MPRI 2-8-2: Systèmes hybrides
(2ème partie)

Laurent Fribourg
03 février 2021
1 Switched systems
2 Numerical integration
3 Euler approximate solutions
4 Control synthesis
Switched systems

A continuous switched system

\[ \dot{x}(t) = f_{\sigma(t)}(x(t)) \]

- state \( x(t) \in \mathbb{R}^n \)
- control rule \( \sigma(\cdot) : \mathbb{R}^+ \rightarrow U \)
- finite set of modes \( U = \{1, \ldots, N\} \)

Focus on time-sampled switched systems:
- given a stepsize (or "sampling period") \( \tau > 0 \), the mode switching occurs at times \( \tau, 2\tau, \ldots \)

The control \( \sigma \) is a piecewise constant function
- with equal steps of length \( \tau \), and height value in \( U \)
Example: Two-room apartment

\[ T_1(t + \tau) = f_1(T_1(t), T_2(t), u_1) \]
\[ T_2(t + \tau) = f_2(T_1(t), T_2(t), u_2) \]

- Modes: \( \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \); sampling period \( \tau \)

- A pattern \( \pi \) is a finite sequence of modes, e.g. \( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \)

- A state dependent control consists in selecting at each \( \tau \) a mode (or a pattern) according to the current value of the state.
Controlled stability

Given a "safety" set $S$ and a "recurrence" set $R \subseteq S$,

**select** at each $t = \tau, 2\tau, \ldots$, a mode $j \in U$ (according to value $x(t)$) in order to satisfy

(R, S)-stability:

$x(t)$ returns to $R$
while never leaving $S$
1 Switched systems

2 Interval-based integration

3 Euler-based integration

4 Application to controlled stability

5 Compositional Euler’s method

6 Final remarks
Interval arithmetic vs. standard arithmetic

- **standard** numerical methods compute approximations to a mathematically correct result (due to finite representation of reals).

- **interval methods** [Moore66] manipulate set-valued real expressions: “interval vectors” or “boxes”

- they give **bounds** that are guaranteed to contain the mathematically correct result, using rules of the form:
  - $[a] + [b] = [a + b, a + b]$
  - $[a] \cdot [b] = [\min\{ab, \overline{ab}, \overline{ab}, \overline{ab}\}, \max\{ab, \overline{ab}, \overline{ab}, \overline{ab}\}]$

- they can account for
  - **rounding** errors
  - **inaccuracies** in measurements of inputs
  - **uncertainty** on parameters, **disturbance**, errors from the model

---

Interval-based integration

For $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we consider the ODE

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

solution denoted by $x(t; x_0)$ (or simply $x(t)$)

**Goal:** Given an interval $l_0$ at $t = t_0$, construct a sequence of intervals:

1. $l_1$ containing at $t_1 = t_0 + \tau$:
   $$x(t_1; l_0) = \{x(t_1; x_0) \mid x_0 \in l_0\}$$

2. $l_2$ containing at $t_2 = t_1 + \tau$:
   $$x(t_2; l_1) = \{x(t_2; x_1) \mid x_1 \in l_1\}$$

3. $\ldots$

\[^2\text{Idem.}\]
Interval-based integration

Given $l_j$ an interval for $t = t_j$, compute a (super)set of solutions $l_{j+1}$ at $t_{j+1} = t_j + \tau$ via a two-step method:

1. **Algorithm I**: compute an a priori enclosure $F_j$:
   \[ x(t; l_j) \subseteq F_j \text{ for all } t \in [t_j, t_{j+1}] \]

2. **Algorithm II**: compute a tighter enclosure $l_{j+1}$:
   \[ x(t; l_j) \subseteq l_{j+1} \subset F_j \text{ at } t = t_{j+1} \]

---

\[ \text{idem.} \]
Algorithm I: a priori enclosure method\textsuperscript{4}

Basic property: If there exists an interval $F$:

1. $l_0 \subseteq F$, and
2. $l_0 + [0, \tau] \cdot f(F) \subseteq F$

then there exists a unique solution $x(t; x_0)$ for all $t \in [0, \tau]$, $x_0 \in l_0$. Furthermore: $x(t; x_0) \in F$.

Proof based on Banach fixed-point th., and Picard-Lindelöf operator

$$(Tu)(t) = x_0 + \int_0^t f(u(s))ds.$$ 

The construction of $F$ relies on fixed-point acceleration heuristics ("widening") using adjustment of stepsize $\tau$.

\textsuperscript{4}idem
Algorithm II: tighter enclosure\(^5\)

Using \( F \), compute a tighter enclosure \( l_1 \) of \( x(t; l_0) \) for \( t = \tau \).

**Approach:** Taylor series + remainder term.

\[
x_1 = x_0 + \sum_{i=1}^{k-1} \tau^i \cdot f^{(i)}(x_0) + \tau^k \cdot f^{(k)}(y), \text{ for some } y \in F.
\]

Hence

\[
l_1 = l_0 + \sum_{i=1}^{k-1} \tau^i \cdot f^{(i)}(l_0) + \tau^k \cdot f^{(k)}(F)
\]

\[\text{NB: with this algo, } |l_1| > |l_0| \]

– even if the true solutions contract!

\[\rightarrow \text{ further refinement needed}\]
INTERVAL-BASED INTEGRATION
Wrapping effect\textsuperscript{6}

A simple rotation:

\[
\dot{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x; \quad x_0 \in l_0
\]

The solution is \( x(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} x_0 \), where \( x_0 \in l_0 \)

\( l_0 \) can be viewed as a parallelepiped.

At each step, the parallelepiped is rotated and has to be wrapped by another one.

At \( t = 2\pi \), the blow up factor is by a factor \( e^{2\pi} \approx 535 \), as the stepsize tends to zero.

\textsuperscript{6}idem
(Dis)advantages of interval methods\(^7\)

**Advantages** over standard numerical methods:

1. ensure a **unique solution** exists
2. provide **guaranteed bounds** on the solution
3. can be efficient for problems with **ranges of parameters**

**Disadvantages**

1. computation is **time consuming**
2. **harder to implement** than standard numerical methods
3. error bounds may be too **large**

\(^7\) idem
1. Switched systems

2. Interval-based integration

3. Euler-based integration

4. Application to controlled stability

5. Compositional Euler’s method

6. Final remarks
Euler’s approximation $\tilde{x}(t)$ of $x(t)$

$$\tilde{x}(t) = \tilde{x}(t_i) + (t - t_i) \cdot f(\tilde{x}(t_i))$$

Piecewise linear fn.:

at each step, constant derivative of $\tilde{x}(t)$ ($= f(\tilde{x}(t_i))$ deriv. at starting pt)
Classical error bound (using Lipschitz constant $L$)

- The error at $t = t_0 + k\tau$ is: $\|x(t) - \bar{x}(t)\|$.

If $f$ is Lipschitz cont. ($\|f(y) - f(x)\| \leq L\|y - x\|$), then:

$$\text{error}(t) \leq \frac{\tau M}{2L} (e^{L(t-t_0)} - 1)$$

where $L$ is the Lipschitz constant of $f$ (and $M$ an upper bound on $f''$).

- In case of "stiff" equations, $L$ can be very big!

Idea: exploit another constant $\lambda$ that will allow for a sharper estimation of Euler’s error
One-sided Lipschitz (OSL) constant $\lambda$

- $\lambda \in \mathbb{R}$ is a constant s.t., for all $x, y \in S$:
  \[ \langle f(y) - f(x), y - x \rangle \leq \lambda \|y - x\|^2 \]
  where $\langle \cdot, \cdot \rangle$ denote the scalar product of two vectors of $\mathbb{R}^n$

- $\lambda$ can be $< 0$ ($\rightarrow$ contractivity)

- even in case $\lambda > 0$, in practice: $\lambda \ll L$
  $\rightarrow$ sharper estimation of Euler error

- $\lambda$ can be computed using constraint optimization algorithms
Hypotheses

(H0) (Lipschitz): for all \( j \in U \), there exists a constant \( L_j > 0 \) such that:

\[
\|f_j(y) - f_j(x)\| \leq L_j \|y - x\| \quad \forall x, y \in S.
\]

(H1) (one-sided Lipschitz): for all \( j \in U \), there exists a constant \( \lambda_j \in \mathbb{R} \) such that

\[
\langle f_j(y) - f_j(x), y - x \rangle \leq \lambda_j \|y - x\|^2 \quad \forall x, y \in T^8,
\]

The constants \( C_j \) for all \( j \in U \) are defined as follows:

\[
C_j = \sup_{x \in S} L_j \|f_j(x)\|.
\]

\(^8 T \) is the one-step expansion of \( S \) under all the modes \( j \) of \( U \)
Computation of the constants

Computation of $L_j, C_j, \lambda_j \ (j \in U)$ realized with constrained optimization algorithms, applied on the following optimization problems:

- **Constant $L_j$:**
  \[
  L_j = \sup_{x, y \in S, \ x \neq y} \frac{\|f_j(y) - f_j(x)\|}{\|y - x\|}
  \]

- **Constant $C_j$:**
  \[
  C_j = \sup_{x \in S} L_j \|f_j(x)\|
  \]

- **Constant $\lambda_j$:**
  \[
  \lambda_j = \sup_{x, y \in T, \ x \neq y} \frac{\langle f_j(y) - f_j(x), y - x \rangle}{\|y - x\|^2}
  \]
Notations

Let $x_j(t)$ the solution at time $t$ of the system under mode $j$ with (implicit) initial point $x^0$

$$
\dot{x}(t) = f_j(x(t)), \\
x(0) = x^0.
$$

Given an (approximate) initial point $\tilde{x}^0 \in S$ and a mode $j \in U$, the Euler approximate, denoted by $\tilde{x}_j(t;\tilde{x}^0)$, is defined by:

$$
\tilde{x}_j(t;\tilde{x}^0) = \tilde{x}^0 + t \cdot f_j(\tilde{x}^0), \quad \text{with } t \in [0,\tau]
$$

We are going to determine an upper bound $\delta_j(t)$ to

$$
\text{error}_j(t) \equiv \|x_j(t; x^0) - \tilde{x}_j(t; \tilde{x}^0)\|,
$$

assuming $\text{error}_j(0) \equiv \|x^0 - \tilde{x}^0\| \leq \delta^0$ for some $\delta^0 \in \mathbb{R}_+$. 
Basic result: local error $\delta_j(t)$ using $\lambda_j$

Theorem

Given a system satisfying (H0-H1), an approximate initial pt $\tilde{x}^0$, a positive real $\delta^0$ and $j \in U$, we have:

For all initial point $x^0 \in B(\tilde{x}^0, \delta^0)$,

$$x_j(t; x^0) \in B(\tilde{x}_j(t; \tilde{x}^0), \delta_j(t)) \quad \text{for all } t \in [0, \tau].$$

with

- if $\lambda_j < 0$ : $\delta_j(t) = \left( (\delta^0)^2 e^{\lambda_j t} + \frac{C_j^2}{\lambda_j^2} \left( t^2 + \frac{2t}{\lambda_j} + \frac{2}{\lambda_j^2} \left( 1 - e^{\lambda_j t} \right) \right) \right)^{\frac{1}{2}}$

- if $\lambda_j = 0$ : $\delta_j(t) = \left( (\delta^0)^2 e^t + C_j^2(-t^2 - 2t + 2(e^t - 1)) \right)^{\frac{1}{2}}$

- if $\lambda_j > 0$ :

  $$\delta_j(t) = \left( (\delta^0)^2 e^{\lambda_j t} + \frac{C_j^2}{3\lambda_j^2} \left( -t^2 - \frac{2t}{3\lambda_j} + \frac{2}{9\lambda_j^2} \left( e^{3\lambda_j t} - 1 \right) \right) \right)^{\frac{1}{2}}$$
Application to one-step controlled safety

Given a ball $B^0 = B(\tilde{x}^0, \delta^0) \subset S$, safely control $B^0$ during one step, select $j \in U$:

$$x_j(t; B^0) \subseteq S, \quad \forall t \in [0, \tau]$$

It suffices to find $j \in U$: $B^1 = B(\tilde{x}^1, \delta_j(\tau)) \subseteq S$ with $\tilde{x}^1 = \tilde{x}^0 + \tau \cdot f_j(\tilde{x}^0)$ provided $\delta_j$ verified to be convex on $[0, \tau]$
Sketch of the proof

Error equation

$$\frac{d}{dt}(x(t) - \tilde{x}(t)) = (f_j(x(t)) - f_j(\tilde{x}^0)),$$

Transformation into a differential inequality

$$\frac{1}{2} \frac{d}{dt} \left( \|x(t) - \tilde{x}(t)\|^2 \right) = \langle f_j(x(t)) - f_j(\tilde{x}^0), x(t) - \tilde{x}(t) \rangle$$
$$\leq \langle f_j(x(t)) - f_j(\tilde{x}(t)), x(t) - \tilde{x}(t) \rangle +$$
$$\|f_j(\tilde{x}(t)) - f_j(\tilde{x}^0)\| \|x(t) - \tilde{x}(t)\|$$
$$\leq \lambda_j \|x(t) - \tilde{x}(t)\|^2 + L_j t \|f_j(\tilde{x}^0)\| \|x(t) - \tilde{x}(t)\|$$

Then integration of the differential inequality, knowing that

$$\|x(t) - \tilde{x}(t)\| \leq \frac{1}{2} \left( \alpha \|x(t) - \tilde{x}(t)\|^2 + \frac{1}{\alpha} \right)$$
for \( \alpha > 0 \)
Application to guaranteed integration

Given a sampled switched system satisfying (H0-H1), consider a point $\bar{x}^0 \in S$, a real $\delta > 0$ and a mode $j \in U$ such that:

1. $B(\bar{x}^0, \delta) \subseteq S$,
2. $B(\tilde{\phi}_j(\tau; \bar{x}^0), \delta_j(\tau)) \subseteq S$, and
3. $\frac{d^2(\delta_j(t))}{dt^2} > 0$ for all $t \in [0, \tau]$.

Then we have, for all $x^0 \in B(\bar{x}^0, \delta)$ and $t \in [0, \tau]$: $\phi_j(t; x^0) \in S$. 
Convexity of the trajectories

Example of a DC-DC converter:
The dynamics is given by the equation $\dot{x}(t) = A_{\sigma(t)}x(t) + B_{\sigma(t)}$ with $\sigma(t) \in U = \{1, 2\}$. The two modes are given by the matrices:

$$A_1 = \begin{pmatrix} -\frac{r_l}{x_l} & 0 \\ 0 & -\frac{1}{x_c} \frac{1}{r_0+r_c} \end{pmatrix} \quad B_1 = \begin{pmatrix} \frac{v_s}{x_l} \\ 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -\frac{1}{x_l} \left( \frac{r_l + \frac{r_0 \cdot r_c}{r_0+r_c}}{r_0+r_c} \right) & -\frac{1}{x_c} \frac{r_0}{r_0+r_c} \\ \frac{1}{x_c} \frac{r_0}{r_0+r_c} & -\frac{1}{x_c} \frac{r_0}{r_0+r_c} \end{pmatrix} \quad B_2 = \begin{pmatrix} \frac{v_s}{x_l} \\ 0 \end{pmatrix}$$

with $x_c = 70$, $x_l = 3$, $r_c = 0.005$, $r_l = 0.05$, $r_0 = 1$, $v_s = 1$.

| $\lambda_1$ | $-0.014215$ |
| $\lambda_2$ | $0.142474$ |
| $C_1$ | $6.7126 \times 10^{-5}$ |
| $C_2$ | $2.6229 \times 10^{-2}$ |
Remarks on the form of $\delta_j(\cdot)$

ex: DC-DC converter

$\lambda_1 = -0.0142 < 0$

$\lambda_2 = 0.142 > 0$

For mode 1 ($\lambda_1 < 0$): optimal stepsize $\tau$ corresponding to minimum of $\delta_1$

For mode 2 ($\lambda_2 > 0$): $\delta_2$ always $\uparrow$

$\rightarrow$ suggests subsampling of $\tau$ for achieving better precision
No wrapping effect in the rotation example

\[ \dot{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x \]

constants: \( \lambda = 0, \ C = 4.2, \ L = 1 \)
initial error: \( \delta^0 = 0.1 \)
stepsize: \( \tau = 0.005 \)
Recap': Interval-based vs. Euler-based method

- input/output: intervals $I_0, I_1$ vs ball $B_0 \equiv B(C_0, \delta_0), B_1 \equiv B(C_1, \delta_1)$
- method: $I_1$ computed from $I_0$ using intermediate structure $F$
  vs. $B_1$ evaluated directly from $C_0$ and $\delta_0$
Euler-based integration (vs. interval integration)

- **Advantages:**
  1. Computationally **very cheap** (standard arithmetic, no need for computation of $f$ derivatives, $\delta_j$ pre-computed)
  2. allows a priori for longer stepsize $\tau$ (often)
  3. reduces wrapping effect (sometimes)
  4. well-suited to controlled safety

- **Limits:**
  - less precise than interval-based integration method
    (1st order Taylor method vs. higher order Taylor method)
Outline

1. Switched systems
2. Interval-based integration
3. Euler-based integration
4. Application to controlled stability
One-step controlled safety

Given a ball \( B^0 \equiv B(\bar{x}^0, \delta^0) \subseteq S \), select a mode \( j \):
\[
\chi_j(t; B^0) \subseteq S \quad \text{for all } t \in [0, \tau]
\]

It suffices to find \( j \):
\[
B^1 \equiv B(\bar{x}^1, \delta^1) \subseteq S \quad \text{with} \quad \bar{x}^1 = \bar{x}^0 + \tau \cdot f_j(\bar{x}^0) \quad \text{and} \quad \delta^1 = \delta_j(\tau)
\]
assuming \( \delta_j(\cdot) \) convex
Multi-step controlled safety

Given a ball $B^0 \equiv B(\bar{x}^0, \delta^0) \subset S$, select a pattern $\pi$ (of length $k$):

$$x(t; B^0) \in S \quad \text{for all } t \in [0, k\tau]$$

It suffices to find a pattern $\pi \equiv j_1 \cdots j_k$:

$$B^1 \equiv B(\bar{x}^1, \delta^1_{j_1}) \subset S, \quad \ldots, \quad B^k \equiv B(\bar{x}^k, \delta^k_{j_k}) \subset S$$
Controlled $(R, S)$-stability

1. Find a set of initial balls $B_i^0 \equiv B(\tilde{x}_i^0, \delta^0) \subset S$ covering $R$.
2. For each $B_i^0$, select a pattern $\pi_i$ of the form $j_1 \cdots j_{k_i}$:
   - **safety**: all the balls $B_i^1 \equiv B(\tilde{x}_i^1, \delta^1), \ldots, B_i^{k_i} \equiv B(\tilde{x}_i^{k_i}, \delta^{k_i})$ are $\subseteq S$, and
   - **recurrence**: the last ball $B_i^{k_i}$ is $\subseteq R$. 

![Diagram showing controlled (R,S)-stability](image)
Euler-based control vs. interval-based control
Example: Building ventilation
[Meyer, Nazarpour, Girard, Witrant, 2014]

Dynamics of a four-room apartment:

\[
\frac{dT_i}{dt} = \sum_{j \in U^*} a_{ij} (T_j - T_i) + \delta_s b_i (T_{si}^4 - T_i^4) + c_i \max \left( 0, \frac{V_i - V_i^*}{V_i - V_i^*} \right) (T_u - T_i).
\]

with \( U^* = \{1, 2, 3, 4, u, o, c\} \)

16 switching modes
(control inputs: \( V_1, V_4 \in \{0V, 3.5V\} \), and \( V_2, V_3 \in \{0V, 3V\} \))
### Building ventilation

<table>
<thead>
<tr>
<th></th>
<th>Euler</th>
<th>DynIBEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$</td>
<td></td>
<td>$[20, 22]^4$</td>
</tr>
<tr>
<td>$S$</td>
<td></td>
<td>$[19, 23]^4$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Time subsampling</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Complete control</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$\max_{j=1,\ldots,16} \lambda_j$</td>
<td>$-6.30 \times 10^{-3}$</td>
<td>$4.18 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\max_{j=1,\ldots,16} C_j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of balls/tiles</td>
<td>4096</td>
<td>252</td>
</tr>
<tr>
<td>Pattern length</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CPU time</td>
<td>63 seconds</td>
<td>249 seconds</td>
</tr>
</tbody>
</table>

Control based on **Euler** (left) and **interval** (right).
Building ventilation

Control based on Euler (left) and interval (right).
Two-tank system

The behavior of $x_1$ is given by $\dot{x}_1 = -x_1 - 2$ when the tank 1 valve is closed, and $\dot{x}_1 = -x_1 + 3$ when it is open. Likewise, $x_2$ is driven by $\dot{x}_2 = x_1$ when the tank 2 valve is closed and $\dot{x}_2 = x_1 - x_2 - 5$ when it is open.
Two-tank system

<table>
<thead>
<tr>
<th></th>
<th>Euler</th>
<th>DynIBEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$</td>
<td>$[-1.5, 2.5]$</td>
<td>$[-0.5, 1.5]$</td>
</tr>
<tr>
<td>$S$</td>
<td>$[-3, 3]$</td>
<td>$[-3, 3]$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Time subsampling</td>
<td>$\tau/10$</td>
<td></td>
</tr>
<tr>
<td>Complete control</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.20711</td>
<td></td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-0.50000</td>
<td></td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.20711</td>
<td></td>
</tr>
<tr>
<td>$\lambda_4$</td>
<td>-0.50000</td>
<td></td>
</tr>
<tr>
<td>$C_1$</td>
<td>11.662</td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>28.917</td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>13.416</td>
<td></td>
</tr>
<tr>
<td>$C_4$</td>
<td>32.804</td>
<td></td>
</tr>
<tr>
<td>Number of balls/tiles</td>
<td>64</td>
<td>10</td>
</tr>
<tr>
<td>Pattern length</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>CPU time</td>
<td>58 seconds</td>
<td>246 seconds</td>
</tr>
</tbody>
</table>
Final remarks

1. Very simple method

2. Very easy to implement (a few hundreds of lines of Octave)

3. Fast, but may lack precision
   w.r.t. sophisticated refinements of interval-based methods

4. Method can be adapted to control reachability (instead of stability)

5. Replacement of forward Euler’s method by better numerical schemes
   (e.g.: backward Euler, Runge-Kutta of order 4)
   does not seem to gain much in the control framework

6. Several examples for which Euler-based control
   - beats state-of-art interval-based control (e.g.: building ventilation)
   - but the converse is also true! (e.g.: DC-DC converter)
APPLICATION TO PARAMETRIZED SYSTEMS WITH LIMIT CYCLES
Euler’s method and error bounds

Let us consider the differential system:

\[ \frac{dx(t)}{dt} = f(x(t)), \]

with states \( x(t) \in \mathbb{R}^n \) and \( x_0 \) a given initial condition.

\( \tilde{x}(t; y_0) \) denotes Euler’s approximate value of \( x(t) \) (defined by \( \tilde{x}(t; y_0) = y_0 + t \times f(y_0) \) for \( t \in [0, \tau] \), where \( \tau \) is the integration time-step).
Proposition

[LC DVCF17] Consider the solution $x(t; y_0)$ of $\frac{dx}{dt} = f(x)$ with initial condition $y_0$ and the approximate Euler solution $\tilde{x}(t; x_0)$ with initial condition $x_0$. For all $y_0 \in B(x_0, \varepsilon)$, we have:

$$\|x(t; y_0) - \tilde{x}(t; x_0)\| \leq \delta_\varepsilon(t).$$

Definition

$\delta_\varepsilon(t)$ is defined as follows for $t \in [0, \tau]$:

if $\lambda < 0$:

$$
\delta_\varepsilon(t) = \left( \varepsilon^2 e^{\lambda t} + \frac{C^2}{\lambda^2} \left( t^2 + \frac{2t}{\lambda} + \frac{2}{\lambda^2} \left( 1 - e^{\lambda t} \right) \right) \right)^{\frac{1}{2}}
$$

if $\lambda = 0$:

$$
\delta_\varepsilon(t) = \left( \varepsilon^2 e^t + C^2 (-t^2 - 2t + 2(e^t - 1)) \right)^{\frac{1}{2}}
$$

if $\lambda > 0$:

$$
\delta_\varepsilon(t) = \left( \varepsilon^2 e^{3\lambda t} + \frac{C^2}{3\lambda^2} \left( -t^2 - \frac{2t}{3\lambda} + \frac{2}{9\lambda^2} \left( e^{3\lambda t} - 1 \right) \right) \right)^{\frac{1}{2}}
$$

where $C$ and $\lambda$ are real constants specific to function $f$, defined as follows:

$$
C = \sup_{y \in S} \|f(y)\|,
$$
Definition

$L$ denotes the Lipschitz constant for $f$, and $\lambda$ is the “one-sided Lipschitz constant” (or “logarithmic Lipschitz constant” [AS14]) associated to $f$, i.e., the minimal constant such that, for all $y_1, y_2 \in S$:

$$\langle f(y_1) - f(y_2), y_1 - y_2 \rangle \leq \lambda \|y_1 - y_2\|^2, \quad (H0)$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product of two vectors of $S$.

The constant $\lambda$ can be computed using a nonlinear optimization solver (e.g., CPLEX [Cpl09]) or using the Jacobian matrix of $f$.

---


Systems with bounded uncertainty

A differential system with bounded uncertainty is of the form

\[
\frac{dx(t)}{dt} = f(x(t), w(t)),
\]

with \( t \in \mathbb{R}_0^+ \), states \( x(t) \in \mathbb{R}^n \), and uncertainty \( w(t) \in \mathcal{W} \subset \mathbb{R}^n \) (\( \mathcal{W} \) is compact, i.e., closed and bounded).

- We suppose (see [LCADSC+17]) that there exist constants \( \lambda \in \mathbb{R} \) and \( \gamma \in \mathbb{R}_{\geq 0} \) such that, for all \( y_1, y_2 \in S \) and \( w_1, w_2 \in \mathcal{W} \):

\[
\langle f(y_1, w_1) - f(y_2, w_2), y_1 - y_2 \rangle \leq \lambda \| y_1 - y_2 \|^2 + \gamma \| y_1 - y_2 \| \| w_1 - w_2 \| \quad (H1).
\]

- Instead of computing \( \lambda \) and \( \gamma \) globally for \( S \), it is advantageous to compute them locally depending on the subregion of \( S \) occupied by the system state during a considered interval of time.

---

Proposition

\( \delta_\varepsilon(t) \) is defined as follows for \( t \in [0, \tau] \):

if \( \lambda < 0 \):
\[
\delta_{\varepsilon,W}(t) = \left( \frac{C^2}{-\lambda^4} \left( -\lambda^2 t^2 - 2\lambda t + 2e^{\lambda t} - 2 \right) + \frac{1}{\lambda^2} \left( \frac{C\gamma|W|}{-\lambda} \left( -\lambda t + e^{\lambda t} - 1 \right) + \lambda \left( \frac{\gamma^2(|W|/2)^2}{-\lambda} (e^{\lambda t} - 1) + \lambda \varepsilon^2 e^{\lambda t} \right) \right) \right)^{1/2}
\]

(1)

if \( \lambda > 0 \):
\[
\delta_{\varepsilon,W}(t) = \frac{1}{(3\lambda)^{3/2}} \left( \frac{C^2}{\lambda} \left( -9\lambda^2 t^2 - 6\lambda t + 2e^{3\lambda t} - 2 \right) + 3\lambda \left( \frac{C\gamma|W|}{\lambda} \left( -3\lambda t + e^{3\lambda t} - 1 \right) + 3\lambda \varepsilon^2 e^{3\lambda t} \right) \right) \right)^{1/2}
\]

(2)

if \( \lambda = 0 \):
\[
\delta_{\varepsilon,W}(t) = \left( C^2 \left( -t^2 - 2t + 2e^t - 2 \right) + \left( C\gamma|W| \left( -t + e^t - 1 \right) + \left( \gamma^2(|W|/2)^2 (e^t - 1) + \varepsilon^2 e^t \right) \right) \right)^{1/2}
\]

(3)
Proposition

Suppose that, for some index $1 \leq j \leq n$, we have $m_+^j < M_-^j$ where $m_+^j$ (resp. $M_-^j$) denotes the minimum (resp. maximum) of $\bar{x}_j^i(t) + \delta_{\varepsilon, \mathcal{W}}(t)$ (resp. $\bar{x}_j^i(t) - \delta_{\varepsilon, \mathcal{W}}(t)$) for $t \in [iT, (i+1)T]$. Then $B[iT, (i+1)T]$ contains no fixed point of $\Sigma'$. 
Consider the Van der Pol (VdP) system $\Sigma_p$ of dimension $n = 2$ with parameter $p \in \mathbb{R}$, and initial condition in $B_0 = B(x_0, \varepsilon)$ for some $x_0 \in \mathbb{R}^2$ and $\varepsilon > 0$ (see [BQ20]):

$$\begin{cases} \frac{du_1}{dt} = u_2 \\ \frac{du_2}{dt} = pu_2 - pu_1^2 u_2 - u_1 \end{cases}$$ (4)
Van der Pol System with uncertainty

Consider now the system $\Sigma'$ with uncertainty $w(\cdot) \in \mathcal{W}_0 = [-0.5, 0.5]$ and initial condition $x_0$:

$$\begin{cases}
\frac{du_1}{dt} = u_2 \\
\frac{du_2}{dt} = (p_0 + w)u_2 - (p_0 + w)u_1^2u_2 - u_1
\end{cases} \tag{5}$$

with $p_0 = 1.1$. It is easy to see that each solution of $\Sigma_p$ with $p \in [p_0 - 0.5, p_0 + 0.5] = [0.6, 1.6]$ is a particular solution of system $\Sigma'$. 

Van der Pol System with uncertainty

VdP system with parameter $p_0 = 1.1$, uncertainty $|\mathcal{W}_0| = 0.5$, initial radius $\varepsilon_0 = 0.2$, initial point $x_0 = (1.7018, -0.1284)$, period $T_0 = 6.746$, time-step $\tau = 10^{-3}$.

- We have: $B((i_0 + 1)T_0) \subset B(i_0T_0)$ for $i_0 = 3$.
- The minimum $m_+^1$ of the upper green curve $\bar{u}_1(t) + \delta_{\mathcal{W}}(t)$ is less than the maximum $M_-^1$ of the lower green curve $\bar{u}_1(t) - \delta_{\mathcal{W}}(t)$.
- Whatever the value of $p \in [p_0 - |\mathcal{W}_0|, p_0 + |\mathcal{W}_0|] = [0.6, 1.6]$, the solution of $\Sigma_p$ never converges to a point of $\mathbb{R}^n$.
- Since the size of the system is $n = 2$, it follows by Poincaré-Bendixson’s theorem that the solution of $\Sigma_p$ converges always towards a limit circle.
Consider now the system $\Sigma'$ with uncertainty $w(\cdot) \in W_1 = [-0.2, 0.2]$ and initial condition $x_0$:

$$
\begin{align*}
\frac{du_1}{dt} &= u_2 \\
\frac{du_2}{dt} &= (p_1 + w)u_2 - (p_1 + w)u_1^2u_2 - u_1
\end{align*}
$$

with $p_1 = 0.4$. It is easy to see that each solution of $\Sigma_p$ with $p \in [p_1 - 0.2, p_1 + 0.2] = [0.2, 0.6]$ is a particular solution of system $\Sigma'$. 

(5)
VdP system with parameter $p_1 = 0.4$, uncertainty $|\mathcal{W}_1| = 0.2$, initial radius $\varepsilon_1 = 0.2$, initial point $x_0 = (1.7018, -0.1284)$, period $T_1 = 6.347$, time-step $\tau = 10^{-3}$.

- We have: $B((i_1 + 1)T_1) \subset B(i_1 T_1)$ for $i_1 = 3$.
- We have $m_1^+ < M_1^-$, this shows that whatever the value of $p \in [p_1 - |\mathcal{W}_1|, p_1 + |\mathcal{W}_1|] = [0.2, 0.6]$, the solution of $\Sigma_p$ never converges to a point of $\mathbb{R}^n$.
- It follows by Poincaré-Bendixson’s theorem that the solution of $\Sigma_p$ converges always towards a limit circle for any $p \in [0.2, 0.6]$ and initial condition in $B(x_0, \varepsilon_1)$. 
Consider now the system $\Sigma'$ with uncertainty $w(\cdot) \in \mathcal{W}_2 = [-0.3, 0.3]$ and initial condition $x_0$:

\[
\begin{align*}
\frac{du_1}{dt} &= u_2 \\
\frac{d u_2}{dt} &= (p_2 + w)u_2 - (p_2 + w)u_1^2 u_2 - u_1
\end{align*}
\]  \tag{5}

with $p_2 = 1.9$. It is easy to see that each solution of $\Sigma_p$ with $p \in [p_2 - 0.3, p_2 + 0.3] = [1.6, 2.2]$ is a particular solution of system $\Sigma'$. 
VdP system with parameter $p_2 = 1.9$, uncertainty $|\mathcal{W}_2| = 0.3$, initial radius $\varepsilon_2 = 0.1$, initial point $x_0 = (1.7018, -0.1284)$, period $T_2 = 7.531$, time-step $\tau = 10^{-3}$.

- We have: $B((i_2 + 1)T_2) \subset B(i_2 T_2)$ for $i_2 = 3$.
- We have $m_1^+ < M_1^-$, then whatever the value of $p \in [p_2 - |\mathcal{W}_2|, p_2 + |\mathcal{W}_2|] = [1.6, 2.2]$, the solution of $\Sigma_p$ never converges to a point of $\mathbb{R}^n$.
- It follows by Poincaré-Bendixson’s theorem that the solution of $\Sigma_p$ converges always towards a limit circle for any $p \in [1.6, 2.2]$ and initial condition in $B(x_0, \varepsilon_2)$.
Let $B_{WW}^u(t) \equiv B(\bar{Y}_{x_0}^{u,0}(t), \delta_{\epsilon,WW}^u(t))$.

Lemma 1. Suppose

\[ (** ) \quad B_{WW}((i + K)\Delta t) \subset B_{WW}(i\Delta t), \text{ for some } i \geq 0. \]

Then we have: $\lambda_i^{i+1} + \cdots + \lambda_i^{i+K} < 0$, where $-\lambda_i^j$ is the local rate of contraction\(^5\) for the region occupied by the system at $t \in [(j-1)\Delta t, j\Delta t]$ \((j = i+1, \ldots, i+K)\). This means that, from $t = i\Delta t$, the unperturbed system is contracting (i.e., the distance between two trajectories decreases exponentially) every $T = K\Delta t$ time-steps), and the unperturbed system converges to a limit cycle.

Proof. (sketch). Ad absurdum: Suppose $\lambda_i^{i+1} + \cdots + \lambda_i^{i+K} \geq 0$. It follows, using (H): $\delta_{\mu,WW}((i + K)\Delta t) \geq e^{(\lambda_i^{i+1} + \cdots + \lambda_i^{i+K})\Delta t} \delta_{\mu,WW}(i\Delta t) \geq \delta_{\mu,WW}(i\Delta t)$. This implies that the radius of $B_{WW}((i + K)\Delta t)$ is greater than or equal to the radius of $B_{WW}(i\Delta t)$, which contradicts (**). So $\lambda_i^{i+1} + \cdots + \lambda_i^{i+K} < 0$, which implies that the unperturbed system converges to an LC (see [20], Theorem 2).\(^6\)
Theorem 2. Let $y_0 \in S$ be a point of $\varepsilon$-representative $z_0 \in \mathcal{X}$ (so $\|y_0 - z_0\| \leq \varepsilon$). Let $T = kT = K\Delta t$. Let $\pi \in U^k$ be the optimal pattern output by PROC$^\varepsilon_k(z_0)$ for the unperturbed system with finite horizon $T$. Let us consider the tube $B_W(t) \equiv B(\tilde{Y}^{\pi^*}_{z_0,0}(t), \delta^\pi_{\mu,\mathcal{W}}(t))$ for some $\mu \geq \varepsilon$. Suppose that the following inclusion condition holds:

\begin{equation}
(*) \quad B_W((i + K)\Delta t) \subseteq B_W(i\Delta t) \text{ for some } i \geq 0.
\end{equation}

Then:

1) The exact solution $Y^\pi_{y_0,0}(t)$ of the unperturbed system under control $\pi^*$ converges to an LC $\mathcal{L}$ when $t \to \infty$.

2) For all $w \in \mathcal{W}$, the exact solution $Y^\pi_{y_0,w}(t)$ of the perturbed system under $\pi^*$ always remains inside the tube $B_W(t)$, which is bounded and contains $\mathcal{L}$.

This reflects the robustness of the perturbed system under $\pi^*$. 


