Relaxation of Euler-Type Discrete-Time Control System[1]

Vladimir Veliov[2]

Abstract

O(h) estimation is obtained for the relaxation of the Euler discretization with step h of some classes of differential inclusions.

Keywords: differential inclusions, discretization, relaxation.

1 Introduction

In 2003 G. Grammel [4] published a result that, roughly reformulated, claims the following:[3] if a differential inclusion $\dot{x} \in F(x)$ with a compact-valued and Lipschitz continuous F is replaced by the Euler discretization of this inclusion, $(x_{k+1} - x_k)/h \in \mathrm{Extr}(F(x_k))$, where only the extreme points of $F(x)$ are taken in the right-side, then the accuracy of approximation of the trajectories is $O(\sqrt{h})$. It is well known from [3] that taking the whole set $F(x_k)$ in the Euler discretization gives accuracy $O(h)$. It seemed that the drop of the accuracy from $O(h)$ to $O(\sqrt{h})$ is the price of the using only the extreme points of $F(x)$ in the Euler discretization. Examples where $O(\sqrt{h})$ is sharp were sought and not found, while in the same time in several special cases the estimation $O(h)$ holds also for the Euler discretization using the extreme points of $F(x)$. The main result in this paper had to be that the estimation $O(h)$ holds in general: under the same conditions us in [4], where only $O(\sqrt{h})$ was proven. However, the author was able so far to prove this only in some special cases and the problem is still open. These special cases cover, in fact, several important applications.

Acknowledgment: The author benefited from the useful discussions with G. Grammel and M. Brokate (TU München).

[1]This work was started during the visit of the author in the Technical University of München, 2001

[2]Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology, Argentinierstrasse 8, A-1040 Vienna, Austria; Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

[3]The claim follows also from the more general earlier paper [2].

2 Preliminaries

Under *Euler-type discrete-time control system* we understand a discrete-time inclusion of the type

$$z_{k+1} \in z_k + hF(z_k), \quad z_0 = z_0, \quad k = 0, \dots, N, \tag{1}$$

where $F : \mathbf{R}^n \Rightarrow \mathbf{R}^n$ is a set-valued mapping, and $h = 1/N$ is presumably a small positive number (N is a large natural number). Clearly, (1) could be viewed as an Euler discretization of the differential inclusion

$$\dot{x}(t) \in F(x(t)), \quad x(0) = z_0, \quad t \in [0, 1], \tag{2}$$

The so-called *relaxation theorem* (see e.g. [1]) relates the set of solutions of (2) to the set of solutions of the convexified differential inclusion

$$\dot{x}(t) \in \operatorname{co} F(x(t)), \quad x(0) = z_0, \quad t \in [0, 1], \tag{3}$$

where "co" stands for the convex hull. The relaxation theorem holds under the following condition, that we assume to be fulfilled everywhere below.

Standing assumptions. $F : \mathbf{R}^n \Rightarrow \mathbf{R}^n$ is compact valued and such that $\operatorname{co} F$ is locally Lipschitz continuous in the Hausdorff metric; there exists a compact set K such that $x(t) \in K$ for every solution $x(\cdot)$ of (2) and every $t \in [0, 1]$.

We remind that a solution of (2), or of (3), is an absolutely continuous function $x(\cdot) : [0, 1] \mapsto \mathbf{R}^n$ that satisfies the respective differential inclusion for almost every $t \in [0, 1]$. The set of all solutions of (2) and of (3) will be denoted by $\mathcal{S}$ and $\mathcal{S}_c$, respectively.

The relaxation theorem clams that under the above assumption the set $\mathcal{S}$ is dense in $\mathcal{S}_c$ with respect to the uniform norm in the space of the continuous functions $[0, 1] \mapsto \mathbf{R}^n$.

The problem that we study in this paper is to estimate the Hausdorff distance, d_H between the set of solutions, $\mathcal{S}^h$, of (1) and that of the relaxed difference inclusion

$$y_{k+1} \in y_k + h\operatorname{co} F(y_k), \quad y_0 = z_0, \quad k = 0, \dots N-1. \tag{4}$$

The set of solutions of (4) is denoted by $\mathcal{S}_c^h$. The Hausdorff distance between $\mathcal{S}^h$ and $\mathcal{S}_c^h$ is meant with respect to the sup-norm in the space of sequences $(z_0, \dots z_N)$.

Under different conditions we shall prove an estimation of the form

$$d_H(\mathcal{S}_c^h, \mathcal{S}^h) \le \text{const.} h^m, \tag{5}$$

where m is either equal to 1, or to 0.5.

It is known that the estimation

$$d_H(\hat{\mathcal{S}}_c^h, \mathcal{S}_c) \le \text{const.} h \tag{6}$$

holds (see [3]), for the Hausdorff distance d_H between $\mathcal{S}_c$ and the set of piece-wise linear interpolations of the elements of $\mathcal{S}_c^h$, the latter denoted by $\hat{\mathcal{S}}_c^h$. Combining this result with (5) we obtain

$$d_H(\mathcal{S}_c, \hat{\mathcal{S}}^h) \leq \text{const.} h^m, \tag{7}$$

where, similarly as above, $\hat{\mathcal{S}}^h$ is the set of piece-wise linear interpolations of the elements of $\mathcal{S}^h$. An estimation of the above type has obvious practical convenience: instead of using in the Euler discretization all admissible "controls" (that is, elements of $\mathrm{co}\, F(x)$) one can use only the extremal "controls" without qualitative loss of accuracy (if $m = 1$).

3 The $O(\sqrt{h})$ estimation

In this section we prove that the estimation (5) with $m = 0.5$ holds under the standing assumption.

Theorem 1 *There are constants N_0 and C such that*

$$d_h(\mathcal{S}_c^h, \mathcal{S}^h) \leq C\sqrt{h}$$

for every $h = 1/N$ with $N \geq N_0$.

Proof. Below we assume that $\mathrm{co}\, F$ is Lipschitz with constant L and bounded by a constant M in the whole space $\mathbf{R}^n$, but this is just for technical convenience – in the final version we shall make it under the standing assumption.

Clearly, $\mathcal{S}^h \subset \mathcal{S}_c^h$. Let an arbitrary solution $\{y_k\}_k$ of (4) be fixed. Inductively, we assume that z_k is already defined. Denote

$$\delta_k = |z_k - y_k|.$$

We have

$$y_{k+1} = y_k + h u_k$$

for some $u_k \in \mathrm{co}\, F(y_k)$. Below we shall specify an element $v_k \in F(z_k)$, which to produce the next

$$z_{k+1} = z_k + h v_k.$$

There exists some $\bar{v}_k \in \mathrm{co}\, F(z_k)$ such that

$$|\bar{v}_k - u_k| \leq L\delta_k.$$

Denote for brevity $l = z_k - y_k$. We have

$$\delta_{k+1} \overset{def}{=} |z_{k+1} - y_{k+1}| = |z_k + h v_k - y_k - h\bar{v}_k - h(u_k - \bar{v}_k)| = |l + h(v_k - \bar{v}_k) + h(\bar{v}_k - u_k)|.$$

Then, using below that $h \le 1$ we obtain

$$\delta_{k+1}^2 = \delta_k^2 + h^2|v_k - \bar v_k|^2 + h^2|\bar v_k - u_k|^2 + 2h\langle l, v_k - \bar v_k\rangle + 2h\langle l, \bar v_k - u_k\rangle + 2h^2\langle v_k - \bar v_k, \bar v_k - u_k\rangle$$
$$\le \delta_k^2 + h^2M^2 + h^2L^2\delta_k^2 + 2h\langle l, v_k - \bar v_k\rangle + 2hL\delta_k^2 + 2h^2ML\delta_k$$
$$\le \delta_k^2 + 2h\langle l, v_k - \bar v_k\rangle + M^2h^2 + L(L+2)h\delta_k^2 + 2h^2ML\delta_k.$$

The value v_k is still to be chosen from the set $F(z_k)$, having in mind that $\bar v_k \in \operatorname{co} F(z_k)$. Since $\bar v_k$ is a convex combination of elements from $F(z_k)$, for at least one of them, say $v \in F(z_k)$, we have $\langle l, v - \bar v_k\rangle \le 0$. We define $v_k = v$. Then from the above estimation we obtain

$$\delta_{k+1}^2 \le \delta_k^2 + Mh^2 + L(L+2)h\delta_k^2 + 2h^2ML\delta_k.$$

The last recursive relation implies in a standard way the claim of the theorem. Q.E.D.

4 Some $O(h)$ estimations

1. The assumption below is somewhat technical, but later we check it in different situations. Moreover, the proof below uses a weaker form of the assumption. Namely, instead of the Lipschitz continuity, below we use the property of bounded variation along the Euler curves. This observation may be significant for obtaining more general results.

Assumption B. For every $z \in \mathbf{R}^n$ there exist linearly independent unit vectors $l_1(z), \ldots, l_n(z)$, depending in a locally Lipschitz way on z, and having the following property. Denote by $\alpha(z; x) = (\alpha_1(z; x), \ldots, \alpha_n(z; x))$ the coordinates of the vector x in the basis $\{l_i(z)\}$. Then for every $z \in \mathbf{R}^n$, for every $\sigma_1, \ldots, \sigma_n \in \{-1, 1\}$, and for every $\bar v \in \operatorname{co} F(z)$ there exists $v \in F(z)$ such that

$$\sigma_i \alpha_i(z; v - \bar v) \le 0, \quad i = 1, \ldots, n.$$

Theorem 2 *Under Assumption* **B** *there are constants* N_0 *and* C *such that*

$$d_h(\mathcal{S}_c^h, \mathcal{S}^h) \le Ch$$

for every $h = 1/N$ *with* $N \ge N_0$*.*

Proof. Again, for simplicity we assume that everything takes place in a bounded domain Z in $\mathbf{R}^n$, where $\operatorname{co} F$ and l_i are Lipschitz with constants L and $\hat L$, respectively, and $F(\cdot)$ is bounded by a constant M.

Denote $b_{ij}(z) = \langle e_i, l_j(z)\rangle$, $i, j = 1, \dots, n$, where $\{e_i\}$ is the standard basis in $\mathbf{R}^n$. Then the matrix

$$B(z) = \{b_{ij}(z)\}$$

is invertible, and its inverse, $B^{-1}(z)$, is locally Lipschitz and bounded in norm on Z by some constant $\bar{M}$. Since we have

$$\alpha(z; x) = B^{-1}(z)x,$$

the mapping $x \longrightarrow \alpha(z; x)$ is Lipschitz in the set Z with the same constant $\bar{L}$ for all $z \in Z$. Moreover,

$$\langle e_i, z\rangle \le \sum_{j=1}^{n} |\alpha_j(z; x)||b_{ij}(z)| \le \|\alpha(z; x)\|_\infty \sum_{j=1}^{n} |b_{ij}(z)|\sqrt{n} \le \|\alpha(z; x)\|_\infty,$$

where $\|\alpha(z; x)\|_\infty \stackrel{def}{=} \max_i\{|\alpha_j(z; x)|\}$. Hence

$$|z| \le n\|\alpha(z; x)\|_\infty.$$

Let us fix an arbitrary trajectory $y_0, \dots, y_N$ of (4), and let z_k be already defined for some $k \ge 0$ (with $z_0 = y_0$, if $k = 0$). Denote

$$\delta_k = \|\alpha(z_k; z_k - y_k)\|_\infty.$$

Similarly as in the proof of Theorem 1, we have

$$y_{k+1} = y_k + hu_k$$

for some $u_k \in \operatorname{co} F(y_k)$. Below we shall specify an element $v_k \in F(z_k)$, which to produce the next

$$z_{k+1} = z_k + hv_k.$$

There exists some $\bar{v}_k \in \operatorname{co} F(z_k)$ such that

$$|\bar{v}_k - u_k| \le L|z_k - y_k| \le nL\|\alpha(z_k; z_k - y_k)\|_\infty = nL\delta_k.$$

Then we have

$$\begin{aligned}
&\alpha_i(z_{k+1}; z_{k+1} - y_{k+1}) \\
&\qquad = \alpha_i(z_{k+1}; z_k - y_k + h(v_k - \bar{v}_k)) + h\alpha_i(z_{k+1}; \bar{v}_k - u_k) \\
&\qquad = \alpha_i(z_k; z_k - y_k) + h\alpha_i(z_k; v_k - \bar{v}_k) \\
&\qquad\quad +[\alpha_i(z_{k+1}; z_k - y_k + h(v_k - \bar{v}_k)) - \alpha_i(z_k; z_k - y_k + h(v_k - \bar{v}_k))] + h\alpha_i(z_{k+1}, \bar{v}_k - u_k).
\end{aligned}$$

We define $v_k \in F(z_k)$ according to Assumption **B**, applied for

$$\bar{v} = \bar{v}_k, \quad z = z_k, \quad \sigma_i = \operatorname{sign}(\alpha_i(z_k; z_k - y_k)).$$

Thus,

$$|\alpha_i(z_{k+1}; z_{k+1} - y_{k+1})| \le \max\{\delta_k, 2\bar{M}Mh\} + \bar{L}|l_i(z_k) - l_i(z_{k+1})||z_k - y_k + h(v_k - \bar{v}_k)| + h\bar{M}|\bar{v}_k - u_k|.$$

Taking the maximum with respect to i we obtain

$$\delta_{k+1} \le \max\{\delta_k, 2\bar{M}Mh\} + \bar{L}\Delta_k(n\delta_k + 2Mh) + nLh\delta_k,$$

where we have denote $\Delta_k \stackrel{def}{=} \max_i\{|l_i(z_{k+1}) - l_i(z_k)|\}$.

Remark: Notice that $\Delta_k \le \hat{L}|z_k - z_{k+1}| \le \hat{L}Mh$. However, we shall not use this local estimate, since there might be some advantage to suppose that the sum of Δ_k is (uniformly) bounded, instead of supposing local Lipschitz continuity. The former means boundedness of the variation of l_i along the Euler curve $\{z_k\}$. In the examples the author has considered, this variation is always bounded, despite that the dimension of $F(x)$ may change. But the author faces difficulties to formalize in a nice way this observation.

We consider two cases:

1) $\delta_k \le 2\bar{M}Mh$. Then we have

$$\delta_{k+1} \le 2\bar{M}Mh + \bar{L}(2n\bar{M}Mh + 2Mh)\Delta_k + 2nL\bar{M}Mh^2 \le C_1 h$$

for an appropriate constant C_1 (Δ_k are uniformly bounded, since their sum is supposed uniformly bounded).

2) $\delta_k > 2\bar{M}Mh$. Then

$$\delta_{k+1} \le (1 + n\bar{L}\Delta_k + nLh)\delta_k + 2\bar{L}Mh\Delta_k.$$

Then the claim of the theorem follows from the following lemma, which can be easily proved by induction.

Lemma 1 *Let Δ_k satisfy the recurrent inequality*

$$\Delta_{k+1} \le (1 + \mu_k)\Delta_k + \lambda_k, \quad k = 0, 1, \ldots.$$

Then for every $k \ge 1$

$$\Delta_k \le e^{\sum_{i=0}^{k-1} \mu_i}(\Delta_0 + \sum_{i=0}^{k-1} \lambda_i).$$

Q.E.D.

Particular cases:

1.1. Let F satisfies

$$F(x) = \partial(\operatorname{co} F(x)) \quad \forall x \in \mathbf{R}^n,$$

where ∂Y denotes the boundary of Y, and co (F) satisfies the standing assumptions. (Of course, this condition makes sense only if $\operatorname{co} F(x)$ has a nonempty interior, otherwise $F(x) = \operatorname{co} F(x)$.) Then Assumption **B** is satisfied. Indeed, we may take an arbitrary fixed orthonormed basis $\{l_i(x) = l_i\}$. Let us take an arbitrary $\bar{v} \in \operatorname{co} F(x)$ and $\sigma_i \in \{-1, 1\}$. If $\bar{v} \notin \partial F(x)$, then moving from $\bar{v}$ along the vector $-(\sigma_1 l_1 + \ldots + \sigma_n l_n)$ we shall reach a point $v \in \partial F(x)$ for which (ii) is obviously satisfied.

1.2. Consider an affine system

$$F(x) = \{f_0(x) + \sum_{i=1}^{r} f_i(x) u_i, \;\; u_i \in \{-1, 1\}\},$$

where the vectors $\{f_i(x)\}_1^r$ are linearly independent for every $x \in \mathbf{R}^n$. Then Assumption **B** is fulfilled. Indeed, one may take

$$l_i(x) = \frac{f_i(x)}{|f_i(x)|}, \;\; i = 1, \ldots, r$$

and extend this collection to a normed basis in $\mathbf{R}^n$ such that the last $n-r$ vectors are also Lipschitz continuous. Then for every $\bar{v} \in \operatorname{co} F(z)$ and $\sigma_1, \ldots, \sigma_n$ we have

$$\alpha_i(z; v - \bar{v}) = \frac{u_i - \bar{u}_i}{|f_i(z)|} \quad \text{for} \quad i = 1, \ldots, r$$

and

$$\alpha_i(z; v - \bar{v}) = 0 \quad \text{for} \quad i > r.$$

In the above equalities $\bar{u}$ and u are the controls corresponding to $\bar{v}$ and $v \in F(z)$, respectively. Assumption **B** is apparently fulfilled by $u_i = -\sigma_i$.

1.3. As in case 2, but with $f_i(x) = a_i(x) l_i$ with linearly independent l_i. Obviously Assumption **B** is satisfied despite that $a_i(x)$ may be zero. This extend the the result from your draft from september, but I think that a more general condition that brings case 2 and case 3 together is also possible.

2. Here we consider another particular case where we can obtain an $O(h)$ estimation, but Assumption **B** does not hold, in general. It concerns the case of a constant mapping $F(x) = F$ whose convex hull consists of finitely many points. Namely, we consider the case

$$F(x) = F \stackrel{def}{=} \{f_1, \ldots, f_s\}, \tag{8}$$

where $f_i \in \mathbf{R}^n$.

Theorem 3 *For the constant mapping F specified in (8) the estimation:*

$$d_h(\mathcal{S}_c^h, \mathcal{S}^h) \leq 2s|F|h,$$

holds for every $h = 1/N$, where $N \in \mathbf{N}$ and $|F| \stackrel{def}{=} \max\{|f_i| : \, i = 1, \ldots, s\}$.

Proof. Let a trajectory $y_0, \ldots, y_N$ of the the convexified inclusion

$$y_{k+1} \in y_k + h\mathrm{co}\, F, \quad k = 0, \ldots N-1$$

be fixed. Then

$$y_{k+1} = y_k + h\xi_k, \quad \text{with} \quad \xi_k = \sum_{j=1}^{s} \alpha_{kj} f_j,$$

where $\alpha_{kj} \geq 0$ and $\sum_{j=1}^{s} \alpha_{kj} = 1$. We have

$$y_k = y_0 + h \sum_{i=0}^{k-1} \sum_{j=1}^{s} \alpha_{ij} f_j = y_0 + h \sum_{j=1}^{s} \beta_{kj} f_j,$$

where $\beta_{kj} = \sum_{i=0}^{k-1} \alpha_{ij}$. Clearly,

$$\sum_{j=1}^{s} \beta_{kj} = \sum_{j=1}^{s} \sum_{i=0}^{k-1} \alpha_{ij} = k.$$

We shall define the sequence

$$z_{k+1} \in z_k + hF, \quad z_0 = y_0,$$

as follows. Let z_k be already defined. Denote

$$\delta_{kj} = \beta_{kj} - \gamma_{kj},$$

where γ_{kj} is the number of times f_j is chosen in the construction of z_k. Clearly,

$$\sum_{j=1}^{s} \gamma_{kj} = k, \quad \text{hence} \quad \sum_{j=1}^{s} \delta_{kj} = 0. \tag{9}$$

Denote also

$$J_k = \{j : \ \delta_{kj} \geq 0\}, \quad \Delta_k = \sum_{j \in J_k} \delta_{kj}.$$

Let $j = \bar{j}$ be the index for which δ_{kj} is maximal. Then we define

$$z_{k+1} = z_k + hf_{\bar{j}}.$$

We shall estimate Δ_{k+1}. First suppose that $\delta_{k+1,\bar{j}} \geq 0$, thus $\bar{j} \in J_{k+1}$. Then we have

$$\Delta_{k+1} = \sum_{j \in J_{k+1}} (\beta_{k+1,j} - \gamma_{k+1,j}) = \beta_{k\bar{j}} + \alpha_{k\bar{j}} - \gamma_{k\bar{j}} - 1 + \sum_{j \in J_{k+1} \setminus \{\bar{j}\}} (\beta_{k+1,j} - \gamma_{k+1,j})$$

$$= \delta_{k,\bar{j}} + \alpha_{k,\bar{j}} - 1 + \sum_{j \in J_{k+1} \setminus \{\bar{j}\}} (\beta_{kj} + \alpha_{kj} - \gamma_{k+1,j}) = \delta_{k,\bar{j}} + \alpha_{k,\bar{j}} - 1 + \sum_{j \in J_{k+1} \setminus \{\bar{j}\}} (\delta_{kj} + \alpha_{kj})$$

$$\leq \delta_{k,\bar{j}} + \alpha_{k,\bar{j}} - 1 + \sum_{j \in J_k \setminus \{\bar{j}\}} \delta_{kj} + \sum_{j=1}^{s} \alpha_{kj} - \alpha_{k,\bar{j}} \leq \Delta_k - 1 + \sum_{j=1}^{s} \alpha_{kj} = \Delta_k.$$

In the last line we have used that $\delta_{kj} < 0$ for $j \notin J_k$.

Now, let us consider the case $\delta_{k+1,\bar{j}} < 0$. Then

$$\beta_{k\bar{j}} + \alpha_{k\bar{j}} - \gamma_{k\bar{j}} - 1 \leq 0,$$

hence

$$\delta_{k\bar{j}} \leq 1 - \alpha_{k\bar{j}} \leq 1.$$

From the maximality of $\delta_{k\bar{j}}$ we obtain $\delta_{k,j} \leq 1$ for all j, hence

$$\Delta_k \leq s.$$

Combining the two cases we obtain that at every step k

$$\text{either} \quad \Delta_{k+1} \leq \Delta_k, \quad \text{or} \quad \Delta_k \leq s.$$

Having in mind that $\Delta_0 = 0$ from here we conclude that

$$\Delta_k \leq s \ \ \forall\, k = 0, \ldots, N.$$

To complete the proof we notice that

$$y_k - z_k = h \sum_{j=1}^{s} \delta_{kj} f_j = h \left(\sum_{j \in J_k} \delta_{kj} f_j + \sum_{j \notin J_k} \delta_{kj} f_j \right),$$

thus

$$|y_k - z_k| \leq h|F| \left(\sum_{j \in J_k} \delta_{kj} + \sum_{j \notin J_k} |\delta_{kj}| \right).$$

Hence, using (9) we obtain

$$|y_k - z_k| \leq h|F|(\Delta_k + \Delta_k) \leq 2s|F|h.$$

Q.E.D.

Remark: The estimation in the last theorem seems to be somewhat inconsistent with the result from **1.1**. Indeed, the constant in Theorem 3 depends on the number of extreme points of F, while we know from **1.1** that the $O(h)$ estimation holds for $F = \partial V$, with any convex compact V, where F may have infinitely many extreme points. This means that the last proof is not quite good.

The conjecture that the O(h) estimation is true under Standing Assumptions remains open.

References

[1] J.-P. Aubin and H. Frankowska. *Set-valued Analysis*. Birkhäuser, Boston, Basel, Berlin, 1990.

[2] T. Donchev. Approximation of lower semicontinuous differential inclusions. *Numer. Funct. Anal. and Optimiz.*, **22**(1&2):55–67, 2001.

[3] A. Dontchev and E. Farkhi. Error estimates for discretized differential inclusion. *Computing*, **41**(4):349–358, 1989.

[4] G. Grammel. Towards Fully Discretized Differential Inclusions, *Set-Valued Analysis*, **11**(1): 1–8, 2003.