Decidability of the Reachability for a Family of Linear Vector Fields

Ting Gan[1], Mingshuai Chen[2], Yangjia Li[2], Bican Xia[1], and Naijun Zhan[2]

[1] LMAM & School of Mathematical Sciences, Peking University
[2] State Key Lab. of Computer Science, Institute of Software, Chinese Academy of Sciences

Aalborg, June 2016

Example : Home Heating

$x_3(t)$ = Temperature in the attic,
$x_2(t)$ = Temperature in the living area,
$x_1(t)$ = Temperature in the basement,
t = Time in hours.

Example : Home Heating

$x_3(t)$ = Temperature in the attic,
$x_2(t)$ = Temperature in the living area,
$x_1(t)$ = Temperature in the basement,
t = Time in hours.

$$\begin{aligned}
\dot{x}_1 &= \frac{1}{2}(45 - x_1) + \frac{1}{2}(x_2 - x_1),\\
\dot{x}_2 &= \frac{1}{2}(x_1 - x_2) + \frac{1}{4}(35 - x_2) + \frac{1}{4}(x_3 - x_2) + 20,\\
\dot{x}_3 &= \frac{1}{4}(x_2 - x_3) + \frac{3}{4}(35 - x_3),
\end{aligned}$$

with the initial set $\mathrm{X} = \{(x_1, x_2, x_3)^T \mid 1 - (x_1 - 45)^2 - (x_2 - 35)^2 - (x_3 - 35)^2 > 0\}$.

Example : Home Heating

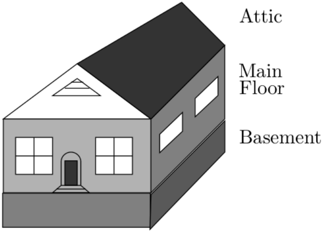
$x_3(t)$ = Temperature in the attic,
$x_2(t)$ = Temperature in the living area,
$x_1(t)$ = Temperature in the basement,
t = Time in hours.

$$\dot{x}_1 = \frac{1}{2}(45 - x_1) + \frac{1}{2}(x_2 - x_1),$$
$$\dot{x}_2 = \frac{1}{2}(x_1 - x_2) + \frac{1}{4}(35 - x_2) + \frac{1}{4}(x_3 - x_2) + 20,$$
$$\dot{x}_3 = \frac{1}{4}(x_2 - x_3) + \frac{3}{4}(35 - x_3),$$

with the initial set $\mathrm{X} = \{(x_1, x_2, x_3)^T \mid 1 - (x_1 - 45)^2 - (x_2 - 35)^2 - (x_3 - 35)^2 > 0\}$.

Is it possible for the temperature x_2 getting over than $70^\circ F$ (unsafe) ?

Example : Home Heating



$x_3(t)$ = Temperature in the attic,
$x_2(t)$ = Temperature in the living area,
$x_1(t)$ = Temperature in the basement,
t = Time in hours.

$$\dot{x_1} = \frac{1}{2}(45 - x_1) + \frac{1}{2}(x_2 - x_1),$$
$$\dot{x_2} = \frac{1}{2}(x_1 - x_2) + \frac{1}{4}(35 - x_2) + \frac{1}{4}(x_3 - x_2) + 20,$$
$$\dot{x_3} = \frac{1}{4}(x_2 - x_3) + \frac{3}{4}(35 - x_3),$$

with the initial set $\mathrm{X} = \{(x_1, x_2, x_3)^T \mid 1 - (x_1 - 45)^2 - (x_2 - 35)^2 - (x_3 - 35)^2 > 0\}$.

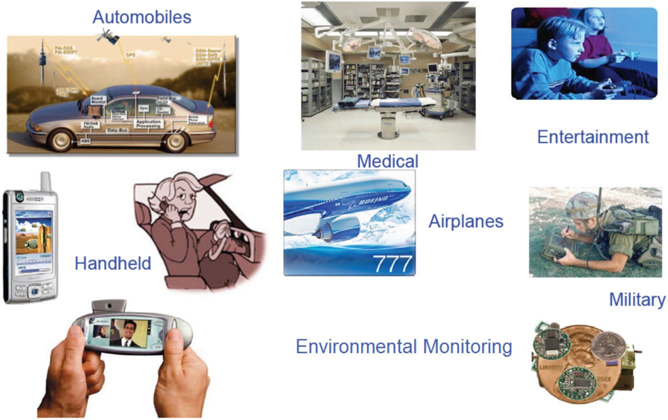
Is it possible for the temperature x_2 getting over than $70^\circ F$ (unsafe) ? **UNBOUNDED.**

Outline

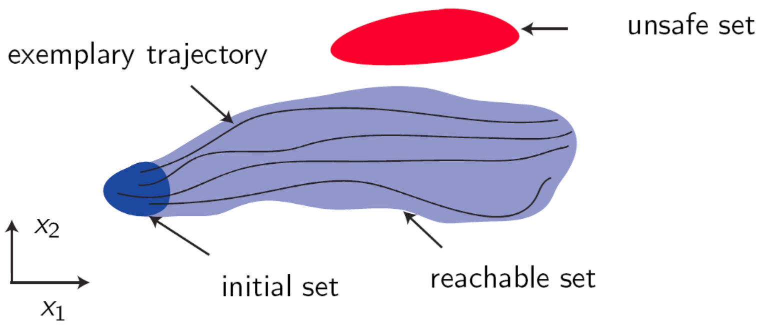
Outline

Hybrid Systems

Hybrid systems exhibit combinations of discrete jumps and continuous evolution, many of which are Safety-critical.



Safety Verification Using Reachable Set [1]



- System is safe, if no trajectory enters the unsafe set.

1. The figure is taken from [M. Althoff, 2010].

LDSs with Inputs

- Linear dymamical systems (LDSs) with inputs :

$$\dot{\xi} = A\xi + \mathbf{u}, \tag{1}$$

where $\xi(t) \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, and $\mathbf{u} : \mathbb{R} \to \mathbb{R}^n$.

LDSs with Inputs

- Linear dymamical systems (LDSs) with inputs :

$$\dot{\xi} = A\xi + \mathbf{u}, \quad (1)$$

where $\xi(t) \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, and $\mathbf{u} : \mathbb{R} \to \mathbb{R}^n$.

- **Reachability problem** (Unbounded) :

$$\mathcal{F}(\mathrm{X}, \mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \Phi(\mathbf{x}, t) = \mathbf{y}.$$

LDSs with Inputs

- Linear dymamical systems (LDSs) with inputs :

$$\dot{\xi} = A\xi + \mathbf{u}, \tag{1}$$

where $\xi(t) \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, and $\mathbf{u} : \mathbb{R} \to \mathbb{R}^n$.

- **Reachability problem** (Unbounded) :

$$\mathcal{F}(\mathrm{X}, \mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \Phi(\mathbf{x}, t) = \mathbf{y}.$$

with initial set :

$$\mathrm{X} = \{\mathbf{x} \in \mathbb{R}^n \mid p_1(\mathbf{x}) \geq 0, \cdots, p_{J_1}(\mathbf{x}) \geq 0\},$$

and unsafe set :

$$\mathrm{Y} = \{\mathbf{y} \in \mathbb{R}^n \mid p_{J_1+1}(\mathbf{y}) \geq 0, \cdots, p_J(\mathbf{y}) \geq 0\}.$$

Decidability Results of the Reachability of LDSs

In [LPY 2001], Lafferriere *et al.* proved the decidability of the reachability problems of the following three families of LDSs :

1. A is *nilpotent*, i.e. $A^n = 0$, and each component of $\mathbf{u}$ is a polynomial ;
2. A is *diagonalizable* with rational eigenvalues, and each component of $\mathbf{u}$ is of the form $\sum_{i=1}^{m} c_i e^{\lambda_i t}$, where λ_is are rationals and c_is are subject to semi-algebraic constraints ;
3. A is *diagonalizable* with purely imaginary eigenvalues, and each component of $\mathbf{u}$ of the form $\sum_{i=1}^{m} c_i \sin(\lambda_i t) + d_i \cos(\lambda_i t)$, where λ_is are rationals and c_is and d_is are subject to semi-algebraic constraints.

Main Contributions

- Generalization of case 2 and case 3 :

  2. A has real eigenvalues, and each component of $\mathbf{u}$ is of the form $\sum_{i=1}^{m} c_i e^{\lambda_i t}$, where λ_is are reals and c_is are subject to semi-algebraic constraints ; [Gan *et al.* 15]

  3. A has purely imaginary eigenvalues, and each component of $\mathbf{u}$ of the form $\sum_{i=1}^{m} c_i \sin(\lambda_i t) + d_i \cos(\lambda_i t)$, where λ_is are reals and c_is and d_is are subject to semi-algebraic constraints.

- Abstraction of general dynamical systems where A may have complex eigenvalues, by reducing the problem to the reachability in the case 2.

Outline

Tarski Algebra and Quantifier Elimination

- Tarski Algebra ($T(\mathbb{R})$)= real numbers with arithmetic and ordering.

Example

$$\varphi := \forall x \exists y : x^2 + xy + b > 0 \wedge x + ay^2 + b \leq 0$$

Tarski Algebra and Quantifier Elimination

- Tarski Algebra ($T(\mathbb{R})$)= real numbers with arithmetic and ordering.

Example

$$\varphi := \forall x \exists y : x^2 + xy + b > 0 \wedge x + ay^2 + b \leq 0$$

- Quantifier Elimination :

$$T(\mathbb{R}) \models \varphi \longleftrightarrow \varphi'$$

Example

$$T(\mathbb{R}) \models \underbrace{\forall x \exists y (x^2 + xy + b > 0 \wedge x + ay^2 + b \leq 0)}_{\varphi} \longleftrightarrow \underbrace{a < 0 \wedge b > 0}_{\varphi'}$$

LDSs with Trigonometric Function Inputs ($\mathrm{LDS}_{\mathrm{TMF}}$)

Definition (TMF)

A term is called a trigonometric function (TMF) w.r.t. t if it can be written as

$$\sum_{l=1}^{r} c_l cos(\mu_l t) + d_l sin(\mu_l t),$$

where $r \in \mathbb{N}$, $c_l, d_l, \mu_l \in \mathbb{R}$.

Definition ($\mathrm{LDS}_{\mathrm{TMF}}$)

An LDS is a linear dynamical system with trigonometric function input ($\mathrm{LDS}_{\mathrm{TMF}}$) if every component of $\mathbf{u}$ is a TMF.

Computing Reachable Set

Given an $\mathrm{LDS}_{\mathrm{TMF}}$ whose system matrix A has purely imaginary eigenvalues, the reachability can be reformulated as :

The Reachability Problem

$$\mathcal{F}(\mathrm{X},\mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \bigwedge_{i=1}^{n} y_i = \sum_{k=1}^{K_i} z_{ik}^{c}(\mathbf{x}) \cos(\gamma_{ik} t) + z_{ik}^{s}(\mathbf{x}) \sin(\gamma_{ik} t). \quad (2)$$

where $z_{ik}^{c}(\mathbf{x}), z_{ik}^{s}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ and $\gamma_{ik} \in \mathbb{R}$.

Decidability by Reduction to Tarski's Algebra

Theorem (Reduction to Tarski's Algebra)

$$\mathcal{F}(\mathrm{X},\mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge$$
$$\bigwedge_{i=1}^{n} y_i = \sum_{k=1}^{K_i} z_{ik}^c(\mathbf{x}) \cos(\gamma_{ik} t) + z_{ik}^s(\mathbf{x}) \sin(\gamma_{ik} t)$$

$$\Updownarrow$$

$$\exists \mathbf{x} \exists \mathbf{y} \exists \mathbf{u} \exists \mathbf{v} : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge \bigwedge_{j=1}^{N} u_j^2 + v_j^2 = 1 \wedge$$
$$\bigwedge_{i=1}^{n} y_i = \sum_{k=1}^{K_i} \left(\begin{array}{c} z_{ik}^c(\mathbf{x}) f_{ik}^c(u_1, v_1, \ldots, u_N, v_N) \\ + z_{ik}^s(\mathbf{x}) f_{ik}^s(u_1, v_1, \ldots, u_N, v_N) \end{array} \right),$$

where f_{ik}^c *and* f_{ik}^s *are polynomials, and* X, Y *are open sets.*

Proof.

Built on the density results given by ***Kronecker's Theorem*** in number theory.

An Example of the Reduction

Example

Given an $\mathrm{LDS}_{\mathrm{TMF}}$ as

$$\begin{pmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -3 & -2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} cos(t) \\ sin(t) \end{pmatrix},$$

with an initial point $\xi(0) = (x_1, x_2)$. The solution is

$$\Phi((x_1, x_2), t) = \begin{pmatrix} (x_1 + 2)\alpha_1 + \sqrt{2}(x_1 + x_2)\beta_1 - 2\alpha_2 - \beta_2 \\ (x_2 - 2)\alpha_1 - \sqrt{2}(\frac{3}{2}x_1 + x_2 + 1)\beta_1 + 2\alpha_2 + 2\beta_2 \end{pmatrix},$$

where $\alpha_1 = cos(\sqrt{2}t), \beta_1 = sin(\sqrt{2}t), \alpha_2 = cos(t), \beta_2 = sin(t)$.

An Example of the Reduction

- For $\mathrm{X} = \{(x_1, x_2) \mid x_1^2 + x_2^2 < 1\}$, $\mathrm{Y} = \{(y_1, y_2) \mid y_1 + y_2 > 4\}$, the reachability is equivalently reduced to

$$\begin{aligned} \mathcal{F} \quad := \quad & x_1^2 + x_2^2 < 1 \wedge \alpha_1^2 + \beta_1^2 = 1 \wedge \alpha_2^2 + \beta_2^2 = 1 \\ & \wedge (x_1 + x_2)\alpha_1 - \sqrt{2}(\frac{1}{2}x_1 + 1)\beta_1 + \beta_2 > 4. \end{aligned}$$

$\nexists\, x_1, x_2, \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ s.t. $\mathcal{F}$ holds. Thus, the system is safe.

An Example of the Reduction

- For $\mathrm{X} = \{(x_1, x_2) \mid x_1^2 + x_2^2 < 1\}$, $\mathrm{Y} = \{(y_1, y_2) \mid y_1 + y_2 > 4\}$, the reachability is equivalently reduced to

$$\begin{aligned}\mathcal{F} \quad := \quad & x_1^2 + x_2^2 < 1 \wedge \alpha_1^2 + \beta_1^2 = 1 \wedge \alpha_2^2 + \beta_2^2 = 1 \\ & \wedge (x_1 + x_2)\alpha_1 - \sqrt{2}(\frac{1}{2}x_1 + 1)\beta_1 + \beta_2 > 4.\end{aligned}$$

$\nexists\, x_1, x_2, \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ s.t. $\mathcal{F}$ holds. Thus, the system is safe.

- While if Y is replaced by $\mathrm{Y}' = \{(y_1, y_2) \mid y_1 + y_2 > 3\}$, then

$$\begin{aligned}\mathcal{F}' \quad := \quad & x_1^2 + x_2^2 < 1 \wedge \alpha_1^2 + \beta_1^2 = 1 \wedge \alpha_2^2 + \beta_2^2 = 1 \\ & \wedge (x_1 + x_2)\alpha_1 - \sqrt{2}(\frac{1}{2}x_1 + 1)\beta_1 + \beta_2 > 3.\end{aligned}$$

Let $x_1 = 0.99, x_2 = 0, \alpha_1 = \frac{\sqrt{5}}{5}, \beta_1 = -\frac{2\sqrt{5}}{5}, \alpha_2 = 0, \beta_2 = 1$, then $(x_1 + x_2)\alpha_1 - \sqrt{2}(\frac{1}{2}x_1 + 1)\beta_1 + \beta_2 \approx 3.334 > 3$, indicating that the system becomes unsafe.

Outline

Decidability of an Extension of Tarski Algebra

$\mathrm{LDS}_{\mathrm{PEF}}$ is decidable due to [Gan *et al.* 15]

$$\mathcal{F}(\mathrm{X},\mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \bigwedge_{i=1}^{n} y_i = \sum_{j=1}^{s_i} \phi_{ij}(\mathbf{x}, t) \mathrm{e}^{\nu_{ij} t}$$

where ϕ_{ij} are polynomials.

LDSs with Polynomial-exponential-trigonometric Function Inputs ($\mathrm{LDS_{PETF}}$)

Definition (PETF)

A term is called a polynomial-exponential-trigonometric function (PETF) w.r.t. t if it can be written as

$$\sum_{k=0}^{r} p_k(t) \mathrm{e}^{\alpha_k t} \cos(\beta_k t + \gamma_k),$$

where $r \in \mathbb{N}, \alpha_k, \beta_k, \gamma_k \in \mathbb{R}$, and $p_k(t) \in \mathbb{R}[t]$.

Definition ($\mathrm{LDS_{PETF}}$)

An LDS is a linear dynamical system with polynomial-exponential-trigonometric function input ($\mathrm{LDS_{PETF}}$) if every component of $\mathbf{u}$ is a PETF.

Computing Reachable Set

Given an $\mathrm{LDS}_{\mathrm{PETF}}$ with the system matrix with complex eigenvalues, the reachability can be reformulated, due to Jordan decomposition, as :

The Reachability Problem

$$\mathcal{F}(\mathrm{X}, \mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \bigwedge_{k=1}^{n} y_k = \sum_{\gamma \in \Gamma} g_{\gamma,k}(\mathbf{x}, t) \cos(\gamma t) + h_{\gamma,k}(\mathbf{x}, t) \sin(\gamma t). \quad (3)$$

where $g_{\gamma,k}$ and $h_{\gamma,k}$ are linear on $\mathbf{x}$, and are polynomial-exponential functions w.r.t. t.

Abstraction by Eliminating trigonometric functions

Theorem (Overapproximation of the Reachable Set)

$$\mathcal{F}(\mathrm{X}, \mathrm{Y}) := \exists \mathbf{x} \exists \mathbf{y} \exists t : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge$$

$$\bigwedge_{k=1}^{n} y_k = \sum_{\gamma \in \Gamma} g_{\gamma,k}(\mathbf{x}, t) \cos(\gamma t) + h_{\gamma,k}(\mathbf{x}, t) \sin(\gamma t)$$

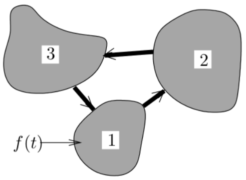
$$\Downarrow$$

$$\exists \mathbf{x} \exists \mathbf{y} \exists u_\gamma \exists v_\gamma : \mathbf{x} \in \mathrm{X} \wedge \mathbf{y} \in \mathrm{Y} \wedge t \geq 0 \wedge \bigwedge_{\gamma} u_\gamma^2 + v_\gamma^2 = 1 \wedge$$

$$\bigwedge_{k=1}^{n} y_k = \sum_{\gamma} g_{\gamma,k}(\mathbf{x}, t) u_\gamma + h_{\gamma,k}(\mathbf{x}, t) v_\gamma .$$

Illustrating Examples

Example (Pond Pollution)



$x_1(t)$ = Amount of pollutant in pond 1,
$x_2(t)$ = Amount of pollutant in pond 2,
$x_3(t)$ = Amount of pollutant in pond 3,
t = Time in minutes.

$$\dot{x}_1(t) = 0.001x_3(t) - 0.001x_1(t) + 0.01,$$
$$\dot{x}_2(t) = 0.001x_1(t) - 0.001x_2(t),$$
$$\dot{x}_3(t) = 0.001x_2(t) - 0.001x_3(t),$$

with the initial set $\mathrm{X} = \{(x_1, x_2, x_3)^T \mid (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 < 1\}$ and the unsafe set $\mathrm{Y} = \{(y_1, y_2, y_3)^T \mid y_2 - y_3 + 6 < 0\}$.

Illustrating Examples

1. $X \cap Y = \emptyset$.

Illustrating Examples

1. $\mathrm{X} \cap \mathrm{Y} = \emptyset$.

2. Note that the system matrix is diagonalizable with complex eigenvalues 0, $(-3 - \mathbf{i}\sqrt{3})/2000$, and $(-3 + \mathbf{i}\sqrt{3})/2000$. By using the solution of this system, the reachability thus becomes

$$\mathcal{F} := \exists x_1 \exists x_2 \exists x_3 \exists t : t > 0 \wedge (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 - 1 < 0$$
$$\wedge\ a + b \cos\left(\frac{\sqrt{3}t}{2000}\right) + c \sin\left(\frac{\sqrt{3}t}{2000}\right) < 0,$$

with $a = 28\mathrm{e}^{3t/2000}$, $b = 3x_2 - 3x_3 - 10$, and $c = \sqrt{3}\,(2x_1 - x_2 - x_3 - 10)$.

Illustrating Examples

3 Reduction to Tarski's algebra by abstracting the second constraint as

$$a + bu + cv < 0 \wedge u^2 + v^2 = 1.$$

Illustrating Examples

3. Reduction to Tarski's algebra by abstracting the second constraint as

$$a + bu + cv < 0 \wedge u^2 + v^2 = 1.$$

4. The reduced reachability problem is then verified as safe in *LinR*.

Illustrating Examples

3. Reduction to Tarski's algebra by abstracting the second constraint as

$$a + bu + cv < 0 \wedge u^2 + v^2 = 1.$$

4. The reduced reachability problem is then verified as safe in *LinR*.

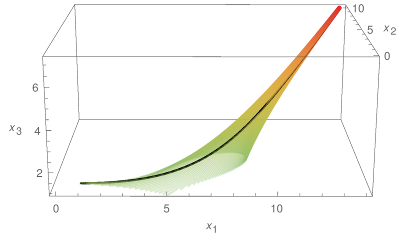
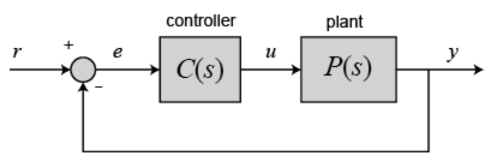


Figure : Overapproximation (the tube) of one single trajectory (the curve) starting from $(1, 1, 1)^T$ initially

Illustrating Examples

Example (PI Controller)

Consider a proportional-integral (PI) controller which is used to control a plant.



$$\underbrace{M\ddot{x} + b\dot{x} + kx}_{plant} = \underbrace{K_d(r \dot{-} x) + K_p(r - x) + K_i \int (r - x)}_{controller}$$

Safety property :

$$\mathbf{G}(t > 0.5 \Rightarrow x \geq 0.9 \wedge x \leq 1.1).$$

Proving of this case was failed in [Tiwari *et al.* 13].

Illustrating Examples

- Let $\mathbf{x} = [\int x, x, \dot{x}, t]^{\mathrm{T}}$, then $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{u}$, where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -300 & -370 & -10 & 300 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and $\mathbf{u} = [0, 0, 350, 1]^{\mathrm{T}}$. The initial value is $\mathbf{x}(0) = [0, 0, 0, 0]$ and unsafe set is $Y = \{\mathbf{x} \mid t > 0.5 \wedge (x < 0.9 \vee x > 1.1)\}$.

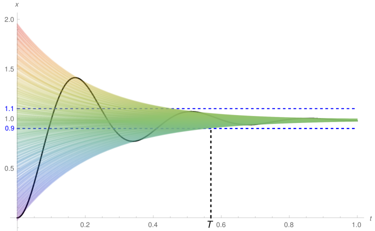


Figure : Overapproximation (the "broom") of the trajectory of x (the curve) starting from 0

Outline

Concluding Remarks

- The decidability of the reachability problem of $\mathrm{LDS}_{\mathrm{TMF}}$ by reduction to the decidability of Tarski's Algebra.
- A more precise abstraction that overapproximates the reachable sets of general linear dynamical systems ($\mathrm{LDS}_{\mathrm{PETF}}$).
- On-going work : extension of the results to solvable dynamical systems.
- Question : is the abstraction complete (δ-decidable) for unbounded verification ?