A Stability Proof for the Inverted Pendulum

Damien Rouhling

Université Côte d'Azur, Inria, France

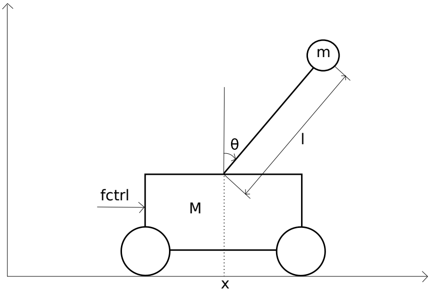
January 8, 2018

Motivations

- Safety is critical in many applications of robotics.
- We focus here on control theory: a program, or control function, operates a robot in order to achieve a goal.
- We want to bring formal guarantees on this control function: the goal is achieved, no safety condition is violated.

The inverted pendulum

- The inverted pendulum is a standard example for testing control techniques.



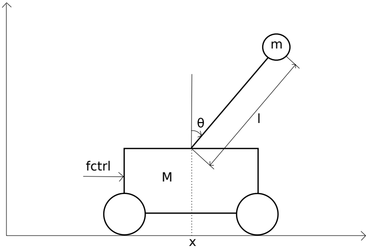
- Goal: stabilize the pendulum on its unstable equilibrium thanks to the control function fctrl.

Context

- Control function and stability proof from [Lozano et al., 2000].
- Proof based on LaSalle's invariance principle [LaSalle, 1960], generalized and formalized in [Cohen and Rouhling, 2017].
- Principle: qualitative analysis of the solutions of a first-order autonomous differential equation:

$$\dot{y} = \mathrm{F} \circ y.$$

Homoclinic orbit



- Lozano et al. prove the convergence of solutions to a homoclinic orbit:

$$\frac{1}{2}ml^2\dot{\theta}^2 = mgl\,(1 - \cos\theta)\,.$$

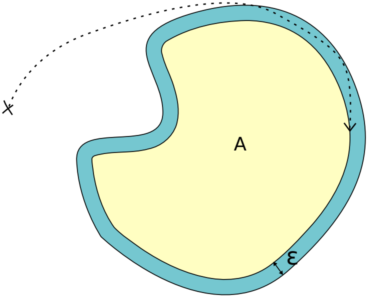
- This is done by an energy approach: the homoclinic orbit is characterised by $E = 0$ and $\dot{x} = 0$.
- They also want the cart to stop at its initial position:

$$x = 0 \text{ and } \dot{x} = 0.$$

Preliminary definitions

- A function of time $y(t)$ converges to a set A as t goes to infinity, denoted by $y(t) \to A$ as $t \to +\infty$, if

$$\forall \varepsilon > 0, \exists T > 0, \forall t > T, \exists p \in A, \|y(t) - p\| < \varepsilon.$$



Preliminary definitions

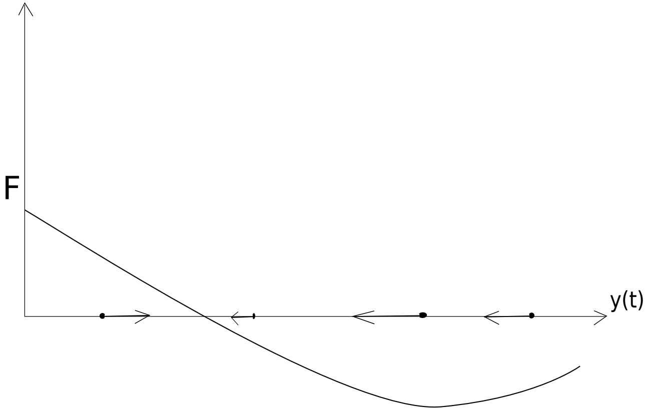
- A function of time $y(t)$ converges to a set A as t goes to infinity, denoted by $y(t) \to A$ as $t \to +\infty$, if

$$\forall \varepsilon > 0, \exists T > 0, \forall t > T, \exists p \in A, \|y(t) - p\| < \varepsilon.$$

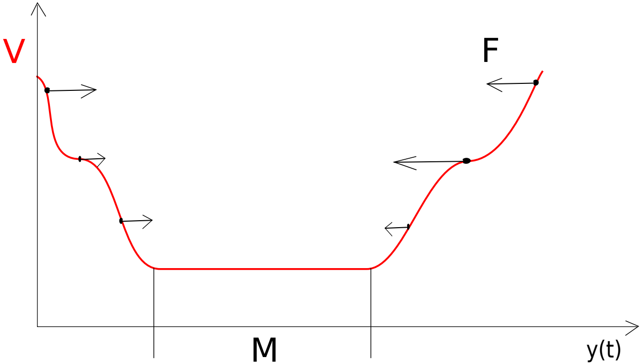
- A set A is said to be invariant if every solution of $\dot{y} = \mathrm{F} \circ y$ starting in A (i.e. $y(0) \in A$) remains in A.
- The positive limiting set of a function of time $y(t)$, denoted by $\Gamma^+(y)$, is the set of limit points of y:

$$\Gamma^+(y) = \{p \mid \forall \varepsilon > 0, \forall T > 0, \exists t > T, \|y(t) - p\| < \varepsilon\}.$$

LaSalle's invariance principle for real functions



LaSalle's invariance principle for real functions



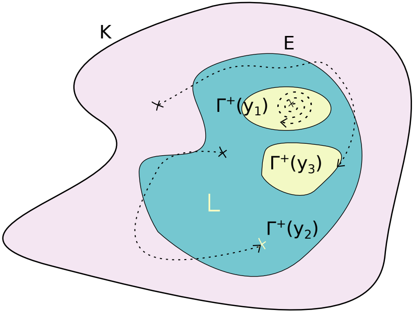
LaSalle's invariance principle improved

Assume

- F is such that we have the existence and uniqueness of solutions to $\dot{y} = \mathrm{F} \circ y$ and the continuity of solutions relative to initial conditions in K
- K compact and invariant
- $V : \mathbb{R}^n \to \mathbb{R}$ is differentiable in K
- $\tilde{V}(p) \leqslant 0$ in K where $\tilde{V}(p) := (dV_p \circ F)(p)$

Then, for $L := \bigcup_{\substack{y \text{ solution} \\ \text{starting in } K}} \Gamma^+(y)$

and $E := \left\{ p \in K \mid \tilde{V}(p) = 0 \right\}$, L is an invariant subset of E and for all solution y starting in K, $y(t) \to L$ as $t \to +\infty$.

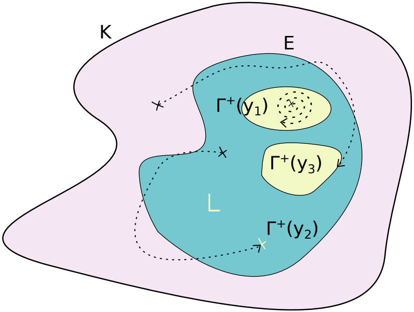


LaSalle's invariance principle improved

Assume

- F is such that we have the existence and uniqueness of solutions to $\dot{y} = \mathrm{F} \circ y$ and the continuity of solutions relative to initial conditions in K
- K compact and invariant
- $V : \mathbb{R}^n \to \mathbb{R}$ is continuous in K and differentiable along trajectories of solutions starting in K
- $\tilde{V}(p) \leqslant 0$ in K where $\tilde{V}(p)$ is the directional derivative of V at point p along the trajectory of the solution starting at p

Then, for $L := \bigcup_{\substack{y \text{ solution} \\ \text{starting in } K}} \Gamma^+(y)$ and $E := \left\{ p \in K \mid \tilde{V}(p) = 0 \right\}$, L is an invariant subset of E and for all solution y starting in K, $y(t) \to L$ as $t \to +\infty$.



LaSalle's invariance principle for the inverted pendulum

- The Lyapunov function V is minimised along trajectories. Our goal is $E = 0$, $x = 0$ and $\dot{x} = 0$.

$$V = \frac{k_E}{2} E^2 + \frac{k_v}{2} \dot{x}^2 + \frac{k_x}{2} x^2$$

- The laws of Physics give a second-order differential equation. We transform the equation on (x, θ) into a first-order equation on

$$p = (p_0, p_1, p_2, p_3, p_4) = \left(x, \dot{x}, \cos\theta, \sin\theta, \dot{\theta} \right).$$

- We lose pieces of information. The invariant compact set K will help keeping them as invariants.

$$K = \left\{ p \in \mathbb{R}^5 \mid p_2^2 + p_3^2 = 1 \text{ and } V(p) \leqslant k_0 \right\}.$$

Formalization

- Formalization in Coq + SSReflect.
- Libraries: Mathematical Components and Coquelicot [Boldo et al., 2015].
- Around 2500 lines of code on top of our formalization of LaSalle's invariance principle:
  - 1000 lines for the definition of the system and the stability proof.
  - 800 lines for topological results.
  - 500 lines for the interface between Mathematical Components and Coquelicot.
  - 100 lines for automatic differentiation.

Differential equations

Differential equation $\dot{y} = \mathrm{F} \circ y$.

- First try:

```
Definition is_sol (y : R -> U) :=
  forall t, is_derive y t (F (y t)).
```

  Issue: in Physics we do not consider negative times.

Differential equations

Differential equation $\dot{y} = \mathrm{F} \circ y$.

- First try:

```
Definition is_sol (y : R -> U) :=
  forall t, is_derive y t (F (y t)).
```

  Issue: in Physics we do not consider negative times.

- Second try:

```
Definition is_sol (y : R -> U) :=
  forall t, 0 <= t -> is_derive y t (F (y t)).
```

  Issue: incompatible with our formal definition of the existence and uniqueness of solutions.

Existence and uniqueness of solutions

- We represent all solutions by a single function.

```
Variable sol : U -> R -> U.
```

- Its first argument is the initial condition.

```
Hypothesis sol0 : forall p, sol p 0 = p.
```

- Existence and uniqueness are expressed with one hypothesis.

```
Hypothesis solP :
  forall y, K (y 0) -> is_sol y <-> y = sol (y 0).
```

Existence and uniqueness of solutions

- We represent all solutions by a single function.

```
Variable sol : U -> R -> U.
```

- Its first argument is the initial condition.

```
Hypothesis sol0 : forall p, sol p 0 = p.
```

- Existence and uniqueness are expressed with one hypothesis.

```
Hypothesis solP :
  forall y, K (y 0) -> is_sol y <-> y = sol (y 0).
```

- However, it is not satisfiable with the following definition of solution.

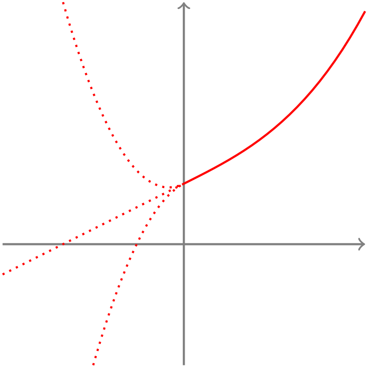
```
Definition is_sol (y : R -> U) :=
  forall t, 0 <= t -> is_derive y t (F (y t)).
```

Differential equations (cont.)

Differential equation $\dot{y} = \mathrm{F} \circ y$.

- Second try:

```
Definition is_sol (y : R -> U) :=
  forall t, 0 <= t -> is_derive y t (F (y t)).
```

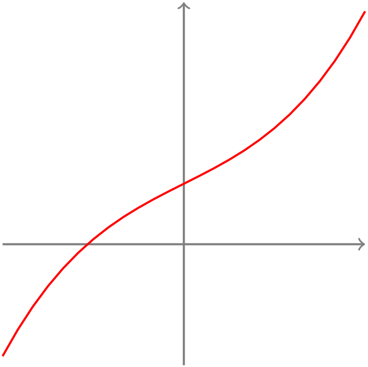


Differential equations (cont.)

Differential equation $\dot{y} = \mathrm{F} \circ y$.

- Last try:

```
Definition is_sol (y : R -> U) :=
  (forall t, 0 <= t -> is_derive y t (F (y t))) /\
  forall t, t < 0 -> y t = 2 (y 0) - (y (- t)).
```



Discussion

```
Definition is_sol (y : R -> U) :=
  (forall t, t < 0 -> y t = 2 (y 0) - (y (- t))) /\
  forall t, 0 <= t -> is_derive y t (F (y t)).
Hypothesis solP :
  forall y, K (y 0) -> is_sol y <-> y = sol (y 0).
```

Pros:

- Solutions are differentiable at any time, without considering restrictions.
- Remove all hypotheses `is_sol y`.
- Closer to SSReflect and pen and paper proof-styles.

Cons:

- Very specific notion of solution.
- Shifted solutions `fun t => y (t + s)` are not solutions anymore.

Computing differentials and derivatives

- How to prove that the derivative at time t of $V \circ \mathrm{sol}_p$ is $-k_d \left(\mathrm{sol}_p^2(t)\right)_1$? In Coq:

```
is_derive (V \o (sol p)) t (- kd * ((sol p t)[1] ^ 2)).
```

- In Coquelicot, as soon as the `auto_derive` tactic is unusable:
  1. Use the `evar_last` tactic to get two goals.

     ```
     is_derive (V \o (sol p)) t ?d,
     ? d = - kd * ((sol p t)[1] ^ 2).
     ```

  2. Use the rules of differentiation in order to instantiate `?d` and close the goal

     ```
     is_derive (V \o (sol p)) t ?d.
     ```

  3. Prove `?d = - kd * ((sol p t)[1] ^ 2)`, using this instantiation.

Automatic differentiation

- In Coquelicot:
  1. Use the `evar_last` tactic to introduce an existential variable for the differential.
  2. Use the rules of differentiation in order to instantiate the existential variable.
  3. Prove that the instantiation is equal to the differential we expected.
- Automation:
  - Step 3 is fairly automated using the `ring` and `field` tactics.
  - Our contribution is an alternative to step 1 which automates step 2.
  - Principle: store a database of differentials/rules of differentiation thanks to type classes.

Type classes for automatic differentiation

- We encapsulate Coquelicot's `filterdiff` predicate in a type class:

```
Class diff (f : U -> V) (F : set (set U))
  (df : U -> V) := diff_prf : filterdiff f F df.
```

- We trigger type class inference through the following lemma.

```
Lemma diff_eq (f f' df : U -> V) (F : set (set U)) :
  diff f F f' -> f' = df -> diff f F df.
```

- We provide the same mechanism for derivatives.

```
Class deriv (f : K -> V) (x : K) (l : V) :=
  deriv_prf : is_derive f x l.
```

Differentiation rules database

- All the hard work (i.e. the proofs) is already done in COQUELICOT.
- We only have to transform COQUELICOT's *lemmas* in type class *instances*.
- For example, the rules for constant functions and for the sum of two functions:

```
Instance diff_const (p : V) (F : set (set U)) :
  Filter F ->
  diff (fun _ => p) F (fun _ => zero).
Instance diff_plus (f g df dg : U -> V)
  (F : set (set U)) :
  Filter F -> diff f F df -> diff g F dg ->
  diff (fun p => plus (f p) (g p)) F
    (fun p => plus (df p) (dg p)).
```

Conclusion

- Our formalization of LaSalle's invariance principle improved.
- The inverted pendulum formalized.
- A quite convenient way of dealing with solutions.
- Automatic derivation/differentiation via type classes.

Bibliography

Boldo, S., Lelay, C., and Melquiond, G. (2015).
Coquelicot: A User-Friendly Library of Real Analysis for Coq.
Mathematics in Computer Science, 9(1):41–62.

Cohen, C. and Rouhling, D. (2017).
A Formal Proof in Coq of LaSalle's Invariance Principle.
In Ayala-Rincón, M. and Muñoz, C. A., editors, *Interactive Theorem Proving - 8th International Conference, ITP 2017, Brasília, Brazil, September 26-29, 2017, Proceedings*, volume 10499 of *Lecture Notes in Computer Science*, pages 148–163. Springer.

LaSalle, J. (1960).
Some Extensions of Liapunov's Second Method.
IRE Transactions on Circuit Theory, 7(4):520–527.

Lozano, R., Fantoni, I., and Block, D. (2000).
Stabilization of the inverted pendulum around its homoclinic orbit.
Systems & Control Letters, 40(3):197–204.