, following [15]. Some definitions are simplified for the example at hand.

A *hybrid time domain* is a subset of $\mathbb{R}^2$ given by

$$\bigcup_{i=0}^{\infty} [t_i, t_{i+1}] \times \{i\},$$

where $t_0 = 0$ and $\{t_i\}$ is a nondecreasing sequence of nonnegative numbers. Recall the set C, let $D = D_1 \cup D_2$, consider $f : C \to \mathbb{R}^2$ given by $f(x) = (1,0)$ for every $x \in C$, and $g : D \to \mathbb{R}^2$

given, for $x \in D$, by

$$g(x) = \begin{cases} (2, x_2^2) & \text{if} \quad x \in D_1 \\ (0, x_2) & \text{if} \quad x \in D_2 \end{cases}.$$

A *solution* to the dynamics described above is a function ϕ defined on a hybrid time domain, denoted from now on by $\operatorname{dom}\phi$, satisfying the following:

(a) f $(t, j) \in \operatorname{dom}\phi$ is such that, for some $\varepsilon > 0$, $(t-\varepsilon, t+\varepsilon) \times \{j\} \subset \operatorname{dom}\phi$, then $t \mapsto \phi(t, j)$ is differentiable on $(t-\varepsilon, t+\varepsilon)$ and

$$\frac{d}{dt}\phi(t) = f(\phi(t));$$

(b) if $(t, j) \in \operatorname{dom}\phi$ and $(t, j+1) \in \operatorname{dom}\phi$ then

$$\phi(t, j+1) = g\left(\phi(t, j)\right).$$

Hence, for example, the solution ϕ with initial condition $(0, 1)$ has the hybrid time domain $\operatorname{dom}\phi = \bigcup_{i=0}^{\infty}[i, i+1] \times \{i\}$ and is given by $\phi(t, 0) = (t, 1)$ if $t \in [0, 1]$, $\phi(t, 1) = (t+1, 1)$, and $\phi(t, j) = \phi(t-2, j-2)$ for $t \geq 2$, $j \geq 2$ with $(t, j) \in \operatorname{dom}\phi$.

With this definition, there is no ambiguity about the solution from $(0, 0)$. This solution ϕ has $\operatorname{dom}\phi = \bigcup_{i=0}^{\infty}[t_i, t_{i+1}] \times \{i\}$ with $t_0 = 0$, $t_1 = 1$, $t_2 = 1$, $t_3 = 1$, $t_4 = 2$, $t_5 = 2$, $t_6 = 2$, $t_7 = 3$ etc., and is given by $\phi(t, 0) = (t, 0)$ for $t \in [0, 1]$, $\phi(1, 1) = (2, 0)$, $\phi(t, 2) = (t-1, 0)$ for $t \in [1, 2]$, $\phi(2, 3) = (2, 0)$, etc.

A natural way to define Ω-limits of solutions is

$$\omega(\phi) = \left\{ x \in \mathbb{R}^n \mid \exists\, (t_i, j_i) \in \operatorname{dom}\phi,\ \lim_{i\to\infty} t_i + j_i = \infty \ \text{ such that } \ \lim_{i\to\infty} \phi(t_i, j_i) = x \right\}.$$

For the remainder of this section, let ϕ be the solution from $(0, 0.5)$. Then

$$\omega(\phi) = [0, 1] \times \{0\} \cup \{(2, 0)\},$$

and the forward invariance of this set can be verified by inspection. For example, the previous paragraph described the solution from $(0, 0) \in \omega(\phi)$. Backward invariance is not discussed, to avoid the trouble with defining "backward" hybrid time domains. The question now is: whether forward invariance of $\omega(\phi)$ can be established without inspection, by using an approach similar to the standard one outlined in Section 4.2.

Take any $x \in \omega(\phi)$, let $(t_i, j_i) \in \operatorname{dom}\phi$ be such that $t_i + j_i \to \infty$ and $\phi(t_i, j_i) \to x$. Define $\phi_i(t, j) = \phi(t+t_i, j+j_i)$ with an appropriate change of domains and note that ϕ_i form a sequence of solutions with $\phi_i(0, 0) \to x$. Let

$$\operatorname{gph}\phi_i = \left\{ (t, j, y) \in \mathbb{R}^{n+2} \mid (t, j) \in \operatorname{dom}\phi_i,\ y = \phi_i(t, j) \right\}$$

be the graph of ϕ_i. Because $\phi_i(0, 0) \to x$, Fact 4.3 implies that the sequence $\{\operatorname{gph}\phi_i\}$ has a convergent subsequence $\{\operatorname{gph}\phi_{i_k}\}$. That is, $\operatorname{gph}\phi_{i_k}$ converge, as sets, to some set $L \subset \mathbb{R}^{n+2}$. If L is a graph of a solution, say ψ, with $\psi(0, 0) = x$ and if $\psi(t, j) \subset \omega(\phi)$ for all $(t, j) \in \operatorname{dom}\phi$, then ψ verifies forward invariance of $\omega(\phi)$ from x.

The key steps to showing that L is a graph of a solution ψ, in other words that the *graphical limit* of the sequence of solutions $\{\phi_{i_k}\}$ is a solution, are as follows:

- Show that the projection of L onto the (t, j)-space, is a hybrid time domain. It is important here that the sequence $\{\phi_{i_k}\}$ is bounded.

- Given that L is the graph of some (possibly set-valued) mapping ψ with $\operatorname{dom}\psi$ being a hybrid time domain, show that ψ is in fact a solution. This can be done directly from the definition of a solution, and may require passing from convergence of graphs of solutions to a differential equation to uniform convergence.
- Argue that ϕ converges to $\omega(\phi)$ and use this to claim that the range of ψ is contained in $\omega(\phi)$.

This is left as an exercise to the reader. Details can be found in [16] or [15], where a more general case, of systems combining differential inclusions and difference inclusions, was treated. Note that Theorem 3.14 and Theorem 3.6, Corollary 3.7 are what are needed in such a general setting.

4.4 Tangent cones, invariance, and viability

The tangent cone $T_C(x)$ to the set C at the point x was defined in (19). For every point $x \in \overline{S}$ which is not isolated, $T_C(x)$ is nonempty, closed, and contains a nonzero element. Then, automatically, $T_C(x)$ is unbounded because it is a cone. For each $x \notin \overline{C}$, $T_C(x) = \emptyset$. When C is a smooth manifold, the tangent cone T_C amounts to the concept of a tangent space. More interesting examples are given below.

Example 4.8 Let $C = [a, b] \subset \mathbb{R}$. Then

$$T_C(x) = \begin{cases} (-\infty, 0] & \text{if} \quad x = a, \\ (-\infty, \infty) & \text{if} \quad x \in (a, b), \\ [0, \infty) & \text{if} \quad x = b. \end{cases}$$

For boxes in $\mathbb{R}^n$, i.e., sets given by products of intervals

$$C = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$

the tangent cone $T_C(x)$ can be found coordinate-wise:

$$T_C(x) = T_{[a_1,b_1]}(x_1) \times T_{[a_2,b_2]}(x_2) \times \cdots \times T_{[a_n,b_n]}(x_n).$$

Some further examples are

$$C = \left\{ x \in \mathbb{R}^2 \,|\, x_1 \geq 0,\, x_2 \leq \sqrt{x_1} \right\}, \qquad T_C(0) = \left\{ w \in \mathbb{R}^2 \,|\, w_1 \geq 0 \right\}.$$

$$C = \left\{ x \in \mathbb{R}^2 \,|\, x_1 \geq 0,\, 0 \leq x_2 \leq \sin x_1 \right\}, \qquad T_C(0) = \left\{ w \in \mathbb{R}^2 \,|\, w_1 \geq 0,\, w_2 \leq w_1 \right\}.$$

△

Tangent cones play a role in the analysis of constrained differential equations and inclusions. Given a set $C \subset \mathbb{R}^n$ and a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, consider

$$\dot{x} \in F(x), \quad x \in C, \tag{30}$$

with $\phi : [0, T] \to \mathbb{R}^n$ being a *solution* to (30) if ϕ is absolutely continuous, satisfies $\dot{\phi}(t) \in F(\phi(t))$ for almost all $t \in [0, T]$, and $\phi(t) \in C$ for every $t \in [0, T]$. *Viability* (also called *weak invariance*) is the property that for every $\xi \in C$ there exists a solution to (30). Fact 4.9 gives necessary and sufficient conditions for this property. For details, see [2, Propositions 3.4.1, 3.4.2].

Fact 4.9 *Suppose that $C \subset \mathbb{R}^n$ is closed and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded.*

(a) If $\phi : [0, T] \to \mathbb{R}^n$ *for some* $T > 0$ *is a solution to* (30) *with* $\phi(0) = \xi$*, then*

$$F(\xi) \cap T_C(\xi) \neq \emptyset.$$

(b) Given $\xi \in C$*, if there exists a neighborhood* U *of* ξ *such that for every* $x \in U \cap C$*,*

$$F(x) \cap T_C(x) \neq \emptyset,$$

then there exists $T > 0$ *and a solution* $\phi : I := [0, T] \to \mathbb{R}^n$ *to* (30) *with* $\phi(0) = \xi$*.*

The set C is *invariant* (also called *strongly invariant*) under $\dot{x} \in F(x)$ if for every $\xi \in C$, every solution to $\dot{x} \in F(x)$ remains in C. In the situations where solutions to $\dot{x} \in F(x)$ are unique, invariance and viability are the same. For the fact below, see [2, Theorem 4.3.6, Corollary 5.2.3]. Some assumptions can be weakened.

Fact 4.10 *Suppose that* $C \subset \mathbb{R}^n$ *is closed and* $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is outer semicontinuous and locally bounded.*

(a) C *is invariant under* F *if* $\operatorname{dom} F = C$ *and, for every* $x \in C$*,*

$$F(x) \subset T_C(x).$$

(b) If the interior of C *is nonempty and, for every* x *on the boundary of* C*,*

$$F(x) \subset \mathbb{R}^n \setminus T_{\mathbb{R}^n \setminus C}(x),$$

then for every solution ϕ *to* $\dot{x} \in F(x)$ *with* $\phi(0)$ *on the boundary of* C *there exists* $T > 0$ *such that* $\phi(t)$ *is in the interior of* C *for every* $t \in (0, T]$*.*

References

[1] Z. Artstein. Stabilization with relaxed controls. *Nonlinear Anal.*, 7(11):1163–1173, 1983.

[2] J.-P. Aubin. *Viability theory.* Birkhauser, 1991.

[3] J.-P. Aubin and A. Cellina. *Differential Inclusions.* Springer-Verlag, 1984.

[4] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis.* Birkhauser, 1990.

[5] N.E. Barabanov. On the Lyapunov exponent of discrete inclusions. II. *Avtomat. i Telemekh.*, (3):24–29, 1988. translation in Automat. Remote Control 49 (1988), no. 3, part 1, 283-287.

[6] N.E. Barabanov. Stability of inclusions of linear type. In *Proc. 14th American Control Conference*, volume 5, pages 3366 – 3370, 1995.

[7] J.M. Borwein and A.S. Lewis. *Convex analysis and nonlinear optimization.* Springer-Verlag, 2000.

[8] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004. available online.

[9] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, 1983.

[10] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, and P.R. Wolenski. *Nonsmooth Analysis and Control Theory.* Springer-Verlag, 1998.

[11] W.P. Dayawansa and C.F. Martin. A converse Lyapunov theorem for a class of dynamical systems which undergo switching. *IEEE Trans. Automat. Control*, 44(4):751–760, 1999.

[12] A.L. Dontchev and R.T. Rockafellar. *Implicit functions and solution mappings.* Springer, 2009. A view from variational analysis.

[13] R. Goebel, T. Hu, and A.R. Teel. Dual matrix inequalities in stability and performance analysis of linear differential/difference inclusions. In *Current trends in nonlinear systems and control.* Birkhauser, 2006.

[14] R. Goebel, R.G. Sanfelice, and A.R. Teel. Hybrid dynamical systems. robust stability and control for systems that combine continuous-time and discrete-time dynamics. *IEEE Control Systems Magazine*, 29:28–93, 2009.

[15] R. Goebel, R.G. Sanfelice, and A.R. Teel. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness.* Princeton University Press, 2012.

[16] R. Goebel and A.R. Teel. Solutions to hybrid inclusions via set and graphical convergence with stability theory applications. *Automatica*, 42:573–587, 2006.

[17] R. Goebel, A.R. Teel, T. Hu, and Z. Lin. Conjugate convex Lyapunov functions for dual linear differential inclusions. *IEEE Trans. Automat. Contr.*, 51:661–666, 2006.

[18] O. Hájek. Discontinuous differential equations, I. *J. Diff. Eq.*, 32:149–170, 1979.

[19] H. Hermes. Discontinuous vector fields and feedback control. In *Differential Equations and Dynamical Systems*, pages 155–165. Academic Press, 1967.

[20] N. N. Krasovskiĭ. *Nekotorye zadachi teorii ustoichivosti dvizheniya.* Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1959.

[21] J. P. LaSalle. Some extensions of Liapunov's second method. *IRE Trans. Circuit Theory*, 7(4):520–527, 1960.

[22] D. Liberzon. *Switching in Systems and Control*. Systems and Control: Foundations and Applications. Birkhauser, 2003.

[23] P.D. Loewen. *Optimal Control via Nonsmooth Analysis*. American Mathematical Society, 1993.

[24] A. M. Lyapunov. The general problem of the stability of motion. *Internat. J. Control*, 55(3):521–790, 1992. Translated by A. T. Fuller from Édouard Davaux's French translation (1907) of the 1892 Russian original.

[25] A.P. Molchanov and Ye.S. Pyatnitskiy. Criteria of asymptotic stability of differential and difference inclusions endountered in control theory. *System Control Lett.*, 13:59–64, 1989.

[26] T. Monovich and M. Margaliot. Analysis of discrete-time linear switched systems: a variational approach. *SIAM J. Control Optim.*, 49(2):808–829, 2011.

[27] B.S. Mordukhovich. *Variational analysis and generalized differentiation. I.* Springer-Verlag, 2006. Basic theory.

[28] B.S. Mordukhovich. *Variational analysis and generalized differentiation. II.* Springer-Verlag, 2006. Applications.

[29] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[30] R.T. Rockafellar. *Conjugate Duality and Optimization.* Number 16 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, 1974.

[31] R.T. Rockafellar and R. J-B Wets. *Variational Analysis.* Springer, 1998.

[32] R. Shorten, F. Wirth, O. Mason, K. Wulff, and C. King. Stability criteria for switched and hybrid systems. *SIAM Rev.*, 49(4):545–592, 2007.

[33] G.V. Smirnov. *Introduction to the theory of differential inclusions*, volume 41 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2002.

[34] E.D. Sontag. Comments on integral variants of ISS. *Systems Control Lett.*, 34(1-2):93–100, 1998.

[35] A. van der Schaft and H. Schumacher. *An Intruduction to Hybrid Dynamical Systems*, volume 251 of *Lect. Notes in Contr. and Inform. Sci.* Springer, 2000.

[36] R. Vinter. *Optimal control.* Systems & Control: Foundations & Applications. Birkhauser, 2000.