A geometric preconditioner for indirect optimal control method and application to hybrid electric vehicle.

Rémy Dutto [1,2,3]

[1]IRIT: Institut de Recherche en Informatique de Toulouse

[2]IMT: Institut de Mathématiques de Toulouse

[3]Vitesco Technologies

FGS Conference on Optimization, Gijón, Spain.

Introduction

In collaboration with:

- Olivier Cots, IRIT, Toulouse,
- Olivier Flebus, Vitesco Technologies, Toulouse,
- Sophie Jan, IMT, Toulouse,
- Serge Laporte, IMT, Toulouse,
- Mariano Sans, Vitesco Technologies, Toulouse.

Table of Contents

Cycle

We consider a Hybrid Electric Vehicle (HEV) on a predefined cycle, i.e. speed and slope trajectories are prescribed.

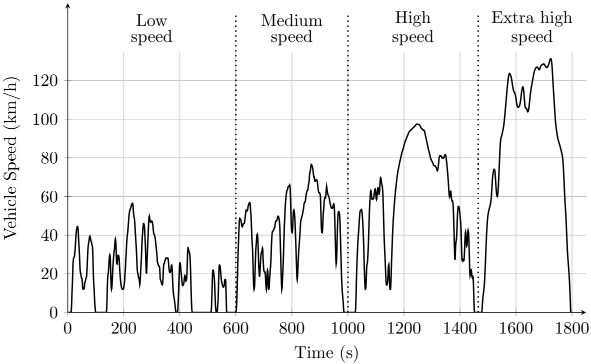


Figure: Worldwide harmonized Light vehicles Test Cycle (WLTC).

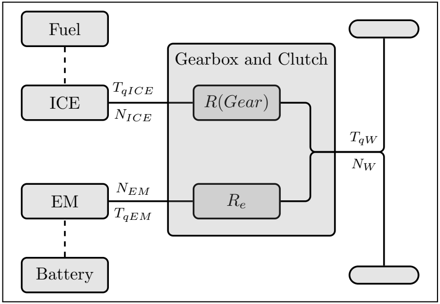
Requested wheels torque $T_{qW}(t)$ and rotation speed $N_W(t)$ are obtained with the information of our vehicle (mass, wheel diameter, aerodynamic coefficient. . .).

Static model of HEV

Inputs of our static model:

Name	Description	Unit
Cost		
m_{Fuel}	Fuel consumption	g
State		
SOC	Battery state of charge	
Commands		
$Gear$	Gearbox selector	
T_{qICE}	ICE torque	N.m
External inputs		
T_{qW}	Wheels torque	N.m
N_W	Wheels rotation speed	RPM

Figure: Schema of the HEV.



Outputs: $\dot{m}_{Fuel}$ and $\dot{SOC}$, where $\dot{}$ stands for $\dfrac{\mathrm{d}}{\mathrm{d}t}$.

HEV torque split and gear shift problem

The HEV torque split and gear shift problem can be formulated as a classical Lagrange optimal control problem

$$
\text{(OCP)} \quad \begin{cases} V(x_0, x_T) = \displaystyle\min_{x,u} \int_{t_0}^{t_f} f^0\big(t, x(t), u(t)\big)\, \mathrm{d}t, & \\ \text{s.t. } \dot{x}(t) = f\big(t, x(t), u(t)\big), & t \in [t_0, t_f] \text{ a.e.}, \\ \qquad u(t) \in \mathrm{U}(t), & t \in [t_0, t_f], \\ \qquad x(t_0) = x_0, \quad x(t_f) = x_T, & \end{cases}
$$

where:

- $x \in \mathrm{AC}([t_0, t_f], \mathbb{R})$ corresponds to the *SOC*,
- $u \in \mathrm{L}^\infty([t_0, t_f], \mathbb{R}^2)$ corresponds to the pair $\big(T_{qICE}, \mathit{Gear}\big)$,
- functions f^0 and f are $\mathcal{C}^1$ w.r.t. x and u,
- $\mathrm{U}(t) \subset \mathbb{R}^2$ is a nonempty closed set for every $t \in [t_0, t_f]$, with regularity assumptions.[1]

[1](cf. [Cesari, 1983, Chapter 4.2, Remark 5] for more information)

Assumptions from previous work

Motivated by [Cots et al., 2023a],

- we only consider the first 100s of the cycle ($[t_0, t_f] = [0, 100]$),
- we have constructed a database $\mathbb{D}$ of the value function V evaluations by an efficient method[2],
- we have trained a neural network C on $\mathbb{D}$ to approximate V.

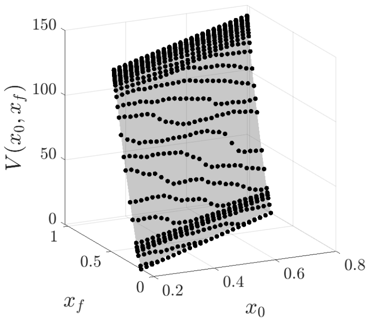


Figure: The points correspond to $\mathbb{D}$ and the surface to the trained neural network C.

[2] cf. [Cots et al., 2023b] for more information

Augmented system

We propose to consider the augmented formulation of (OCP)

$$
\text{(AOCP)} \qquad \begin{cases} \min\limits_{\hat{x},u} x^0(t_f) \\ \text{s.t. } \dot{\hat{x}}(t) = \hat{f}\big(t, \hat{x}(t), u(t)\big) & t \in [t_0, t_f] \text{ a.e.}, \\ \qquad u(t) \in \mathrm{U}(t) & t \in [t_0, t_f], \\ \qquad \hat{x}(t_0) = \hat{x}_0, \quad x(t_f) = x_f, \end{cases}
$$

where $\hat{f} \colon \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ is the augmented system

$$
\hat{f}(t, \hat{x}, u) = \big(f^0(t, x, u), f(t, x, u)\big)
$$

and where $\hat{x} = (x^0, x)$ corresponds to the cost-state pair, with $\hat{x}_0 = (0, x_0)$.

Pontryagin's Maximum Principle

If $(\hat{x}, u)$ is solution of (AOCP), there exists a non trivial augmented costate $\hat{p} = (p^0, p) \in \mathrm{AC}([t_0, t_f], \mathbb{R}^2) \neq 0$ with $p^0 \leq 0$ such that the Hamilton's dynamic is satisfied for almost every $t \in [t_0, t_f]$

$$\begin{cases} \dot{\hat{x}}(t) = \quad \nabla_{\hat{p}} h\big(t, \hat{x}(t), \hat{p}(t), u(t)\big), \\ \dot{\hat{p}}(t) = -\nabla_{\hat{x}} h\big(t, \hat{x}(t), \hat{p}(t), u(t)\big), \end{cases} \tag{1}$$

as well as the maximization condition for almost every $t \in [t_0, t_f]$

$$h\big(t, \hat{x}(t), \hat{p}(t), u(t)\big) = \max_{w \in \mathrm{U}(t)} h\big(t, \hat{x}(t), \hat{p}(t), w\big), \tag{2}$$

where $h(t, \hat{x}, \hat{p}, u) = \big(\hat{p} \,\big|\, \hat{f}(t, \hat{x}, u)\big)$ is the pseudo-Hamiltonian of the augmented system.

Notations

For the following presentation, we denote

$$\hat{x} = (x^0, x) \quad \text{and} \quad \hat{p} = (p^0, p).$$

Moreover, we denote

$$\hat{z} = (\hat{x}, \hat{p}) \quad \text{and} \quad z = (x, p).$$

These notations can be used for absolutely continuous functions or for vectors.

Remark

Since $\hat{f}$ does not depend on the cost x^0, we obtain $\dot{p}^0(\cdot) = 0$ and thus $p^0(\cdot)$ is constant.

Definitions - Extremals

An extremal is a couple $(\hat{z}, u) \in \mathrm{AC}([t_0, t_f], \mathbb{R}^4) \times \mathrm{L}^\infty([t_0, t_f], \mathbb{R}^2)$ which satisfies the Hamilton's dynamic (1) and the maximization condition (2).

A BC-extremal is an extremal which satisfies the boundary conditions given by $\hat{x}(t_0) = \hat{x}_0$ and $x(t_f) = x_T$.

An extremal is said normal$^-$ if $p^0 < 0$, normal$^+$ if $p^0 > 0$ and abnormal if $p^0 = 0$.

Framework

Let us denote $\exp_{\vec{h}}(\hat{z}_0)$ a solution at time t_f of

$$\begin{cases} \dot{\hat{z}}(t) = \vec{h}\left(t, \hat{z}(t), u(t)\right), & t \in [t_0, t_f] \text{ a.e.} \\ h\left(t, \hat{z}(t), u(t)\right) = \max_{w \in \mathrm{U}(t)} h\left(t, \hat{z}(t), w\right), & t \in [t_0, t_f] \text{ a.e.} \\ \hat{z}(t_0) = \hat{z}_0, \end{cases}$$

where $\vec{h}$ is the pseudo-Hamiltonian vector field of the augmented system, defined by

$$\vec{h}(t, \hat{x}, \hat{p}, u) = \left(\nabla_{\hat{x}} h(t, \hat{x}, \hat{p}, u), -\nabla_{\hat{p}} h(t, \hat{x}, \hat{p}, u)\right).$$

Hypothesis 1

The possibly multivalued function $\exp_{\vec{h}}(\hat{x}_0, \hat{p}_0)$ is an application, defined for all $\hat{x}_0 \in \mathbb{R}^2$ and for all non trivial $\hat{p}_0 \in \mathbb{R}^2$.

Simple shooting method

Under the previous hypothesis, the maximum principle leads to the resolution of

$$\text{(TPBVP)} \qquad \begin{cases} \pi_x\big(\exp_{\vec{h}}(\hat{z}_0)\big) = x_T, \\ \pi_{\hat{x}}(\hat{z}_0) = \hat{x}_0, \quad \pi_{p^0}(\hat{z}_0) \leq 0, \end{cases}$$

where $\pi_x(\cdot)$ is the classical x-space projection.

The simple shooting method aims to find a non-trivial zero $\hat{p}_0 = (p^0, p_0)$ of the shooting function

$$\begin{array}{rcccc} S & : & \mathbb{R}^- \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & \hat{p}_0 & \longmapsto & \pi_x\big(\exp_{\vec{h}}(\hat{x}_0, \hat{p}_0)\big) - x_T \end{array}$$

Normalization of the shooting function

Let us remark that if $\hat{p}_0 \neq 0$ satisfies $S(\hat{p}_0) = 0$ then for all $k > 0$, $S(k\hat{p}_0) = 0$ (due to homogeneity of BC-extremals on $\hat{p}$).

We propose two normalizations of the shooting function S.

- Method 1: if we assume that the extremals associated to a solution are normal$^-$ ($p^0 < 0$), then we can fix $p^0 = -1$ and consider $S_1 \colon \mathbb{R} \to \mathbb{R}$ defined by

$$S_1(p_0) = S(-1, p_0),$$

- Method 2: without the above assumption, we can fix $\|\hat{p}_0\|_2 = 1$ and consider $S_2 \colon [-1, 1] \to \mathbb{R}$ defined by

$$S_2(p_0) = S\big(\eta(p_0), p_0\big), \quad \text{where} \quad \eta(p_0) = -\sqrt{1 - p_0^2}.$$

A solution of the shooting method is found by a Newton-like solver.

Thanks to [Cots et al., 2023a], a natural initial guess for S_1 is given by

$$p_* = -\nabla_{x_0} C(x_0, x_f).$$

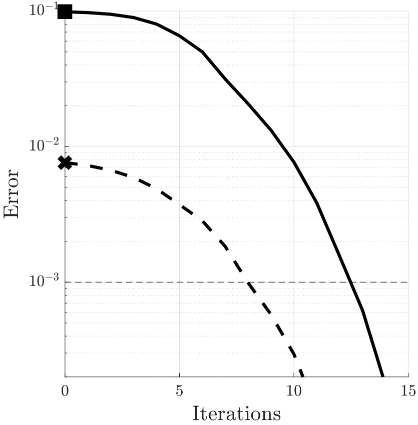


Figure: Evolution of the error $|S_1(\cdot) - x_f|$ w.r.t the number of iterations (with 100 different initial and final states).
——: fixed initialization $p = 500$ (■) - - -: natural initialization p_* (✖)
---: industrial tolerance 10^{-3}

Goal

<u>**Goal**</u>: Reducing the number of iterations of the solver

<u>**Main idea**</u>[3]: Preconditioning method of the shooting function based on

- a geometric interpretation of the costate
- and the Mathieu transformation.

[3] cf. [Cots et al., 2024] for more information

Geometric interpretation of the costate

The proof of the maximum principle is constructive.

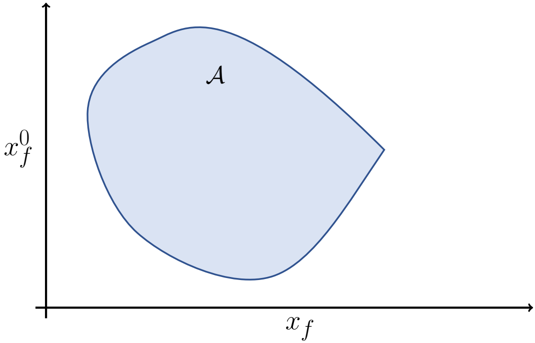


Figure: Illustration of the accessible augmented state set $\mathcal{A}$, which is the set of reachable augmented states $\hat{x}_f = (x_f^0, x_f)$ at t_f from $\hat{x}_0$ at t_0.

Geometric interpretation of the costate

The proof of the maximum principle is constructive. The final augmented costate $\hat{p}_f = \left(p^0, p(t_f)\right)$ is taken in the polar of the proper convex Boltyanskii cone $\mathcal{K}^\circ$.

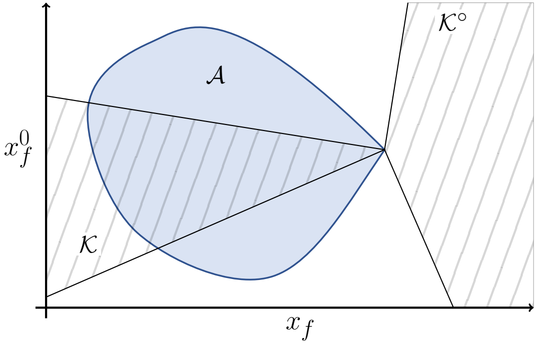


Figure: Illustration of the Botlyanskii cone $\mathcal{K}$ and its polar $\mathcal{K}^\circ$ at an augmented final state $\hat{x}_f \in \partial\mathcal{A}$.

Geometric interpretation of the costate

The proof of the maximum principle is constructive. The final augmented costate $\hat{p}_f = \left(p^0, p(t_f)\right)$ is taken in the polar of the proper convex Boltyanskii cone $\mathcal{K}^\circ$.

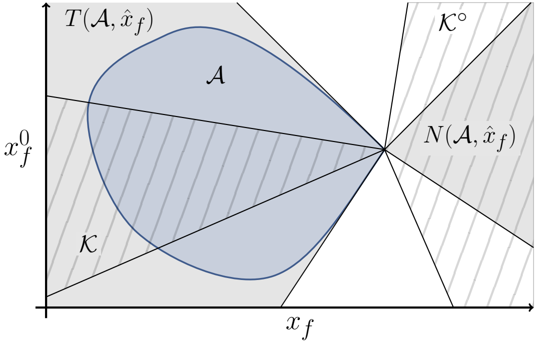
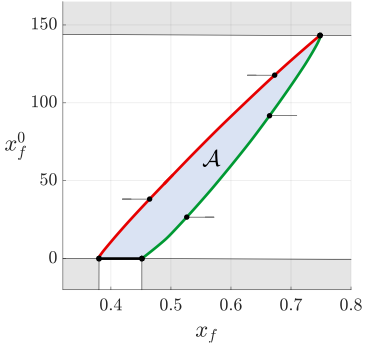
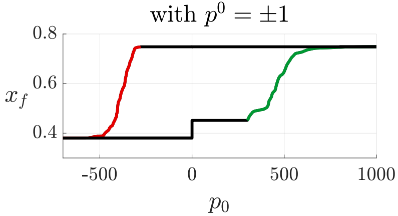


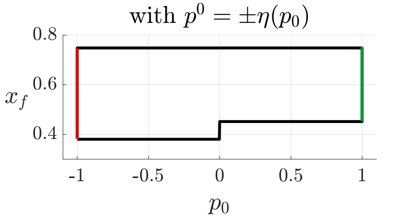
Figure: Illustration of the link between $\mathcal{K}^\circ$ and the normal cone $N\left(\mathcal{A}, \hat{x}_f\right)$ of the set $\mathcal{A}$ at the point $\hat{x}_f$.

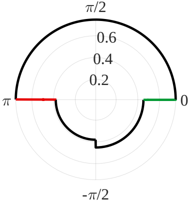
If $\mathcal{A}$ is closed and convex, we can take $\hat{p}(t_f) \in N\left(\mathcal{A}, \hat{x}_f\right)$.

Accessible augmented set and shooting functions

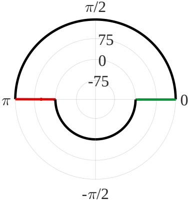








State x_f



Cost x_f^0

Figure: With $x_0 = 0.5$.

Mathieu transformation

A diffeomorphism $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ on the augmented state is lifted into a diffeomorphism $\Phi\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^2$ on the augmented state-costate that preserves the Hamiltonian dynamics

$$\Phi(\hat{x}, \hat{p}) = \left(\phi(\hat{x}), J_\phi(\hat{x})^{-\top} \hat{p}\right),$$

which is called Mathieu transformation.

This diffeomorphism transforms $\hat{z} = (\hat{x}, \hat{p})$ into $\hat{w} = (\hat{y}, \hat{q})$:

$$\hat{z} = \begin{pmatrix} \hat{x} \\ \hat{p} \end{pmatrix} \underset{\Phi^{-1}}{\overset{\Phi}{\rightleftarrows}} \begin{pmatrix} \hat{y} \\ \hat{q} \end{pmatrix} = \hat{w}.$$

Moreover, we denote $\hat{y} = (y^0, y)$ and $\hat{q} = (q^0, q)$.

Construction of the transformation

Main idea: fitting an ellipse on $\partial \mathcal{A}$ and creating the linear diffeomorphism $\phi(\hat{x}) = A\hat{x} + b$ that transforms this ellipse into the unit circle.

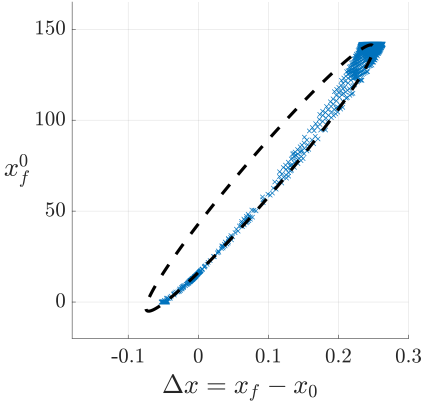


Figure: Original coordinates

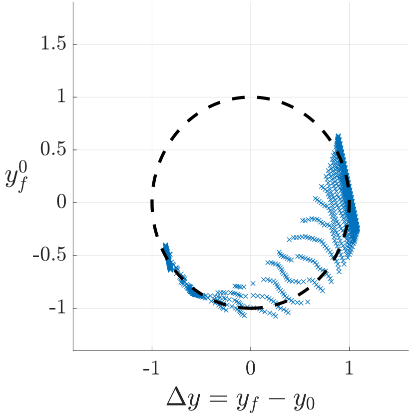


Figure: New coordinates

Construction of the transformation

Main idea: fitting an ellipse on $\partial\mathcal{A}$ and creating the linear diffeomorphism $\phi(\hat{x}) = A\hat{x} + b$ that transforms this ellipse into the unit circle.

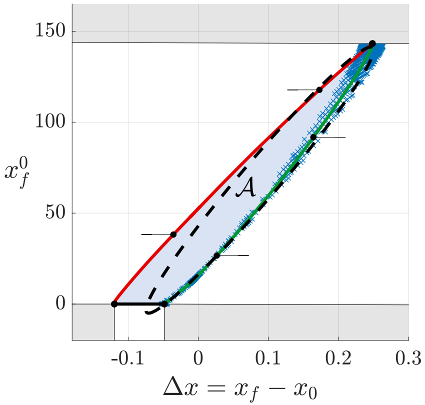


Figure: Original coordinates

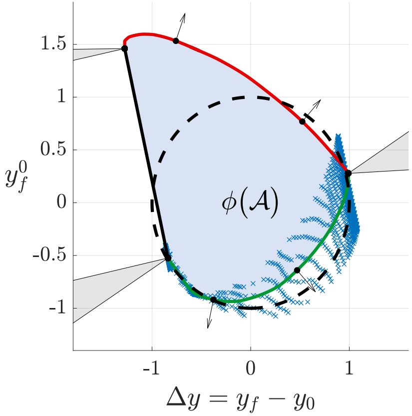
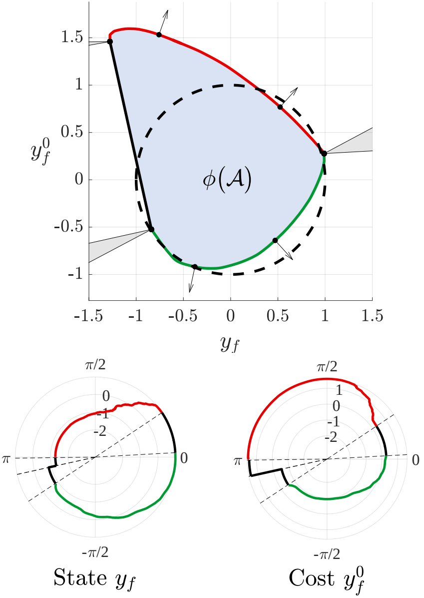


Figure: New coordinates

Geometric interpretation of the costate



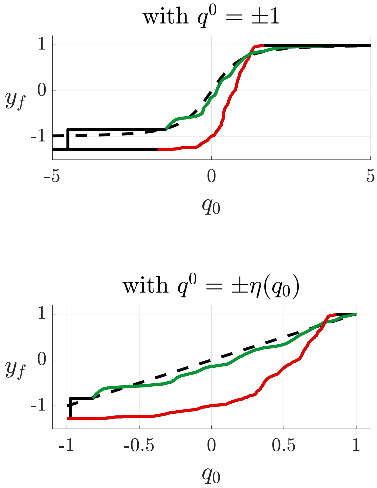


Figure: With $x_0 = 0.5$.

Definition of the shooting functions

In the new coordinates, the shooting function $T\colon \mathbb{R}^- \times \mathbb{R} \to \mathbb{R}$ is given by

$$T(\hat{q}_0) = \pi_y\big(\hat{y}_f(\hat{q}_0)\big) - y_T$$

where π_y is the classical y-space projection, and the function $\hat{y}_f(\cdot)$ is constructed by

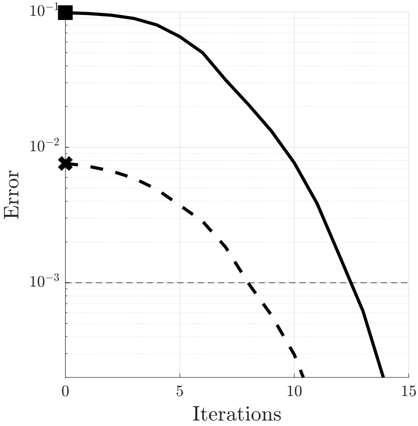
$$\begin{array}{ccc}
(p^0, p_0) = \hat{p}_0 & \xleftarrow[\hat{p}_0 = A^\top \hat{q}_0]{\hat{p}_0 = J_\phi(\hat{x}_0)^\top \hat{q}_0} & \hat{q}_0 = (q^0, q_0) \\
\hat{x}_f = \pi_{\hat{x}}\big(\exp_{\vec{h}}(\hat{x}_0, \hat{p}_0)\big) \Big\downarrow & & \\
(x_f^0, x_f) = \hat{x}_f & \xrightarrow[A x_f + b = \hat{y}_f]{\phi(\hat{x}_f) = \hat{y}_f} & \hat{y}_f = (y_f^0, y_f)
\end{array}$$

The functions T_1 and T_2 are defined from T similarly as S_1 and S_2 from S.

Results

Figure: Evolution of the error w.r.t the number of iterations (with 100 different initial and final states).

Error \ Init	Fixed[4]	Natural
	■	✖
$\lvert S_1(\cdot) \rvert$	———	- - -

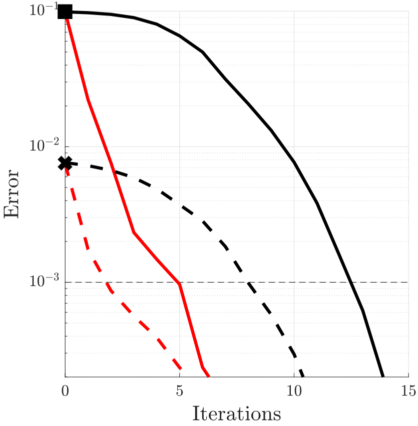


[4] $p = 500$ for S_1

Results

Figure: Evolution of the error w.r.t the number of iterations
(with 100 different initial and final states).

Init / Error	Fixed[4]	Natural
	■	✖
$\lvert S_1(\cdot) \rvert$	—	- - -
for T_2	—	- - -



The error for T_2 is converted into the original coordinates.

[4] $p = 500$ for S_1 and $q = 0$ for T_2.

Conclusion

We propose a new geometric preconditioner of the shooting function :

- based on a geometric interpretation of the costate and on the Mathieu transformation,
- which only needs 2 iterations in average of the solver to find a zero in our application,
- which is non-intrusive with respect to the model,
- no additional computational cost (since we have $\mathbb{D}$ and C).

Bibliography

Cesari, L. (1983).
Statement of the Necessary Condition for Mayer Problems of Optimal Control.
In
Optimization—Theory and Applications: Problems with Ordinary Differential Equations, chapter 4, pages 159–195. Springer New York.

Cots, O., Dutto, R., Jan, S., and Laporte, S. (2023a).
A bilevel optimal control method and application to the hybrid electric vehicle.
Submitted to Optim. Control Appl. Methods.

Cots, O., Dutto, R., Jan, S., and Laporte, S. (2023b).
Generation of value function data for bilevel optimal control method.
Proceeding submitted for the Thematic Einstein Semester 2023 conference.

Cots, O., Dutto, R., Jan, S., and Laporte, S. (2024).
Geometric preconditioner for indirect shooting and application to hybrid vehicle.
Proceeding submitted to the IFAC MICNON 2024 conference.

Main property on the transformation

If $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ is a diffeomorphism then

$$\left\{\begin{array}{l} \min\limits_{\hat{x}=(x^0,x)} \; x^0, \\ \text{s.t. } \hat{x} \in \mathcal{A}, \\ \qquad x = x_T, \end{array}\right. \iff \left\{\begin{array}{l} \min\limits_{\hat{y}=(y^0,y)} \; \pi_{x^0}\big(\phi^{-1}(\hat{y})\big), \\ \text{s.t. } \hat{y} \in \phi\big(\mathcal{A}\big), \\ \qquad \pi_x\big(\phi^{-1}(\hat{y})\big) = x_T, \end{array}\right.$$

where π_{x^0} is the x^0-space projection. Moreover, if ϕ satisfy

$$\frac{\partial \phi}{\partial x^0} = \begin{pmatrix} k \\ 0 \end{pmatrix}, \quad k > 0, \tag{3}$$

then $\phi(\hat{x}) = \big(\phi_0(\hat{x}), \phi_x(x)\big)$ and

$$\left\{\begin{array}{l} \min\limits_{\hat{x}=(x^0,x)} \; x^0, \\ \text{s.t. } \hat{x} \in \mathcal{A}, \\ \qquad x = x_f, \end{array}\right. \iff \left\{\begin{array}{l} \min\limits_{\hat{y}=(y^0,y)} \; y^0, \\ \text{s.t. } \hat{y} \in \phi\big(\mathcal{A}\big), \\ \qquad y = y_T, \end{array}\right.$$

where $y_T = \phi_x(x_T)$.

Results

Figure: Evolution of the error

Black: $|S_1(\cdot)|$

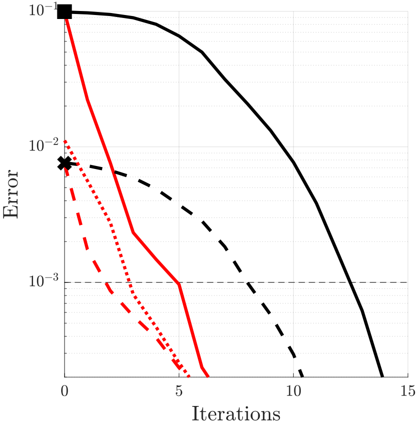
Red: for $T_2(\cdot)$

w.r.t. the number of iterations (with 100 different final states).

—— / ——: fixed init (■) ($p = 500$ / $q = 0$)

- - - / - - -: natural init (✖)

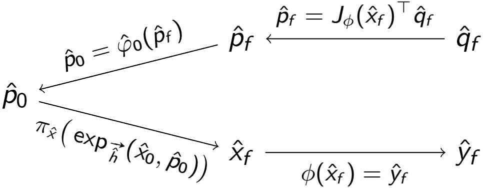
····· : $q = y_T$



The error for T_2 is calculated for each iterate on the initial coordinates.

Definition in the general case

In a general case, the function $\hat{y}_f(\cdot)$ is constructed by



where the function $\hat{\varphi}_0$ is an approximation of the map $\hat{p}_f \mapsto \hat{p}_0$. In our case, this approximation is the identity: