STABILITY AND CONVERGENCE OF EULER'S METHOD FOR STATE-CONSTRAINED DIFFERENTIAL INCLUSIONS

ROBERT BAIER*, ILYES AÏSSA CHAHMA†, AND FRANK LEMPIO‡

Abstract. A discrete stability theorem for set-valued Euler's method with state constraints is proven. This theorem is combined with known stability results for differential inclusions with so-called smooth state constraints. As a consequence, order of convergence equal to 1 is proven for set-valued Euler's method, applied to state-constrained differential inclusions.

Key words. Filippov theorem, set-valued Euler's method, differential inclusions with state constraints, stability and convergence of discrete approximations

AMS subject classifications. 49J24, 65L20, 34K28, 34A60

:

1. Introduction and Preliminaries. Differential inclusions appear in various fields of applications, e.g. in the study of (deterministic) perturbations of differential equations, in dynamical systems with discontinuous system equations, optimal control problems, viability theory, especially climate impact research, cf. e.g. [2, 3, 14, 10, 1, 6].

An important subclass consists of differential inclusions with additional monotonicity properties which, in general, guarantee uniqueness of the solution of the initial value problem (cf. e.g. [2, 3, 4, 5, 20, 21]). Differential inclusions with Lipschitz right-hand sides (with respect to Hausdorff distance) in the usual sense form another important subclass. The latter class is the principal focus of this paper which deals with stability and convergence properties of set-valued Euler's method for differential inclusions with state constraints.

The main result of this paper is the proof of a discrete stability theorem for a difference inclusion with state constraints in Section 3, which serves as a basis for the convergence analysis for set-valued Euler's method in Section 4. Intrinsically, this result is a variant of Gronwall-Filippov-Wazewski's theorem and in fact an existence theorem as well. Whereas the proofs for explicit difference inclusions with appropriate Lipschitz properties offer no difficulties, additional state constraints cause essential problems.

Fortunately, since some years there are remarkable stability results for state-constrained differential inclusions available in the literature, cf. [22, 15, 17, 18, 7, 8, 23]. But discrete analogues for the approximation of all feasible trajectories under comparably weak conditions are still missing. Therefore, we concentrate on the so-called smooth case where the state constraint is described by a single scalar inequality resp. by a smooth signed distance function. This case has already been treated in [6], but contrary to [6] we allow time-dependent state constraints and improve the final error estimate.

In Section 3 we give a rather complete analysis of the discrete situation which heavily relies on the proof strategy in [15, Theorem 4.1] for the continuous problem.

*University of Bayreuth, Chair of Applied Mathematics, D-95440 Bayreuth, Germany, e-mail: robert.baier@uni-bayreuth.de

†Banco Cetelem, S.A., C/ Retama, 3, 3ª Planta, 28045 Madrid, Spain, e-mail: aissa.chahma@cetelem.es

‡University of Bayreuth, Chair of Applied Mathematics, D-95440 Bayreuth, Germany, e-mail: frank.lempio@uni-bayreuth.de

In some respects, the discrete analysis is rather technical, and some additional difficulties have to be overcome. Especially, a discrete solution might not hit exactly the boundary of the state constraints, neighboring continuous solutions of feasible discrete solutions could violate the state contraints outside the grid, and additional error terms appear in Taylor expansions.

But, we want to emphasize urgently the fact, that only both stability results, the continuous **and** the discrete one together, will give us convergence results for discrete approximations of state-constrained differential inclusions. This is the essential subject of Section 4, where order of convergence $\mathcal{O}(h)$ with respect to the step-size h is proven for set-valued Euler's method in the presence of state constraints.

In Section 5, the results are applied to a differential inclusion resulting from a state-constrained bilinear control problem which originally served as an academic test example for unconstrained problems and was communicated by Petar Kenderov. The order of convergence of the reachable sets of Euler's difference inclusion with state constraints to the corresponding reachable sets of the differential inclusion is visualized by computer tests. For a more detailed discussion and applications to climate impact research cf. [6].

Hence, the main objective of this paper is the discrete approximation of the **whole** solution set of state-constrained differential inclusions, especially the whole feasible set of state-constrained optimal control problems. But, in addition, the authors are convinced that this methodology, if combined with sufficient optimality conditions, could turn out to be another conceptual approach to order of convergence proofs for numerical methods for the direct computation of optimal solutions, cf. e.g. [13, 12].

Naturally, convergence of the whole set of discrete solutions to the solution set of the continuous differential inclusion, implies the convergence of the corresponding reachable sets. Hence, at least for set-valued Euler's method we need not distinguish between these two aspects, but cf. in this connection the papers [24, 25] which extends the results in [11] for set-valued Euler's method to Runge-Kutta methods of order at least equal to 2 for problems without state constraints.

We denote by $\mathrm{AC}(I)$ the set of all absolutely continuous functions $y : I \to \mathbb{R}^n$ and by $\Theta : I \rightrightarrows \mathbb{R}^n$ a set-valued map with nonempty subsets of $\mathbb{R}^n$ as images.

PROBLEM 1.1. *Given an interval $I = [t_0, T]$, a nonempty set $Y_0 \subset \mathbb{R}^n$, set-valued maps $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $\Theta : I \rightrightarrows \mathbb{R}^n$ with nonempty images.*

Find all absolutely continuous solutions $y(\cdot)$ of the state-constrained differential inclusion (DIC)

$$y'(t) \in F(t, y(t)) \qquad (a.e.\ t \in I), \tag{1.1}$$

$$y(t) \in \Theta(t) \qquad (t \in I), \tag{1.2}$$

$$y(t_0) = y_0 \in Y_0. \tag{1.3}$$

Clearly, $y_0 \in \Theta(t_0)$ must be demanded as well.

The unconstrained problem (DI) is given by (1.1),(1.3). *The set of solutions of (DI) and (DIC) is denoted by $\mathcal{Y}[T, t_0, Y_0]$ resp. $\mathcal{Y}^{\Theta}[T, t_0, Y_0]$.*

ALGORITHM 1.2. *Euler's method for (DIC) in Problem 1.1 with $N \in \mathbb{N}$ as number of subintervals and step-size $h = \frac{T-t_0}{N}$ is given by*

$$\mathcal{Y}_N^\Theta[t_0, t_0, Y_0] := Y_0 \cap \Theta(t_0), \tag{1.4}$$

$$\mathcal{Y}_N^\Theta[t_{j+1}, t_0, Y_0] := \bigcup_{\eta_j \in \mathcal{Y}_N^\Theta[t_j, t_0, Y_0]} \big(\eta_j + hF(t_j, \eta_j)\big) \cap \Theta(t_{j+1}) \tag{1.5}$$

for $j = 0, ..., N-1$.

Problem (DDIC) describes the solution of (1.4)–(1.5), its set of solutions is denoted by $\mathcal{Y}_N^\Theta[T, t_0, Y_0]$. In the absence of state constraints, the problem is called (DDI) and $\mathcal{Y}_N[T, t_0, Y_0]$ denotes the corresponding set of solutions.

To measure distances, we define for $\eta = (\eta_j)_{j=0,...,N} \in \mathcal{Y}_N^\Theta[T, t_0, Y_0]$

$$\operatorname{dist}_\infty(y(\cdot), \mathcal{Y}_N^\Theta[T, t_0, Y_0]) := \inf\{\sup_{j=0,...,N} \|y(t_j) - \eta_j\| \,:\, \eta \in \mathcal{Y}_N^\Theta[T, t_0, Y_0]\},$$

$$\operatorname{dist}_\infty(\eta, \mathcal{Y}^\Theta[T, t_0, Y_0]) := \inf\{\sup_{j=0,...,N} \|\eta_j - y(t_j)\| \,:\, y(\cdot) \in \mathcal{Y}^\Theta[T, t_0, Y_0]\},$$

$$\operatorname{d}_{H,\infty}(\mathcal{Y}^\Theta[T, t_0, Y_0], \mathcal{Y}_N^\Theta[T, t_0, Y_0]) := \max\Big\{ \sup_{y(\cdot) \in \mathcal{Y}^\Theta[T, t_0, Y_0]} \operatorname{dist}_\infty(y(\cdot), \mathcal{Y}_N^\Theta[T, t_0, Y_0]),$$

$$\sup_{\eta \in \mathcal{Y}_N^\Theta[T, t_0, Y_0]} \operatorname{dist}_\infty(\eta, \mathcal{Y}^\Theta[T, t_0, Y_0]) \Big\}.$$

Here, the Euclidean vector norm on $\mathbb{R}^n$ is denoted by $\|\cdot\|$. For a subset $U \subset \mathbb{R}^n$, we denote by $\operatorname{dist}(x, U)$ the infimum of all Euclidean distances of the point $x \in \mathbb{R}^n$ to the points in U. $\operatorname{d}(U, V) = \sup_{u \in U} \operatorname{dist}(u, V)$ is the one-sided Hausdorff distance from a subset $U \subset \mathbb{R}^n$ to another subset $V \subset \mathbb{R}^n$, and $\operatorname{d_H}(U, V)$ is the Hausdorff-distance defined as

$$\operatorname{d_H}(U, V) = \max\{\operatorname{d}(U, V), \operatorname{d}(V, U)\}.$$

We pose some of the following basic assumptions on the right-hand side:

(H1) F satisfies a linear growth condition, i.e. there exists $C \geq 0$ with

$$\|F(t, x)\| := \sup_{y \in F(t,x)} \|y\| \leq C(\|x\| + 1) \quad (t \in I,\, x \in \mathbb{R}^n).$$

(H2) F has nonempty, compact, convex images in $\mathbb{R}^n$.

(H3) F is Lipschitz in (t, x) for all $t \in I$, $x \in \mathbb{R}^n$ with constant $L \geq 0$, i.e.

$$\operatorname{d_H}(F(s, x), F(t, y)) \leq L \cdot (|s - t| + \|x - y\|) \quad (s, t \in I,\, x, y \in \mathbb{R}^n).$$

The linear growth condition (H1) gives locally a boundedness of the images $F(t, x)$. A sufficient condition for (H1) is (H3) together with one bounded set $F(\widehat{t}, \widehat{x})$ (or (H2)). Condition (H2) is needed, since we want to apply the results from [11] for the unconstrained case. For practical applications, e.g. the Lipschitz condition could be restricted onto a compact set in which all values of all trajectories stay.

The following assumptions are required for the state constraints:

(C1) $\Theta : I \rightrightarrows \mathbb{R}^n$ has nonempty images explicitely given as

$$\Theta(t) := \{x \in \mathbb{R}^n : \ g(t,x) \le 0\}$$

by a single scalar function $g : I \times \mathbb{R}^n \to \mathbb{R}$ which fulfills $g(\cdot,\cdot) \in \mathcal{C}^{1,L}(I \times \mathbb{R}^n)$, i.e. the derivative $\nabla g(\cdot,\cdot)$ is Lipschitz on $I \times \mathbb{R}^n$.
Furthermore, points $x \in \partial\Theta(t)$ with $t \in I$ are characterized by $g(t,x) = 0$.

(C2) The boundary of $\Theta(\cdot)$ fulfills the "strict inwardness condition" (cf. [15, 17, 18, 7]), i.e. there exists $\alpha, \mu > 0$ such that for all $(t,x) \in B_\mu(\operatorname{graph}\partial\Theta(\cdot)) \cap (I \times \mathbb{R}^n)$ it follows that

$$\min_{v \in F(t,x)} \langle \nabla g(t,x), \binom{1}{v} \rangle \le -\alpha,$$

where

$$B_\mu(\operatorname{graph}\partial\Theta(\cdot)) = \big\{ \tbinom{t}{x} \in \mathbb{R}^{1+n} \ : \ \operatorname{dist}(\tbinom{t}{x}, \operatorname{graph}\partial\Theta(\cdot)) \le \mu \big\}.$$

From (C1) it follows that the images of $\Theta(\cdot)$ are closed. Existence of viable solutions could be proven under weaker assumptions, cf. in this respect e.g. [16]. But since we are interested mainly in stability results, which require stronger assumptions anyway and imply existence as well, we will not discuss weaker existence results for the continuous and the discrete case in this paper.

For the discrete situation in Section 2, it is sufficient to pose weaker assumptions on $F(\cdot,\cdot)$:

(H1') F satisfies a linear growth condition in integrable form, i.e. there exists a non-negative function $C(\cdot) \in \mathcal{L}_1(I,\mathbb{R})$ with

$$\|F(t,x)\| := \sup_{y \in F(t,x)} \|y\| \le C(t) \cdot (\|x\| + 1) \quad (t \in I,\, x \in \mathbb{R}^n).$$

(H2') F has nonempty, closed images in $\mathbb{R}^n$.

(H3') F is $L(t)$-Lipschitz in x for all $t \in I$ with $L(\cdot) \in \mathcal{L}_1(I,\mathbb{R})$, i.e.

$$\mathrm{d_H}(F(t,x), F(t,y)) \le L(t) \cdot \|x - y\| \quad (x, y \in \mathbb{R}^n).$$

Usually, uniform boundedness of $C(\cdot)$ is assumed in (H1'), i.e. (H1). The same remark applies to $L(\cdot)$ in (H3').

2. Stability for the Unconstrained Case. The essential stability result for differential inclusions without state constraints is given by (for a complete proof cf. [9, Lemma 8.3])

THEOREM 2.1 (Gronwall-Filippov-Wazewski's Theorem). *Let $F(\cdot,\cdot)$ have closed images in $\mathbb{R}^n$, and let $Y_0 \subset \mathbb{R}^n$ be nonempty, closed. For a given $\eta(\cdot) \in \mathrm{AC}(I)$ with*

$$\begin{aligned} \operatorname{dist}(\eta(t_0), Y_0) &\le \delta_0, \\ \operatorname{dist}(\eta'(t), F(t,\eta(t))) &\le \delta(t) \quad (\textit{a.e. } t \in I) \end{aligned}$$

with $\delta_0 \ge 0$ and non-negative $\delta(\cdot) \in \mathcal{L}_1(I,\mathbb{R})$, assume that

$$S := \{(t,x) \in I \times \mathbb{R}^n \ : \ \|x - \eta(t)\| \le \gamma\} \subset \operatorname{dom}(F)$$

for some $\gamma > \delta_0$. Let $F(\cdot, x)$ be measurable in t for all $x \in S$ and fulfill (H3') on S.

Let $z(\cdot)$ be the solution of

$$\begin{aligned} z'(t) &= L(t)z(t) + \delta(t) \quad (a.e.\ t \in I), \\ z(t_0) &= \delta_0. \end{aligned}$$

Then for all $\widetilde{T} \in I$ with $z(\widetilde{T}) \le \gamma$ there exists a solution $y(\cdot)$ on $[t_0, \widetilde{T}] \subset I$ with

$$\begin{aligned} y'(t) &\in F(t, y(t)) \qquad (a.e.\ t \in [t_0, \widetilde{T}]), \\ y(t_0) &= y_0 \in Y_0, \end{aligned}$$

fulfilling the estimates

$$\begin{aligned} \|y(t) - \eta(t)\| &\le z(t) \qquad (t \in [t_0, \widetilde{T}]), \\ \|y'(t) - \eta'(t)\| &\le L(t)z(t) + \delta(t) \quad (a.e.\ t \in [t_0, \widetilde{T}]), \end{aligned}$$

where

$$z(t) = e^{\int_{t_0}^{t} L(\sigma)\ d\sigma} \cdot \delta_0 + \int_{t_0}^{t} e^{\int_{\tau}^{t} L(\sigma)\ d\sigma} \cdot \delta(\tau)\ d\tau.$$

It will turn out in Section 3 that Theorem 2.1 together with the following discrete analogue is essential for the proof of stability for state-constrained differential inclusions.

THEOREM 2.2 (Discrete Gronwall-Filippov-Wazewski's Theorem).
Let $F : [t_0, T] \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$ fulfill (H2') and (H3').

Consider the discrete difference inclusion

$$\frac{y_{k+1} - y_k}{h} \in F(t_k, y_k) \qquad (k = 0, \dots, N-1), \tag{2.1}$$

$$y_0 \in Y_0 \tag{2.2}$$

for a given $N \in \mathbb{N}$, the step-size $h = \frac{T - t_0}{N}$ and a closed, nonempty starting set $Y_0 \subset \mathbb{R}^n$.

Let $(\eta_k)_{k=0,\dots,N}$ be a grid function with values in $\mathbb{R}^n$ and

$$\begin{aligned} \operatorname{dist}(\eta_0, Y_0) &\le \delta_0, \\ \operatorname{dist}\Big(\frac{\eta_{k+1} - \eta_k}{h}, F(t_k, \eta_k)\Big) &\le \delta_{k+1} \qquad (k = 0, \dots, N-1). \end{aligned}$$

Abbreviate $L_k = L(t_k)$, $k = 0, \dots, N$, and let $(z_k)_{k=0,\dots,N} \subset \mathbb{R}$ be the solution of

$$\frac{z_{k+1} - z_k}{h} = L_k z_k + \delta_{k+1} \qquad (k = 0, \dots, N-1), \tag{2.3}$$

$$z_0 = \delta_0.$$

Then there exists a solution $(y_k)_{k=0,\dots,N}$ of the discrete problem (2.1)–(2.2) with

$$\begin{aligned} \|\eta_k - y_k\| &\le z_k \qquad (k = 0, \dots, N), \\ \Big\| \frac{\eta_{k+1} - \eta_k}{h} - \frac{y_{k+1} - y_k}{h} \Big\| &\le L_k z_k + \delta_{k+1} \quad (k = 0, \dots, N-1). \end{aligned}$$

Proof. Since $Y_0 \subset \mathbb{R}^n$ is nonempty, there exists $y \in Y_0$ with $\operatorname{dist}(\eta_0, Y_0) \leq \|\eta_0 - y\| =: r$. Hence, the best approximation y_0 of η_0 in Y_0 coincides with that in the compact set $Y_0 \cap B_r(\eta_0)$, i.e.

$$\|\eta_0 - y_0\| = \operatorname{dist}(\eta_0, Y_0) \leq \delta_0 = z_0.$$

Assume that the assertion is true for $j = 0, \dots, k$, $k \in \{0, \dots, N-1\}$. Arguing as in the case $k = 0$, there exists $\xi_k^y \in F(t_k, y_k)$ for $\xi_k^\eta = \frac{1}{h}(\eta_{k+1} - \eta_k)$ with

$$\begin{aligned}
&\|\xi_k^\eta - \xi_k^y\| = \operatorname{dist}(\xi_k^\eta, F(t_k, y_k)),\\
&\|\xi_k^\eta - \xi_k^y\| \leq \operatorname{dist}(\xi_k^\eta, F(t_k, \eta_k)) + \mathrm{d_H}(F(t_k, \eta_k), F(t_k, y_k)) \leq L_k\|\eta_k - y_k\| + \delta_{k+1}.
\end{aligned}$$

Setting $y_{k+1} := y_k + h\xi_k^y$ yields

$$\begin{aligned}
\|\eta_{k+1} - y_{k+1}\| &= \|(\eta_k + h\xi_k^\eta) - (y_k + h\xi_k^y)\| \leq \|\eta_k - y_k\| + h\|\xi_k^\eta - \xi_k^y\|\\
&\leq (1 + hL_k)\|\eta_k - y_k\| + h\delta_{k+1} \leq (1 + hL_k)z_k + h\delta_{k+1} = z_{k+1}. \qquad \square
\end{aligned}$$

The explicit solution formula for the linear difference equation (2.3) yields immediately the following more specific estimates of the growth of the error bounds z_k $(k = 0, \dots, N)$.

COROLLARY 2.3. *With the assumptions as in Theorem 2.2 and for a Riemann integrable $L(\cdot)$ in (H3'), we can estimate the error bounds z_k for $k = 0, \dots, N$ as*

$$\begin{aligned}
z_k &= \delta_0 \cdot \prod_{\mu=0}^{k-1}(1 + hL_\mu) + h\sum_{j=1}^{k}\delta_j \cdot \prod_{\mu=j}^{k-1}(1 + hL_\mu),\\
\prod_{\mu=j}^{k-1}(1 + hL_\mu) &\leq \prod_{\mu=j}^{k-1} e^{hL_\mu} = e^{h\sum_{\mu=j}^{k-1} L_\mu} \leq e^{C_L} \quad (j = 0, \dots, k),
\end{aligned} \tag{2.4}$$

where C_L is an upper bound for the Riemann sums of the integral $\int_{t_0}^{T} L(t)\,dt$.

If furthermore $L_k = L$ for $k = 0, \dots, N$, then $(1 + hL)^k \leq e^{Lkh}$ and for $L > 0$

$$z_k \leq e^{Lkh}\delta_0 + \begin{cases} \frac{1}{L}(e^{Lkh} - 1) \cdot \max\limits_{j=1,\dots,k} \delta_j, \\ e^{L(k-1)h} \cdot h\sum\limits_{j=1}^{k} \delta_j. \end{cases} \tag{2.5}$$

The following lemmas are simple consequences of the growth condition and well-known in the literature (cf. e.g., [11, 19, 6]). They exhibit interesting connections between the continuous situation and the discrete situation in case $N \to \infty$.

LEMMA 2.4. *Let $F(\cdot, \cdot)$ satisfy (H1'). Then all solutions $y(\cdot)$ of (DI) in Problem 1.1 with bounded starting set $Y_0 \subset \mathbb{R}^n$ are uniformly bounded by $M := (\|Y_0\| + C_L) \cdot (1 + C_L e^{C_L})$ with $C_L := \|C(\cdot)\|_{\mathcal{L}_1(I)}$ and stay in a compactum $S \subset \mathbb{R}^n$.*

LEMMA 2.5. *Let $F(\cdot, \cdot)$ satisfy (H1). Then all solutions $y(\cdot)$ of (DI) in Problem 1.1 with bounded starting set $Y_0 \subset \mathbb{R}^n$ have a uniform Lipschitz constant.*

LEMMA 2.6. *Let $F(\cdot, \cdot)$ satisfy (H1') with Riemann integrable $C(\cdot)$, and let C_R denote an upper bound for the Riemann sums. Then all solutions $(\eta_k)_{k=0,\dots,N}$ of (DDI) in Euler's method 1.2 with bounded starting set $Y_0 \subset \mathbb{R}^n$ are bounded uniformly in $N \in \mathbb{N}$ by $M := (\|Y_0\| + C_R) \cdot (1 + C_R e^{C_R})$ and stay in a compactum $S \subset \mathbb{R}^n$.*

Choosing $C_R = \|C(\cdot)\|_{\mathcal{L}_1(I)} + \varepsilon$ for all $N \geq N_0(\varepsilon)$, emphasizes the similarity to Lemma 2.4.

LEMMA 2.7. *Let $F(\cdot,\cdot)$ satisfy (H1). Then all solutions $(\eta_k)_{k=0,\dots,N}$ of (DDI) in Euler's method 1.2 with bounded starting set $Y_0 \subset \mathbb{R}^n$ have a Lipschitz constant uniformly in $N \in \mathbb{N}$.*

Proof. Let M be the bound for all discrete solutions $(\eta_k)_{k=0,\dots,N}$ according to Lemma 2.6. Then it follows for $N \in \mathbb{N}$ and $j,k \in \{0,1,\dots,N\}$ with $j \leq k$

$$\begin{aligned}\|\eta_k - \eta_j\| &= \|\sum_{\mu=j}^{k-1}(\eta_{\mu+1}-\eta_\mu)\| \leq h\sum_{\mu=j}^{k-1}\|\frac{1}{h}(\eta_{\mu+1}-\eta_\mu)\| \leq h\sum_{\mu=j}^{k-1}\|F(t_\mu,\eta_\mu)\| \\ &\leq h\sum_{\mu=j}^{k-1}C(\|\eta_\mu\|+1) \leq C(M+1)(k-j)h = C(M+1)(t_k-t_j). \qquad \square\end{aligned}$$

3. Stability Analysis for the State-Constrained Case. There are several variants of the Gronwall-Filippov-Wazewski's Theorem for the continuous state-constrained case in the literature (cf. [15, Theorems 4.1 and 4.2], [17, Lemmata 3.3 and 4.4], [18, Theorem 3.1], as well as [7, Lemma 3.9], [8], [23, Lemma 2.2(b)] based on Soner's work in [22]). They were also denoted as theorems on the "existence of feasible neighboring trajectories" or as "tracking lemma". Exemplarily, we treat here the so-called "smooth" case, where the function $g(t,x)$ determines the state constraints $\Theta(t)$ and $g(\cdot,\cdot) \in \mathcal{C}^{1,L}(I \times \mathbb{R}^n)$.

A typical result for the continuous situation is given in the following

THEOREM 3.1. *Consider Problem 1.1 with time-dependent state constraint $\Theta(\cdot)$. Assume the conditions (H2)–(H3) on the right-hand side $F(\cdot,\cdot)$ and conditions (C1), (C2) on the state constraints.*

Then for every $y_0 \in \Theta(t_0)$ there exists a positive constant C such that for every $\eta(\cdot) \in \mathcal{Y}[T,t_0,y_0]$ there exists $y(\cdot) \in \mathcal{Y}^\Theta[T,t_0,y_0]$ with

$$\sup_{t\in[t_0,T]}\|\eta(t)-y(t)\| \leq C \sup_{t\in[t_0,T]}\operatorname{dist}(\eta(t),\Theta(t)).$$

We will omit the proof of this theorem, since it exploits a similar strategy as [15, Theorem 4.1], using in addition a result from [6, Theorem 3.2.4].

The reader should be aware that under considerably weaker assumptions, e.g. no convexity is needed, Lipschitz with respect to both variables can be weakened, analoguous results for the continuous situation hold. But, the proof of the discrete analogue presented here could be given only under stronger assumptions until now. Contrary to the assumptions (HC_1)–(HC_4) in [6], we allow time-dependent state constraints even in the discrete situation and simplify the conditions for the error estimate.

In any case, we want to emphasize the fact that **both** stability results for the continuous and the discrete case are needed for convergence of discrete approximations of state-constrained differential inclusions, described in Section 4.

We now present a rather detailed analysis of the discrete analogue of Theorem 3.1 following partly [6], but admitting time-dependent state constraints. We want to stress that this discrete analysis is in some respects rather technical, but nevertheless essential for the convergence analysis in the following Section 4. It would be very desirable to have available the discrete analogues of all those refined results [15, Theorem 4.2], [17, Lemma 3.3], [18, Theorem 3.1] (smooth case) resp. [15, Theorem

4.1], [17, Lemma 4.4] (non-smooth case), for the continuous situation. Cf. [17] for a detailed discussion of the smooth and non-smooth case.

THEOREM 3.2. *Consider Problem (DDIC) in* (1.4)–(1.5) *with time-dependent state constraint* $\Theta(\cdot)$. *Assume the conditions (H2)–(H3) on the right-hand side* $F(\cdot,\cdot)$ *and conditions (C1), (C2) on the state constraints.*

Then for every $y_0 \in \Theta(t_0)$ *there exist* $N_0 \in \mathbb{N}$ *and a positive constant* C *such that for all* $N \geq N_0$ *and for all discrete solutions* $(\eta_k)_{k=0,\dots,N} \in \mathcal{Y}_N[T, t_0, y_0]$ *there exists a discrete solution* $(y_k)_{k=0,\dots,N} \in \mathcal{Y}_N^{\Theta}[T, t_0, y_0]$ *with*

$$\max_{k=0,\dots,N} \|\eta_k - y_k\| \leq C(h + \max_{k=0,\dots,N} \operatorname{dist}(\eta_k, \Theta(t_k))).$$

Proof. Consider an arbitrary, in general non-feasible solution $(\eta_k)_{k=0,\dots,N}$ and set

$$\delta_N := \max_{k=0,\dots,N} \operatorname{dist}(\eta_k, \Theta(t_k)).$$

Case A: *solution* η_k *is feasible for* $k \in \mathcal{I} = \{0, \dots, N\}$

Clearly, $\delta_N = 0$ and the assertion is valid for $y_k := \eta_k$, $k \in \mathcal{I}$.

Case B: *solution* η_k *is not feasible for some* $k \in \mathcal{I}$

In this case, $\delta_N > 0$. On a small index set $\mathcal{I}_0 = \{0, \dots, k_1\}$ with k_1 independent from $(\eta_k)_{k\in\mathcal{I}}$ the result will be proven as a first step.

Denote by L_η the uniform Lipschitz constant for all discrete solutions according to Lemma 2.7, by L resp. $L_{\nabla g}$ the Lipschitz constant of $F(\cdot,\cdot)$ resp. $\nabla g(\cdot,\cdot)$, and choose the constants μ and α as in (C2). Without loss of generality, $L > 0$. Let M_2 be the maximum of $\|\nabla g(t,x)\|$ for $(t,x) \in I \times S$, S being the compactum according to Lemma 2.6.

Define

$$\tau_1 := \max \Bigg\{ t \in [t_0, T] \,:\, t \leq t_0 + \frac{\mu}{2(L_\eta + 1)}, \tag{3.1}$$

$$L_{\nabla g}(t - t_0) \leq \frac{M_2}{2(L_\eta + 1)}, \tag{3.2}$$

$$\max\{M_2(L_\eta + 1), (L_\eta + 1)^2 \cdot \frac{L_{\nabla g}}{L}\} \cdot (e^{L(t-t_0)} - 1) \leq \frac{\alpha}{12} \Bigg\} \tag{3.3}$$

which is independent of all discrete solutions and all $N \in \mathbb{N}$. [1]

For the discrete case additional assumptions on the step-size are necessary to construct a viable solution.

Choose $N_0 \in \mathbb{N}$ with

$$h_{N_0} = \frac{T - t_0}{N_0} \leq \tau_1 - t_0, \tag{3.4}$$

$$h_{N_0} \leq \frac{\mu}{2(L_\eta + 1)}, \tag{3.5}$$

$$h_{N_0} L_{\nabla g} \leq \frac{\alpha}{2(L_\eta + 1)^2}, \tag{3.6}$$

$$h_{N_0} L_{\nabla g} \leq \frac{M_2}{L_\eta + 1}, \tag{3.7}$$

[1] inequalities (3.1)–(3.3) are used in (3.14), (3.27) resp. in (3.25), (3.26)

determining the maximal allowed step-size h_{N_0}. [2]

(3.4) is needed to guarantee that at least one step of Euler's method can be performed to reach a time not exceeding τ_1. (3.5) follows from (3.1) and (3.4). It ensures that a discrete solution, before violating the state constraints at the next index, will be sufficiently near to the boundary such that there exists a direction which steers the solution into the interior. (3.6)–(3.7) are needed to show the viability of the solution in this phase and control the error of Taylor expansions.

From now on, let $N \geq N_0$, $h = \frac{T-t_0}{N}$, and define in view of (3.4)

$$k_1 := \lfloor \frac{\tau_1 - t_0}{h} \rfloor \geq 1, \tag{3.8}$$
$$\hat{k}_1 := \min\{k \in \mathcal{I} \,:\, \eta_{k+1} \notin \Theta(t_{k+1})\} < N,$$

where k_1 is the biggest natural number not exceeding $\frac{\tau_1 - t_0}{h}$.

It is clear that $t_{k_1} \leq \tau_1$ also satisfies the requirements in (3.1)–(3.3).

Case B, (i): $k_1 \leq \hat{k}_1$, i.e. the *solution* η_k *is feasible for* $k \in \widetilde{\mathcal{I}}_0 := \{0, \ldots, \hat{k}_1\} \supset \mathcal{I}_0$

Define

$$y_k := \eta_k \quad (k \in \mathcal{I}_0)$$

which fulfills the assertion on $\mathcal{I}_0$.

Case B, (ii): $k_1 > \hat{k}_1$, i.e. the *solution* η_k *is feasible for* $k \in \widetilde{\mathcal{I}}_0 \subsetneq \mathcal{I}_0$

In the first phase, set

$$y_k := \eta_k \quad (k \in \widetilde{\mathcal{I}}_0). \tag{3.9}$$

Since $\eta_{\hat{k}_1} \in \partial\Theta(t_{\hat{k}_1})$ cannot be guaranteed in the discrete case (only $\eta_{\hat{k}_1} \in \Theta(t_{\hat{k}_1})$), the distance to the boundary must be estimated and should not exceed $\frac{\mu}{2}$ to guarantee an inward steering direction. The function $\varphi(s) = g(t_{\hat{k}_1} + s, \eta_{\hat{k}_1} + s\frac{\eta_{\hat{k}_1+1} - \eta_{\hat{k}_1}}{h})$ is continuous on $[0, h]$ with

$$\varphi(0) = g(t_{\hat{k}_1}, \eta_{\hat{k}_1}) \leq 0, \quad \varphi(h) = g(t_{\hat{k}_1+1}, \eta_{\hat{k}_1+1}) > 0.$$

Therefore, there exists a zero $\bar{s} \in [0, h]$ of the function $\varphi(\cdot)$. Now, use (3.5) and (C1) to show

$$\begin{aligned} \operatorname{dist}(\tbinom{t_{\hat{k}_1}}{\eta_{\hat{k}_1}}, \operatorname{graph}\partial\Theta(\cdot)) &\leq \| \tbinom{t_{\hat{k}_1}}{\eta_{\hat{k}_1}} - \begin{pmatrix} t_{\hat{k}_1} + \bar{s} \\ \eta_{\hat{k}_1} + \bar{s}\frac{\eta_{\hat{k}_1+1} - \eta_{\hat{k}_1}}{h} \end{pmatrix} \| \\ &\leq \bar{s}(1 + \frac{1}{h} \cdot \|\eta_{\hat{k}_1+1} - \eta_{\hat{k}_1}\|) \leq (1 + L_\eta)h \leq \frac{\mu}{2}. \end{aligned} \tag{3.10}$$

Define (without loss of generality, the Lipschitz constant L_g of $g(\cdot)$ is greater 0)

$$\kappa_1 := \min\{\frac{k_1 - \hat{k}_1}{1 + \frac{\delta_N}{h}}, \frac{3}{\alpha}(L_g + 3M_2(L_\eta + 1))\}, \tag{3.11}$$
$$\bar{\delta}_1 := \lfloor \kappa_1(1 + \frac{\delta_N}{h}) + 1 \rfloor \geq 1, \tag{3.12}$$
$$\bar{k}_1 := \hat{k}_1 + \bar{\delta}_1$$

[2] inequalities (3.4)–(3.7) are used in (3.8), (3.10), (3.16) resp. (3.24)

which determines the length of the inward steering phase $\widehat{\mathcal{I}}_0 := \{\hat{k}_1, \hat{k}_1+1, \dots, \bar{k}_1\} \subset \mathcal{I}_0$. [3] κ_1 controls that the corresponding time interval either reaches t_{k_1} or guarantees the feasibility on the second time interval, $\bar{\delta}_1$ is the number of steps in the second phase in Case (ii.1) resp. (ii.2). Notice that κ_1 and $\bar{k}_1$ depend on the individual solution.

Consider the solution $(\widehat{y}_k)_{k\in\widehat{\mathcal{I}}_0}$ of the discrete inclusion

$$\begin{aligned}\frac{1}{h}(x_{k+1}-x_k) &\in Y(t_k,x_k) \quad (k\in\widehat{\mathcal{I}}_0\setminus\{\bar{k}_1\}),\\ x_{\hat{k}_1} &= y_{\hat{k}_1}\end{aligned}$$

on the second index set $\widehat{\mathcal{I}}_0$. Here, $Y(t,x)$ is defined as follows:

$$\begin{aligned}\varphi(t,x) &= \min_{v\in F(t,x)} \langle \nabla g(t,x), \tbinom{1}{v}\rangle, \\ Y(t,x) &= \{v\in F(t,x) \,:\, \langle \nabla g(x), \tbinom{1}{v}\rangle = \varphi(t,x)\},\end{aligned} \tag{3.13}$$

where $\varphi(\cdot,\cdot)$ is continuous on $\operatorname{graph}\Theta(\cdot)$ by [3, Theorem 1.4.16] and $Y(t,x)$ has compact, nonempty images and is upper semi-continuous by [2, §1.2, Theorem 6].

$\hat{k}_1$ is chosen so that inward steering is possible. We show that this is the case for all $k\in\widehat{\mathcal{I}}_0$ as well. From the Lipschitz continuity of all discrete solutions by Lemma 2.7 and (3.10) we get for $k\in\widehat{\mathcal{I}}_0$:

$$\begin{aligned}\|\widehat{y}_k-\widehat{y}_{\hat{k}_1}\| &\le L_\eta(k-\hat{k}_1)h,\\ \operatorname{dist}(\tbinom{t_k}{\widehat{y}_k}, \operatorname{graph}\partial\Theta(\cdot)) &\le \|\tbinom{t_k}{\widehat{y}_k}-\tbinom{t_{\hat{k}_1}}{\widehat{y}_{\hat{k}_1}}\| + \operatorname{dist}(\tbinom{t_{\hat{k}_1}}{\widehat{y}_{\hat{k}_1}}, \operatorname{graph}\partial\Theta(\cdot))\\ &\le |t_k-t_{\hat{k}_1}| + \|\widehat{y}_k-\widehat{y}_{\hat{k}_1}\| + \frac{\mu}{2}.\end{aligned}$$

Estimate $(k-\hat{k}_1)h$ by $t_{k_1}-t_0$ and use (3.1) to show

$$\operatorname{dist}(\tbinom{t_k}{\widehat{y}_k}, \operatorname{graph}\partial\Theta(\cdot)) \le (L_\eta+1)(k-\hat{k}_1)h + \frac{\mu}{2} \le \mu. \tag{3.14}$$

The proof of the feasibility of $(\widehat{y}_k)_{k\in\widehat{\mathcal{I}}_0}$ is not as simple as in the continuous case. Since $\widehat{y}_{\hat{k}_1}\in\Theta(t_{\hat{k}_1})$ per definition, we have $g(t_{\hat{k}_1},\widehat{y}_{\hat{k}_1})\le 0$, and with the telescopic sum

$$g(t_k,\widehat{y}_k) \le g(t_k,\widehat{y}_k) - g(t_{\hat{k}_1},\widehat{y}_{\hat{k}_1}) = \sum_{j=\hat{k}_1}^{k-1} (g(t_{j+1},\widehat{y}_{j+1}) - g(t_j,\widehat{y}_j)).$$

Set $\psi(s) = g(t_j+sh, \widehat{y}_j + s(\widehat{y}_{j+1}-\widehat{y}_j))$ for $s\in[0,1]$ and some $j\in\widehat{\mathcal{I}}_0$, then Taylor expansion up to terms of order 1 yields by the Lipschitz continuity of $\nabla g(\cdot,\cdot)$

$$g(t_{j+1},\widehat{y}_{j+1}) \le g(t_j,\widehat{y}_j) + \langle \nabla g(t_j,\widehat{y}_j), \tbinom{h}{\widehat{y}_{j+1}-\widehat{y}_j}\rangle + L_{\nabla g}(L_\eta+1)^2h^2. \tag{3.15}$$

[3] The first term in (3.11) is used in (3.19), the second one in (3.29), while (3.12) is used in (3.28) and (3.33).

Hence, due to (3.6) it follows

$$g(t_k,\widehat{y}_k) \le \sum_{j=\hat{k}_1}^{k-1} h\langle \nabla g(t_j,\widehat{y}_j), \left(\tfrac{1}{\frac{\widehat{y}_{j+1}-\widehat{y}_j}{h}}\right)\rangle + (L_{\nabla g}(L_\eta+1)^2 h)\cdot(k-\hat{k}_1)h$$
$$\le \sum_{j=\hat{k}_1}^{k-1} h\langle \nabla g(t_j,\widehat{y}_j), \left(\tfrac{1}{\frac{\widehat{y}_{j+1}-\widehat{y}_j}{h}}\right)\rangle + \frac{\alpha}{2}\cdot(k-\hat{k}_1)h. \tag{3.16}$$

Using (C2) due to (3.14) and $\frac{\widehat{y}_{j+1}-\widehat{y}_j}{h} \in Y(t_j,\widehat{y}_j)$ together with (3.13) we progress to the inequalities

$$g(t_k,\widehat{y}_k) \le h\sum_{j=\hat{k}_1}^{k-1} \varphi(t_j,\widehat{y}_j) + \frac{\alpha}{2}\cdot(k-\hat{k}_1)h \le -\frac{\alpha}{2}\cdot(k-\hat{k}_1)h. \tag{3.17}$$

Therefore, we have finally proven that $\widehat{y}_k \in \Theta(t_k)$ and

$$\|\widehat{y}_k-\eta_k\| \le \|\widehat{y}_k - y_{\hat{k}_1}\| + \|\eta_{\hat{k}_1}-\eta_k\| \le 2L_\eta(k-\hat{k}_1)h \le 2L_\eta\bar{\delta}_1 h \quad (k\in\widehat{\mathcal{I}}_0). \tag{3.18}$$

Case B, (ii.1): *inward steering phase reaches end of index set* $\mathcal{I}_0$

If $\bar{k}_1 = \hat{k}_1+\bar{\delta}_1 = k_1$, then the definition of the constructed solution is continued to $\widehat{\mathcal{I}}_0$ as

$$y_k := \widehat{y}_k \quad (k\in\widehat{\mathcal{I}}_0\setminus\{\hat{k}_1\}),$$

so that the claim is verified on $\widehat{\mathcal{I}}_0$ and therefore also on $\mathcal{I}_0$.

Case B, (ii.2): *Filippov solution follows time-delayed solution for the rest of indices in* $\mathcal{I}_0\setminus\widehat{\mathcal{I}}_0$

Now $\bar{k}_1 = \hat{k}_1+\bar{\delta}_1 < k_1$, set $\bar{\mathcal{I}}_0 := \{\bar{k}_1,\bar{k}_1+1,\dots,k_1\}$. From $\kappa_1(1+\frac{\delta_N}{h}) < \bar{\delta}_1$ follows that $\kappa_1 = \frac{3}{\alpha}(L_g+3M_2(L_\eta+1))$, since

$$\kappa_1 < \frac{k_1-\hat{k}_1}{1+\frac{\delta_N}{h}}. \tag{3.19}$$

Consider the Filippov solution $(\bar{y}_k)_{k\in\bar{\mathcal{I}}_0}$ of

$$\frac{1}{h}(x_{k+1}-x_k)\in F(t_k,x_k) \quad (k\in\bar{\mathcal{I}}_0\setminus\{k_1\}),$$
$$x_{\bar{k}_1} = y_{\bar{k}_1}$$

following the solution $(\eta_{k-\bar{\delta}_1})_{k\in\bar{\mathcal{I}}_0}$. Since the discrete version of Filippov's Theorem 2.2 will be applied, we study the following error terms:

$$\|\bar{y}_{\bar{k}_1}-\eta_{\bar{k}_1-\bar{\delta}_1}\| = \|y_{\bar{k}_1}-\eta_{\hat{k}_1}\| = \|y_{\bar{k}_1}-y_{\hat{k}_1}\| \le L_\eta\bar{\delta}_1 h, \tag{3.20}$$
$$\operatorname{dist}(\underbrace{\frac{1}{h}(\eta_{k+1-\bar{\delta}_1}-\eta_{k-\bar{\delta}_1})}_{\in F(t_{k-\bar{\delta}_1},\eta_{k-\bar{\delta}_1})}, F(t_k,\eta_{k-\bar{\delta}_1})) \le L\bar{\delta}_1 h$$

The time delay $\bar\delta_1$ does not only help in (3.20), since $\eta_{\bar k_1-\bar\delta_1}$ coincides with $y_{\hat k_1}$, but also allows to reuse the estimates on the second index set $\widehat{\mathcal{I}}_0$ (namely (3.18)) for the starting values on the third index set. For the distance to the right-hand side of the difference inclusion, the Lipschitz continuity of $F(\cdot,\cdot)$ with respect to t was used. The discrete Filippov's Theorem 2.2 together with Corollary 2.3 finally establishes the estimates

$$\begin{aligned}\|\bar y_k-\eta_{k-\bar\delta_1}\| &\le (1+hL)^{k-\bar k_1}L_\eta\bar\delta_1 h+((1+hL)^{k-\bar k_1}-1)\bar\delta_1 h\\ &=((L_\eta+1)(1+hL)^{k-\bar k_1}-1)\bar\delta_1 h, \qquad (3.21)\\ \|\frac{1}{h}(\eta_{k+1-\bar\delta_1}-\eta_{k-\bar\delta_1})-\frac{1}{h}(\bar y_{k+1}-\bar y_k)\|&\\ &\le L(L_\eta+1)(1+hL)^{k-\bar k_1}\bar\delta_1 h \qquad (3.22)\end{aligned}$$

on $\bar I_0$. They are used twice, first to estimate the deviation of the feasible solution to the given one in

$$\begin{aligned}\|\bar y_k-\eta_k\| &\le \|\bar y_k-\eta_{k-\bar\delta_1}\|+\|\eta_{k-\bar\delta_1}-\eta_k\|\\ &\le \Big((L_\eta+1)e^{L(k-\bar k_1)h}+L_\eta-1\Big)\bar\delta_1 h \qquad (3.23)\end{aligned}$$

and secondly, to show feasibility. To this purpose, the state constraint is splitted into four terms for each $k\in\bar{\mathcal{I}}_0$. Hereby, the Taylor expansion as in (3.15) will be used:

$$\begin{aligned}g(t_k,\bar y_k) &= \underbrace{g(t_{\bar k_1},\bar y_{\bar k_1})}_{=T_A}+\underbrace{g(t_{k-\bar\delta_1},\eta_{k-\bar\delta_1})-g(t_{\bar k_1-\bar\delta_1},\eta_{\bar k_1-\bar\delta_1})}_{=T_B}\\ &\quad+\sum_{j=\bar k_1}^{k-1}\big(g(t_{j+1},\bar y_{j+1})-g(t_j,\bar y_j)\big)\\ &\quad-\sum_{j=\bar k_1}^{k-1}\big(g(t_{j+1-\bar\delta_1},\eta_{j+1-\bar\delta_1})-g(t_{j-\bar\delta_1},\eta_{j-\bar\delta_1})\big)\\ &\le T_A+T_B+h\sum_{j=\bar k_1}^{k-1}\langle\nabla g(t_j,\bar y_j),\big(\tfrac{1}{h}(\bar y_{j+1}-\bar y_j)\big)\rangle+L_{\nabla g}(L_\eta+1)^2(k-\bar k_1)h^2\\ &\quad-h\sum_{j=\bar k_1}^{k-1}\langle\nabla g(t_{j-\bar\delta_1},\eta_{j-\bar\delta_1}),\big(\tfrac{1}{h}(\eta_{j+1-\bar\delta_1}-\eta_{j-\bar\delta_1})\big)\rangle+L_{\nabla g}(L_\eta+1)^2(k-\bar k_1)h^2\\ &= T_A+T_B+\underbrace{h\sum_{j=\bar k_1}^{k-1}\langle\nabla g(t_j,\bar y_j),\big(\tfrac{1}{h}(\bar y_{j+1}-\bar y_j)\big)-\big(\tfrac{1}{h}(\eta_{j+1-\bar\delta_1}-\eta_{j-\bar\delta_1})\big)\rangle}_{=T_C}\\ &\quad+\underbrace{h\sum_{j=\bar k_1}^{k-1}\langle\nabla g(t_j,\bar y_j)-\nabla g(t_{j-\bar\delta_1},\eta_{j-\bar\delta_1}),\big(\tfrac{1}{h}(\eta_{j+1-\bar\delta_1}-\eta_{j-\bar\delta_1})\big)\rangle}_{=T_D}\\ &\quad+\underbrace{2L_{\nabla g}(L_\eta+1)^2(k-\bar k_1)h^2}_{=T_E}=T_A+T_B+T_C+T_D+T_E.\end{aligned}$$

The next task will be to estimate each term separately. We estimate

$$T_A = g(t_{\bar{k}_1}, \widehat{y}_{\bar{k}_1}) \leq -\frac{\alpha}{2}\bar{\delta}_1 h$$

by (3.17), the corresponding inequality on the second index set.

The treatment of the second term is slightly more complicated as in the continuous case, since we can not assume that $g(t_{\hat{k}_1}, \eta_{\hat{k}_1}) = 0$. Nevertheless, we know that at index $\hat{k}_1$ we are close to the boundary and at the next index $\hat{k}_1 + 1$ the iterate violates the state constraints so that

$$T_B = g(t_{k-\bar{\delta}_1}, \eta_{k-\bar{\delta}_1}) - g(t_{\hat{k}_1}, \eta_{\hat{k}_1}) < g(t_{k-\bar{\delta}_1}, \eta_{k-\bar{\delta}_1}) + \underbrace{g(t_{\hat{k}_1+1}, \eta_{\hat{k}_1+1})}_{>0} - g(t_{\hat{k}_1}, \eta_{\hat{k}_1}).$$

The difference of the last two terms could be estimated as in (3.15):

$$\begin{aligned} g(t_{\hat{k}_1+1}, \eta_{\hat{k}_1+1}) - g(t_{\hat{k}_1}, \eta_{\hat{k}_1}) &\leq h\|\nabla g(t_{\hat{k}_1}, \eta_{\hat{k}_1})\| \cdot (1 + \|\frac{\eta_{\hat{k}_1+1} - \eta_{\hat{k}_1}}{h}\|) \\ &+ L_{\nabla g}(L_\eta + 1)^2 h^2 \leq \underbrace{\max_{(t,x)\in I\times S} \|\nabla g(t,x)\|}_{=M_2} \cdot (1 + L_\eta)h + L_{\nabla g}(L_\eta + 1)^2 h^2, \end{aligned}$$

where we used again that all discrete solutions are contained within a compactum S by Lemma 2.6 and that all discrete solutions have a uniform Lipschitz constant L_η by Lemma 2.7. Mimicing the proof in the continuous case, we distinguish two cases to treat the first term in T_B.

If $\eta_{k-\bar{\delta}_1} \in \Theta(t_{k-\bar{\delta}_1})$, then $g(t_{k-\bar{\delta}_1}, \eta_{k-\bar{\delta}_1}) \leq 0$ so that this first term has an advantageous sign. Otherwise, we introduce the projection $\eta^\pi_{k-\bar{\delta}_1} \in \partial\Theta(t_{k-\bar{\delta}_1})$ and estimate by using the definition of δ_N:

$$\begin{aligned} |g(t_{k-\bar{\delta}_1}, \eta_{k-\bar{\delta}_1}) - g(t_{k-\bar{\delta}_1}, \eta^\pi_{k-\bar{\delta}_1})| &\leq L_g \|\eta_{k-\bar{\delta}_1} - \eta^\pi_{k-\bar{\delta}_1}\| \\ &= L_g \operatorname{dist}(\eta_{k-\bar{\delta}_1}, \Theta(t_{k-\bar{\delta}_1})) \leq L_g \delta_N. \end{aligned}$$

In both cases, due to (3.7)

$$T_B \leq L_g\delta_N + M_2 \cdot (1 + L_\eta)h + L_{\nabla g}(L_\eta + 1)^2 h^2 \leq L_g\delta_N + 2M_2 \cdot (1 + L_\eta)h \quad (3.24)$$

In term T_C, the difference quotient of both solutions is compared, which was estimated in (3.22) by the discrete Filippov Theorem. Moreover, the boundedness of the discrete solutions and the continuity of $\nabla g(\cdot,\cdot)$ are used, yielding

$$\begin{aligned} T_C &\leq h \sum_{j=\bar{k}_1}^{k-1} \|\nabla g(t_j, \bar{y}_j)\| \cdot \|\frac{\bar{y}_{j+1} - \bar{y}_j}{h} - \frac{\eta_{j+1-\bar{\delta}_1} - \eta_{j-\bar{\delta}_1}}{h}\| \\ &\leq M_2 h \sum_{j=\bar{k}_1}^{k-1} \left(L(L_\eta + 1)(1 + hL)^{j-\bar{k}_1}\bar{\delta}_1 h \right) = M_2(L_\eta + 1)((1 + hL)^{k-\bar{k}_1} - 1)\bar{\delta}_1 h. \end{aligned}$$

Since $(1 + hL)^{k-\bar{k}_1}$ can be estimated by Corollary 2.3 as $e^{L(k-\bar{k}_1)h} \leq e^{Lk_1 h} \leq e^{L(\tau_1 - t_0)}$, we can exploit that τ_1 was suitably chosen by (3.3), and we get

$$T_C \leq \frac{\alpha}{12}\bar{\delta}_1 h. \quad (3.25)$$

The same estimate will be reached for the term T_D. The main keys are the Lipschitz continuity of $\nabla g(\cdot,\cdot)$, the uniform Lipschitz constant for all discrete solutions and the estimates (3.21) from the discrete Filippov Theorem together with the one in (2.5):

$$\begin{aligned}
T_D &\le h\sum_{j=\bar{k}_1}^{k-1} \|\nabla g(t_j,\bar{y}_j)-\nabla g(t_{j-\bar{\delta}_1},\eta_{j-\bar{\delta}_1})\|\cdot(1+\|\frac{\eta_{j+1-\bar{\delta}_1}-\eta_{j-\bar{\delta}_1}}{h}\|)\\
&\le h\sum_{j=\bar{k}_1}^{k-1} L_{\nabla g}(|t_j-t_{j-\bar{\delta}_1}|+\|\bar{y}_j-\eta_{j-\bar{\delta}_1}\|)\cdot(1+L_\eta)\\
&\le (L_\eta+1)L_{\nabla g}h\sum_{j=\bar{k}_1}^{k-1}(1+(L_\eta+1)(1+hL)^{j-\bar{k}_1}-1)\cdot\bar{\delta}_1 h\\
&\le (L_\eta+1)L_{\nabla g}\frac{L_\eta+1}{L}hL\sum_{j=\bar{k}_1}^{k-1}(1+hL)^{j-\bar{k}_1}\cdot\bar{\delta}_1 h\\
&\le (L_\eta+1)^2\frac{L_{\nabla g}}{L}((1+hL)^{k-\bar{k}_1}-1)\cdot\bar{\delta}_1 h.
\end{aligned}$$

Now, the reasoning is the same as for the term T_C, hence

$$T_D \le \frac{\alpha}{12}\bar{\delta}_1 h. \tag{3.26}$$

For the estimation of T_E we need (3.2):

$$\begin{aligned}
T_E &= 2L_{\nabla g}(L_\eta+1)^2(k-\hat{k}_1)h^2 \le 2L_{\nabla g}(L_\eta+1)^2(t_k-t_{\hat{k}_1})h\\
&\le 2L_{\nabla g}(L_\eta+1)^2(\tau_1-t_0)h \le M_2(L_\eta+1)h.
\end{aligned} \tag{3.27}$$

Now, we put all estimates together to show the feasibility. We have

$$\begin{aligned}
g(t_k,\bar{y}_k) &\le T_A+T_C+T_D+T_B+T_E \le -\frac{\alpha}{2}\bar{\delta}_1 h+2\cdot\frac{\alpha}{12}\bar{\delta}_1 h+T_B+T_E\\
&\le -\frac{\alpha}{3}\bar{\delta}_1 h+T_B+T_E.
\end{aligned}$$

The definition (3.12) for $\bar{\delta}_1$ and $\kappa_1=\frac{3}{\alpha}(L_g+3M_2(L_\eta+1))$ yield

$$\frac{\alpha}{3}\bar{\delta}_1 h \ge \frac{\alpha}{3}\kappa_1(1+\frac{\delta_N}{h})h \tag{3.28}$$

$$= (L_g+3M_2(L_\eta+1))(h+\delta_N) \ge L_g\delta_N+3M_2(L_\eta+1)h \tag{3.29}$$

and hence, the problematic term $L_g\delta_N$ could be eliminated by

$$\begin{aligned}
g(t_k,\bar{y}_k) \le &-L_g\delta_N-3M_2(L_\eta+1)h+L_g\delta_N+2M_2(L_\eta+1)h\\
&+M_2(L_\eta+1)h\le 0.
\end{aligned} \tag{3.30}$$

Extend the feasible solution in the third phase to $\mathcal{I}_0$ by

$$y_k := \bar{y}_k \quad (k\in\bar{\mathcal{I}}_0\setminus\{\bar{k}_1\}). \tag{3.31}$$

For all $k \in \mathcal{I}_0$, (3.9) and the estimates (3.18),(3.23) yield altogether

$$\|y_k - \eta_k\| \leq \max\{2L_\eta, \underbrace{(L_\eta+1)e^{L(\tau_1-t_0)} + L_\eta - 1}_{=:M_3 \geq 2L_\eta}\} \cdot \bar{\delta}_1 h. \tag{3.32}$$

In the last inequality, $(k_1 - \bar{k}_1)h$ was estimated by $k_1 h \leq \tau_1 - t_0$. Moreover,

$$\begin{aligned} \bar{\delta}_1 h &= \lfloor \kappa_1(1+\frac{\delta_N}{h}) + 1 \rfloor \cdot h \leq (\kappa_1(1+\frac{\delta_N}{h}) + 1)h \\ &\leq (\frac{3}{\alpha}(L_g + 3M_2(L_\eta+1))(h+\delta_N) + h = \mathcal{O}(h+\delta_N), \end{aligned} \tag{3.33}$$

$$\|y_k - \eta_k\| \leq M_3 \bar{\delta}_1 h \leq \underbrace{M_3(1 + \frac{3}{\alpha}(L_g + 3M_2(L_\eta+1)))}_{=:\widetilde{M}}(h+\delta_N) = \mathcal{O}(h+\delta_N).$$

Extension to the whole index set $\mathcal{I}$:

This process is well explained in the proof of [6, Theorem 3.2.6]: Divide the index set in J subsets with k_1 elements and set $\mathcal{I}_j := \{k_j, k_j+1, \ldots, k_{j+1}\} \cap \{0, \ldots, N\}$ with $k_j = jk_1$, $j = 0, \ldots, J$.

(i) *first index set*

For $j = 0$ the solution y_k is already constructed for $\mathcal{I}_0$. Set $\widetilde{C}_0 := 1 + \frac{\delta_N}{h}$ and $\Delta_0 = \lfloor \kappa_1 \widetilde{C}_0 + 1 \rfloor$.

(ii) *recursive approach*

For $j > 0$ start the process by taking the end value of the feasible solution $y_{j \cdot k_1}$ on $\mathcal{I}_{j-1}$ as starting value for the next iteration. Now, apply again the discrete Filippov Theorem to construct the (in general, non-feasible) solution $(z_k^{(j)})_{k \in \mathcal{I}_j}$ of

$$\begin{aligned} \frac{1}{h}(x_{k+1} - x_k) &\in F(t_k, x_k) \quad (k \in \mathcal{I}_j), \\ x_{k_j} &= y_{k_j}, \end{aligned}$$

that follows the non-feasible one $(\eta_k)_{k \in \mathcal{I}_j}$. The error term is governed by the difference of the starting values. Now, construct a feasible solution $(y_k)_{k \in \mathcal{I}_j}$ from $(z_k^{(j)})_{k \in \mathcal{I}_j}$. Then show that the deviation from $(y_k)_{k \in \mathcal{I}_j}$ to $(\eta_k)_{k \in \mathcal{I}_j}$ could be estimated by

$$\|y_k - \eta_k\| \leq \widetilde{M} \sum_{\nu=0}^{j} e^{(j-\nu)Lk_1 h} \Delta_\nu h \quad (k \in \mathcal{I}_j),$$

where for $j = 1, \ldots, J$

$$\widetilde{C}_j = \widetilde{C}_0 + \widetilde{M} \sum_{\nu=0}^{j-1} e^{(j-\nu)Lk_1 h}, \quad \Delta_j = \lfloor \kappa_1 \widetilde{C}_j + 1 \rfloor.$$

Estimate J uniformly for all $N \in \mathbb{N}$ by $\lfloor \frac{T-t_0}{\tau_1 - h_{N_0}} + 1 \rfloor$ so that finally we have proven the overall order $\mathcal{O}(h + \delta_N)$. ∎

REMARK 3.3. *Assume that $\Theta : I \Rightarrow \mathbb{R}^n$ with images in $\mathcal{C}(\mathbb{R}^n)$ has a $\mathcal{C}^{1,L}$-signed distance function*

$$\widetilde{d}(t,x) := \begin{cases} \operatorname{dist}(x, \partial\Theta(t)), & \text{if } x \in \Theta(t), \\ -\operatorname{dist}(x, \partial\Theta(t)) = -\operatorname{dist}(x, \Theta(t)), & \text{if } x \in \mathbb{R}^n \setminus \Theta(t). \end{cases}$$

Then $\Theta(t) = \{x \in \mathbb{R}^n : -\widetilde{d}(t,x) \leq 0\}$ fulfills the assumptions of Theorem 3.2.

4. Convergence Analysis. Combining the stability results from Section 3 for the continuous and discrete situation, we are now in a position to prove order of convergence results for the discrete approximation of the set of all viable solutions of the differential inclusion by all viable discrete solutions.

An essential tool is the following result for differential inclusions without state constraints, cf. [11, 1. Theorem] which we formulate under stronger assumptions, needed later on anyway. The convexity is an important assumption for the convergence of Euler's method.

PROPOSITION 4.1. *Choose a compactum $S \subset \mathbb{R}^n$ containing all solutions of* (1.1), (1.3). *Let $F(\cdot,\cdot)$ fulfill (H2)–(H3) on S and let $Y_0 = \{y_0\}$.*

Then there exists a positive constant C such that for all $N \in \mathbb{N}$

$$\mathrm{d}_{H,\infty}(\mathcal{Y}[T,t_0,y_0], \mathcal{Y}_N[T,t_0,y_0]) \leq Ch.$$

The stability results from Section 3 (Theorem 3.1 for the continuous case and Theorem 3.2 for the discrete case) are essential for the convergence proof of Euler's discretization of differential inclusions with state constraints.

THEOREM 4.2. *Assume hypotheses (H2)–(H3) together with (C1)–(C2) and let $Y_0 = \{y_0\}$ with $y_0 \in \Theta(t_0)$.*

Then there exist a positive constant C and $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$

$$\mathrm{d}_{H,\infty}(\mathcal{Y}^{\Theta}[T,t_0,y_0], \mathcal{Y}_N^{\Theta}[T,t_0,y_0]) \leq Ch.$$

Proof. This proof will use the notation of some constants from the proof of Theorem 3.2. Choose $N_0 \in \mathbb{N}$ from this theorem and $N \geq N_0$ so that additionally $h_{N_0} \leq \mu$ and $(C(M+1)+1)^2 L_{\nabla g} h_{N_0} \leq \frac{\alpha}{2}$, where M is the bound in Lemma 2.6 and α, μ originate from (C2).

Let us first construct a close discrete solution to a given $y(.) \in \mathcal{Y}^{\Theta}[T,t_0,y_0]$ to estimate the one-sided distance. According to Proposition 4.1, there exists $(\widetilde{\eta}_k)_{k=0,\dots,N} \in \mathcal{Y}_N[T,t_0,y_0]$ with

$$\max_{k=0,\dots,N} \|y(t_k) - \widetilde{\eta}_k\| \leq \widetilde{C}_1 h.$$

Since

$$\operatorname{dist}(\widetilde{\eta}_k, \Theta(t_k)) \leq \|\widetilde{\eta}_k - y(t_k)\| + \operatorname{dist}(y(t_k), \Theta(t_k)) \leq \widetilde{C}_1 h,$$

a solution $(\eta_k)_{k=0,\dots,N} \in \mathcal{Y}_N^{\Theta}[T,t_0,y_0]$ can be constructed by Theorem 3.2 with

$$\max_{k=0,\dots,N} \|\eta_k - \widetilde{\eta}_k\| \leq \widetilde{C}_2 h.$$

Hence, the grid function $y_N := (y(t_k))_{k=0,\dots,N}$ fulfills

$$\begin{aligned}\|\eta_k - y(t_k)\| &\le \|\eta_k - \widetilde{\eta}_k\| + \|\widetilde{\eta}_k - y(t_k)\| \le (\widetilde{C}_1 + \widetilde{C}_2)h,\\ \operatorname{dist}_\infty(y_N, \mathcal{Y}_N^\Theta[T,t_0,y_0]) &\le (\widetilde{C}_1 + \widetilde{C}_2)h.\end{aligned}$$

On the other hand, for a given discrete solution $\eta := (\eta_k)_{k=0,\dots,N} \in \mathcal{Y}_N^\Theta[T,t_0,y_0]$ one has to estimate the other one-sided distance. Proposition 4.1 shows the existence of $\widetilde{y}(\cdot) \in \mathcal{Y}[T,t_0,y_0]$ with

$$\max_{k=0,\dots,N} \|\eta_k - \widetilde{y}(t_k)\| \le \widetilde{C}_1 h.$$

The reasoning is now more complicated, since we need to estimate the following distance for all $t \in [t_k, t_{k+1}]$ and all $k \in \{0,\dots,N-1\}$,

$$\operatorname{dist}(\widetilde{y}(t), \Theta(t)) \le \|\widetilde{y}(t) - \widetilde{y}(t_k)\| + \|\widetilde{y}(t_k) - \eta_k\| + \operatorname{dist}(\eta_k, \Theta(t)). \tag{4.1}$$

Since $\eta_k \in \Theta(t_k)$, the inequality $g(t_k,\eta_k) \le 0$ holds.

(i) If $\binom{t_k}{\eta_k} \in B_\mu(\operatorname{graph}\partial\Theta(\cdot))$, then there exists $v_k \in F(t_k,\eta_k)$ by (C2) with

$$\langle \nabla g(t_k,\eta_k), \tbinom{1}{v_k}\rangle \le -\alpha.$$

For $t \in [t_k, t_{k+1}]$, we set $\eta(t) := \eta_k + (t-t_k)v_k$ and consider

$$\begin{aligned}g(t,\eta(t)) &= g(t_k,\eta_k) + \int_{t_k}^t \frac{\mathrm{d}}{\mathrm{ds}} g(s,\eta(s))\, ds \le \int_{t_k}^t \langle \nabla g(s,\eta(s)), \tbinom{1}{v_k}\rangle\, ds\\ &= \int_{t_k}^t \langle \nabla g(t_k,\eta_k), \tbinom{1}{v_k}\rangle\, ds + \int_{t_k}^t \langle \nabla g(s,\eta(s)) - \nabla g(t_k,\eta_k), \tbinom{1}{v_k}\rangle\, ds\\ &\le -\alpha(t-t_k) + \int_{t_k}^t \|\nabla g(s,\eta(s)) - \nabla g(t_k,\eta_k)\| \cdot (1+\|v_k\|)\, ds.\end{aligned}$$

Let us estimate both terms using (H1) and Lemma 2.6 by

$$\begin{aligned}1 + \|v_k\| &\le 1 + \|F(t_k,\eta_k)\| \le 1 + C(\|\eta_k\|+1) \le C(M+1)+1,\\ \|\nabla g(s,\eta(s)) - \nabla g(t_k,\eta_k)\| &\le L_{\nabla g}(|s-t_k| + \|\eta(s)-\eta_k\|)\\ &\le L_{\nabla g}(1+\|v_k\|)(s-t_k) \le (C(M+1)+1)L_{\nabla g}h\end{aligned}$$

and continue the inequality with

$$g(t,\eta(t)) \le -\alpha(t-t_k) + (C(M+1)+1)^2 L_{\nabla g} h(t-t_k) \le -\frac{\alpha}{2}(t-t_k) \le 0.$$

Therefore, $\eta(t) \in \Theta(t)$ is close to η_k with

$$\operatorname{dist}(\eta_k, \Theta(t)) \le \|\eta_k - \eta(t)\| = (t-t_k)\|v_k\| \le C(M+1)(t-t_k).$$

(ii) If $\binom{t_k}{\eta_k} \notin B_\mu(\operatorname{graph}\partial\Theta(\cdot))$, then $\binom{t_k}{\eta_k} \notin \operatorname{graph}\partial B_\mu(\Theta(\cdot))$ and $\operatorname{dist}(\eta_k, \partial\Theta(t_k))$ is greater than μ. Let us assume that $g(t,\eta_k) > 0$. With the continuous function $\varphi(s) := g(s,\eta_k)$ on $[t_k,t_{k+1}]$, we will soon arrive at a contradiction. Since the inequalities

$$\begin{aligned}\varphi(t_k) &= g(t_k,\eta_k) < 0,\\ \varphi(t) &= g(t,\eta_k) > 0\end{aligned}$$

hold, there exists $\bar{t} \in (t_k, t) \subset (t_k, t_{k+1}]$ with $\varphi(\bar{t}) = 0$. Then $g(\bar{t}, \eta_k) = 0$ and $\eta_k \in \partial\Theta(\bar{t})$ such that $\binom{\bar{t}}{\eta_k} \in \operatorname{graph} \partial\Theta(\cdot)$. The following inequality shows the contradiction

$$\operatorname{dist}(\tbinom{t_k}{\eta_k}, \operatorname{graph} \partial\Theta(\cdot)) \le \|\tbinom{t_k}{\eta_k} - \tbinom{\bar{t}}{\eta_k}\| = |\bar{t} - t_k| \le h \le \mu.$$

Hence, the assumption was wrong which yields now $g(t, \eta_k) \le 0$ so that $\eta_k \in \Theta(t)$.

In both cases (i)–(ii), $\operatorname{dist}(\eta_k, \Theta(t)) \le C(M+1)(t - t_k)$. Using (4.1), we get

$$\operatorname{dist}(\widetilde{y}(t), \Theta(t)) \le L_y|t - t_k| + \widetilde{C}_1 h + C(M+1)(t - t_k) \le (C(M+1) + \widetilde{C}_1 + L_y)h,$$

where L_y is the uniform Lipschitz constant from Lemma 2.5. Therefore, a solution $y(\cdot) \in \mathcal{Y}^{\Theta}[T, t_0, y_0]$ exists by Theorem 3.1 with

$$\sup_{t \in I} \|y(t) - \widetilde{y}(t)\| \le \widetilde{C}_3 h.$$

Hence,

$$\|\eta_k - y(t_k)\| \le \|\eta_k - \widetilde{y}(t_k)\| + \|\widetilde{y}(t_k) - y(t_k)\| \le (\widetilde{C}_1 + \widetilde{C}_3)h,$$
$$\operatorname{dist}_\infty(\eta, \mathcal{Y}^{\Theta}[T, t_0, y_0]) \le (\widetilde{C}_1 + \widetilde{C}_3)h. \qquad \square$$

5. Example. The dynamical sytem, underlying the following two test examples, is due to P. Kenderov. It serves as a model problem for the illustration of first order convergence. We restrict ourselves to the visualization of the convergence of reachable sets. The visualization of the convergence of the whole discrete solution sets would require much more space and the choice of more appropriate data structures.

Naturally, the realization of set-valued Euler's method (1.4)–(1.5) on a computer amounts to an additional perturbation of the set-valued right-hand side of order 1 and an evaluation of the set union with a local error of order 2 (with respect to Hausdorff distance, uniformly in $t \in I$), for computational details cf. [6].

EXAMPLE 5.1. *Consider the following differential inclusion*

$$y'(t) \in F(t, y(t)) = \{Ay(t) + uBy(t) \in \mathbb{R}^2 \,:\, 0 \le u \le 1\} \quad (a.e.\ t \in [0, 8]),$$
$$y(t) \in \Theta := \{y \in \mathbb{R}^2 \,:\, g(y) \le 0\},$$
$$y(0) = y_0 = \tbinom{2}{2},$$

where

$$A = \begin{pmatrix} \sigma^2 - 1 & \sigma\sqrt{1-\sigma^2} \\ -\sigma\sqrt{1-\sigma^2} & \sigma^2 - 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -2\sigma\sqrt{1-\sigma^2} \\ 2\sigma\sqrt{1-\sigma^2} & 0 \end{pmatrix},$$
$$g(y) := -\frac{1}{2}(y_1 - 2)^2 + 2 - y_2, \qquad y = \tbinom{y_1}{y_2}$$

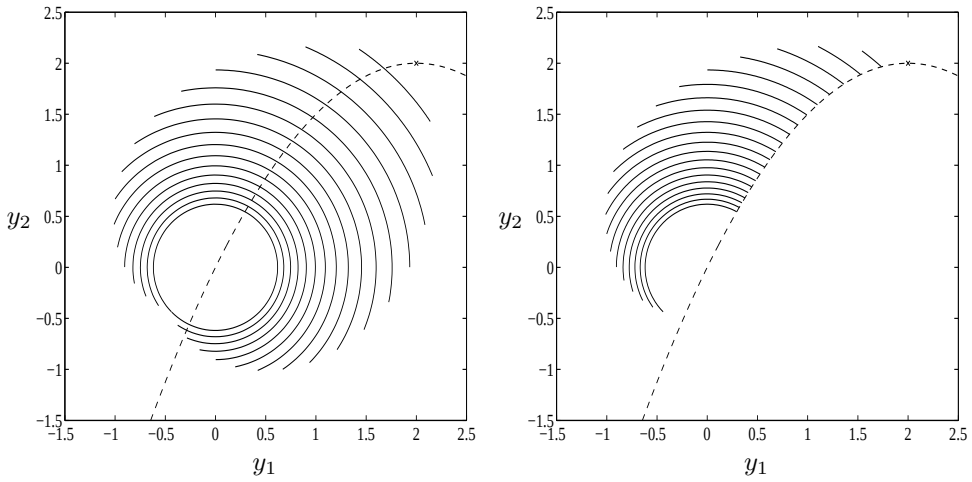
and $\sigma \in (0, 1)$ is a fixed parameter.

The reachable set for the unconstrained case can be expressed by representing its points with polar coordinates,

$$\mathcal{R}(t, t_0, r_0\tbinom{\cos(\phi_0)}{\sin(\phi_0)}) = \{r(t)\tbinom{\cos(\phi(t))}{\sin(\phi(t))} \,:\, r(t) = r_0 e^{(\sigma^2-1)t},$$
$$\phi(t) = \phi_0 + \sigma\sqrt{1-\sigma^2}(2u - 1)t, \quad 0 \le u \le 1\},$$

where the initial point y_0 has polar coordinates $(r_0, \phi_0) = (2\sqrt{2}, \frac{\pi}{4})$. Further on, we fix $\sigma = \frac{9}{10}$.

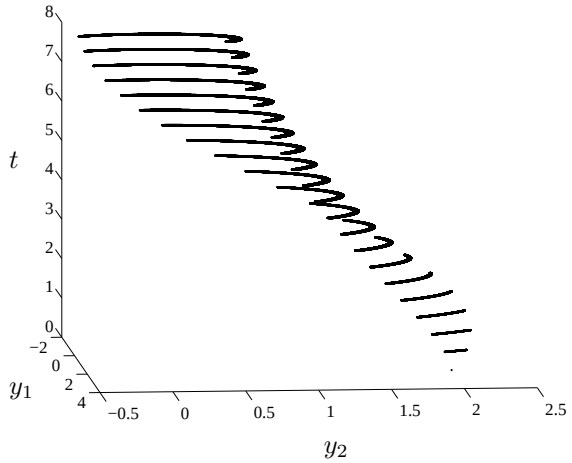
FIGURE 5.1. *Reachable sets for different end times t (without resp. with state constraints)*



In Figure 5.1 (left picture), the exact reachable sets for the unconstrained problem with varying end time $t_i = i \cdot \frac{1}{2}$, $i = 0, \ldots, 16$, and the boundary of the (quadratic) state constraint (dotted line) are illustrated. For $t = 0$, the starting set is just the upper right point in this figure (marked by the cross), for increasing time t the reachable set moves to the lower left of the figure and the two ends of the arcs approach each other. Approximately for $t \geq 8$, the two end points of the arc will overlap and the reachable sets form the boundary of a circle. In the right picture of Figure 5.1, the reachable sets for the state-constrained problem are visualized for the same times. In contrast to [6, Example 5.2.2] with a linear constraint, the reachable set cannot be gained by the intersection $\mathcal{R}(t, t_0, Y_0) \cap \Theta$, as the comparison of both pictures shows. For $t \geq 7.2$, the small part of the circle that moves out of the interior of Θ (everything below the quadratic function) originates from points that were already cut off by the quadratic state constraint at an earlier time.

Figure 5.2 shows the integral funnel with state constraints.

FIGURE 5.2. *Integral funnel of Euler's method with $N = 240$ and state constraints*



This second figure was calculated with set-valued Euler's method for $N = 240$ on the interval $[0, 8]$ (cf. [6, §4] for details on the implementation of Euler's method for the approximation of nonlinear differential inclusions).

Let us check, whether Theorem 4.2 for state-constrained Euler's method can be applied. Observe that $F(t, y) = \{f(t, y, u) : u \in [0, 1]\}$ with $f(t, y, u) = Ay + uBy$ is Lipschitz with respect to (t, y) and has nonempty, compact, convex images. Clearly, (H1) and (C1) are also fulfilled. Furthermore,

$$\langle \nabla g(y)^\top, v\rangle = -(\sigma^2 - 1) \cdot \frac{1}{2} y_1^2 + \sigma\sqrt{1-\sigma^2} \cdot (1 - 2u) \cdot \frac{1}{2} y_1 \cdot (y_1^2 - 6y_1 + 10)$$

for all $y \in \partial\Theta$ and $v = f(t, y, u)$.

For $y_1 < 0$, the choice of $u = 0$ yields

$$\langle \nabla g(y)^\top, v\rangle = \frac{1}{2} y_1 \cdot \sigma\sqrt{1-\sigma^2} \cdot \underbrace{(y_1^2 - (6 - \frac{1}{\sigma}\sqrt{1-\sigma^2})y_1 + 10)}_{=:h(y_1)}.$$

A discussion of the function h shows $h(y_1) \geq (y_1 - 4)^2 - 6 \geq 10$ so that the scalar product is less than zero.

For $y_1 \in (0, \frac{5}{2}]$, $u = 1$ is chosen such that

$$\langle \nabla g(y)^\top, v\rangle = -\frac{1}{2} y_1 \cdot \sigma\sqrt{1-\sigma^2} \cdot (y_1^2 - (6 + \frac{1}{\sigma}\sqrt{1-\sigma^2})y_1 + 10).$$

The quadratic function in this term could be strictly estimated from below by the function $\widetilde{h}(y_1) = y_1^2 - \frac{13}{2} y_1 + 10$ which is strictly decreasing and is not less than $\widetilde{h}(\frac{5}{2}) = 0$. Hence, the scalar product is also negative.

Let us note that the final reachable set is a circle avoiding the origin, cf. Figure 5.1. Therefore, all discrete reachable sets for small step-sizes have a positive distance to the origin so that on a compactum containing all Euler solutions and near to the boundary of Θ we have a positive distance to the origin. A compactness argument yields therefore the validity of (C2). Hence, order of convergence 1 with respect to the step-size h holds by Theorem 4.2.

For the state-constrained case, Tables 5.1 and 5.2 visualize the order of convergence for the approximation of the reachable set $\mathcal{R}(0.5, 0, \binom{2}{2})$ resp. $\mathcal{R}(7.5, 0, \binom{2}{2})$. The tables are calculated by using the theoretical reachable set as reference set.

TABLE 5.1
Estimated order of convergence for $T = 0.5$ (state-constrained problem)

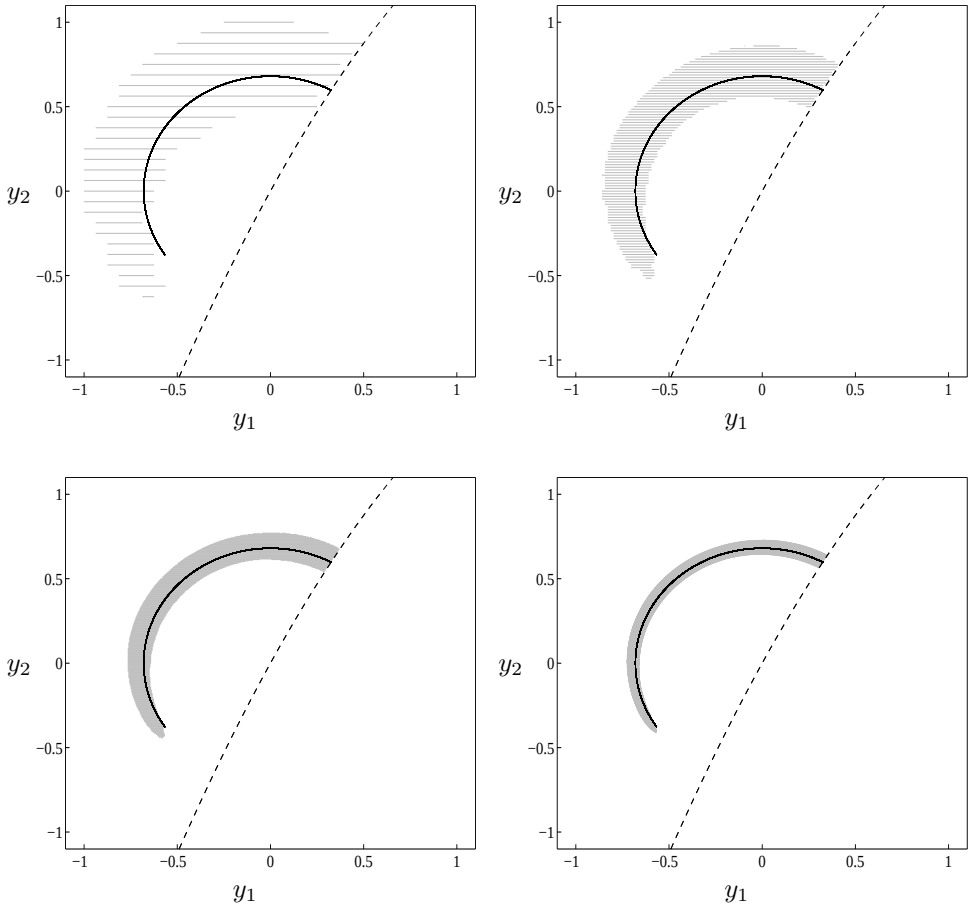
N	estimated Hausdorff distance from the reference set	difference to Ch^p
16	0.0897488	-1.1E-02
32	0.0280925	9.2E-03
64	0.0182812	-6.6E-04
128	0.0104471	-2.1E-03
256	0.0036226	3.2E-04
512	0.0018178	4.8E-05

TABLE 5.2
Estimated order of convergence for $T = 7.5$ (state-constrained problem)

N	estimated Hausdorff distance from the reference set	difference to Ch^p
16	0.3275207	1.1E-02
32	0.1842108	-7.3E-03
64	0.0923952	-1.2E-04
128	0.0483326	-2.0E-04
256	0.0250892	2.0E-05
512	0.0129622	1.4E-04

Based on these data, a least squares problem with the function $\log(Ch^p)$ with unknowns $C, p \geq 0$ yields the values $p = 1.0800$ and $C = 1.1812$ resp. $p = 0.9388$ and $C = 1.9156$. The estimated order of convergence for $T = 7.5$ is slightly worse than for $T = 0.5$ due to possible increasing rounding errors.

FIGURE 5.3. *Discrete reachable sets for $T = 7.5$ and various step-sizes $N = 16, 32, 64, 128$*



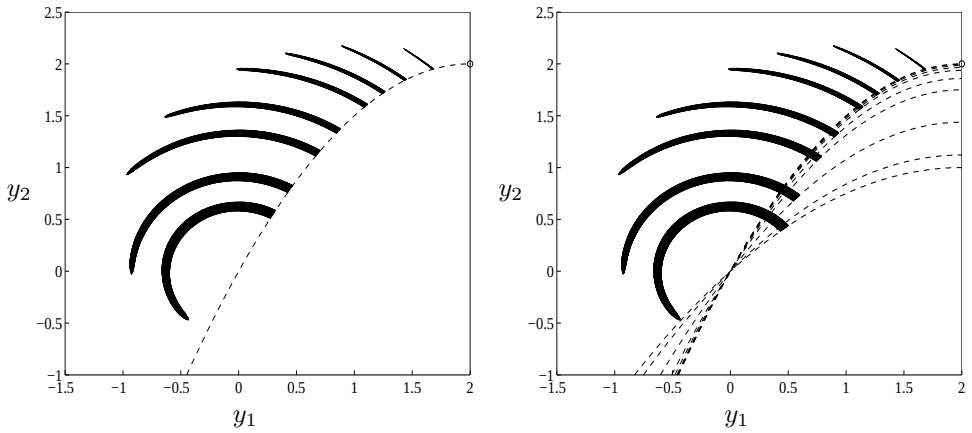
In Figure 5.3 the difference between the discrete reachable set generated by Euler's method (gray shaded set) and the theoretical one (arc with black solid line, almost included in the gray set) is depicted. The pictures show the approximations of the reachable set with state constraints at time $t = 7.5$ for several numbers N of subintervals: $N = 16$ (left upper picture), 32 (right upper one), 64 (left lower one) and 128 (right lower one).

EXAMPLE 5.2. *Consider the modified Example 5.1 in which the state constraint is now time-dependent, i.e.*

$$y(t) \in \Theta(t) := \{y \in \mathbb{R}^2 \,:\, g(t,y) \le 0\},$$
$$g(t,y) := -\frac{1}{4}\cdot(2-\frac{t^2}{64})\cdot(y_1-2)^2 + (2-\frac{t^2}{64}) - y_2, \quad y = \binom{y_1}{y_2}$$

Observe that $g(0,y)$ equals the time-independent state constraint in Example 5.1. From Figure 5.4, it is clear that in the case of time-dependent constraints (right picture), the reachable sets are bigger than in the time-independent case (left picture). This figure shows the discrete reachable sets for the constrained problem at the times $t \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, 3, 4, 6, 8\}$. For these times, the boundary of the state constraints $g(t,\cdot) = 0$ are depicted in the right picture with dotted lines.

FIGURE 5.4. *Discrete reachable sets from Euler's method with $N = 128$ for both examples*



With considerable more effort, it is even possible to show the validity of (C2) by choosing the same values u depending on the sign of y_1 as in Example 5.1.

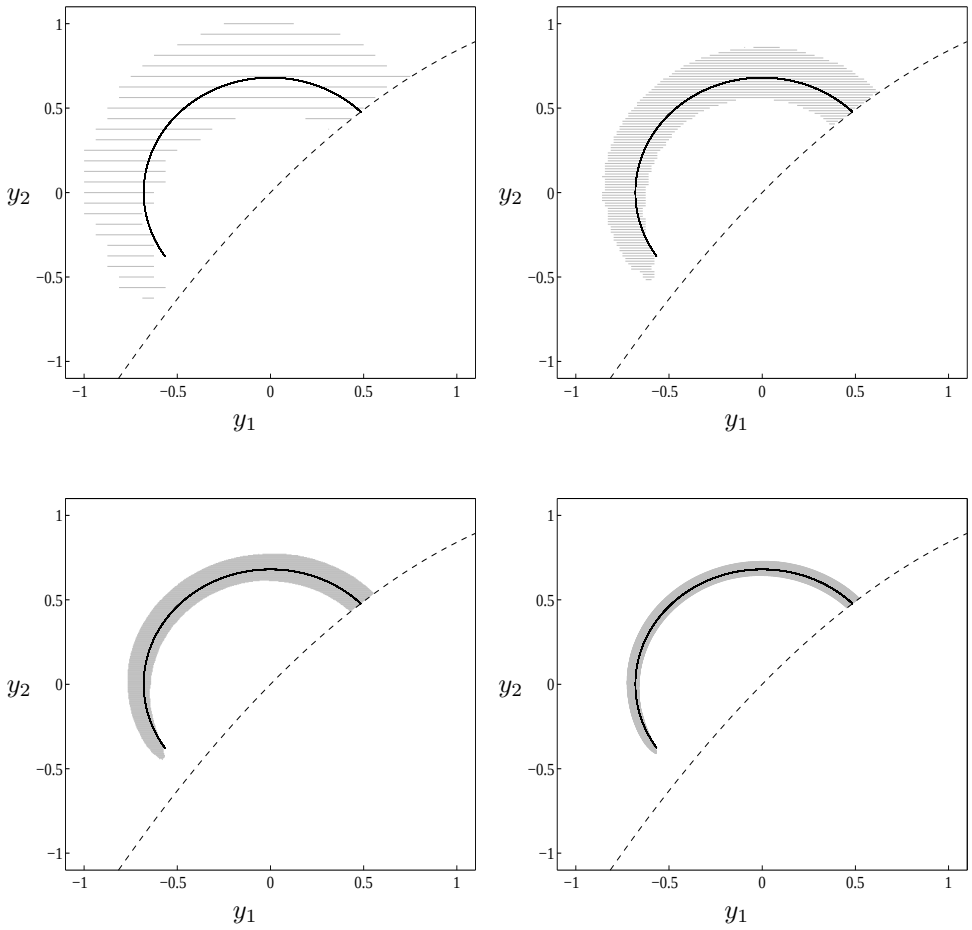
Table 5.3 is created for the time $T = 7.5$ similarly to the tables for the previous example, but include the data for the time-dependent state constraint. A least squares approximation with $\log(Ch^p)$ yields the values $p = 0.9431$ and $C = 1.9387$.

Figure 5.5 visualizes how the discrete reachable sets generated by Euler's method (gray shaded sets) approximate the theoretical reachable sets.

TABLE 5.3
Estimated order of convergence for $T = 7.5$ (time-dep. state-constrained problem)

N	estimated Hausdorff distance from the reference set	difference to Ch^p
16	0.3371631	7.2E-03
32	0.1842111	-5.1E-03
64	0.0931844	-4.1E-05
128	0.0483323	1.1E-04
256	0.0250866	1.1E-04
512	0.0130925	1.2E-05

FIGURE 5.5. *Discrete reachable sets for $T = 7.5$ and various step-sizes $N = 16, 32, 64, 128$*



REFERENCES

[1] J.-P. Aubin, *Viability Theory*, Systems & Control: Foundations & Applications, Birkhäuser, Boston, MA, 1991.

[2] J.-P. Aubin and A. Cellina, *Differential Inclusions*, vol. 264 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1984.

[3] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, vol. 2 of Systems & Control: Foundations and Applications, Birkhäuser, Boston–Basel–Berlin, 1990.

[4] J. Bastien and M. Schatzman, *Numerical precision for differential inclusions with uniqueness*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 427–460.

[5] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, vol. 50 of North-Holland Mathematics Studies, No. 5. Notas de Matemática (50), North-Holland Publishing Co., Amsterdam–London, 1973.

[6] I. A. Chahma, *Set-valued discrete approximation of state-constrained differential inclusions*, Bayreuth. Math. Schr., 67 (2003), pp. 3–162.

[7] F. H. Clarke, L. Rifford, and R. J. Stern, *Feedback in state constrained optimal control*, ESAIM Control Optim. Calc. Var., 7 (2002), pp. 97–133.

[8] F. H. Clarke and R. J. Stern, *State constrained feedback stabilization*, SIAM J. Control Optim., 42 (2003), pp. 422–441.

[9] K. Deimling, *Multivalued Differential Equations*, vol. 1 of De Gruyter Series in Nonlinear Analysis and Applications, Walter de Gruyter, Berlin–New York, 1992.

[10] A. Dontchev and F. Lempio, *Difference methods for differential inclusions: A survey*, SIAM Rev., 34 (1992), pp. 263–294.

[11] A. L. Dontchev and E. M. Farkhi, *Error Estimates for Discretized Differential Inclusions*, Computing, 41 (1989), pp. 349–358.

[12] A. L. Dontchev and W. W. Hager, *The Euler approximation in state constrained optimal control.*, Math. Comp., 70 (2001), pp. 173–203.

[13] A. L. Dontchev, W. W. Hager, and K. Malanowski, *Error bounds for Euler approximation of a state and control constrained optimal control problem.*, Numer. Funct. Anal. Optim., 21 (2000), pp. 653–682.

[14] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, Mathematics and Its Applications (Soviet Series), Kluwer Academic Publishers, Dordrecht–Boston–London, 1988.

[15] F. Forcellini and F. Rampazzo, *On nonconvex differential inclusions whose state is constrained in the closure of an open set. Applications to dynamic programming*, Differ. Integral Equ., 12 (1999), pp. 471–497.

[16] H. Frankowska, S. Plaskacz, and T. Rzeżuchowski, *Measurable Viability Theorems and the Hamilton-Jacobi-Bellman Equation*, J. Differ. Equ., 116 (1995), pp. 265–305.

[17] H. Frankowska and F. Rampazzo, *Filippov's and Filippov-Wazewski's theorems on closed domains*, J. Differ. Equ., 161 (2000), pp. 449–478.

[18] H. Frankowska and R. B. Vinter, *Existence of neighboring feasible trajectories: Applications to dynamic programming for state-constrained optimal control problems*, J. Optim. Theory Appl., 104 (2000), pp. 21–40.

[19] F. Lempio, *Difference methods for differential inclusions*, in Modern Methods of Optimization. Proceedings of a Summer School at the Schloß Thurnau of the University of Bayreuth (Germany), FRG, October 1–6, 1990, vol. 378 of Lecture Notes in Econom. and Math. Systems, Springer-Verlag, Berlin–Heidelberg–New York, 1992, pp. 236–273.

[20] ———, *Euler's method revisited*, Proc. Steklov Inst. Math., 211 (1995), pp. 429–449.

[21] F. Lempio and D. Silin, *Generalized differential equations with strongly one-sided Lipschitzian right-hand side*, Differ. Equ., 32 (1996), pp. 1485–1491.

[22] H. M. Soner, *Optimal control with state-space constraints*, SIAM J. Control Optim., 24 (1986), pp. 552–561.

[23] R. J. Stern, *Characterization of the state constrained minimal time function*, SIAM J. Control Optim., 43 (2003), pp. 697–707.

[24] V. M. Veliov, *Second order discrete approximations to strongly convex differential inclusions*, Systems Control Lett., 13 (1989), pp. 263–269.

[25] ———, *Second Order Discrete Approximation to Linear Differential Inclusions*, SIAM J. Numer. Anal., 29 (1992), pp. 439–451.