HAL
open science

An algorithmic guide for finite-dimensional optimal control problems

Jean-Baptiste Caillau, Roberto Ferretti, Emmanuel Trélat, Hasnaa Zidani

To cite this version:

Jean-Baptiste Caillau, Roberto Ferretti, Emmanuel Trélat, Hasnaa Zidani. An algorithmic guide for finite-dimensional optimal control problems. Handbook of numerical analysis: Numerical control, Part B, 24, North-Holland; Elsevier, pp.559-626, 2023, Handbook of Numerical Analysis, 10.1016/bs.hna.2022.11.006 . hal-03883822

HAL Id: hal-03883822

https://hal.science/hal-03883822v1

Submitted on 6 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An algorithmic guide for finite-dimensional optimal control problems

Jean-Baptiste Caillau[*], Roberto Ferretti[†], Emmanuel Trélat[‡], Hasnaa Zidani[§]

Abstract

We survey the main numerical techniques for finite-dimensional nonlinear optimal control. The chapter is written as a guide to practitioners who wish to get rapidly acquainted with the main numerical methods used to efficiently solve an optimal control problem. We consider two classical examples, simple but significant enough to be enriched and generalized to other settings: Zermelo and Goddard problems. We provide sample of the codes used to solve them and make these codes available online. We discuss direct and indirect methods, Hamilton–Jacobi approach, ending with optimistic planning. The examples illustrate the pros and cons of each method, and we show how these approaches can be combined into powerful tools for the numerical solution of optimal control problems for ordinary differential equations.

Keywords: optimal control, dynamical systems, Pontryagin maximum principle, direct and indirect methods, Hamilton-Jacobi-Bellman equation, optimistic planning

Contents

1 Introduction and statement of the problem **1**

2 Direct methods: nonlinear programming **6**

3 Indirect approaches: the shooting method **10**

4 Hamilton–Jacobi–Bellman approach **24**

5 Optimistic planning algorithms **46**

1 Introduction and statement of the problem

1.1 Brief overview

Optimal control theory has been widely developed since many decades. The theoretical and numerical achievements, motivated by a body of diverse applications in various domains, provide

[*]Université Côte d'Azur, CNRS, Inria, LJAD (jean-baptiste.caillau@univ-cotedazur.fr)

[†]Dipartimento di Matematica e Fisica, Università Roma Tre, Roma (ferretti@mat.uniroma3.it)

[‡]Sorbonne Université, CNRS, Université de Paris, Inria, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France (emmanuel.trelat@sorbonne-universite.fr).

[§]Insa Rouen Normandie, LMI (hasnaa.zidani@insa-rouen.fr)

valuable insights into the nature of optimal controls and the corresponding optimal trajectories. In this chapter our objective is to provide practitioners with a guide to the most powerful but however easy-to-use methods and algorithms to solve efficiently a given nonlinear optimal control problem in finite dimension.

The most intuitive and popular numerical methods for solving an optimal control problem, called *direct methods*, consist of *first discretizing then optimizing*. Such approaches have been investigated in a number of contributions (see, e.g., [17]). From the theoretical point of view, the efficiency of these methods has been established for some classes of control problems (see [26, 57, 58, 75, 77]). From the numerical point of view, the direct approach benefits from the tremendous advances in numerical optimization methods achieved in the last decades. As shown in Section 2, the direct approach is extremely easy to implement for general control problems with constraints on both the control variable and the state variable. However, in general, direct methods may provide only locally optimal solutions and may require a fair initialization of the iterative process in optimization algorithms. Also, they may lack numerical accuracy, and may become computationally demanding in high dimension.

A major breakthrough in optimal control theory were achieved in the 1950's by Pontryagin's research group, who successfully generalized and extended to a general nonlinear optimal control setting the classical Euler-Lagrange and Weierstrass conditions of the Calculus of Variations. The early optimality conditions, called Pontryagin Maximum principle (in short, PMP), were subsequently strengthened and extended using methods of convex and non-smooth analysis, or methods of differential geometry (see [72]). The PMP inspired effective computational schemes, like the shooting method presented in Section 3. As illustrated on some classical examples, the shooting method provides very accurate numerical solutions. The major drawback of this method is that it requires an *a priori* knowledge of the structure of the solution, as well as a good approximation of the adjoint state. In addition, shooting methods are difficult to implement in state-constrained problems, in particular for a large number of state constraints.

Another major advance in optimal control was achieved in the 1950's by Richard Bellman, who provided a description of how the minimum cost depends on initial conditions, as the solution to the so-called Hamilton–Jacobi–Bellman (HJB) partial differential equation. Whenever applicable, the Hamilton–Jacobi approach yields a global solution to the optimal control problem, and provides the optimal control in a feedback form, suitable for many engineering applications. Despite these advantages, this approach suffers from the difficulty of computing the solution to the Hamilton–Jacobi equation in higher dimensions. The numerical simulations presented in Section 4 show that the approximation of the HJB equation, even on coarse grids, provides solutions which depict well the qualitative structure of the optimal trajectories and may thus be used to guess the structure of optimal solutions. However, if an accurate computation of the optimal trajectories is required, the use of fine grids causes a strong increase in computational complexity.

More recently, other global methods have been developed for control problems, such as the optimistic planning (OP) algorithms introduced in Section 5. These methods are based on a discretization of the control space and do not require any discretization of the state space. As a consequence, OP methods are particularly efficient in control problems with a low-dimensional control space; moreover, this approach seems also well suited for problems where the dynamics and the cost are given by learning models. A preliminary analysis of the complexity of OP methods is now well established, but further developments are expected to make them more accurate, to possibly integrate the knowledge of the structure of trajectories, and also for a more

efficient implementation.

Throughout the chapter we consider two well known but representative optimal control examples, the Zermelo and the Goddard problem: both are quite simple but can be extended towards more intricate models. We use these two examples to illustrate the numerical methods and show how they can be rapidly and efficiently implemented, with up-to-date existing solvers. We argue that the various approaches are complementary rather than in competition, and can be suitably combined to exploit their peculiarities. For instance, HJB approach and direct methods can be used to obtain a rough estimate of adjoint state and cost, which could provide a good initial guess for the more accurate shooting method. Moreover, HJB solvers may allow to rule out local minima as it will be illustrated on the Zermelo problem with obstacle.

We hope that the codes that we provide, which are also available on the web, can serve as templates to readers interested in adapting them to their specific setting.

Notations. Throughout the chapter, $\mathbb{R}$ denotes the set of real numbers, $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote respectively the Euclidean inner product and the norm on $\mathbb{R}^N$(for any $N \geq 1$), $\mathbb{B}_N = \{x \in \mathbb{R}^N : \|x\| \leq 1\}$ is the closed unit ball (also denoted $\mathbb{B}$ if there is no ambiguity) and $\mathbb{B}(x;r) = x + r\mathbb{B}$. For any set $S \subseteq \mathbb{R}^N$, $\overset{\circ}{S}, \overline{S}$, ∂S, co$S$ denote its interior, closure, boundary, and convex envelope, respectively. For any $a, b \in \mathbb{R}$, we define $a \bigvee b := \max(a, b)$. Similarly, for $a_1, \cdots, a_m \in \mathbb{R}$, we define $\bigvee_{i=1}^m a_i := \max(a_1, \cdots, a_m)$. The notation $W^{1,1}([a,b])$ stands for the usual Sobolev space $\{f \in L^1([a,b]), f' \in L^1([a,b])\}$. Finally, the abbreviation "w.r.t." stands for "with respect to", and "a.e." means "almost everywhere".

1.2 Formulation of the optimal control problem

Let $d, r \in \mathbb{N}^*$, let $T > 0$ be a fixed final time horizon and let U be a compact subset of $\mathbb{R}^r$ (with $r \geq 1$). We consider the finite-dimensional control system in $\mathbb{R}^d$ (for $0 \leq t < T$)

$$\dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), \boldsymbol{u}(s)), \text{ a.e. } s \in (t, T), \tag{1}$$

where the control input $\boldsymbol{u} : [0,T] \longrightarrow \mathbb{R}^r$ $(r \geq 1)$ is a measurable function such that $\boldsymbol{u}(s) \in U$ for almost every $s \in [0,T]$. Throughout the paper, we assume that

$(\mathbf{H_0})$ $\quad U$ is a compact subset of $\mathbb{R}^r$.

We denote by $\mathcal{U}$ the set of all admissible controls

$$\mathcal{U} := \{\boldsymbol{u} : [0,T] \longrightarrow \mathbb{R}^r \text{ measurable, and } \boldsymbol{u}(s) \in U \text{ a.e.}\}.$$

The dynamics $f : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^d$ satisfies:

$$(\mathbf{H_{1a}}) \quad \begin{cases} (i) \quad f \text{ is continuous on } [0,T] \times \mathbb{R}^d \times U; \\ (ii) \quad x \to f(s,x,u) \text{ is locally Lipschitz continuous in the following sense:} \\ \qquad \forall R > 0,\ \exists k_R \geq 0,\ \forall (x,y) \in (\mathbb{B}_d(0;R))^2,\ \forall (s,u) \in [0,T] \times U \\ \qquad\qquad \|f(s,x,u) - f(s,y,u)\| \leq k_R \|x - y\|; \\ (iii)\ \exists c_f > 0 \text{ such that } \quad \|f(s,x,u)\| \ \leq\ c_f(1 + \|x\|) \ \ \forall (s,x,u) \in [0,T] \times \mathbb{R}^d \times U. \end{cases}$$

The assumptions (ii)-(iii) of $(\mathbf{H_{1a}})$ guarantee, for every $\boldsymbol{u} \in \mathcal{U}$ and $(t,x) \in [0,T] \times \mathbb{R}^d$, the existence of an absolutely continuous curve $\mathbf{x} : [t,T] \to \mathbb{R}^N$ which satisfies (1) and the initial

condition $\mathbf{x}(t) = x$. By the Gronwall lemma,

$$1 + \|\mathbf{x}(s)\| \leq (1 + \|x\|)e^{c_f(s-t)} \qquad \forall s \in [t, T]. \tag{2}$$

Given any $t \in [0, T]$ and any $x \in \mathbb{R}^d$, the set of all admissible pairs control-and-trajectories starting at x at time t is denoted by

$$\mathbb{X}_{[t,T]}(x) := \{(\mathbf{x}, \boldsymbol{u}) \in W^{1,1}(t, T) \times \mathcal{U} \mid \ (\mathbf{x}, \boldsymbol{u}) \text{ satisfies (1) with } \mathbf{x}(t) = x\}.$$

Throughout the chapter, we consider the (Bolza) optimal control problem

$$\begin{cases} \text{minimize} \ \ \varphi(\mathbf{x}(T)) + \displaystyle\int_t^T \ell(s, \mathbf{x}(s), \boldsymbol{u}(s))\, ds, \\ \qquad (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x), \\ \qquad g(\mathbf{x}(s)) \leq 0 \quad \text{for } s \in [t, T], \\ \qquad g_f(\mathbf{x}(T)) \leq 0, \end{cases} \tag{3}$$

with the convention that $\inf \emptyset = +\infty$. The distributed cost $\ell : [0, T] \times \mathbb{R}^d \times U \to \mathbb{R}$, the final cost $\varphi : \mathbb{R}^d \to \mathbb{R}$, and the constraint functions g and g_f are given functions satisfying:

$$(\mathbf{H_{1b}}) \qquad \begin{cases} (i) \quad \ell \text{ is continuous on } [0, T] \times \mathbb{R}^d \times U; \\ (ii) \quad x \to \ell(s, x, u) \text{ is locally Lipschitz continuous in the following sense:} \\ \qquad \forall R > 0, \ \exists k_R \geq 0, \ \forall (x, y) \in (\mathbb{B}_d(0; R))^2, \ \forall (s, u) \in [0, T] \times U \\ \qquad\qquad |\ell(s, x, u) - \ell(s, y, u)| \leq k_R \|x - y\|; \\ (iii) \ \exists c_\ell > 0 \text{ such that} \quad |\ell(s, x, u)| \leq \ c_\ell(1 + \|x\|) \ \ \forall (s, x, u) \in [0, T] \times \mathbb{R}^d \times U. \end{cases}$$

$(\mathbf{H_2})$ $\varphi : \mathbb{R}^d \to \mathbb{R}$ is locally Lipschitz continuous and there exists a constant $c_\varphi \geq 0$ such that

$$|\varphi(y)| \leq c_\varphi(1 + \|y\|).$$

$(\mathbf{H_3})$ The constraint functions $g : \mathbb{R}^d \to \mathbb{R}^{m_g}$ and $g_f : \mathbb{R}^d \to \mathbb{R}^{m_f}$ are locally Lipschitz continuous (with $m_g, m_f \geq 1$) and there exists a constant $c_g \geq 0$ such that, for every $y \in \mathbb{R}^d$,

$$\|g(y)\| + \|g_f(y)\| \leq c_g(1 + \|y\|).$$

The so-called augmented control system

$$\dot{\mathbf{x}}(s) \ = \ f(x, \mathbf{x}(s), \boldsymbol{u}(s)) \quad \text{a.e. } s \in (t, T), \tag{4a}$$

$$\dot{\mathbf{z}}(s) \ = \ -\ell(s, \mathbf{x}(s), \boldsymbol{u}(s)) \quad \text{a.e. } s \in (t, T), \tag{4b}$$

is usually considered in optimal control theory to recast the Bolza problem in the Mayer form

$$\begin{cases} \text{minimize} \ \ \varphi(\mathbf{x}(T)) - \mathbf{z}(T), \\ \qquad (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x_0), \\ \qquad g(\mathbf{x}(s)) \leq 0 \quad \text{for } s \in [t, T], \\ \qquad g_f(\mathbf{x}(T)) \leq 0, \end{cases} \tag{5}$$

In what follows, we will **sometimes** assume that the augmented dynamics satisfies the following assumption (convex epigraph):

$(\mathbf{H_4})$ For any $s \in [0, T]$ and every $x \in \mathbb{R}^d$,

$$\left\{ \begin{pmatrix} f(s, x, u) \\ -\ell(s, x, u) + \eta \end{pmatrix}, \ \ u \in U, \ \ -c_\ell(1 + \|x\|) + \ell(s, x, u) \leq \eta \leq 0 \right\} \quad \text{is a convex set.}$$

Remark 1.1. Note that if $\ell \equiv 0$ (Mayer problem), $(\mathbf{H_4})$ reduces to

$$f(s, x, U) \text{ convex for all } (s, x) \in [0, T] \times \mathbb{R}^d.$$

Remark 1.2. For $0 \leq a < b \leq T$ and $(x, z) \in \mathbb{R}^d \times \mathbb{R}$, consider the set of all trajectories satisfying (4) on the time interval $[a, b]$, for a control input $u \in \mathcal{U}$, and starting at a position (x, z) at time a:

$$\mathcal{S}_{[a,b]}(x, z) := \{(\mathbf{x}, \mathbf{z}) \in W^{1,1}([a, b]) \mid \exists \boldsymbol{u} \in \mathcal{U} \text{ such that } (\mathbf{x}, \mathbf{z}, \boldsymbol{u}) \text{ satisfies (4)}, \\ \text{with } \mathbf{x}(a) = x,\ \mathbf{z}(a) = z\}.$$

Under assumptions $(\mathbf{H_0})$, $(\mathbf{H_{1a}})$, $(\mathbf{H_{1b}})$, by (2), the set $\mathcal{S}_{[a,b]}(x, z)$ is bounded in $W^{1,1}([a, b])$. Moreover, if $(\mathbf{H_4})$ is satisfied then $\mathcal{S}_{[a,b]}(x, z)$ is a compact set in $W^{1,1}([a, b])$ endowed with the $C^0([a, b])$-topology (see [8, Theorem 1.4.1]). Therefore, if there exists a trajectory $\mathbf{x} \in \mathcal{S}_{[0,T]}(x_0, 0)$ that satisfies the constraints $g(\mathbf{x}(s)) \leq 0$ for all $s \in [t, T]$ and $g(\mathbf{x}(T)) \leq 0$, then the control problem (5) has an optimal solution.

1.3 Examples

In this section, we present two well-known examples that we are going to consider throughout. In our opinion, they illustrate nicely the most classical difficulties encountered in theoretical and numerical optimal control, and lend themselves to a number of more complicated variants. They are expected to serve as "templates" to the reader who aims at getting acquainted with the main issues in numerical optimal control.

Example 1: Zermelo problem A boat with coordinates $\mathbf{x}(t) = (\mathbf{x}_1(t), \mathbf{x}_2(t))$ navigates through a canal $\mathbb{R} \times [a, b]$, starting at $\mathbf{x}(0) = x = (x_1, x_2)$, and wants to reach an island $\mathcal{C}$ with minimal cost. The cost function may be an energy, the final time, etc. The control system is

$$\dot{\mathbf{x}}_1(t) = \boldsymbol{v}(t) \cos(\boldsymbol{u}(t)) + h(\mathbf{x}(t)), \tag{6a}$$

$$\dot{\mathbf{x}}_2(t) = \boldsymbol{v}(t) \sin(\boldsymbol{u}(t)), \tag{6b}$$

where $\boldsymbol{u}(t) \in [0, 2\pi]$ is the first control (angle), $\boldsymbol{v}(t) \in [0, V_{max}]$ is a second control (speed of the boat), and $h(\mathbf{x}(t))$ is the current drift (along the x_1-axis). Because of the drift term (which can be greater than $V_{\max}$), the system may not be controllable. Consider a target $\mathcal{C} := \mathbb{B}(c, r_0)$ that is a ball with radius $r_0 \geq 0$ and centered at a given point c located in the canal. The target represented by a function g_f defined by $g_f(x) := \|x - c\| - r_0$ as

$$x \in \mathcal{C} \Longleftrightarrow g_f(x) \leq 0.$$

Consider also a set of constraints $\mathcal{K} := \{x \in \mathbb{R}^2,\ g(x) \leq 0\}$ where g is a given function that is non-positive in a region where the boat can move and g is positive in the location of the obstacles that the boat should avoid. In this example, the cost function could be the time, or the energy, required to steer the boat from a given position x to the target $\mathcal{C}$.

The Zermelo problem has several variants depending on the choice of the dynamic h, on the expression of the constraints, as well as on the values of the different parameters entering the model (a, b, and V_{max}). In the next sections, we will consider different settings to better illustrate the pros and cons of each numerical method.

Example 2: Goddard Problem We consider the optimal control problem associated with a vertical ascension flight of a rocket, known as Goddard problem. The dynamics of the rocket is defined with three state variables: r, the altitude, v the relative speed and m, the total mass. In general, a dimensionless version for the motion equations is considered in the literature:

$$\dot{\mathbf{r}}(t) = \boldsymbol{v}(t) \tag{7a}$$

$$\dot{\boldsymbol{v}}(t) = \frac{1}{m}(T_{max}\,\boldsymbol{u}(t) - D(\mathbf{r}(t), \boldsymbol{v}(t))) - \frac{1}{\mathbf{r}(t)^2} \tag{7b}$$

$$\dot{\mathbf{m}}(t) = -T_{max}\,b\,\boldsymbol{u}(t) \tag{7c}$$

where $D(r, v)$ is the drag force and $T_{max} \cdot u$ is the thrust force. The dimensionless initial state of the rocket is given by $\mathbf{r}(0) = 1$, $\boldsymbol{v}(0) = 0$, $\mathbf{m}(0) = 1$. Here $\mathbf{r}(0) = 1$ corresponds to the Earth's ground level and $m = 1$ to the initial total mass of the rocket. The motion of the rocket is controlled by the thrust factor $\boldsymbol{u}(s) \in [0, 1]$ so that the thrust force is on the interval $T_{max} \cdot \boldsymbol{u}(t) \in [0, T_{max}]$. The drag force is a nonlinear function of $\mathbf{r}$ and $\boldsymbol{v}$. Its expression depends on the choice of a model for the atmosphere and on the structure of the rocket: $D(r, v) = C_D \rho(r)|v|v$, where C_D is the drag coefficient of the rocket and ρ is the atmospheric density. In this section we consider the case of constant drag coefficient (it depends in general on the Mach number) and exponential model of the atmospheric density:

$$D(r, v) = C_D\ v^2\ e^{-\beta\,(r-1)} \quad \forall r \geq 1, v \geq 0.$$

This definition of the rocket's model has been widely studied in the literature. We take $C_D = 310.0$, $\beta = 500.0$, $T_{max} = 3.5$, $b = 2.0$, as in [82] and other references. The corresponding model is then an approximation of real flight conditions for some rockets. The optimal control problem consists of maximizing the final altitude $\mathbf{r}(t_f)$ at free final time $t_f > 0$ under a final constraint on the fuel consumption, and a pointwise constraint on the velocity. Finally, the optimal control problem is formulated as follows

$$\begin{cases} \text{minimize} \ \Big(-\mathbf{r}(t_f) \Big), \\ \quad (\mathbf{r}, \boldsymbol{v}, \mathbf{m}) \text{ satisfies (7a)–(7c) with } \mathbf{r}(0) = 1,\ \boldsymbol{v}(0) = 0,\ \mathbf{m}(0) = 1, \\ \quad \boldsymbol{v}(s) \leq v_{\max} \ \text{ a.e. on } (0, t_f) \\ \quad m^* - \mathbf{m}(t_f) \leq 0. \end{cases} \tag{8}$$

In the numerical simulations that will be presented in the next sections, the limit of fuel consumption is $m^* = 0.6$ and the maximum velocity is $v_{\max} = 0.1$.

2 Direct methods: nonlinear programming

2.1 Principle

We call *direct methods* all numerical methods consisting of numerically solving the optimal control problem as follows: without applying a priori any first-order necessary condition for optimality, we choose a discretization for the state and for the control, we choose a numerical scheme to discretize the control system (differential equation) and the cost functional (integral quadrature),

so that the discretized optimal control problem is expressed as a family of nonlinear optimization problems in finite dimension, indexed by a discretization parameter N, of the form

$$\min_{Z \in \mathcal{C}} F(Z) \tag{9}$$

where $Z = (x_1, \ldots, x_N, u_1, \ldots, u_N)$ and

$$\mathcal{C} = \{Z \mid g_i(Z) = 0,\ i \in 1, \ldots, r, \quad g_j(Z) \leq 0,\ j \in r+1, \ldots, m\}. \tag{10}$$

This is a classical optimization problem under constraints in finite dimension, with a dimension growing as the discretization is refined. Of course, there exist an infinite number of variants to discretize the problem and ending up with a problem of the form (9). We discuss hereafter several classes of discretizations. Once this transcription has been done, one can then implement a number of various optimization routines to solve (9).

Let us first explain hereafter one of the simplest possible discretizations. Consider the optimal control problem (3) with $t = 0$ as initial time. Consider a subdivision $0 = t_0 < t_1 < \cdots < t_N = T$ of the interval $[0, T]$. Controls are discretized on U-valued piecewise constant functions on this subdivision. To discretize the control system, we choose the explicit Euler method: setting $h_i = t_{i+1} - t_i$, we have $x_{i+1} = x_i + h_i f(t_i, x_i, u_i)$ for $i = 0, \ldots, N-1$. To discretize the integral cost, we choose the left rectangle method (which is equivalent to the explicit Euler method for the augmented system). We obtain the nonlinear programming problem (x_0 being known)

$$\begin{aligned}
&\min C(x_1, \ldots, x_N, u_0, \ldots, u_{N-1}),\\
&x_{i+1} = x_i + h_i f(t_i, x_i, u_i), \quad u_i \in U,\\
&g(x_{i+1}) \leq 0, \qquad i = 0, \ldots, N-1,\\
&g_f(x_N) \leq 0.
\end{aligned}$$

We note that this approach is flexible and robust insofar it is not much sensitive to the model (contrarily to the shooting method, described further): it is very easy to add various constraints to the optimal control problem. This is why direct methods are often privileged when the model is not completely fixed. The resulting numerical simulations often give an interesting feedback that may lead to change or adapt the optimal control model under consideration.

2.2 Practical numerical implementation

The numerical implementation of such a nonlinear programming problem is standard and can be done in a number of ways, for instance using a penalty method or a sequential quadratic programming (SQP) method or dual methods (like Uzawa's). A survey on the use of direct methods in optimal control, with a special interest to applications in aerospace, can be found in [17].

From the point of view of practical implementation, in the last years much progress has been done in the direction of combining automatic differentiation softwares (such as the modelling language `AMPL`, see [53], or `CasADi`, see [6]) with expert optimization routines (such as the open-source package `Ipopt`, see [85], providing an interior point optimization algorithm for large-scale differential algebraic systems combined with a filter line-search method). With such tools, it has become very simple to implement with only few lines of code difficult (nonacademic) optimal control problems, with success and within a reasonable time of computation. Websites such as

NEOS (www.neos-server.org/neos/solvers) propose to launch online such kinds of computation: codes can be written in a modelling language such as [53] (or others) and can be combined with many optimization routines (specialized either for linear problems, nonlinear, mixed, discrete, etc). The advantage of using NEOS is that one has nothing to install on his own machine, and moreover one can test a large number of possible optimization routines.

Note that there exist a large number (open-source or not) of automatic differentiation softwares and of optimization routines. It is not our aim to provide a list of them, since they are easy to find on the web. Note also that AMPL, which is very easy and friendly to use, is however not free of charge (although the licence is not expensive) and that CasADi offers a very good and efficient free alternative, see web.casadi.org.

2.3 Variants

As alluded above, there exist many possible approaches to discretize an optimal control problem, see [17] where the important sparsity issues are also discussed. Among those various approaches, we quote the following.

Collocation methods consist of choosing specific points or nodes on every subinterval of a given subdivision of the time interval. Such methods approximate the trajectories and the controls by polynomials on each subinterval. The collocation conditions state that the derivatives of the approximated state match with the dynamics at each node.

Spectral and pseudospectral methods are another class in which the above nodes are chosen as the zeros of specific polynomials such as Gauss-Legendre or Gauss-Lobatto polynomials. Such polynomials are used as a basis to approximate trajectories and controls in appropriate approximation spaces. Since they share nice orthogonality properties, the collocation conditions turn into constraints that are easily tractable for numerical purposes. We refer the reader to [46, 57, 75] and to the references therein for more details.

There exist also some probabilistic approaches, such as the method described in [64] which consists of first relaxing the optimal control problem in measure spaces and then of seeking the optimal control as an occupation measure, which is approximated by a finite number of its moments (see [62]). This approach relies on algebraic geometry tools and reduces the optimal control problem to some finite dimensional optimization problem involving linear matrix inequalities (LMI). On this topic involving Sums-of-Square (SoS) considerations, we refer the reader to another chapter of the Handbook, [63].

Remark 2.1. Direct methods are characterized by first discretizing and then optimizing, i.e., optimality conditions are applied in a second step to the discretized model; in contrast to this approach, indirect methods (to be discussed in the next section) consist of applying first optimality conditions (the Pontryagin Maximum Principle) and then discretizing the resulting boundary value problem. While the latter method clearly falls in the classical Lax scheme, "consistency plus stability imply convergence", there is a serious gap there in the direct approaches: to ensure convergence using the Lax scheme, one would a priori need a *uniform* (with respect to N) consistency property, which fails in general because the optimal control problem is an optimization problem in infinite dimensions. Surprisingly simple examples of divergence are provided in [58]. In this same paper, it is shown that convergence is obtained for "smooth enough" optimal control problems without any final state constraint, discretized with Runge–Kutta methods with positive coefficients. We also refer to [26, 77] for further comments and considerations on symplectic in-

tegrators. Convergence has also been established for classes of Legendre pseudospectral methods (see [46, 57, 75]).

2.4 Goddard problem by a direct approach

To illustrate the method, we present a treatment of the Goddard case presented Section 1.3 using the Crank-Nicolson scheme. This scheme, which is *dual* to the midpoint rule (in the sense defined in [59]), has the advantage of using only gridpoints (contrary to the midpoint scheme). It is very easy to implement and the code below is a basic "do it yourself" direct solver that can easily be adapted to other problems. We use the nice `JuMP` interface in `julia` to define the discretization of the problem and call the celebrated interior point solver `Ipopt` previously mentioned. The symbolic-numeric framework put forward by `Julia` allows to efficiently and transparently use AD (automatic differentiation / differentiable programming) and sparse linear algebra (which is important for structured constraints stemming from one-step like methods to discretize the dynamics). In the spirit of reproducible research [31], the code itself is available and executable online on the gallery of the `ct: control toolbox` project.[1] Discretizing the dynamics boils down to the following lines (note that nonlinear expressions encoding the right-hand side are first defined):

```
# Dynamics
@NLexpressions(sys, begin
    # D = Cd v^2 exp(-β(r-1))
    D[i = 1:N+1], Cd * v[i]^2 * exp(-β * (r[i] - 1.0))
    # r'= v
    dr[i = 1:N+1], v[i]
    # v' = (Tmax.u-D)/m - 1/r^2
    dv[i = 1:N+1], (Tmax*u[i]-D[i])/m[i] - 1/r[i]^2
    # m' = -b.Tmax.u
    dm[i = 1:N+1], -b*Tmax*u[i]
end)

# Crank-Nicolson scheme
@NLconstraints(sys, begin
    con_dr[i = 1:N], r[i+1] == r[i] + Δt * (dr[i] + dr[i+1])/2.0
    con_dv[i = 1:N], v[i+1] == v[i] + Δt * (dv[i] + dv[i+1])/2.0
    con_dm[i = 1:N], m[i+1] == m[i] + Δt * (dm[i] + dm[i+1])/2.0
end)
```

Note that the constraints have been labeled so that the corresponding Lagrange multipliers can be retrieved. They are indeed approximations of the costate of the continuous problem and will be used as such to initialize successfully a subsequent shooting method (see Section 3).

A strong benefit of direct methods is that state constraints are very easy to handle. We may for instance add a state constraint on the velocity (note that some other constraints have been added, directly when defining the unknowns of the problem, to improve convergence of the solver; what would be a complication for indirect methods is actually an asset here):

[1] ct.gitlabpages.inria.fr/gallery

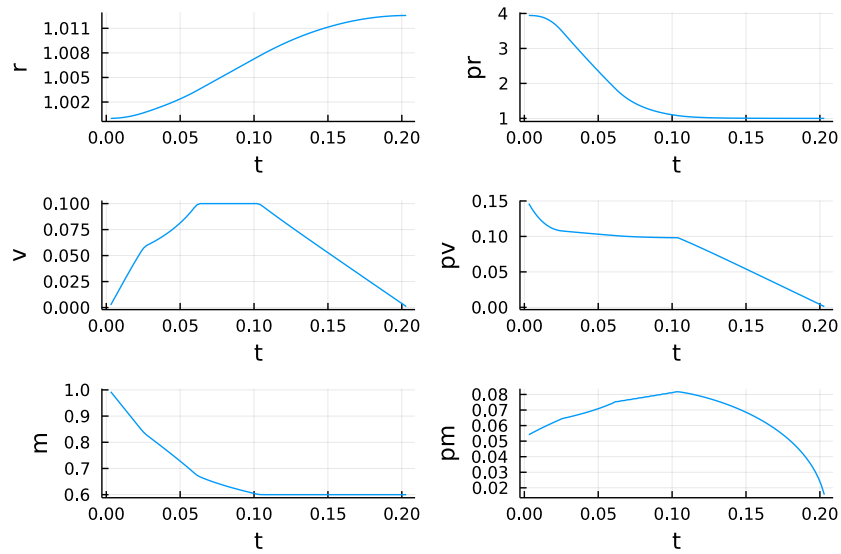

Figure 1: Goddard problem: result of the direct code (Crank-Nicolson scheme), states and Lagrange multipliers associated with the equality constraints that discretize the dynamics. These multipliers approximate the adjoint states and will be used as such to initialise an indirect method. A boundary arc is observed as the state constraint on the velocity is saturated.

```
# As final time is free, time step Δt is unknown
@variables(sys, begin
    0.0 ≤ Δt
    r[1:N+1] ≥ r0
    0 ≤ v[1:N+1] ≤ vmax
    mf ≤ m[1:N+1] ≤ m0
    0.0 ≤ u[1:N+1] ≤ 1.0
end)
```

Boundary constraints are obviously added in the same way, and for $N = 100$ gridpoints we get the results given on Figures 1 and 2. The resulting optimization problem solved has about 400 variables (the state is of dimension 3 and the control is scalar) and 300 equality constraints (the discretized dynamics), plus box constraints. One could use a finer grid (up to the price of a larger problem), but the result we obtain turns to be precise enough to trigger convergence of much more accurate solver by shooting in the next section. The bang-singular-constrained-bang structure of the solution has indeed been captured by the direct solver, and this is essentially all we need to resort to indirect methods to complete the computation.

3 Indirect approaches: the shooting method

3.1 Pontryagin maximum principle

In all this section, we assume that hypotheses $(\mathbf{H_0})$-$(\mathbf{H_3})$ are satisfied, and that the functions φ, f, ℓ and g_f are of class C^1 w.r.t. the space variable x. Consider the optimal control problem (3) with $t = 0$ as initial time, and with $g = 0$, i.e., for the moment, with no state constraint.

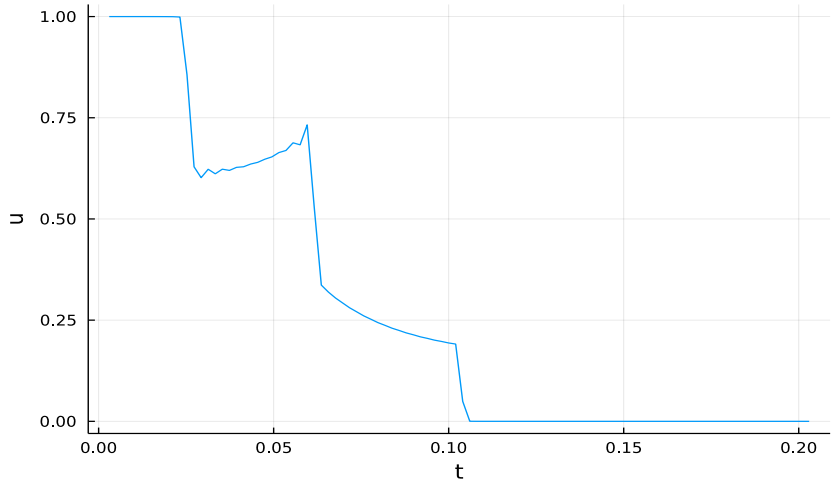

Figure 2: Goddard problem: result of the direct code, control values. A structure with four different subarcs is observed. A simple inspection indicates a concatenation of a bang arc ($u \simeq 1$), possibly of a singular arc interior to the control bounds ($u \in (0,1)$), of a boundary arc due to the state constraint on the velocity, and of a final bang arc ($u \simeq 0$). This structure information, of combinatorial nature, is obtained here without any *a priori* knowledge on the solution. It is the key to define an appropriate shooting function and solve very accurately the problem thanks to an indirect method.

Recall that the *end-point mapping* of the system is defined by $E(x_0, T, \boldsymbol{u}) = \mathbf{x}(x_0, T, \boldsymbol{u})$, where $t \mapsto \mathbf{x}(x_0, t, \boldsymbol{u})$ is the trajectory solution of the control system, corresponding to the control $\boldsymbol{u}$, such that $\mathbf{x}(x_0, 0, \boldsymbol{u}) = x_0$. The end-point mapping is well defined and C^1 for $\boldsymbol{u}$ belonging to a neighbourhood in $L^\infty([0,T], \mathbb{R}^r)$ of the reference control (and C^k whenever the dynamics are C^k). Denoting by C the cost functional, the optimal control problem can be written as the infinite dimensional minimization problem of minimizing C over the considered set of controls, under the constraint $g_f(E(x_0, T, \boldsymbol{u})) \leq 0$ on the final state.

We first assume that we are in the simple case where the initial point x_0 is fixed, the final point x_1 is fixed, the final time T is fixed, $U = \mathbb{R}^r$ and there are no state constraints along the trajectory. Then, the optimal control problem is exactly equivalent to the optimization problem

$$\min_{E(x_0,T,\boldsymbol{u})=x_1} C(T, \boldsymbol{u}). \tag{11}$$

If $\boldsymbol{u}$ is optimal, then there must exist Lagrange multipliers $(\psi, \psi^0) \in \big(\mathbb{R}^d \times \mathbb{R}\big) \setminus \{0\}$ such that

$$\psi \cdot dE_{x_0,T}(\boldsymbol{u}) = -\psi^0 dC_T(\boldsymbol{u}). \tag{12}$$

Defining the Lagrangian $L_T(\boldsymbol{u}, \psi, \psi^0) = \psi E_{x_0,T}(\boldsymbol{u}) + \psi^0 C_T(\boldsymbol{u})$, the first-order condition (12) is written as

$$\frac{\partial L_T}{\partial u}(\boldsymbol{u}, \psi, \psi^0) = 0. \tag{13}$$

The Pontryagin maximum principle (PMP), recalled hereafter, is a far-reaching statement of these conditions (see [72], see also [65]). We recall here the statement of the PMP in the case without pointwise state constraints (that is for $g = 0$).

Theorem 3.1. *If the trajectory* $\mathbf{x}$*, associated to the optimal control* $\boldsymbol{u}$ *on* $[0,T]$*, is optimal, then it must be the projection of an extremal* $(\mathbf{x},\boldsymbol{p},p^0,u)$ *(called extremal lift), where* $p^0 \leq 0$ *and* $\boldsymbol{p}:[0,T]\to\mathbb{R}^d$ *is an absolutely continuous mapping called adjoint vector, with* $(\boldsymbol{p},p^0)\neq(0,0)$*, such that*

$$\dot{\mathbf{x}}(t)=\frac{\partial H}{\partial p}(t,\mathbf{x}(t),\boldsymbol{p}(t),p^0,\boldsymbol{u}(t)),\ \dot{\boldsymbol{p}}(t)=-\frac{\partial H}{\partial x}(t,\mathbf{x}(t),\boldsymbol{p}(t),p^0,\boldsymbol{u}(t)),$$

almost everywhere on $[0,T]$*, where* $H(t,x,p,p^0,u)=\langle p,f(t,x,u)\rangle+p^0\ell(t,x,u)$ *is the Hamiltonian, and the condition*

$$H(t,\mathbf{x}(t),\boldsymbol{p}(t),p^0,\boldsymbol{u}(t))=\max_{v\in U}H(t,\mathbf{x}(t),\boldsymbol{p}(t),p^0,v) \tag{14}$$

holds almost everywhere on $[0,T]$*. If moreover the final time* T *is not fixed, then*

$$\max_{v\in U}H(T,\mathbf{x}(T),\boldsymbol{p}(T),p^0,v)=0. \tag{15}$$

Additionally, if $g_f(x(T))=0$*, then the adjoint vector can be built in order to satisfy the transversality condition*

$$\boldsymbol{p}(T)-p^0\nabla\varphi(\mathbf{x}(T))\ \perp\ \ker dg_f(\mathbf{x}(T)), \tag{16}$$

where $dg_f(x)$ *stands for the Jacobian of* g_f *at point* x*.*

The adjoint vector of the Pontryagin maximum principle can be constructed so that, up to a scaling, $(\boldsymbol{p}(T),p^0)=(\psi,\psi^0)$ from (12). In particular, the Lagrange multiplier ψ is unique (up to a multiplicative scalar) if and only if the trajectory $\mathbf{x}$ admits a unique extremal lift (up to scaling). The scalar p^0 is a Lagrange multiplier associated with the cost. When $p^0<0$, the extremal is said to be *normal*, and in this case, since the Lagrange multiplier is defined up to scaling, it is usual to normalize it so that $p^0=-1$. When $p^0=0$, the extremal is said to be *abnormal*. In many situations, where some qualification conditions hold, abnormal extremals do not exist in the problem under consideration, but in general it is difficult to guess whether, given some initial and final conditions, these qualification conditions hold or not. In lack of control constraints, i.e., when $U=\mathbb{R}^r$, abnormal extremals project exactly onto singular trajectories. Recall that a couple $(\mathbf{x},\boldsymbol{u})$ defined on $[0,T]$ is said to be singular when the linearized control system along it is not controllable in time T; equivalently, in terms of the end-point mapping, the Fréchet differential $dE_{x_0,T}$ is not surjective. In the normal case, $\psi=\boldsymbol{p}(T)$ coincides (up to a scaling) with the gradient of the value function (solution of the Hamilton–Jacobi equation). This point is further discussed in Section 4.4.

Remark 3.1. Generically, the solution of the optimal control problem is unique, and moreover it has a unique extremal lift. This well known fact is related to the differentiability properties of the value function (see for instance [9, 40], and see [34, 73, 74, 79] for results on the size of the set where the value function is differentiable).

Remark 3.2. The fact that $p^0\leq 0$ is a convention due to Pontryagin, which leads to the maximum principle. The choice $p^0\geq 0$ would lead to a minimum condition, instead. The component p^0 appears in the transversality condition (16). Note that, if the final point $x(T)$ is let free (i.e., there is no g_f) then this condition leads to $\boldsymbol{p}(T)=p^0\nabla\varphi(\mathbf{x}(T))$, and then necessarily $p^0\neq 0$ and we can normalize to $p^0=-1$.

Remark 3.3. When there are some state constraints $g(\mathbf{x}(t)) \leq 0$ along the trajectory, the Pontryagin maximum principle is modified as follows. We keep the same definition for the Hamiltonian H. If $\mathbf{x}$ is optimal then there must exist $p^0 \leq 0$, an absolutely continuous adjoint vector $\boldsymbol{p}(\cdot)$ and a nonnegative Radon measure μ, the triple $(\boldsymbol{p}, p^0, \mu)$ being nontrivial, such that the adjoint equation is

$$\dot{\boldsymbol{p}}(t) = -\frac{\partial H}{\partial x}(t, \mathbf{x}(t), \boldsymbol{q}(t), p^0, \boldsymbol{u}(t))$$

almost everywhere, with

$$\boldsymbol{q}(t) = \boldsymbol{p}(t) + \int_{[0,t)} \frac{\partial g}{\partial x}(\mathbf{x}(s))\, d\mu(s),$$

(the interval being closed when $t = T$) and the maximization condition becomes

$$H(t, \mathbf{x}(t), \boldsymbol{q}(t), p^0, \boldsymbol{u}(t)) = \max_{v \in U} H(t, \mathbf{x}(t), \boldsymbol{q}(t), p^0, v)$$

almost everywhere. Finally, in the transversality condition on the final adjoint vector, one replaces $\boldsymbol{p}(T)$ with $\boldsymbol{q}(T)$. Note that, taking $\boldsymbol{q}$ as a new adjoint, we have $d\boldsymbol{q}(t) = d\boldsymbol{p}(t) + \frac{\partial g}{\partial x}(\mathbf{x}(t))\, d\mu(t)$ and, since $\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial x}$ and provided μ is absolutely continuous w.r.t. Lebesgue measure, one has

$$\dot{\boldsymbol{q}} = -\boldsymbol{q}\frac{\partial f}{\partial x} - p^0 \frac{\partial f^0}{\partial x} + \nu \frac{\partial g}{\partial x}$$

for some nonnegative ν ($d\mu = \nu dt$). This is the formulation that one can also find in the existing literature (possibly with an opposite sign for the state constraint, and obvious changes in the previous expressions).

3.2 Shooting method

To compute optimal trajectories thanks to the Pontryagin Maximum Principle, the first step is to make explicit the maximization condition, at least when this is possible (otherwise this can be done numerically). A usual assumption, often satisfied, is the strict Legendre assumption, requiring that $\frac{\partial^2 H}{\partial u^2}(t, x, p, p^0, u)$ is negative definite along the reference extremal. Under this assumption, an implicit function argument gives, locally, a control expressed as a function of the state and of the adjoint. Let us assume, in what follows, that we are in the normal case, $p^0 = -1$. Plugging the resulting expression of the control in the Hamiltonian equations, and defining the *reduced (normal) Hamiltonian* by $H_r(t, x, p) = H(t, x, p, -1, u(x, p))$, we obtain that every normal extremal is solution of the reduced Hamiltonian system

$$\dot{\mathbf{x}}(t) = \frac{\partial H_r}{\partial p}(t, \mathbf{x}(t), \boldsymbol{p}(t)), \quad \dot{\boldsymbol{p}}(t) = -\frac{\partial H_r}{\partial x}(t, \mathbf{x}(t), \boldsymbol{p}(t)). \tag{17}$$

Definition 3.1. *Denoting by* $(\mathbf{x}(t, x_0, p_0), \boldsymbol{p}(t, x_0, p_0))$ *the solution of* (17) *starting at* (x_0, p_0) *for* $t = 0$*, the exponential mapping is defined by* $\exp_{x_0}(t, p_0) = \mathbf{x}(t, x_0, p_0)$.

The exponential mapping parametrizes the normal extremal flow. The abnormal extremal flow can be parametrized as well, provided an appropriate Legendre assumption holds in the abnormal case.

Remark 3.4. Let us give an example where the Hessian of the Hamiltonian is degenerate: the minimal time problem for single-input control affine systems $\dot{\mathbf{x}}(t) = F_0(\mathbf{x}(t)) + \boldsymbol{u}(t)F_1(\mathbf{x}(t))$ without constraints on controls. In that case, the maximization condition implies that the bracket $\langle \boldsymbol{p}(t), F_1(\mathbf{x}(t))\rangle$ vanishes along the extremals, and the optimal control $\boldsymbol{u}$ is singular. To compute it, we perform two successive derivations in time of the latter relation, obtaining $\langle \boldsymbol{p}(t), [F_0, F_1](\mathbf{x}(t))\rangle = 0$ and $\langle \boldsymbol{p}(t), [F_0, [F_0, F_1]](\mathbf{x}(t))\rangle + \boldsymbol{u}(t)\langle \boldsymbol{p}(t), [F_1, [F_0, F_1]](\mathbf{x}(t))\rangle = 0$, where $[\cdot, \cdot]$ is the Lie bracket of vector fields. We thus retrieve the control as a function of x and p, provided that

$$\langle \boldsymbol{p}(t), [F_1, [F_0, F_1]](\mathbf{x}(t))\rangle > 0,$$

which is the so-called strong generalized Legendre–Clebsch condition (see, e.g., [28]). Actually, under generic conditions on the vector fields, the above computation can always be performed (see [36, 37]). See Goddard example in this section for an example of this computation.

Remark 3.5. When an abnormal flow can be well defined, we then have to deal with two extremal flows (and two exponential mappings). In general, however, the abnormal flow "does not fill much space". For example, in [1, 74] it is proved that for control-affine systems without drift (satisfying the Hörmander assumption), with quadratic cost, the image of the abnormal exponential mapping has an empty interior in the state space, and is even of zero Lebesgue measure under some additional assumptions.

Remark 3.6. The Pontryagin Maximum Principle is a first-order necessary condition for optimality, asserting that if a trajectory is optimal then it should be sought among projections of extremals joining the initial point to the final target. This yields the shooting method that is described hereafter. But, before coming to the description of that method, it is interesting to observe that, conversely, the projection of a given extremal may not be (locally or globally) optimal. This important observation has led to develop second-order optimality conditions in optimal control.

In terms of the Lagrangian (in the simplified setting), considering the intrinsic second order derivative Q_T of the Lagrangian, given by

$$Q_T = \frac{\partial^2 L_T}{\partial^2 u}(\boldsymbol{u}, \psi, \psi^0)_{|\ker \frac{\partial L_T}{\partial u}},$$

a second-order necessary condition for optimality is that Q_T be nonpositive, and a second-order sufficient condition for *local* optimality is that Q_T be negative definite. Such conditions admit a number of generalizations for optimal control problems involving control and/or state constraints. It happens that, given a fixed extremal starting at (x_0, p_0), in the simplified context and under appropriate generic assumptions, the quadratic form Q_T is not degenerate (i.e., its kernel is trivial) if and only if the exponential mapping $\exp_{x_0}(t_c, \cdot)$ is not an immersion at p_0 (that is, its differential is not injective). This result, coming from symplectic considerations and generalizing the Riccati theory (see [2, 30]), yields to algorithms for computing the first conjugate time along a given extremal (see [27] for a survey). By definition, the first conjugate time along an extremal is the first time t_c at which the quadratic form Q_{t_c} has a nontrivial kernel. This means that the trajectory $x(\cdot)$ under consideration is locally optimal (in L^∞ topology) on $[0, t]$ if and only if $t < t_c$. Computing a first conjugate time amounts to computing the first time at which some determinant along the extremal vanishes. More generally, the fact that the exponential mapping is not an immersion can be translated in terms of Jacobi fields, like in Riemannian geometry.

Let us now describe the contents of the shooting method. The Pontryagin Maximum Principle states that every optimal trajectory is the projection of an extremal. After making explicit the maximization condition, the problem is reduced (for instance, in the normal case) to an extremal system of the form $\dot{\mathbf{z}}(t) = F(t, \mathbf{z}(t))$, where $\mathbf{z}(t) = (\mathbf{x}(t), \boldsymbol{p}(t))$, and initial, final, transversality conditions, are written as $R(\mathbf{z}(0), \mathbf{z}(T)) = 0$. We thus end up with a boundary value problem (BVP) of the form

$$\dot{\mathbf{z}}(t) = F(t, \mathbf{z}(t)), \quad R(\mathbf{z}(0), \mathbf{z}(T)) = 0. \tag{18}$$

Denote by $\mathbf{z}(t, z_0)$ the solution of the Cauchy problem $\dot{\mathbf{z}}(t) = F(t, \mathbf{z}(t))$, $\mathbf{z}(0) = z_0$, and set $G(z_0) = R(z_0, \mathbf{z}(T, z_0))$. The boundary value problem (18) is then equivalent to solving $G(z_0) = 0$, i.e., to finding a zero of the function G. By definition, the (single) shooting method consists of implementing a Newton-like method to find a zero of G.

The feasibility of the shooting method relies on the fact that the Jacobian of the mapping G is nonzero. According to the above remark, in the simplified case, this determinant is nonzero, i.e., the (single) shooting method is well-posed at time T, locally around p_0, if and only if the exponential mapping $\exp_{x_0}(t, \cdot)$ is an immersion at p_0, that is, if and only if T is not a conjugate time. Although this result admits generalizations to a number of contexts (see, e.g., [24, 25]), there still misses a complete conjugate time theory involving state and control constraints, in which the trajectories may have bang, singular, boundary arcs.

The single shooting method can be refined to the multiple shooting method, in which one may add a number of intermediate nodes, thus incorporating new (matching) conditions in the shooting function G. This can be useful for instance to face numerical instability issues, or to implement the shooting method in bang-bang situations where one knows in advance the number of switchings. Efficient shooting methods are available in the `HamPath` package[2], now encapsulated in the `Python` package `nutopy` of the `ct: control toolbox` project.[3] These open-source packages also contain implementations of conjugate time computations and of several homotopy routines that are particularly useful in a number of contexts (see [80] for a survey on the use of continuation methods in optimal control).

Remark 3.7. Numerically, the shooting method is the combination of a numerical integration of a differential equation with a Newton method for finding a zero of a map (the shooting function). It thus inherits of the main features of a Newton method: when it converges, the convergence is extremely fast and the result is very accurate. However, it may be difficult to initialize successfully: finding a good initial guess for z_0, in the above notations, may be a real challenge. To face with this difficulty, several possible remedies are known, such as the following, surveyed in [80] (see also the references therein):

- Since direct methods are less sensitive to the initialization, it is often successful to first run a direct approach (even with a quite rough mesh) so that, if can obtain its convergence, then the corresponding optimal solution and Lagrange multiplier can be used as an approximation of the searched extremal.

- Continuation and homotopy methods can be combined with the shooting method (and also, by the way, with direct methods): when a given problem happens to be difficult to solve, or quite ill-posed, because of some too restricted parameters or because of too constraining

[2] hampath.org
[3] ct.gitlabpages.inria.fr/gallery

terms in the dynamics, one can try to relax the optimal control problem by adding some continuation parameters in front of those terms, then run a series of shooting methods with the continuation parameters ranging iteratively from 0 to 1 (with adaptive steps if necessary).

- Geometric control gives useful information on the local or global structure of the optimal controls. For instance, one can guess in advance the number of switchings in a bang-bang strategy under some appropriate assumptions. This knowledge can then be combined with the Pontryagin maximum principle in order to drastically reduce the complexity of the shooting problem.

Much more could be said on these classical issues, but since they have already been surveyed in [80] we do not elaborate more. We next describe another powerful remedy that has emerged recently, although it relies on an old concept.

3.3 Turnpike property

Assume that f and ℓ do not depend on time. In few words, the turnpike property stipulates that, for optimal control problems in large time, under mild assumptions it is expected that the optimal solution, the optimal control and the associated adjoint remain essentially close to static values, except at the beginning and at the end of the time interval. Moreover, these static values correspond to the optimal solution of a static optimization problem. The idea is very easy to understand. When the time T is large, setting $\varepsilon = 1/T$ and $\tau = t/T = \varepsilon t$, the optimal control problem consists of determining a trajectory $y(\tau) = \mathbf{x}(t)$ and a control $v(\tau) = \boldsymbol{u}(t)$ solution of $\varepsilon y'(\tau) = f(y(\tau), v(\tau))$ and minimizing the cost $\int_0^1 \ell(y(\tau), v(\tau))\, d\tau$, under various constraints. At the formal level, we see that, when $\varepsilon \to 0$, at the limit we find the static optimization problem consisting of minimizing $\ell(y, v)$ under the constraint $f(y, v) = 0$, i.e., under the constraint of being an equilibrium of the controlled dynamics.

To give a more precise insight, let us establish the so-called exponential turnpike phenomenon in the linear-quadratic case. Let $x_0, x_1, \hat{x} \in \mathbb{R}^d$ and $\hat{u} \in \mathbb{R}^r$ be fixed. We consider the optimal control problem in fixed final time $T > 0$:

$$\begin{aligned} &\min \int_0^T \Big((\mathbf{x}(t) - \hat{x})^\top Q (\mathbf{x}(t) - \hat{x}) + (\boldsymbol{u}(t) - \hat{u})^\top R (\boldsymbol{u}(t) - \hat{u}) \Big)\, dt \\ &\quad \dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\boldsymbol{u}(t) \\ &\quad \mathbf{x}(0) = x_0,\ \mathbf{x}(T) = x_1 \end{aligned} \tag{19}$$

where $\mathbf{x}(t) \in \mathbb{R}^d$ and $\boldsymbol{u}(t) \in \mathbb{R}^r$, and where Q et R are real-valued symmetric positive definite matrices. By strict convexity, there exists a unique optimal solution $(\mathbf{x}_T, \boldsymbol{u}_T(\cdot))$ of (19). We assume that the pair (A, B) satisfies the Kalman condition. The Hamiltonian of the problem is $H(x, p, p^0, u) = \langle p, Ax \rangle + \langle p, Bu \rangle + p^0((x - \hat{x})^\top Q (x - \hat{x}) + (u - \hat{u})^\top R (u - \hat{u}))$. Let us prove that $p^0 \neq 0$. By contradiction, if $p^0 = 0$ then the condition $\frac{\partial H}{\partial u} = 0$ yields $\langle \boldsymbol{p}(t), B \rangle = 0$, and by successive derivations and using the fact that $\dot{\boldsymbol{p}}(t) = -A^\top \boldsymbol{p}(t)$, we obtain $\langle \boldsymbol{p}(t), A^k B \rangle = 0$, which raises a contradiction with the Kalman condition since $\boldsymbol{p}(t) \neq 0$. We choose then to normalize the adjoint so that $p^0 = -1/2$. The condition $\frac{\partial H}{\partial u} = 0$ yields $\boldsymbol{u}_T(t) = \hat{u} + R^{-1} B^\top \boldsymbol{p}_T(t)$, and the

extremal system is

$$\begin{aligned}\dot{\mathbf{x}}_T(t) &= A\mathbf{x}_T(t) + BR^{-1}B^\top \boldsymbol{p}_T(t) + B\hat{u}\\ \dot{\boldsymbol{p}}_T(t) &= Q\mathbf{x}_T(t) - A^\top \boldsymbol{p}_T(t) - Q\hat{x},\end{aligned}$$

i.e.,

$$\frac{d}{dt}\begin{pmatrix}\mathbf{x}_T(t)\\ \boldsymbol{p}_T(t)\end{pmatrix} = M\begin{pmatrix}\mathbf{x}_T(t)\\ \boldsymbol{p}_T(t)\end{pmatrix} + \begin{pmatrix}B\hat{u}\\ -Q\hat{x}\end{pmatrix},$$

where

$$M = \begin{pmatrix}A & BR^{-1}B^\top\\ Q & -A^\top.\end{pmatrix}.$$

Besides, the static optimization problem is

$$\min_{(x,u)\in\mathbb{R}^d\times\mathbb{R}^r,\ Ax+Bu=0}\Big((x-\hat{x})^\top Q(x-\hat{x}) + (u-\hat{u})^\top R(u-\hat{u})\Big).$$

This strictly convex problem has a unique solution $(\bar{x},\bar{u})$, associated with a normal Lagrange multiplier $(\bar{p},-1)$ (the problem is qualified because $\ker(A^\top)\cap\ker(B^\top)=\{0\}$, as a consequence of the Kalman condition). According to the Lagrange multiplier rule, which is here a necessary and sufficient condition for optimality, there exists $\bar{p}\in\mathbb{R}^d\setminus\{0\}$ such that $\bar{u}=\hat{u}+R^{-1}B^*\bar{p}$ and

$$\begin{aligned}A\bar{x} + BR^{-1}B^*\bar{p} + B\hat{u} &= 0\\ Q\bar{x} - A^*\bar{p} - Q\hat{x} &= 0,\end{aligned} \tag{20}$$

i.e.,

$$M\begin{pmatrix}\bar{x}\\ \bar{p}\end{pmatrix} + \begin{pmatrix}B\hat{u}\\ -Q\hat{x}\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}.$$

We have the following exponential turnpike property.

Proposition 3.1. *There exist constants $C>0$ and $\nu>0$, not depending on T, such that*

$$\|\mathbf{x}_T(t)-\bar{x}\| + \|\boldsymbol{u}_T(t)-\bar{u}\| + \|\boldsymbol{p}_T(t)-\bar{p}\| \le C(e^{-\nu t} + e^{-\nu(T-t)}) \qquad \forall t\in[0,T]. \tag{21}$$

Proof. Using the above optimality systems, we have

$$\frac{d}{dt}\begin{pmatrix}\mathbf{x}_T(t)-\bar{x}\\ \boldsymbol{p}_T(t)-\bar{p}\end{pmatrix} = M\begin{pmatrix}\mathbf{x}_T(t)-\bar{x}\\ \boldsymbol{p}_T(t)-\bar{p}\end{pmatrix}$$

In order to prove (21), the crucial observation is that the matrix M, which is Hamiltonian, is hyperbolic, i.e., all its eigenvalues have a nonzero real part (actually, the number of unstable modes is equal to the number of stable modes). To prove this hyperbolicity property, we start by noting that, as a consequence of the Kalman condition on (A,B), we have

$$\ker(A^\top - \xi I_n)\cap\ker(B^\top) = \{0\} \qquad \forall \xi\in\mathbb{C}. \tag{22}$$

It follows that the matrix M has no purely imaginary eigenvalue. Indeed, let $(z_1,z_2)\in\mathbb{C}^n\times\mathbb{C}^n$and let $\mu\in\mathbb{R}$ be such that $(M-i\mu)\begin{pmatrix}z_1\\ z_2\end{pmatrix}=0$. Then,

$$\begin{aligned}(A-i\mu)z_1 + BR^{-1}B^\top z_2 &= 0\\ Qz_1 - (A^\top + i\mu)z_2 &= 0,\end{aligned}$$

hence $z_1 = Q^{-1}(A^\top + i\mu)z_2$ and thus $(A - i\mu)Q^{-1}(A^\top + i\mu)z_1 + BR^{-1}B^\top z_2 = 0$. Multiplying to the left by $\bar{z}_2^\top$, we obtain $\|Q^{-1/2}(A^\top + i\mu)z_2\|^2 + \|R^{-1/2}B^\top z_2\|^2 = 0$ and hence $(A^\top + i\mu)z_2 = 0$ and $B^\top z_2 = 0$. We infer that $z_2 = 0$ by using (22). The claim is proved.

Since M is hyperbolic, there exists a real-valued square matrix P of size $2n$ such that

$$P^{-1}MP = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}$$

where all eigenvalues of the matrix M_1 have a negative real part, and all eigenvalues of M_2 have a positive real part. Now, setting $\begin{pmatrix} \mathbf{x}_T - \bar{x} \\ \boldsymbol{p}_T - \bar{p} \end{pmatrix} = P \begin{pmatrix} v \\ w \end{pmatrix}$, the extremal system gives

$$\frac{d}{dt}\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = P^{-1}MP \begin{pmatrix} v(t) \\ w(t) \end{pmatrix}$$

so $\dot{v}(t) = M_1 v(t)$ and $w(t) = M_2 w(t)$. All eigenvalues of M_1 have a negative real part, hence there exist constants $C_1 > 0$ and $\nu_1 > 0$, not depending on T, such that $\|v(t)\| \le C_1 e^{-\nu_1 t}$. For $w(t)$, we reverse time and we apply the same argument, hence $\|w(t)\| \le C_2 e^{-\nu_2(T-t)}$. Finally, (21) follows by noting that $\mathbf{x}_T(t)$ et $\boldsymbol{p}_T(t)$ are linear combinations of $v(t)$ and $w(t)$. □

The exponential turnpike property (21) says that, except near $t = 0$ and $t = T$, the optimal state, control and adjoint are exponentially close to static values, themselves corresponding to the solution of the associated static optimization problem. In [81], the proof of the above proposition is a bit different and relies on the use of the Riccati theory: actually, the matrix P is built by considering the minimal and maximal solutions of the Riccati algebraic equation, which gives an interpretation of the constants C and ν in terms of these matrices. Anyway, the argument remains very easy and it withstands a number of generalizations: to infinite dimension, to nonlinear dynamics, and also to situations where the turnpike is not restricted, as above, to a singleton but may even consist of a nontrivial set of trajectories (an example being the periodic turnpike). We refer to the chapter [50] of the Volume 1 of the present Handbook, for a recent survey on the turnpike property in optimal control, containing a number of references and commenting also on the important related notion of dissipativity. What we want to point out here is that, when the turnpike property is satisfied, we can use it to successfully initialize a variant of the shooting method.

Remark 3.8 (Variant of the shooting method)**.** In the turnpike context, we know that, in the middle of the trajectory, $\mathbf{x}_T(T/2)$ and $\boldsymbol{p}_T(T/2)$ are exponentially close to $\bar{x}$ and $\bar{p}$. In such conditions, if this is feasible, it is convenient to first compute the solution of the static problem, then to implement a variant of the shooting method by initializing it "in the middle", as follows. Using the notations of the previous section, the unknown is now $\mathbf{z}(T/2) \in \mathbb{R}^{2d}$, which is initialized at $(\bar{x}, \bar{p})$, the steady-state solution of the static optimal control problem. Then, integrating backwards the extremal system over $[0, T/2]$, we compute $\mathbf{z}(0)$; integrating forward the extremal system over $[T/2, T]$, we compute $\mathbf{z}(T)$. Finally, the unknown $\mathbf{z}(T/2)$ is tuned so that $G(\mathbf{z}(0), \mathbf{z}(T)) = 0$, thanks to a Newton method. It has been observed in [81] that this variant of the single shooting method is very efficient.

3.4 Solving the Zermelo problem by the shooting method

We consider the navigation problem of Zermelo presented in Section 1.3, with $a = 0$ and $b = 1$. We take $V_{\max} = 1$, $y_f = (20, 1)$, $h(y_2) = 3 + 0.2y_2(1 - y_2)$, where the target is the point y_f

(a slightly simpler problem than the general previously presented), while the initial point is $y_0 = (0, 0)$. We treat the minimum time case so that the velocity can be set to $V_{\max}$ and the only control is the angle. An obstacle is inserted along the unconstrained optimal trajectory:

$$\frac{(y_1 - y_{1,f}/2)^2}{a_1^2} + \frac{(y_2 - y_{2,f}/2.5)^2}{a_2^2} \leq 1$$

with $a_1 = 2$ and $a_2 = 0.1$. We use a logarithmic barrier to penalize internally the state constraint and we consider the augmented cost

$$T - \alpha \int_0^1 \ln\left(\frac{(y_1(t) - y_{1,f}/2)^2}{a_1^2} + \frac{(y_2(t) - y_{2,f}/2.5)^2}{a_2^2} - 1\right) dt \to \min .$$

For a detailed study of internal penalization, we refer to [67]. For the computation below, α is set to 10^{-3}. To obtain more accurate results, one could perform a numerical continuation on the penalty parameter α as is customary for interior methods [69]. We apply the maximum principle which leads to the (normal) maximized Hamiltonian

$$H(y, p) = V_{\max}\sqrt{p_1^2 + p_2^2} + p_1 h(y_2) + \alpha \ln\left(\frac{(y_1 - y_{1,f}/2)^2}{a_1^2} + \frac{(y_2 - y_{2,f}/2.5)^2}{a_2^2} - 1\right) - 1. \quad (23)$$

The shooting problem then consists in finding the initial value $\boldsymbol{p}(0) \in \mathbb{R}^2$ so that, integrating the flow of the maximized Hamiltonian, the target y_f is reached. Moreover, the equation $H = 0$ is added to accomodate the fact that the final time T is free. Depending on the initialization of the shooting method, two solutions are obtained, see Figure 3. One of them is clearly a local minimizer as one can check comparing the numerical final times. An interesting approach in such a situation is to rely on an HJB solver (see Section 4) to retrieve an initial guess of the adjoint state (and of the final time) that avoid strict local minima. It can indeed be checked numerically that the HJB solution on the previous data allows shooting to converge towards the global minimizer on the internally penalized problem. This illustrates how one can leverage the strengths of both HJB and shooting: while HJB might not be able to produce a high precision numerical control, it will allow to select a proxy for the adjoint of the global minimum and for the associated value, good enough to ensure convergence of the shooting method (the difficult issue for indirect methods) towards a precise numerical solution.

Remark 3.9. A refinement of this computation could involve a continuation (or differential homotopy) on the size of the obstacle. Path following methods would indeed allow to track the two branches associated with the global and strictly local minimum, then to decide which one is optimal for each size of the obstacle by comparing the associated final times.

Another relevant observation on this problem is related to the turnpike property described in Section 3.3. Although, the problem is not linear-quadratic, one can easily guess the role played by this property for "large" (in terms of y_1) final conditions. We keep the same data for $V_{\max}$ and for the current, remove the obstacle, and now target $y_f = (200, 1)$. For such a distant target, one expects the optimal control to use the "fast lane" so that, most of the time, y_2 remains close to $1/2$ where the current h is maximum. So the guess would be that, for a large part of the trajectory, $y_2 \simeq 1/2$, $u_2 = \dot{y}_2 \simeq 0$, $u_1 \simeq 1$, and $T \simeq (y_{1,f} - y_{1,0})/(1 + h(1/2))$. The maximized (normal) Hamiltonian is (compare (23))

$$H(y, p) = V_{\max}\sqrt{p_1^2 + p_2^2} + p_1 h(y_2) - 1,$$

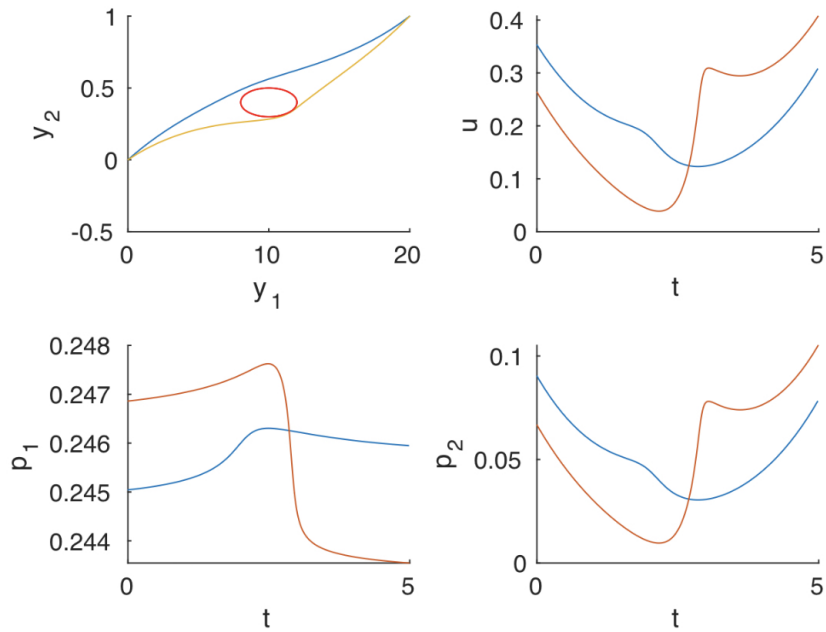

Figure 3: Zermelo problem with obstacle: result of the shooting method. The state, costate and control are portrayed for the two solutions (penalisation parameter $\alpha = 10^{-3}$). The first solution (in blue) is slightly shorter ($T \simeq 4.98$) than the second one (yellow, $T \simeq 4.99$).

so that, on $H = 0$, $p \simeq (1/(1 + h(1/2, 0))$. These approximations can be used to initialize the variant of the shooting method described in Remark 3.8. It is straightforwardly implemented in `Julia` according to[4]

```
# Regular maximized Hamiltonians and associated flow
H(y, p, u) = -1.0 + Vmax*p'*u + p[1]*h(y[2])
ur(p) = p / sqrt(p[1]^2+p[2]^2)
Hr(y, p) = H(y, p, ur(p))
fr = Flow(Hr)

# Shooting function
function shoot(y1, p1, tf)

    yy0, p0 = fr(tf/2.0, y1, p1, t0)
    yyf, pf = fr(tf/2.0, y1, p1, tf)
    s = zeros(eltype(y1), 5)
    s[1:2] =  yy0-y0
    s[3:4] = (yyf-yf) ./ yf
    s[5] = Hr(y1, p1)

    return s
```

[4]The code is available and executable online at ct.gitlabpages.inria.fr/gallery

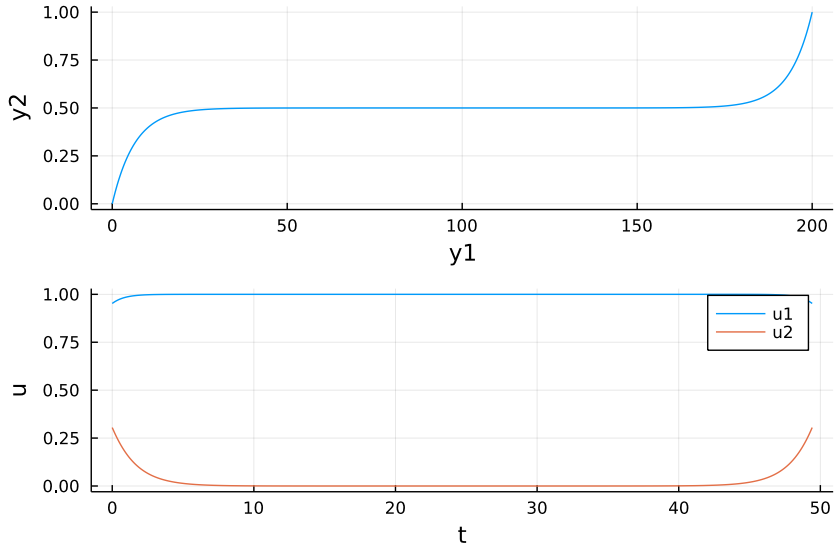

Figure 4: Zermelo problem, turnpike property: result of the variant of the shooting method. The state and control are portrayed, illustrating that most of the time is spent close to $y_2 = 1/2$ where the current is maximum. Accordingly, $u_2 \simeq 0$ on the corresponding control subarc.

```
end
```

The structure of the optimal solution is as expected, see Figure 4.

Remark 3.10. Further analysis is required to tackle the case of "strong" currents. When there exist zones where the drift cannot be compensated by the control, abnormal trajectories come into play and discontinuities of the value function (minimum time) are observed. See [29] for a detailed treatment of such cases.

3.5 Solving the Goddard problem by the shooting method

We chain the result of Section 2.4, obtained by a direct method, with a shooting method. The previous numerical solution, although not very accurate, has captured the structure of the solution (more precisely, of what one may hope to be a local minimizer, at least). This knowledge allows us to define the appropriate shooting function, namely one that combines four arcs (bang-singular-constrained-bang), each one being the flow of a relevant Hamiltonian. On can then leverage the accurate knowledge (including the Hamiltonian character) gained on each subarc by means of Pontryagin maximum principle to obtain a very accurate solver. In general, knowing the structure alone is not enough to actually solve the problem. One also needs a good initial guess for the zero of the shooting function, which turns to be also provided by the previous direct solver. In order to set up our shooting, we rely on the maximum principle and observe that the control can be either bang, singular of boundary. Indeed, system (7a)-(7c) is affine in the control and can be written according to

$$\dot{\mathbf{x}}(t) = F_0(\mathbf{x}(t)) + \boldsymbol{u}(t)F_1(\mathbf{x}(t)), \quad \boldsymbol{u}(t) \in [0,1],$$

with $\mathbf{x} = (\mathbf{r}, \boldsymbol{v}, \mathbf{m}) \in \mathbb{R}^3$ and vector fields that we shall define in `Julia`, completing the code started Section 2.4:

```
# Dynamics
function F0(x)
    r, v, m = x
    D = Cd * v^2 * exp(-β*(r-1.0))
    F = [ v, -D/m-1.0/r^2, 0.0 ]
    return F
end

function F1(x)
    r, v, m = x
    F = [ 0.0, Tmax/m, -b*Tmax ]
    return F
end
```

In order to deal with the state constraint $g(\mathbf{x}(t)) := v_{\max} - \boldsymbol{v}(t) \geq 0$, we follow Remark 3.3 (note that we use the opposite sign for the constraint) and introduce the Hamiltonian (where the constraint has been directly adjoined, see [60])

$$H(x, p, u, \mu) = H_0(x, p) + uH_1(x, p) + \mu g(x),$$

where $H_i(x, p) := \langle p, F_i(x)\rangle$ are the Hamiltonian lifts of the aforementioned fields. The maximization condition implies that u is bang (0 or 1) whenever H_1 is not zero. Besides, whenever the state constraint is not active, the associated non-negative multiplier μ vanishes (complementarity condition). As a result, bang arcs are obtained by computing the flow of either H_0 (case $H_1 < 0$, $u = 0$) or $H_0 + H_1$ (case $H_1 > 0$, $u = 1$). Conversely, when H_1 vanishes identically, assuming the state constraint is not active, as indicated in Remark 3.4 one can differentiate a.e. two times H_1 to retrieve the singular control provided the length three Poisson bracket $H_{101} := \{H_1, \{H_0, H_1\}\}$ is not zero (singular of *order one*; see, e.g., [28]):

$$u_s(x, p) = -\frac{H_{001}}{H_{101}}.$$

(Same notation used for H_{001}.) Plugging this dynamic feedback control into the original Hamiltonian (with $\mu = 0$) defines the singular Hamiltonian

$$H_s(x, p) := H(x, p, u_s(x, p), \mu = 0)$$

whose flow coincides with the extremal flow on $\Sigma' := \{H_1 = H_{01} = 0\}$ (see, e.g., [2]). Along a boundary arc where the state constraint is activated, if the control is interior ($\boldsymbol{u}(t) \in (0, 1)$), H_1 must also vanish. Moreover, if the constraint is of order one (that is if the control appears when the equality $g(\mathbf{x}(t)) = 0$ is differentiated once) which gives in this case (Lie derivative of the constraint along F_1) $F_1 \cdot g \neq 0$, the interiority of the boundary control implies that the contact with the constraint is transverse at the exit time t_3 (where, in our case, a bang arc $u = 0$ is joined): $\dot{g}(t_3+) \neq 0$, and there is no jump on the adjoint [60]. Denoting t_2 the entry point, on (t_1, t_2) one retrieves the boundary control by differentiating once $g(\mathbf{x}(t)) = 0$ as $0 = \dot{g} = F_0 \cdot g + \boldsymbol{u} F_1 \cdot g$, which implies that $u_b(x) = -\frac{F_0 \cdot g}{F_1 \cdot g}$ under the two previous assumptions. Similarly, differentiating once $H_1 = 0$ allows to compute the multiplier μ as $0 = \dot{H}_1 = \{H_0 + \boldsymbol{u} H_1 + \mu g, H_1\} = H_{01} - \mu F_1 \cdot g$, so that $\mu_b(x, p) = \frac{H_{01}}{F_1 \cdot g}$. The relevant flow is an integral curve of the boundary Hamiltonian

$$H_b(x, p) := H(x, p, u_b(x), \mu_b(x, p)).$$

All in all, the symbolic-numeric framework allows to define everything in terms of the vector fields and of the constraint:

```
# Computation of singular control of order 1
H0(x, p) = p' * F0(x)
H1(x, p) = p' * F1(x)
H01 = Poisson(H0, H1)
H001 = Poisson(H0, H01)
H101 = Poisson(H1, H01)
us(x, p) = -H001(x, p)/H101(x, p)

# Computation of boundary control
g(x) = vmax-x[2] # vmax - v ≥ 0
ub(x) = -Lie(F0, g)(x) / Lie(F1, g)(x)
μb(x, p) = H01(x, p) / Lie(F1, g)(x)

# Hamiltonians (regular, singular, boundary) and associated flows
H(x, p, u, μ) = H0(x, p) + u*H1(x, p) + μ*g(x)
Hr(x, p) = H(x, p, 1.0, 0.0)
Hs(x, p) = H(x, p, us(x, p), 0.0)
Hb(x, p) = H(x, p, ub(x), μb(x, p))
```

Then, to integrate the Hamiltonians to obtain the flows and define the shooting function in terms of the initial adjoint p_0, the entry point of the singular arc t_1, the entry point of the boundary arc t_2, the exit point t_3, and the free final time T:

```
f0 = Flow(H0)
fr = Flow(Hr)
fs = Flow(Hs)
fb = Flow(Hb)

# Shooting function
function shoot(p0, t1, t2, t3, tf)

    x1, p1 = fr(t0, x0, p0, t1)
    x2, p2 = fs(t1, x1, p1, t2)
    x3, p3 = fb(t2, x2, p2, t3)
    xf, pf = f0(t3, x3, p3, tf)
    s = zeros(eltype(p0), 7)
    s[1:2] = pf[1:2] - [ 1.0, 0.0 ]
    s[3] = xf[3] - mf
    s[4] = H1(x1, p1)
    s[5] = H01(x1, p1)
    s[6] = g(x2)
    s[7] = H0(xf, pf)

    return s
end
```

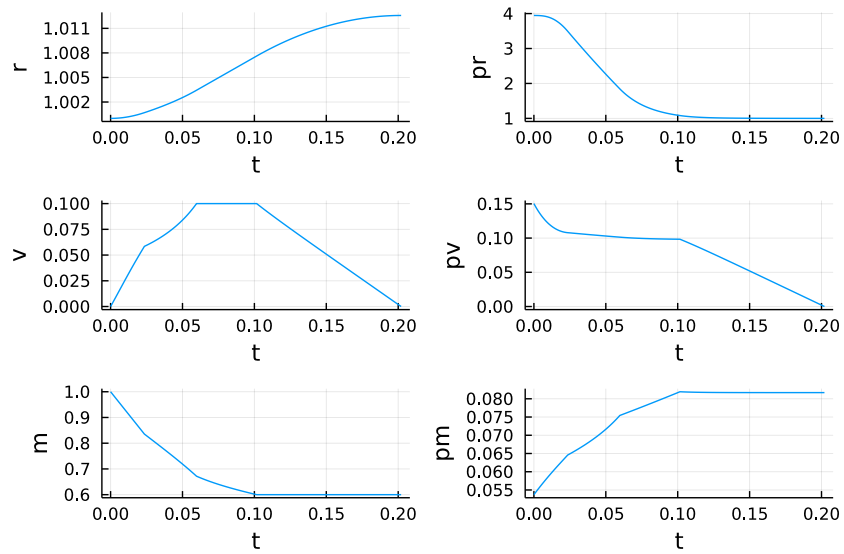

Figure 5: Goddard problem: result of the shooting method, states and costates. The values are in line with those obtained by the direct method. They are computed by concatenating the integration of the four Hamiltonian flows involved.

It is straightforward to retrieve initial guesses for these unknowns from the direct code solution, most notably by retrieving the Lagrange multipliers from the optimizer through `JuMP` interface (note the minus sign to accomodate the convention on the adjoint state in contrast with the one on Lagrange multipliers):

```
p = -[ [ dual(con_dr[i]), dual(con_dv[i]), dual(con_dm[i]) ] for i in 1:N ]
```

Moreover, in view of these results, the constraint on the final mass is assumed to be active (which is in accordance with the final zero bang arc). Automatic differentiation can also be used to compute the derivative of the shooting function (several `Julia` backends to do this are available and include differentiating properly calls to ODE solvers). See Figures 5 and 6 for the associated numerical simulations, while the code itself is available and executable online.[5]

4 Hamilton–Jacobi–Bellman approach

The systematic study of optimal control problems dates back to the late 1950s, and one major tool is Dynamical Programming and Hamilton–Jacobi–Bellman (HJB) approach. This approach describes the optimal control problem via the so-called *value function* $\mathcal{V} : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$, which associates to any initial condition the optimal value of the control problem, and is defined accordingly as

$$\begin{aligned} \mathcal{V}(t,x) \quad := \quad & \inf \Big\{ \varphi(\mathbf{x}(T)) + \int_t^T \ell(s,\mathbf{x}(s),\boldsymbol{u}(s))ds \;\Big| \\ & (\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x),\; g(\mathbf{x}(s)) \leq 0 \quad \text{for all } \tau \in [s,T], \quad \text{and } g_f(\mathbf{x}(T)) \leq 0 \Big\}. \end{aligned} \tag{24}$$

[5] ct.gitlabpages.inria.fr/gallery

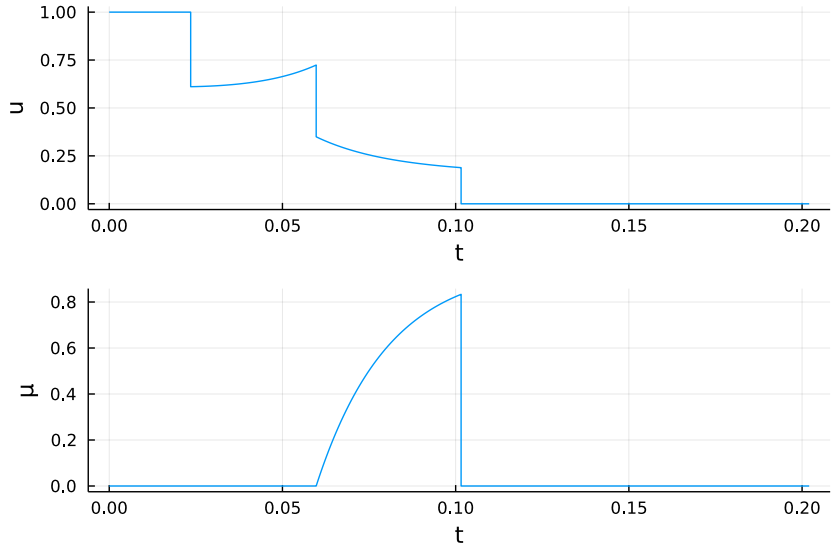

Figure 6: Goddard problem: result of the shooting method, control and state constraint multiplier. The bang-singular-constrained-bang structure of the control is very accurately determined, while the state constraint multiplier is positive only along the boundary arc.

It is known that the value function $\mathcal{V}$ can be characterized as the unique solution, in a suitably weak sense, of a Hamilton–Jacobi type equation [12]. Starting from the knowledge of the value function, which is typically obtained via numerical approximation, it is possible to reconstruct the optimal solution in feedback form, i.e., with an optimal control given as a function of the current state.

4.1 Unconstrained Bolza control problems

In this section we present first some classical results of HJB approach when the optimal control problems is free of state constraints (i.e., $g = g_f \equiv 0$). The control problem is described by the unconstrained value function

$$\mathcal{V}(t,x) = \inf \left\{ \varphi(\mathbf{x}(T)) + \int_t^T \ell(s, \mathbf{x}(s), \boldsymbol{u}(s)) ds \;\middle|\; (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x) \right\}. \tag{25}$$

When $(\mathbf{H_4})$ is satisfied, the control problem admits a solution. The Gronwall estimate on the trajectories and the Lipschitz regularity of the cost functions ensure that the value function, although in general non-differentiable, enjoys itself a Lipschitz continuity property.

Proposition 4.1. *Under* $(\mathbf{H_0})$-$(\mathbf{H_2})$*, the value function* $\mathcal{V}$ *is locally Lipschitz continuous.*

To deal with the lack of smoothness, two important tools have been developed: the theory of viscosity solutions and the non-smooth analysis. The theory of viscosity solutions for non-linear Hamilton–Jacobi equations, introduced in the early 1980s by Crandall–Lions [43, 44] and Crandall–Evans–Lions [42], allows to define generalized solutions to broad classes of nonlinear partial differential equations, including the HJB equations of optimal control problems. We refer also to the books [13, 12] for a more complete introduction to this theory. Another important

tool is the non-smooth analysis, which addresses to differential analysis for non-smooth functions. We refer the reader to [8, 39, 38, 84] for an introduction of the theory and its applications.

4.1.1 Dynamic programming and Hamilton–Jacobi–Bellman equation

The fundamental idea of Dynamic Programming is that the value function $\mathcal{V}$ satisfies a functional equation, often called the *Dynamic Programming Principle* (DPP).

Proposition 4.2 (Dynamic Programming Principle)**.** *Assume* $(\mathbf{H_0})$–$(\mathbf{H_2})$*, and denote by* $\mathbf{x}^{\boldsymbol{u}}(s)$ *the solution of* (1) *for a given control* $\boldsymbol{u}$*, and such that* $\mathbf{x}^{\boldsymbol{u}}(t) = x$*. Then, for all* $x \in \mathbb{R}^d$*,* $t \in [0, T]$ *and* $h \in [0, T - t]$*, the value function* $\mathcal{V}$ *satisfies the equality*

$$\mathcal{V}(t, x) = \inf_{\boldsymbol{u} \in \mathcal{U}} \left\{ \int_t^{t+h} \ell(s, \mathbf{x}^{\boldsymbol{u}}(s), \boldsymbol{u}(s)) ds + \mathcal{V}(t + h, \mathbf{x}^{\boldsymbol{u}}(t + h)) \right\}. \tag{26}$$

This principle provides two properties, called sub- and super-optimality, which are defined as follows. For any function $V : [0, T] \times \mathbb{R}^d \to \mathbb{R}$,

1. we say that V satisfies the *super-optimality principle* if for any $t \in [0, T]$, $x \in \mathbb{R}^d$, there exists $(\mathbf{x}^{\boldsymbol{u}}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x)$ such that

$$V(t, x) \geq V(t + h, \mathbf{x}^{\boldsymbol{u}}(t + h)) + \int_t^{t+h} \ell(s, \mathbf{x}^{\boldsymbol{u}}(s), \boldsymbol{u}(s))\, ds, \ \forall\, h \in [0, T - t];$$

2. we say that V satisfies the *sub-optimality principle* if for any $t \in [0, T]$, $x \in \mathbb{R}^d$, and $(\mathbf{x}^{\boldsymbol{u}}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x)$,

$$V(t, x) \leq V(t + h, \mathbf{x}^{\boldsymbol{u}}(t + h)) + \int_t^{t+h} \ell(s, \mathbf{x}^{\boldsymbol{u}}(s), \boldsymbol{u}(s))\, ds, \ \forall\, h \in [0, T - t].$$

In principle, once chosen a "small" time increment h, the DPP allows to compute the value function at the point (t, x) by splitting the trajectories at time $t + h$ and starting with the position of the trajectory $y_{t,x}$ at time $t + h$. As it will be seen later on, it is possible to construct numerical schemes based on this idea to compute an approximation of the value function.

Under the assumption of differentiabillity for the function $\mathcal{V}$, we can derive from the DPP its infinitesimal version, the *Hamilton–Jacobi–Bellman* equation

$$-\mathcal{V}_t(t, x) + H(t, x, D\mathcal{V}(t, x)) = 0 \quad \text{for } (t, x) \in (0, T) \times \mathbb{R}^d, \tag{27a}$$

$$\mathcal{V}(T, x) = \varphi(x) \qquad \text{for } x \in \mathbb{R}^d, \tag{27b}$$

where the Hamiltonian is given by

$$H(t, x, q) = \sup_{u \in U} \big\{ - f(t, x, u) \cdot q - \ell(t, x, u) \big\}. \tag{28}$$

In general, as mentioned before, neither $\mathcal{V}$ is differentiable, nor the nonlinear equation (27) is expected to admit a classical solution. These problems are circumvented by the theory of viscosity solutions and the non-smooth analysis, see [38, 12, 84].

4.1.2 Theory of viscosity solutions for Hamilton–Jacobi–Bellman equations

First, we recall the definition of viscosity solution for HJB equations (see [43, 44, 13, 12]).

Definition 4.1 (Viscosity solution)**.** *Let $V : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$.*

(i) We say that V is a viscosity supersolution if V is lower semicontinuous (lsc) and for any $\phi \in C^1((0,T) \times \mathbb{R}^d)$ and $(t_0, x_0) \in (0,T) \times \mathbb{R}^d$ local minimum point of $V - \phi$, we have

$$-\phi_t(t_0, x_0) + H(t_0, x_0, D\phi(t_0, x_0)) \geq 0.$$

(ii) We say that V is a viscosity subsolution if V is upper semicontinuous (usc) and for any $\phi \in C^1((0,T) \times \mathbb{R}^d)$ and $(t_0, x_0) \in (0,T) \times \mathbb{R}^d$ local maximum point of $V - \phi$, we have

$$-\phi_t(t_0, x_0) + H(t_0, x_0, D\phi(t_0, x_0)) \leq 0.$$

(iii) We say that V is a viscosity solution if it is both a viscosity supersolution and a viscosity subsolution and the final condition is satisfied:

$$V(T, x) = \varphi(x) \text{ in } \mathbb{R}^d.$$

Remark 4.1. There are also some equivalent definitions which are more local using the super- and sub-differentials, this meaning that the differentials of the test functions can be replaced by some weak differentials of the viscosity solution. See [13, 12] for the definition using the Dini-differentials and [38] for the definition using the proximal differentials.

Theorem 4.1. *Suppose that* $(\mathbf{H_0})$-$(\mathbf{H_2})$ *hold. Then the value function $\mathcal{V}$ is the unique viscosity solution of* (27) *in the sense of Definition 4.1.*

The first statement of this theorem is that $\mathcal{V}$ is a viscosity solution of (27). The proof of this claim relies on the regularity of the value function (continuity) and on the dynamic programming principle [12]. The theorem also claims that the value function is the unique solution of (27). This is the consequence of the following equivalences that can be established by non-smooth analysis (see [39, 84] for the proof):

$$\begin{cases} V : [0,T] \times \mathbb{R}^d \text{ is usc,} \\ V \text{ satisfies the sub-optimality principle} \end{cases} \iff V \text{ is a viscosity sub-solution of (27a),}$$

$$\begin{cases} V : [0,T] \times \mathbb{R}^d \text{ is lsc,} \\ V \text{ satisfies the super-optimality principle} \end{cases} \iff V \text{ is a viscosity super-solution of (27a).}$$

From the point of view of PDEs and viscosity theory, one can also obtain uniqueness by using a general theorem on the comparison principle which can be stated as follows.

Theorem 4.2. *Let V_1 be a subsolution of* (27) *and V_2 be a supersolution of* (27) *with $V_1(T,x) \leq \varphi(x) \leq V_2(T,x)$ for $x \in \mathbb{R}^d$. Then for any $t \in [0,T]$, $x \in \mathbb{R}^d$,*

$$V_1(t,x) \leq V_2(t,x).$$

The classical proof of the comparison principle is based on the variable doubling technique. We refer to [43, 44, 13, 12] for more details.

4.2 Other unconstrained control problems and their HJB formulation

In addition to the Bolza problem, which has been taken here as a model, various other formulations have been considered for optimal control problems, in particular without a final time, or in which the final time is itself a parameter to be optimally chosen. We briefly review some of these formulations, while a more extensive discussion can be found, for example, in [12].

Minmax control problems. In this class of control problems, the cost is not defined in integral form. More precisely, the control problem reads as

$$\mathcal{V}^{\#}(t,x) := \min\Big\{\varphi(\mathbf{x}(T)) \bigvee \max_{s\in[t,T]} \Psi(s,\mathbf{x}(s)) \Big| (\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x)\Big\}.$$

Here, the cost function is the maximum between the final cost and a maximum running cost along the trajectory. Relation between minmax control problems and state constrained control problems have been noticed and exploited to derive necessary conditions of optimal trajectories, see [84, Chapter 9]. In Hamilton–Jacobi approach, minmax control problems have been also analyzed in [15]. It has been shown that the value function is Lipschitz continuous and can be characterized by the HJB variational inequality

$$\min\big(\partial_t \mathcal{V}^{\#}(t,x) + H^{\#}(t,x,D\mathcal{V}^{\#}(t,x)), \mathcal{V}^{\#}(t,x) - \Psi(t,x)\big) = 0 \quad \text{on } [0,T)\times\mathbb{R}^d, \tag{29a}$$

$$\mathcal{V}^{\#}(T,x) = \varphi(x) \vee \Psi(T,x) \quad \text{for } x\in\mathbb{R}^d. \tag{29b}$$

In this inequality, the running cost function plays the role of an "obstacle". The value function satisfies $\mathcal{V}^{\#}(t,x) \geq \Psi(t,x)$. It is also a super-solution of the equation:

$$\partial_t v(t,x) + H^{\#}(t,x,D_x v(t,x)) = 0 \quad \text{on } [0,T)\times\mathbb{R}^d. \tag{30}$$

However, $\mathcal{V}^{\#}$ is a sub-solution of (30) only in open sets where $\mathcal{V}^{\#} < \Psi$.

Infinite horizon. *Infinite horizon problems* are intended to model optimal control strategies in the long-time behaviour. Assume that both the dynamics f and the running cost ℓ do not depend on time, so that

$$\dot{\mathbf{x}}(s) = f(\mathbf{x}(s),\boldsymbol{u}(s)), \tag{31}$$

with the initial condition $\mathbf{x}(0) = x$. The *discounted infinite horizon* cost functional is defined as

$$J(x,\boldsymbol{u}) = \int_0^\infty \ell(\mathbf{x}(s),\boldsymbol{u}(s))e^{-\lambda s}ds,$$

where, in addition to the basic assumptions, we require that $\lambda > 0$, and, for simplicity, that ℓ is uniformly bounded. According to the definition of the cost functional, it is possible to define a value function, which will depend in this case on x alone:

$$\mathcal{V}^{\infty}(x) := \inf\Big\{J(x,\boldsymbol{u}) \Big| (\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,\infty)}(x)\Big\}.$$

The value function can still be characterized as the viscosity solution of a (stationary) Hamilton–Jacobi–Bellman equation, which takes the form

$$\lambda\mathcal{V}^{\infty}(x) + H(x,D\mathcal{V}^{\infty}(x)) = 0, \tag{32}$$

in which $x \in \mathbb{R}^d$, and the Hamiltonian function H is defined by (28). As for the regularity of the value function, the basic assumptions imply uniform continuity of the value function. Hölder regularity of $\mathcal{V}^\infty$ holds under the assumptions of boundedness for ℓ, and global Lipschitz continuity for both ℓ and f; in addition, $\mathcal{V}^\infty$ is itself globally Lipschitz continuous if λ is larger than the Lipschitz constant of f (see [12]).

Free final time. In *free final time problems*, also termed as *optimal stopping time problems*, the endtime θ of the control interval is itself a free parameter to be chosen in an optimal way. In the simplest case, the dynamics is set in the form (31) and the cost functional in the form

$$J(x, (\boldsymbol{u}, \theta)) = \int_0^\theta \ell(\mathbf{x}(s), \boldsymbol{u}(s)) e^{-\lambda s} ds + e^{-\lambda\theta} \varphi(\mathbf{x}(\theta)).$$

Accordingly, the value function is defined as

$$\mathcal{V}^f(x) := \inf \left\{ J(x, (\boldsymbol{u}, \theta)) \,\middle|\, (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[0,\infty)}(x), \theta \geq 0 \right\}.$$

In this case, the HJB equation is stationary, but comes in the form of an obstacle problem:

$$\max\left(\lambda \mathcal{V}^f(x) + H(x, D\mathcal{V}^f(x)), \mathcal{V}^f(x) - \varphi(x)\right) = 0.$$

The state space is then split in two (possibly overlapping) sets: in the first one the first argument of the max vanishes, so that

$$\lambda \mathcal{V}^f(x) + H(x, D\mathcal{V}^f(x)) = 0,$$

and the optimal control requires to keep the system evolving; in the second set the second argument of the max vanishes, and therefore

$$\mathcal{V}^f(x) = \varphi(x).$$

As soon as the state of the system enters this set, the optimal strategy requires to stop the system, paying the stopping cost $\varphi(\mathbf{x}(\theta))$. Under the basic assumptions, the value function is uniformly continuous (see [12]).

Minimum time. For simplicity, we assume again in this paragraph that the dynamics does not depend explicitly on time, that is $f(s, x, u) = f(x, u)$. For minimum time function with time-dependent dynamics, we refer to [18] and the references therein.

In *minimum time problems*, the goal is to drive the state of the system, in the shortest time, to the final closed set $\mathcal{C} := \{x \in \mathbb{R}^d, g_f(x) \leq 0\}$ (called the target). In general, the possibility of driving the state to the target in finite time may not be ensured for each initial state; this leads to define the so-called *backward reachable set* $\mathcal{R}$ as the set of initial states x which can be driven to the target in finite time. The minimum time control problem is formulated as

$$\mathcal{T}(x) := \inf \left\{ t \,\middle|\, g_f(\mathbf{x}(t)) \quad (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[0,\infty)}(x) \right\}, \tag{33}$$

with the convention that $\mathcal{T}(x) = +\infty$ when there is no trajectory that starts from x and reaches the target in finite time. It is not difficult to show that the value function satisfies a dynamic programming principle (DPP) in the form

$$\mathcal{T}(x) = \inf \left\{ \mathcal{T}(\mathbf{x}(h)) + h, \quad (\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[0,\infty)}(x) \right\}.$$

When the backward reachable set is open and the minimum time function is continuous, then the DPP leads to a characterization of $\mathcal{T}$ by the HJB equation

$$\sup_{u\in U}\big\{-f(x,u)\cdot D\mathcal{T}(x)\big\}=1,$$

complemented with the boundary condition

$$\mathcal{T}(x)=0 \ \text{ for } x\in\mathcal{C} \qquad \text{and} \qquad \lim_{x\to z}\mathcal{T}(x)=+\infty \text{ for } z\in\partial\mathcal{R}.$$

The drawback of this formulation is that the reachable set should be known in advance. An alternative formulation makes use of the so-called *Kružkov transformation*

$$\mathcal{V}(x)=\begin{cases}1 & \text{if } \mathcal{T}(x)=+\infty\\ 1-e^{-\mathcal{T}(x)} & \text{else,}\end{cases}$$

which solves (yet when the time function $\mathcal{T}$ is continuous) an auxiliary infinite horizon problem of the form (32), with $\lambda=1$, $\ell\equiv 1$, and the boundary condition

$$\mathcal{V}(x)=0 \quad \text{for } x\in\mathcal{C}.$$

In this case, the reachable set is obtained as a byproduct of the computation of the value function $\mathcal{V}$, as

$$\mathcal{R}=\big\{x\in\mathbb{R}^d \mid \mathcal{V}(x)<1\big\}.$$

Continuity of the function $\mathcal{T}$ is closely related to controllability properties satisfied by the system in a neighbourhood of $\mathcal{C}$ and more precisely, to the so-called *Small-Time Local Controllability*, see [12, Chapter IV]. When the target is smooth enough (for instance, assume here that $g_f:\mathbb{R}^d\to\mathbb{R}$ is C^2) with a compact boundary, a necessary condition for the continuity of the minimum time function is given by the condition

$$\min_{u\in U} f(x,u)\cdot\nabla g_f(x)\le 0. \tag{34}$$

This condition is restrictive and excludes a large class of systems with drift. An equivalent condition to the Lipschitz continuity of the minimal time function is the *Petrov condition*

$$\min_{u\in U} f(x,u)\cdot\nabla g_f(x)< 0, \tag{35}$$

which is even more restrictive than (34). From a geometrical point of view, Petrov condition states that at every point of a neighborhood of the target there exists an admissible control such that the corresponding trajectory points towards the target.

In general, when g_f satisfies assumption ($\mathbf{H_3}$), the set $\mathcal{C}$ is closed and the controllability conditions (34)-(35) might not be satisfied. A practical approach to compute the minimum time function and the corresponding optimal trajectories is based on the *level set method*, introduced by Osher and Sethian [70]. Consider the final cost function as

$$\Phi(x):=\max\Big(g_{1,f}(x),\cdots,g_{m_f,f}(x)\Big) \quad \text{for } x\in\mathbb{R}^d,$$

where $(g_{1,f}, \cdots, g_{m_f,f})$ are the components of the function g_f. The level set approach consists of considering the value function V associated to the *Mayer problem* with final cost Φ, defined by

$$V(t,x) := \min\{\Phi(\mathbf{x}(t)),\ (\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,t]}(x)\}. \tag{36}$$

The value function V can be characterized by an HJB equation as in Theorem 4.1. This function is Lipschitz continuous, while the minimum time function may be discontinuous. Besides, one can notice immediately that if $V(t,x) \leq 0$, then there exists an admissible pair $(\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,t]}(x)$ such that $\Phi(\mathbf{x}_x^{\boldsymbol{u}}(t))$, which means that $(\mathbf{x},\boldsymbol{u})$ satisfies the final constraint. More precisely, the set of positions from where it is possible to reach the target at time t is given by

$$\mathcal{R}(t) := \{x \in \mathbb{R}^d \ \mid\ \exists(\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,t]}(x), g_f(\mathbf{x}(t)) \leq 0\} = \{x \in \mathbb{R}^d \mid V(t,x) \leq 0\}.$$

Therefore, the value function V gives valuable information on the minimum time function and the corresponding trajectories, without any controllability assumption.

Theorem 4.3. *Suppose that* $(\mathbf{H_0})$-$(\mathbf{H_3})$ *hold. Let* V *be the value function associated to problem* (36)*. Then, for every* $x \in \mathbb{R}^d$*,*

$$\mathcal{T}(x) := \min\{t \mid V(t,x) \leq 0\}. \tag{37}$$

Furthermore, for $x \in \mathbb{R}^d$*, any optimal trajectory for the minimal time problem* (33) *is also an optimal trajectory of the control problem* (36) *where* t *is fixed to the minimum time given by* (37)*.*

The level set approach provides an effective way to compute the minimum time to reach a target without assuming any specific regularity. Moreover, it has been shown that the level set method can be generalized to minimum time problems with state constraints (for instance, the case when the trajectory should avoid some obstacles, see Example 1). In this case, the minimum time function is defined as

$$\mathcal{T}(x) := \inf\left\{t \ \middle|\ \exists(\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,\infty)}(x) \ \text{ with } g_f(\mathbf{x}(t)) \leq 0, \ \text{ and } g(\mathbf{x}(s)) \leq 0 \text{ on } [0,t]\right\}.$$

To use the level set approach in this context, the definition of the value function should be adapted and defined as

$$V(t,x) := \min\left\{\Phi(\mathbf{x}(t)) \bigvee \max_{s\in[0,t]} \Psi(\mathbf{x}(s)) \ \mid\ (\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,t]}(x)\right\},$$

with $\Psi(x) := \max(g_1(x), \ldots, g_{m_g}(x))$ and $\Phi(x) := \max(g_{f,1}(x), \ldots, g_{f,m_f}(x))$ for every $x \in \mathbb{R}^d$. Here, the value function V is again Lipschitz continuous, while the minimum time function may be discontinuous. Besides, if $V(t,x) \leq 0$, then there exists an admissible pair $(\mathbf{x},\boldsymbol{u}) \in \mathbb{X}_{[0,t]}(x)$ such that $\Phi(\mathbf{x}(t)) \leq 0$ and $\max_{s\in[0,t]} \Psi(\mathbf{x}(s)) \leq 0$, which means that $(\mathbf{x},\boldsymbol{u})$ satisfies the final and pointwise state constraints. With this new definition of the value function V, the statement of Theorem 4.3 remains valid in the case with obstacles. In particular, the minimum time value and the corresponding optimal trajectories can be obtained form the value function V without assuming any controllability hypothesis. Finally, notice that the value function V corresponds to a minmax problem, and its characterization is given by the HJB inequality (29).

4.3 Constrained Bolza problems

In this section, we consider a control problem with state constraints. We denote by $\mathcal{K}$ the set of constraints $\mathcal{K} := \{x \in \mathbb{R}^d, g(x) \leq 0\}$, and define the set of admissible trajectories by

$$\mathbb{X}^g_{[t,T]}(x) := \big\{(\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[t,T]}(x) \ \mid \ g(\mathbf{x}(s)) \leq 0 \text{ for } s \in [t,T]\big\}.$$

We adopt the convention $\mathcal{V}(t,x) = +\infty$, when the set of admissible trajectories is empty, i.e., $\mathbb{X}^g_{[t,T]}(x) = \emptyset$. Similarly to the unconstrained case, the value function $\mathcal{V}$ satisfies a dynamic programming principle that can be stated as follows.

i) For all $x \in \mathcal{K}$,

$$\mathcal{V}(T,x) = \varphi(x).$$

ii) Dynamic programming principle: for all $x \in \mathcal{K}$, $\tau \in [0,T]$ and $h \in [0, T-\tau]$, we have:

$$\mathcal{V}(t,x) = \inf_{(\mathbf{x},\boldsymbol{u}) \in \mathbb{X}^g_{[t,T]}(x)} \mathcal{V}(t+h, \mathbf{x}(t+h)) + \int_t^{t+h} \ell(s, \mathbf{x}(s), \boldsymbol{u}(s))\, ds. \tag{38}$$

4.3.1 Inward pointing condition

To analyze the properties of the value function $\mathcal{V}$, it is important first to understand the structure of the set the admissible trajectories. This structure depends on an interplay between the dynamics of the state equation and the set of constraints $\mathcal{K}$. Assume in this section that $g : \mathbb{R}^d \to \mathbb{R}$ is $C^{1,1}$ function, and its zero-level set is suitably smooth. Consider the following controllability assumption:

(HK1) Inward pointing qualification (IPQ) condition: For every $R > 0$, there exists $\beta > 0$ and $\rho > 0$ such that for every $t \in [0,T]$,

$$\min_{u \in U} f(t,y,u) \cdot \nabla g(y) < -\beta, \qquad \forall y \in \partial\mathcal{K} \cap \mathbb{B}(0,R). \tag{39}$$

The IPQ condition states that the set of constraints $\mathcal{K}$ has a smooth structure and that on each point of the boundary $\mathcal{K}$ it is possible to find an admissible control that allows the trajectory to stay in the set $\mathcal{K}$. So, the IPQ condition implies that the set $\mathcal{K}$ is weakly invariant. Moreover, the IPQ condition guarantees even a nicer property, called **(NFT) Neighbouring feasible trajectories** principle, whose proof can be found in [16, Theorem 2.1].

Lemma 4.1. *Assume* **(HK1)**. *Let* $(t_0, x_0) \in [0,T] \times \mathcal{K}$ *and let* $(\mathbf{y}, \boldsymbol{v}) \in \mathbb{X}_{[t_0,T}(x_0)$. *There exists a constant* C *and a feasible pair* $(\mathbf{x}, \boldsymbol{u}) \in \mathbb{X}_{[t_0,T]}(x_0)$ *such that*

$$\mathbf{x}(t) \in \mathcal{K}, \ \forall\, t \in [t_0,T], \ \|\mathbf{y} - \mathbf{x}\|_{W^{1,1}([t_0,T];\mathbb{R}^d)} \leq C g^+(\mathbf{y}(\cdot)),$$

where $g^+(\mathbf{y}(\cdot)) = \max_{t \in [t_0,T]} \{\max(g(\mathbf{y}(t)), 0)\}$.

The NFT property states the existence of an admissible control-trajectory pair satisfying the state constraints, close to an admissible pair that violates the state constraints. This property is the key point to ensure continuity of the value function and to provide a characterization of the value function in terms of viscosity solutions of the relevant HJB equation on $\mathcal{K}$ (see [78]).

Theorem 4.4. *Assume **(HK1)**. Then, the value function $\mathcal{V}(\cdot,\cdot)$ is uniformly continuous and bounded on $[0,T]\times\mathcal{K}$. Moreover, it is the unique constrained viscosity solution of the following HJB equation:*

$$-\partial_t\mathcal{V}(t,x)+H(t,x,D_x\mathcal{V}(t,x))\geq 0 \quad \text{on } [0,T[\times\mathcal{K}, \tag{40a}$$

$$-\partial_t\mathcal{V}(t,x)+H(t,x,D_x\mathcal{V}(t,x))\leq 0 \quad \text{on } [0,T[\times\overset{\circ}{\mathcal{K}}, \tag{40b}$$

$$\mathcal{V}(T,x)=\varphi(x), \tag{40c}$$

with $\mathcal{V}(t,x)=+\infty$ for every $x\in\mathbb{R}^d\setminus\mathcal{K}$.

It should be noticed that the HJB equation in the above theorem provides only partial information on the boundary of $\mathcal{K}$. Moreover, the function $\mathcal{V}$ takes infinite values outside $\mathcal{K}$ (i.e., $\mathcal{V}(t,x)=+\infty$ for every $x\notin\mathcal{K}$). These two facts make the approximation of $\mathcal{V}$ on $\mathcal{K}$ very challenging and require some penalization techniques.

Finally, let us mention that the IPQ condition $(HK1)$ can be weakened a bit and generalized to the case with several constraints as in [54]. In [41], a "higher order" inward pointing condition involving Lie brackets of the dynamics' vector fields is also analyzed.

4.3.2 Case of state constraints without controllability assumptions

As we mentioned in the previous section, the Lipschitz regularity of the value function requires an interplay between the dynamics f and the set of constraints $\mathcal{K}$. When the controllabity condition **(HK1)** is not satisfied, the value function may be discontinuous and its characterization by a HJB equation becomes very delicate. In this section, we introduce an alternative formulation of state-constrained control problems, in case the controllability assumption is not satisfied.

We set $G(y):=\max(g_1(y),\cdots,g_{m_g}(y))$ and $G_f(y):=\max(g_{f,1}(y),\cdots,g_{f,m_f}(y))$ for every $y\in\mathbb{R}^d$. We introduce an auxiliary control problem and its associated value function $\mathcal{W}$ defined by

$$\mathcal{W}(t,x,z):=\inf_{\mathbf{y}=(\mathbf{x},\zeta)\in\mathcal{S}_{[t,T]}(x,z)}\Big(\varphi(\mathbf{x}(T))-\zeta(T)\Big)\bigvee\max_{\theta\in(t,T)}G(\mathbf{x}(\theta))\bigvee G_f(\mathbf{x}(T)) \tag{41}$$

for $x\in\mathbb{R}^d$, $z\in\mathbb{R}$, $t\in[0,T]$, $a\vee b:=\max(a,b)$, and where the set of trajectories $\mathcal{S}_{[t,T]}(x,z)$ is defined in Remark 1.2. In this auxiliary control problem, the term $\max_{\theta\in[t,T]}g\big(\mathbf{x}(\theta)\big)$ is an exact penalization of the state constraints. Here, we shall use the problem (41) to characterize the epigraph of the value function ϑ without requiring any additional controllability assumption.

Theorem 4.5. *Assume that* $(\mathbf{H_0})$-$(\mathbf{H_4})$ *are satisfied. Then, for any $t\in[0,T]$ and $(x,z)\in\mathbb{R}^d\times\mathbb{R}$,*
(i)

$$\mathcal{V}(t,x)-z\leq 0 \quad\Longleftrightarrow\quad \mathcal{W}(t,x,z)\leq 0.$$

(ii) In addition, the function $\mathcal{V}$ is characterized by $\mathcal{W}$ through the relation

$$\mathcal{V}(t,x)=\min\Big\{z\in\mathbb{R},\ \mathcal{W}(t,x,z)\leq 0\Big\}. \tag{42}$$

Proof. (i) Let us assume that $\mathcal{V}(t,x)\leq z$. So there exists a sequence $(\mathbf{x}_n,\boldsymbol{u}_n)_{n\in\mathbb{N}}$ of admissible pairs in $\mathbb{X}_{[t,T]}(x)$, such that

$$\lim_{n\to+\infty}\int_t^T\ell(s,\mathbf{x}_n(s),\boldsymbol{u}_n(s))\,ds+\varphi(\mathbf{x}_n(T))-z=\mathcal{V}(t,x)-z\leq 0.$$

By admissibility, we have for each $n \geq 0$, $\max_{\theta\in[t,T]} G(\mathbf{x}_n(\theta)) \leq 0$ and $G_f(\mathbf{x}_n(T)) \leq 0$. Hence,

$$\begin{aligned}\mathcal{W}(t,x,z) &\leq \liminf_{n\to\infty}\left[\left(\int_t^T \ell(s,\mathbf{x}_n(s),\boldsymbol{u}_n(s))\,ds + \varphi(y_n(T)) - z\right) \bigvee \max_{\theta\in[t,T]} G(\mathbf{x}_n(\theta)) \bigvee G_f(\mathbf{x}_n(T))\right] \\ &\leq 0.\end{aligned}$$

Conversely, let us assume that $\mathcal{W}(t,x,z) \leq 0$. We know that $\mathcal{S}_{[t,T]}(\xi)$ is a compact set in $C^0([t,T])$, therefore the infimum in $\mathcal{W}(t,x,z)$ is achieved by some trajectory $\mathbf{x} \in \mathcal{S}[t,T]((x,z))$ (with an associated control $\boldsymbol{u} \in \mathcal{U}$). Moreover,

$$0 \geq \mathcal{W}(t,x,z) \quad = \quad \left(\int_t^T \ell(s,\mathbf{x}(s),\boldsymbol{u}(s))\,ds + \varphi(\mathbf{x}(T)) - z\right) \bigvee \max_{\theta\in[t,T]} G(y(\theta)) \bigvee G_f(\mathbf{x}(T)).$$

On the one hand, $\max_{\theta\in[t,T]} G(\mathbf{x}(\theta)) \bigvee G_f(\mathbf{x}(T)) \leq 0$ and $\mathbf{x}$ satisfies the state constraints, and on the other hand,

$$\mathcal{V}(t,x) - z \;\leq\; \int_t^T \ell(s,\mathbf{x}(s),\boldsymbol{u}(s))\,ds + \varphi(\mathbf{x}(T)) - z \;\leq\; 0$$

which is the desired result. Finally, statement (ii) is an immediate consequence of (i). □

Remark 4.2. Should the convexity assumption $(\mathbf{H_4})$ not be satisfied, the statements of the above theorem may not hold. Indeed, in general we have

$$\inf\{z \in \mathbb{R}, \mathcal{W}(t,x,z) \leq 0\} \leq \mathcal{V}(t,x) \leq \inf\{z \in \mathbb{R}, \mathcal{W}(t,x,z) < 0\}.$$

The auxiliary control problem suggests a reformulation of the state-constrained optimal control problem in an augmented state space. In this new formulation, the constraints are integrated into the functional to be minimized. The value function $\mathcal{W}$ is Lipschitz continuous and it can be characterized by an HJB equation without any additional controllability assumption.

Theorem 4.6. *Assume that* $(\mathbf{H_0})$-$(\mathbf{H_3})$ *are satisfied. Then, the auxiliary value function is Lipschitz continuous and it is the unique viscosity solution of the HJB equation*

$$\begin{aligned}&\min\Big(-\partial_t\mathcal{W}(t,x,z) + H(t,x,D_x\mathcal{W}(t,x,z),D_z\mathcal{W}(t,x,z)), \mathcal{W}(t,x,z) - G(x)\Big) = 0, && \text{on } [0,T[\times\mathbb{R}^d\times\mathbb{R}, \\ &\mathcal{W}(T,x,z) = \Big(\varphi(x) - z\Big) \bigvee G(x) \bigvee G_f(x), && \text{on } \mathbb{R}^d\times\mathbb{R},\end{aligned}$$

where the Hamiltonian H is defined by

$$H(t,x,p,q) = \max_{u\in U}\big(-f(t,x,u)\cdot p - \ell(t,x,u)q\big)$$

for every $(t,x,p,q) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}$.

We point out that any optimal trajectory for the original problem is also a solution of the auxiliary problem when $z = \bar{z} := \mathcal{V}(t,x)$. Conversely, any solution of the auxiliary problem with $z = \bar{z}$ is an optimal solution of the original state-constrained problem. As a consequence, the auxiliary problem provides the value of the original state-constrained control problem and also allows to reconstruct the optimal trajectories (see [5, 7]).

4.4 Relationship between HJB and PMP

To explain the relationship between HJB and PMP, we first recall a classical result (see for example [52]), valid under the very restrictive assumption that $\mathcal{V} \in C^2$.

Theorem 4.7. *Consider the optimal control problem* (3) *with* $g = 0$ *and* $g_f = 0$ *(no state constraint), under the assumptions* $(\mathbf{H_0})$–$(\mathbf{H_{1a}})$–$(\mathbf{H_{1b}})$. *Assume moreover that* f *and* ℓ *are continuously differentiable with respect to the space variable, that* $\mathcal{V} \in C^2([0,T] \times \mathbb{R}^d)$, *and that there exists an optimal pair* $(\mathbf{x}^*, \boldsymbol{u}^*)$. *Then, the vector* $\boldsymbol{p}(t)$ *defined by*

$$\boldsymbol{p}(t) := -D\mathcal{V}(t, \mathbf{x}^*(t)) \tag{43}$$

satisfies the Pontryagin maximum principle.

While this relatively simple result works under unrealistically strong assumptions, more recent theory [83] justifies the PMP–HJB relationship (in a suitably weakened form) under general assumptions. In the case of Mayer problems with a locally Lipschitz continuous cost, the sensitivity relations have also been studied in [40, 83]. In these results, the value function is only required to be Lipschitz continuous in a neighborhood of the optimal trajectory. The final cost is not differentiable and therefore, the costate function is not necessarily unique. In this context the sensitivity relations assert that there exists $\boldsymbol{p}$ verifying the PMP, the terminal conditions

$$-\boldsymbol{p}(0) \in \partial_x \mathcal{V}(0, \mathbf{x}^*(0)), \qquad -\boldsymbol{p}(T) \in \partial_x \mathcal{V}(T, \mathbf{x}^*(T)), \tag{44a}$$

and both a *partial sensitivity relation*

$$-\boldsymbol{p}(t) \in \partial_x \mathcal{V}(t, \mathbf{x}^*(t)), \quad \text{a.e. on } (0,T), \tag{44b}$$

and a *global sensitivity relation*

$$(\mathcal{H}(t, \mathbf{x}^*(t), \boldsymbol{p}(t)), -\boldsymbol{p}(t)) \in \partial \mathcal{V}(t, \mathbf{x}^*(t)), \quad \text{for all } t \in [0,T]. \tag{44c}$$

These relations extend (43) by using the generalized gradient of the value function (which is well defined for locally Lipschitz continuous functions). The set of conditions (44) is in essence a strengthened necessary condition, asserting that it is possible to choose a co-state trajectory to satisfy the sensitivity relations.

The relation (44a) can be simply derived by noting that the optimal solution $(\mathbf{x}^*, \boldsymbol{u}^*)$ is also solution of the free initial state problem

$$\min \Big\{ \varphi(\mathbf{x}(T)) + \int_0^T \ell(t, \mathbf{x}(t), \boldsymbol{u}(t))\, dt - \mathcal{V}(0, \mathbf{x}(0)) \;\Big|\; \dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), \boldsymbol{u}(s)) \text{ on } (0,T) \Big\}.$$

Applying necessary optimality conditions to this problem yields a costate arc $\boldsymbol{p}^0$. Relation (44a) turns out to be nothing else than the transversality conditions at the endpoints. With the same reasoning, we can notice that $(\mathbf{x}^*, \boldsymbol{u}^*)$ is also solution of the free initial state problem on $[t,T]$ for every $0 \le t \le T$

$$\min \Big\{ \varphi(\mathbf{x}(T)) + \int_0^T \ell(t, \mathbf{x}(t), \boldsymbol{u}(t))\, dt - \mathcal{V}(t, \mathbf{x}(t)) \;\Big|\; \dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), \boldsymbol{u}(s)) \text{ on } (0,T) \Big\}.$$

Here again, the optimality condition applied to the free intial point asserts the existence of an adjoint arc $\boldsymbol{p}^t$ (which depends on the initial time t). The left-endpoint transversality condition yields the relation $\boldsymbol{p}^t(t) \in \partial\mathcal{V}(t, \mathbf{x}^*(t))$. When the final cost function is C^1-regular, the costate $\boldsymbol{p}^0$ restricted to $[t, T]$ is the unique solution to the costate equation on this interval satisfying the right transversality condition. It follows that $\boldsymbol{p}^0(t)$ coincides with $\boldsymbol{p}^t(t)$; the proof of (44b) is then completed. This analysis breaks down when the final cost φ is non-smooth. Indeed, in that case, co-state trajectory may not be unique. An example is given in [40, Section 4] showing that, in some cases, there are a number of possible choices of co-state trajectories associated with the same optimal control problem, but not all of them satisfy the sensitivity relations.

The original proof of the sensitivity relations (44) is given in [40, 83, 84]. In the case when the control problem is in presence of state constraints, the sensitivity relations can be expressed in term of relations between the adjoint vector $(\boldsymbol{p}, p^0)$ and the value function $\mathcal{W}$ of the auxiliary control problem, defined in Section 4.3.2, see [19, 61].

4.5 Numerical methods for HJB

In order to present the general theory for the approximation of viscosity solutions of HJB equations, we refer to an abstract forward problem

$$\begin{cases} v_t + H(t, x, Dv) = 0 & (t, x) \in (0, T] \times \mathbb{R}^d, \\ v(0, x) = v_0(x) & x \in \mathbb{R}^d, \end{cases} \tag{45}$$

(for some continuous Hamiltonian H) and set ourselves in the usual finite difference scheme framework. Time is discretized with a (fixed) time step Δt, so that $t_k = k\Delta t$; space is discretized with a fixed space step $\Delta x = (\Delta x_1, \cdots, \Delta x_d)$. A generic node will be denoted by $x_j = j\Delta x$, for $j \in \mathbb{Z}^d$. We also define $\Delta = (\Delta x, \Delta t)$. More general options can be considered, in particular variable time steps and unstructured space grids, but we will restrict here to the basic ideas. In the following, we denote by V_i^n the desired approximation of $v(t_n, x_i)$, and by V^n the set of nodal values for the numerical solution $v(t_n, \cdot)$ at time t_n. A scheme may be written in compact form as

$$V^{n+1} = S(\Delta; V^n), \tag{46}$$

where S may be defined in terms of its components S_j, for $j \in \mathbb{Z}^d$.

4.5.1 Monotone schemes

The first and basic convergence theory aimed at approximating HJ equations of the form (45) uses the concept of monotone scheme. Among the various results, we quote here the Barles–Souganidis theory [14], which applies to the widest class of schemes and models, including the possibility of treating second-order, degenerate and singular equations. Roughly speaking, this theory states that any monotone, stable and consistent scheme converges to the exact viscosity solution, provided there exists a comparison principle for the limiting equation. Consider a scheme in the general form (46). We recall the concepts of consistency, monotonicity and stability.

Consistency. Let $\Delta_m = (\Delta x_m, \Delta t_m)$ be a generic sequence of discretization parameters, (t_{j_m}, x_{j_m}) be a generic sequence of nodes in the space–time grid such that, for $m \to \infty$,

$$(\Delta x_m, \Delta t_m) \to 0 \quad \text{and} \quad (t_{n_m}, x_{j_m}) \to (t, x). \tag{47}$$

The scheme S is said to be *consistent* if for any $\phi \in C^\infty((0,T]\times\mathbb{R}^d)$, we have

$$\lim_{m\to\infty} \frac{\phi(t_{n_m}, x_{j_m}) - S_{j_m}(\Delta_m; \phi(t_{n_m-1}, \cdot))}{\Delta t_m} = \phi_t(t,x) + H(t,x,D\phi(t,x)), \tag{48}$$

Monotonicity. The scheme S is said to be *monotone* if, for any couple of vectors V and W such that $V_j \geq W_j$:

$$S_j(\Delta; V) \geq S_j(\Delta; W), \tag{49}$$

for any Δx, Δt satisfying suitable compatibility conditions, that are typically in the form of the so-called *Courant–Friedrich–Levy (CFL) conditions*.

It is also possible to give a generalized form of definition of monotonicity, to treat some case of high-order scheme. We refer the reader to the discussion carried out in [20].

Given a numerical solution V^n, we define its piecewise constant (in time) interpolation V^Δ as

$$V^\Delta(t,x) = \begin{cases} I[V^n](x) & \text{if } t \in [t_n, t_{n+1}), \\ v_0(x) & \text{if } t \in [0, \Delta t), \end{cases} \tag{50}$$

where $I[V^n](x)$ denotes an interpolation of the node values in V^n, computed at x. We remark that the interpolation operator has to satisfy itself a monotonicity property to obtain a monotone scheme (this holds, for example, for a piecewise linear reconstruction).

We can now state (in a slightly rephrased form) the convergence result given in [14]:

Theorem 4.8. *Assume that* (45) *satisfies a comparison principle, and let $v(t,x)$ be the unique viscosity solution of* (45)*. Assume that* (48) *and* (49) *hold. Assume in addition that the family $v^{\Delta t}$ is uniformly bounded in L^∞. Then, $V^\Delta(t,x) \to v(t,x)$ locally uniformly on $\mathbb{R}^d \times [0,T]$ as $\Delta \to 0$.*

This result directly applies to the most classical cases of monotone schemes, as in the examples below.

Finite difference schemes. Given a numerical Hamiltonian $\mathcal{H} : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$, we define an explicit scheme (see [45]) as follows:

$$V_i^{n+1} = V_i^n - \Delta t\, \mathcal{H}(t_n, x_i,\, D^- V_i^n,\, D^+ V_i^n), \tag{51a}$$

$$V_i^0 = v_0(x_i). \tag{51b}$$

Here, $i \in \mathbb{Z}^d$, and the terms $D^- V_i^n$ and $D^+ V_i^n$ represent respectively left and right finite difference approximations of the gradient at x_i, defined as $D^\pm V_i^n = (D_k^\pm V_i^n)_{1\leq k\leq d}$ with

$$D_k^\pm V_i^n := \pm \frac{V_{i\pm e_k}^n - V_i^n}{\Delta x_k},$$

and where $\{e_k\}_{k=1,\dots,d}$ is the canonical basis of $\mathbb{R}^d$ ($(e_k)_k = 1$ and $(e_k)_j = 0$ if $j \neq k$).

For schemes of this form, and assuming that the numerical Hamiltonian $\mathcal{H}$ is Lipschitz continuous with respect to all its arguments, consistency with H comes down to the condition

$$\mathcal{H}(t,x,p,p) = H(t,x,p)$$

and monotonicity is checked in the form

$$\frac{\partial \mathcal{H}}{\partial p_k^-}(t,x,p^-,p^+) \geq 0, \quad \frac{\partial \mathcal{H}}{\partial p_k^+}(t,x,p^-,p^+) \leq 0.$$

These latter conditions typically require a CFL-type compatibility condition between Δt and Δx_i.

Two classical choices for the numerical Hamiltonian are in Upwind and Lax–Friedrichs form:

- If the Hamiltonian H is defined by (28), then an upwind numerical Hamiltonian may be constructed in the form

$$\mathcal{H}^{Up}(t,x,p^-,p^+) = \max_{u\in U}\left[\sum_{i=1}^{d}\left(\max(-f_i(t,x,u),0)\,p_i^- + \min(-f_i(t,x,u),0)\,p_i^+\right) - \ell(t,x,u)\right]. \quad (52)$$

  This form fulfils consistency and monotonicity conditions for Δt satisfying the CFL condition

$$\Delta t\left(\sum_{1\leq i\leq d}\frac{\max_{t,x,u}|f_i(t,x,u)|}{\Delta x_i}\right) \leq 1. \quad (53)$$

- The Lax-Friedrichs scheme can be defined for a generic Hamiltonian H as

$$\mathcal{H}^{LF}\left(t,x,p^-,p^+\right) := H\left(t,x,\frac{p^-+p^+}{2}\right) - \sum_{i=1}^{d} C_i\,\frac{p_i^+ - p_i^-}{2}, \quad (54)$$

  The numerical Hamiltonian $\mathcal{H}^{\mathrm{LF}}$ satisfies the monotonicity condition provided the constants C_i are chosen such that $C_i \geq \max_{u\in U}|f_i(t,x,u)|$ and $(\Delta t, \Delta x)$ satisfies (53).

Semi-Lagrangian schemes. The Semi-Lagrangian (SL) scheme is written here directly in the form suitable for the backward dynamic programming equation (27)–(28). In fact, the SL scheme can be derived by discretizing the Dynamic Programming Principle on a single time step:

$$\begin{cases} v_j^{n-1} = \min\limits_{\alpha\in U}\{\Delta t\ell(t,x,\alpha) + I_1[V^n](x_j + \Delta t f(t,x,\alpha)\} \\ v_j^N = \varphi(x_j), \end{cases} \quad (55)$$

in which $I_1[V](x)$ denotes the $\mathbb{P}_1$ (piecewise linear or multilinear) interpolate of the vector V of node values, computed at the point x. The SL scheme is consistent, and the choice of a linear interpolation as I_1 implies also monotonicity of the scheme.

4.5.2 High-order schemes

While the framework of monotone schemes remains the most classical, in the last decades high-order numerical schemes for HJB equations have been developed and analyzed. Their convergence analysis relies typically on two theoretical tools:

- **ε-monotonicity**

  The Barles–Souganidis theorem allows for an $o(\Delta t)$ monotonicity defect, making it possible to prove convergence for quasi-monotone schemes. This theory has been applied to high-order SL schemes and to filtered schemes, as in [10, 49, 20].

- **Lin–Tadmor theory**

  Lin–Tadmor convergence theory is inspired by the *Lip*'-stability theory for conservation laws. Here, a different concept of stability is singled out, i.e., *uniform semi-concavity* of numerical solutions, along with a suitable definition of consistency. The convergence result, together with a practical application of this theory is presented in [66].

Higher-order FD schemes. The basic strategy for constructing high-order finite difference methods has been first proposed in [71] and uses a TVD Runge–Kutta method for the time discretization combined with high order approximations of the right/left derivatives $D_j^{\pm}[V]$ at the node x_j (for example the ENO approximation, see [71]).

Higher-order SL schemes. The SL scheme (55) is easily extended to a higher consistency rate by replacing the $\mathbb{P}_1$ space interpolation I_1 with an interpolation of higher accuracy [48, 35]. In general, since characteristics are not straight lines, a more accurate method of characteristics tracking is also desirable [47]. In some model cases convergence of high-order SL schemes, for both the evolutive and the stationary case, can be proved by showing their quasi-monotonicity (see [51, 49, 20]).

Filtered schemes. The general idea of filtered schemes (which had previously appeared in the context of conservation laws as *flux-limiter schemes*) is to provide a clever coupling between a monotone and a high-order scheme. Starting from a monotone scheme S^M, a high-order scheme S^{HO} and a bounded *filter function* $F : \mathbb{R} \to \mathbb{R}$, the filtered scheme S^F is defined as

$$v_j^{n+1} = S_j^F(V^n) := S_j^M(V^n) + \epsilon \Delta t F\left(\frac{S_j^{HO}(V^n) - S_j^M(V^n)}{\epsilon \Delta t}\right), \tag{56}$$

where $\varepsilon = \varepsilon(\Delta) > 0$ is a parameter vanishing for $\Delta t, \Delta x \to 0$, which controls the monotonicity defect of the filtered scheme (more hints on the choice of ε and of the filter function can be found in [55, 21]). In constructing the filter function, the basic idea is that "large" values of the ratio $\rho = (S^{HO} - S^M)/(\varepsilon \Delta t)$ indicate a singularity (where the scheme needs to be monotone), while "small" values indicate a smooth region (in which the scheme can be high-order). It can be shown that, for a suitable choice of ϵ, the filtered scheme converges to the viscosity solution by quasi-monotonicity.

Further comments. Several advances have been made to improve the numerical schemes of approximations of HJB equations, in particular in high dimension. Let us mention the resolution techniques on sparse grids [23, 56], on tree structures as in [4], or approximation by sophisticated model reduction techniques as in [3]. We also mention another chapter [11] of the Volume 1 of this Handbook, dedicated to model reduction methods. All these methods aim at providing accurate numerical approaches for solving efficiently HJB equations with reasonable numerical efforts (measured by complexity of algorithms, CPU time and accuracy).

4.5.3 Optimal trajectory reconstruction from the value function

From a control viewpoint, the approximation of the value function $\mathcal{V}$ has a relatively lesser interest with respect to the construction of the (approximate) optimal control. We propose in this section some algorithms that lead, given an approximations of the value function, to construct a quasi-optimal controls in feedback form. The procedure does not depend on the specific scheme used to compute $\mathcal{V}$. For simplicity, we consider the trajectory reconstruction on the time interval $[0,T]$, although all the results remain valid for a reconstruction on any sub-interval $[t,T]$. For $n_h \in \mathbb{N}$ and $h = T/n_h$, consider a partition $s_0 = 0 < s_1 < \cdots < s_{n_h} = T$ of $[0,T]$, with $s_k = kh$. Consider a numerical approximation f_h of the dynamics f such that, for every $R > 0$, we have

$$|f_h(t,x,u) - f(t,x,u)| \leq C_R h, \qquad \forall t \in [0,T],\ |x| \leq R, u \in U, \tag{57}$$

where the constant C_R is independent of $h \in (0,1]$. An approximation scheme for the differential equation $\dot{\mathbf{x}}(t) = f(t,\mathbf{x}(t),\boldsymbol{u}(t))$ (for a constant control u, discrete times s_k and time step h) can be written as

$$y_{k+1} = y_k + h f_h(s_k, y_k, u), \quad k \geq 0. \tag{58}$$

Here and in the sequel we use the notation y_k to denote a state at discrete times. The case of the Euler forward scheme corresponds to the choice $f_h := f$. Higher order Runge-Kutta schemes can also be written as (58) and with a function f_h satisfying (57). For instance, the Heun scheme (with constant control) corresponds to the choice

$$f_h(t,y,u) := \frac{1}{2}(f(t,y,u) + f(t+h, y + hf(t,y,u), u)).$$

Bolza problems. Consider first, the case of Bolza unconstrained problems. Let $\mathcal{V}$ the exact value function defined in (3). Let $\mathcal{V}^h$ be an approximation of $\mathcal{V}$, and define E_h as a uniform bound on the error:

$$|\mathcal{V}^h(t,x) - \mathcal{V}(t,x)| \leq E_h, \quad \forall t \in [0,T], |x| \leq R,$$

with $R > 0$ large enough. The approximate feedback is defined on the basis of the approximate value function with a discrete dynamic programming procedure.

Following some arguments introduced in [76], it can be shown that any cluster point of $(\mathbf{y}^h)_{h>0}$ is an optimal trajectory that realizes a minimum in the definition of the original control problem.

Theorem 4.9. *Assume* $(\mathbf{H_0})$, $(\mathbf{H_1})$, $(\mathbf{H_2})$ *and* $(\mathbf{H_4})$. *Assume also that the approximation* (57) *is valid and the error estimate* $E_h = o(h)$. *Let* (y,z) *be in* $\mathbb{R}^d \times \mathbb{R}$ *and let* (y_k^h) *be the sequence generated by Algorithm (TR).*

(i) The approximate trajectories $(y_k^h)_{k=0,\ldots,n_h}$ *constitute a minimizing sequence in the following sense:*

$$\mathcal{V}(0,x) \quad = \quad \lim_{h\to 0}\left(\varphi(y_{n_h}^h) + h\sum_{k=0}^{n_h} \ell(s_k, y_k^h, u_k^h)\right). \tag{61}$$

(ii) Moreover, the family $(\mathbf{y}^h)_{h>0}$ *admits cluster points, for the* L^∞ *norm, when* $h \to 0$. *Any such cluster point* $\bar{\mathbf{y}}$ *is an admissible trajectory and* $\mathbf{y}$ *is an optimal trajectory for* $\mathcal{V}(0,x)$.

Algorithm (TR) - Trajectory reconstruction algorithm for Bolza problems

Require: First we set $y_0^h := x$.

1: For $k = 0, \dots, n_h - 1$, knowing the state y_k^h we define

(i) an optimal control value $u_k^h \in U$ such that

$$u_k^h \in \operatorname{argmin}_{u \in U} \Big[\mathcal{V}^h\big(s_k, y_k^h + h_k\, f_h(s_k, y_k^h, u)\big) + h\ell(s_k, y_k^h, u) \Big] \tag{59}$$

(ii) a new state position y_{k+1}^h

$$y_{k+1}^h := y_k^h + h_k f_h(s_k, y_k^h, u_k^h). \tag{60}$$

2: **return** a piecewise constant control $\boldsymbol{u}^h(s) := u_k^h$ on $[s_k, s_{k+1}[$, and a piecewise linear trajectory $\mathbf{y}^h$ such that $\mathbf{y}^h(s_k) = y_k^h$.

Let us emphasize that the condition $E_h = o(h)$ indicates that the approximation $\mathcal{V}^h$ should be provided with a given precision. Typically, a numerical scheme would provide an approximation V^Δ for $\Delta = (\Delta t, \Delta x)$. Under a CFL condition, the error estimate is of order $O(\sqrt{\Delta t})$. To ensure that $\|V^\Delta - \mathcal{V}\| = o(h)$, it suffices to take $\sqrt{\Delta t} = o(h)$.

Minmax problems. The same idea as in Algorithm (TR) can be adapted for minmax control problems (see section 4.2). Let $\mathcal{V}^{\#,h}$ be an approximation of the value function $\mathcal{V}^{\#}$. Here, the function $\mathcal{V}^{\#,h}$ could be again a numerical approximation obtained by solving a discretized form of the HJB equation (29).

Algorithm (TRM) - Trajectory reconstruction for Minmax problems

Require: Set $y_0^h = x$.

1: Define the positions $(y_k^h)_{k=0,\dots,n_h}$, and control values $(u_k^h)_{k=0,\dots,n_h-1}$, by recursion as follows. or $k = 0, \dots, n_h - 1$, knowing the state y_k^h we define

(i) an optimal control value $u_k^h \in U$ such that

$$u_k^h \in \operatorname{argmin}_{u \in U} \mathcal{V}^{\#,h}\big(s_k, y_k^h + h_k\, f_h(s_k, y_k^h, u), z\big) \bigvee \Psi(s_k, y_k^h) \tag{62}$$

(ii) a new state position y_{k+1}^h

$$y_{k+1}^h := y_k^h + h_k f_h(s_k, y_k^h, u_k^h).$$

2: **return** a piecewise constant control $\boldsymbol{u}^h(s) := u_k^h$ on $[s_k, s_{k+1}[$, and a piecewise linear trajectory $\mathbf{y}^h$ such that $\mathbf{y}^h(s_k) = y_k^h$.

Note that in (62) the value of u_k^h can also be defined as a minimizer of $u \to \mathcal{V}^{\#,h}\big(s_k, y_k^h + h_k\, f_h(s_k, y_k^h, u), z\big)$, since this will imply in turn to be a minimizer of (62)

Theorem 4.10. *Assume* $(\mathbf{H_0})$, $(\mathbf{H_1})$, $(\mathbf{H_2})$ *and* $(\mathbf{H_4})$ *hold true. Assume also that the approximation* (57) *is valid and that* $\mathcal{V}^{\#,h}$ *is an approximation of* $\mathcal{V}^{\#}$ *with error estimate* $\|\mathcal{V}^{\#} - \mathcal{V}^{\#,h}\| = o(h)$. *let* (y_k^h) *be the sequence generated by Algorithm 2.*

(i) The approximate trajectories $(y_k^h)_{k=0,\dots,n_h}$ *constitute a minimizing sequence in the following sense:*

$$\mathcal{V}^{\#}(0,y) \quad = \quad \lim_{h\to 0}\left(\max_{0\le k\le n_h} \Psi(s_k, y_k^h)\right) \bigvee \varphi(y_{n_h}^h).$$

(ii) Moreover, the family $(\mathbf{y}^h)_{h>0}$ *admits cluster points, for the* L^∞ *norm, when* $h \to 0$. *For any such cluster point* $\bar{\mathbf{y}}$, *we have* $\bar{\mathbf{y}} \in \mathcal{S}_{[0,T]}(y)$ *and* $\bar{\mathbf{y}}$ *is an optimal trajectory for* $\mathcal{V}^{\#}(0,x)$.

4.6 Numerical examples

Test 1 - Unconstrained Zermelo problem In this example, we consider the same setting as in Section 3.4. To compute the minimal time function, we use the level-set approach described in Section 4.2. We consider a domain of computation large enough to contain the initial position $(0,0)$ and the final target $y_f = (20,1)$. More precisely, the computation will be performed on $\mathcal{D} := [-1,21] \times [-0.5,1.5]$. We use the finite difference scheme (51) combined with ENO approximation and the Lax-Friedrichs numerical Hamiltonian (54). Figure 7 shows the optimal solution computed on a uniform grid with 500×100 nodes. The optimal time to steer the system from the initial point $(0,0)$ to the final state $y_f = (20,1)$ is 4.94916.

While the approximation of the optimal trajectory on Figure 7(left) seems quite accurate, the optimal control law in Figure 7(right) presents some oscillations. This happens because the control law is constructed in a "blind" way: in fact, in the reconstruction algorithms, at each time step, the control value is computed to follow a minimal path. It may happen that several values of the control lead to the same position of the trajectory. The reconstruction process picks one of the optimal control values and this arbitrary choice may generate oscillations. Notice also that Theorems 4.9 and 4.10 state the convergence of the reconstructed optimal trajectories but not the optimal control laws (convergence of control laws would require additional assumptions).

In this example, the minimum time function $\mathcal{T}$ is continuous on its domain, as it can be seen on Figure 8, where some level-sets of the $\mathcal{T}$ are presented.

Recall that for this version of Zermelo problem, the shooting method provides a globally optimal trajectory or a locally optimal trajectory depending on the initialization of the adjoint state (see Section 4). By using the HJB approach, the reconstructed trajectory is an approximation of the global optimum, even when performing the calculations on very coarse grids. Note also that an approximation of the derivative of the minimal time function at the initial position (0,0) can be computed. In our simulation, we obtain

$$D_x\mathcal{T}^{\Delta}(0) = (-0.24585, -0.09163)^{\mathsf{T}}, \qquad \text{and } \mathcal{T}(0) = 4.94916.$$

An initialization of the shooting method with this vector allows the method to converge towards the global solution in very few iterations.

Test 2 - Unconstrained Zermelo problem - Case with a strong current. Here, we consider a variant of the Zermelo problem where the dynamics is given by

$$\dot{\mathbf{y}}_1(s) = \boldsymbol{v}(s)\cos(\boldsymbol{u}(s)) + 2 - \frac{1}{2}\mathbf{y}_2^2(s), \tag{63a}$$

$$\dot{\mathbf{y}}_2(s) = \boldsymbol{v}(s)\sin(\boldsymbol{u}(s)), \tag{63b}$$

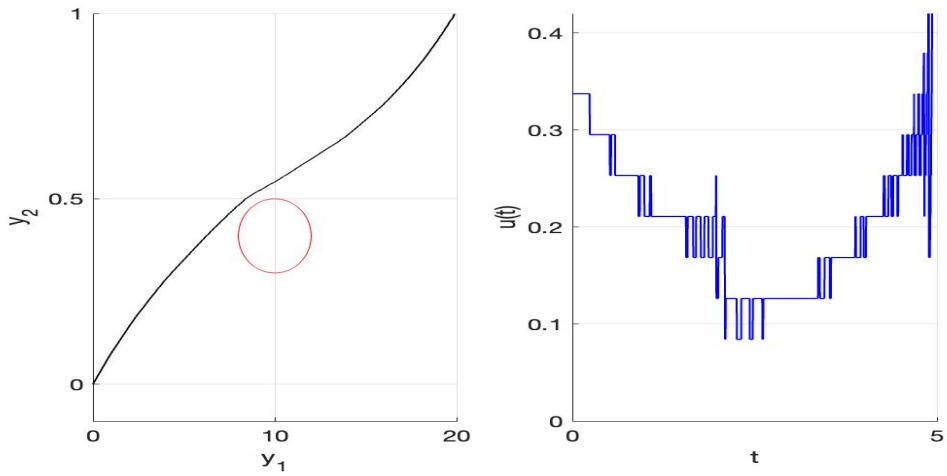

Figure 7: (Test 1) Approximation of the optimal trajectory for Zermelo problem with HJB approach.

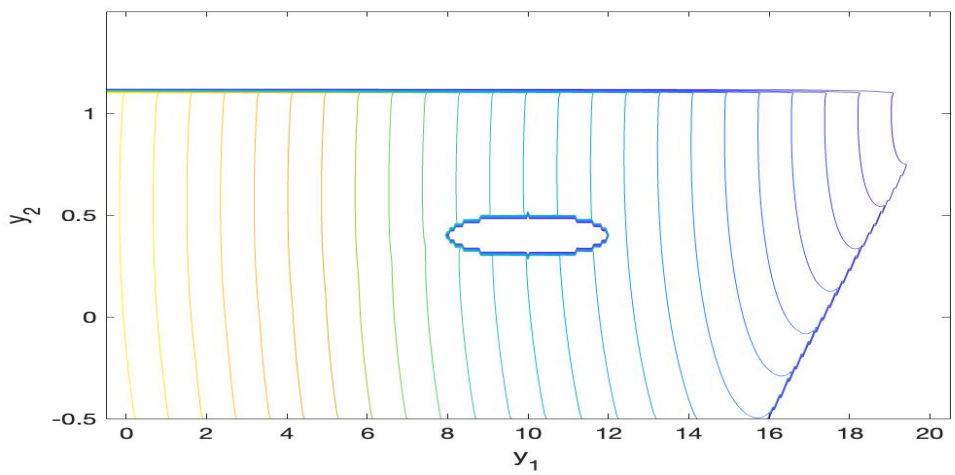

Figure 8: (Test 1) Level sets of the minimum time function on the domain $[-0.5, 20] \times [-0.5, 1.1]$.

where the control inputs are the speed $\boldsymbol{v}(s)$ and the angle of orientation $\boldsymbol{u}(s)$ of the boat. The set of control values is $U := [0, 1] \times [0, 2\pi]$. In this example, the drift is strong in the middle of the channel $\mathbb{R} \times [-2, 2]$ and is zero along the channel banks. The target is a ball centred at the origin and with radius $r = 0.1$. The domain of computation is $\mathcal{D} := [-5, 2] \times [-2, 2]$.

- **Test 2-1**: The numerical simulations are performed on a uniform grid with $N_{x_1} = N_{x_2} = 100$ nodes on each axis. The numerical results of this test are displayed in Figure 9. The left-hand side of Figure 9 shows level sets of the minimum time function. Some trajectories starting from different initial positions are given in the right-hand side of Figure 9.

- **Test 2-2** : Here, the numerical simulations are performed on a uniform grid with $N_{x_1} = N_{x_2} = 500$ nodes on each axis. The numerical results of this test are given in Figure 10.

Numerical convergence can be observed when refining the size of the grid (i.e., at the increase

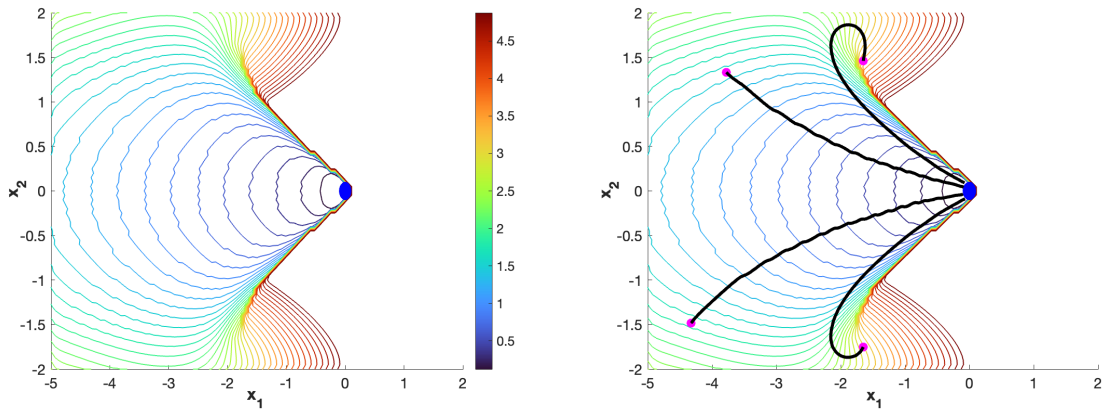

Figure 9: (Test 2-1) Zermelo problem. The figure on the left-hand side shows some level-sets of the minimum time $\{x \mid \mathcal{T}(x) = c\}$ for c between 0 and 5. The figure on the right displays some trajectories (the black curves) starting from different initial positions. The simulations presented in this figure are performed on a grid with $N_{x_1} = N_{x_2} = 100$

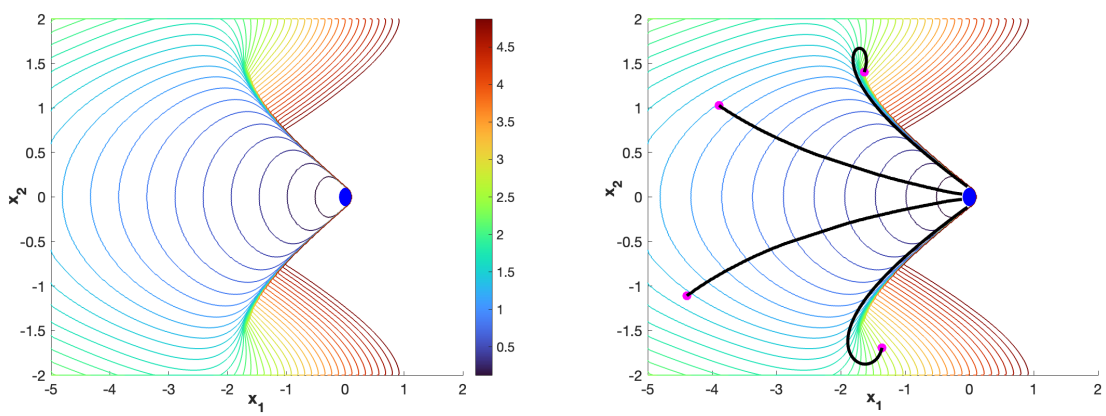

Figure 10: (Test 2-1) Zermelo problem. The figure on the left-hand side shows some level-sets of the minimum time $\{x \mid \mathcal{T}(x) = c\}$ for c between 0 and 5. The figure on the right displays some trajectories starting from different initial positions. The simulations presented in this figure are performed on a grid with $N_{x_1} = N_{x_2} = 500$

of N_{x_1} and N_{x_2}).

Test 3 - Constrained Zermelo problem. Consider again the same dynamics for Zermelo problem as in the previous paragraph. Now, the state is required to avoid the rectangular obstacle $[-2.5, -1.5] \times [-0.5, 0.5]$. Figure 11 shows the level sets of the minimum time function and some samples of optimal trajectories. Figure 11 (left) displays the results obtained with 250×250 nodes on the domain of computation, while Figure 11 (right) corresponds to computation on grid of 500×500 nodes. Notice that the optimal trajectories avoid the obstacle (in red) but they tend to get closer to the central section of the channel, in which the current is stronger.

Moreover, gradient singularities upstream of the obstacles indicate points at which the optimal trajectory is not unique (e.g., it can go either left or right of the obstacle).

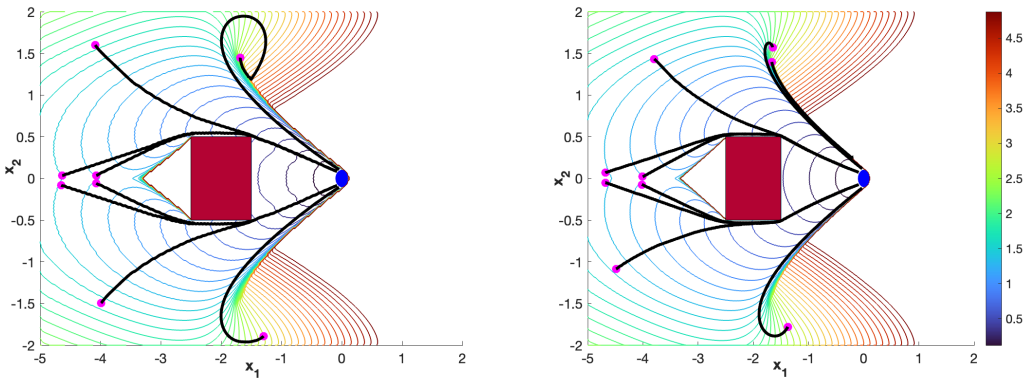

Figure 11: (Test 3) Zermelo problem with an obstacle. The two figures display level sets of the minimum time function $\{x \mid \mathcal{T}(x) = c\}$ for some values of c between 0 and 5, and some optimal trajectories starting from different initial positions. The figure on the left-hand side displays the numerical results obtained on a grid with $N_{x_1} = N_{x_2} = 250$. The results on a grid with $N_{x_1} = N_{x_2} = 500$ are given in the right-hand side of the figure

Test 4 - Goddard Problem. Now, consider the Goddard problem as described in Sections 2.4 and 3.5. The problem is with free final time and state constraints. Since the inward pointing condition is not satisfied, we reformulate the problem as described in Section 4.3.2. We compute the auxiliary value function (in dimension 4) on $\mathcal{D} = [1.0, 1.2] \times [0, 0.12] \times [0, 1] \times [0, 1]$. We use a uniform grid with N_x^4 node points (i.e., N_x points on each axis). In the sequel, we will choose $N_x = 20$ and $N_x = 40$ so that the computation of the value function and the optimal trajectory is performed in less than 1 minute.

Figure 12 corresponds to a simulation with $Nx = 20$, while Figure 13 corresponds to a simulation with $Nx = 40$. In both figures, we show the three state variables in the top line, and the adjoint states in the middle line. In the third line, the control variable is displayed on the left. We present also, in the middle of the third line, an approximation of the derivative of the Hamiltonian with respect to variable u.

In Figures 12–13, we notice that the derivative of the Hamiltonian w.r.t. the control variable u vanishes identically on a time interval whose entry and exit times are indicated by red dots (middle of the third line). On this time interval, the control law is singular. Moreover, we notice that the constraint on the velocity $\mathbf{x}_2 = \boldsymbol{v}$ is saturated on another time interval (the entry and exit times are indicated by red dots in the figure situated in the middle of the first line).

In this example, we observe again that the approximation of the control law presents high oscillations (for the same reasons we mentioned in Test 2). We notice that the oscillations persist even when N_x increases. These oscillations occur mainly during the time interval when the trajectory is singular-constrained. Besides, we observe a numerical convergence of the trajectories when refining the size of the grid (i.e., increasing N_x). At the increase of the number of nodes, CPU time and memory used also increase, but we notice that even a coarse grid calculation, by HJB approach, provides interesting results that can at least serve as an initialization for a

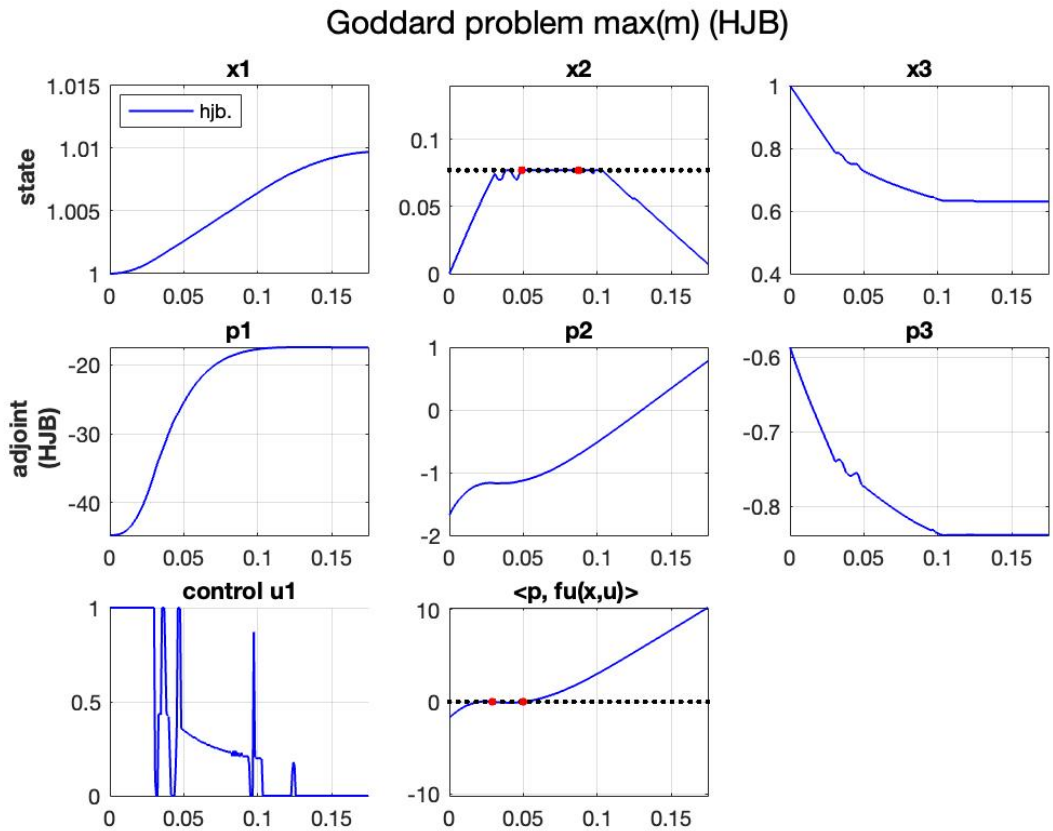

Figure 12: (Test 4) Goddard problem solved by HJB approach, with $N_x = 20$

more precise method like the shooting method. The computation of the value function and the reconstruction of the optimal trajectory give an approximation of

- the co-state $\boldsymbol{p}(0) = (\boldsymbol{p}_1(0), \boldsymbol{p}_2(0), \boldsymbol{p}_3(0))$ (by using the sensitivity relations in Section 4.4);

- the entry point of the singular arc t_1, the entry and exit points of the boundary arc t_2 and t_3, and the free final time t_f.

In Table 1, the second row presents the numerical results obtained by HJB simulations on a grid of $N_x = 20$ points on each axis. The fourth row presents the numerical results obtained by the shooting method (initialized with the values obtained by the HJB simulation). The convergence of the shooting method requires 43 iterations. The adjoint vector being defined up to a multiplicative constant, we give in the third row of Table 1 the results of HJB with a re-normalization of the adjoint vector. This makes it easier to compare the results of the HJB simulation and those of the shooting method.

5 Optimistic planning algorithms

As mentioned previously, direct and indirect methods are quite simple to implement, and provide locally optimal solutions with high accuracy. These methods depend on the initialization – espe-

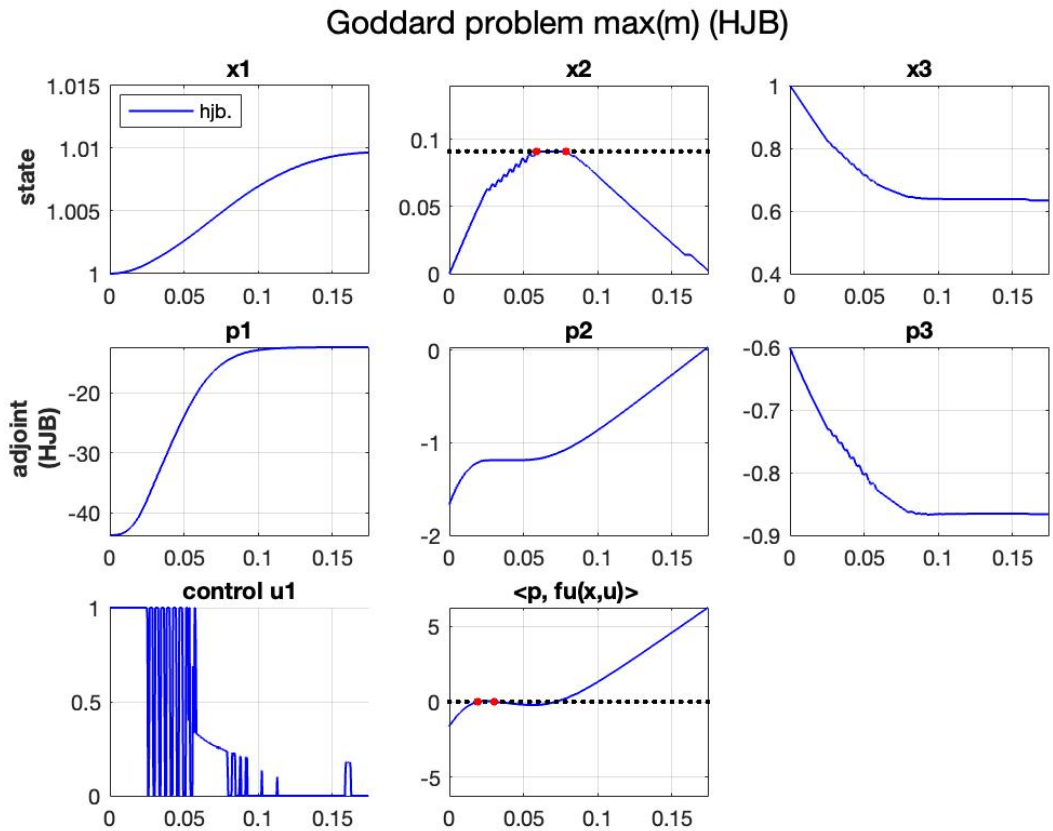

Figure 13: (Test 4) Goddard problem solved by HJB approach, with $N_x = 40$

	$\boldsymbol{p}_1(0)$	$\boldsymbol{p}_2(0)$	$\boldsymbol{p}_3(0)$	t_1	t_2	t_3	t_f
HJB simulation	5.205e1	1.947e0	6.826e-1	2.912e-2	4.980e-2	8.735e-2	1.747e-1
HJB simulation with a re-normalization of $p(0)$	3.945e0	1.476e-1	5.174e-2	2.912e-2	4.980e-2	8.735e-2	1.747e-1
Shooting method	3.945e0	1.504e-1	5.371e-2	2.351e-2	5.974e-2	1.016e-1	2.020e-1

Table 1: Godard problem. The second row presents the results obtained by HJB simulation on a grid with $N_x = 20$ points on each axis. The third row presents the same results of row 2 with a re-normalization of the costate vector. The fourth row presents the results obtained by the shooting method (initialized by the HJB results given in row 2).

cially the shooting method which is particularly sensitive to the initialization and also requires an *a priori* knowledge of the optimal trajectory structure (existence of bang and/or singular and/or saturated arcs). On the other hand, the HJB approach always provides a global optimum, but, if ever feasible, it requires a greater computational effort because of the high dimension of the space in which the value function must be computed.

A further approach that we present here is a global approach based on a discretization in the *space of controls*, combined with *optimistic planning* (OP) algorithms [32, 33] (without requiring any discretization of the state space). This approach is interesting especially for applications where the control dimension r is lower compared to the state dimension d. On a given discretization of the time interval, our approach will seek to identify the best control strategy to apply on each time sub-interval. The OP methods perform the optimal control search by branch and bound on the control set, always refining the region with the best lower bound of the optimal value (this is what justifies the term "optimistic"). An interesting feature of these algorithms is the close relationship between computational resources and quasi-optimality, which exploits some ideas of reinforcement learning [68]. Indeed, for given computational resources, the OP approaches provide a sub-optimal strategy whose performance is *close* to the optimal value (with the available resources).

First, for $N \geq 2$, consider a uniform partition of $[0, T]$ with $N+1$ time steps: $t_k = k\Delta t$, $k = 0, \ldots, N$, where $\Delta t = \frac{T}{N}$ is the step size. For a sequence of actions $\boldsymbol{u} = (u_k)_{0 \leq k \leq N-1} \in U^N$, we consider $(y_k^{x,\boldsymbol{u}})_{0 \leq k \leq N}$ the trajectory solution of the discrete-time dynamical system

$$\begin{cases} y_0 = x, \\ y_{k+1} = F_k(y_k, u_k) \quad k = 0, ..., N-1, \end{cases} \tag{64}$$

where $F_k(x, u)$ is an approximation of the solution to the system $\dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), u)$ on (t_k, t_{k+1}), with the initial condition $\mathbf{x}(t_k) = x$. More precisely, we assume that

$$\left\| F_k(x, u) - x - \int_{t_k}^{t_k+1} f(s, \mathbf{x}(s), u)\, ds \right\| = \mathrm{O}(\Delta t) \quad \text{for all } x \in \mathbb{R}^d,\ u \in U.$$

Consider also an instantaneous cost function L_k that approximates the integral of ℓ over an interval $[t_k, t_{k+1}]$, for $k = 0, ..., N-1$:

$$\left\| L_k(x, u) - \int_{t_k}^{t_k+1} \ell(s, \mathbf{x}(s), u)\, ds \right\| = \mathrm{O}(\Delta t) \quad \text{for all } x \in \mathbb{R}^d,\ u \in U.$$

In this section, we assume that $(\mathbf{H_0})$-$(\mathbf{H_4})$ are satisfied and that f and ℓ are Lipschitz continuous with respect to the control variable. The approximations F_k and L_k are also assumed to be Lipschitz continuous:

$$\begin{aligned} \|F_k(x, u) - F_k(x', u')\| &\leq L_{F,x}\|x - x'\| + L_{F,u}\|u - u'\|, \\ |L_k(x, u) - L_k(x', u')| &\leq L_{L,x}\|x - x'\| + L_{L,u}\|u - u'\|, \end{aligned}$$

for every $x \in \mathbb{R}^d$ and $u \in U$. Moreover, we assume that the Lipschitz constants of F_k and L_k are related to the Lipschitz constants of f and ℓ by the following relations

$$L_{F,x} := 1 + \Delta t L_{f,x} C \quad L_{F,u} := \Delta t L_{f,u} C \ \text{ and} \tag{65a}$$

$$L_{L,x} := \frac{\Delta t}{2} L_{\ell,x} C, \quad L_{L,a} := \Delta t (L_{\ell,a} + C), \tag{65b}$$

where the constant $C > 0$ may depend on Δt and the Lipschitz constants of f and ℓ.

Now, we define the state-constrained optimal control problem

$$V(x) := \inf_{\boldsymbol{u}=(u_k)_k \in U^N} \left\{ \sum_{k=0}^{N-1} L_k(y_k^{x,\boldsymbol{u}}, u_k) + \varphi(y_N^{x,\boldsymbol{u}}) \ \mid \ g(y_k^{x,\boldsymbol{u}}) \leq 0 \ \ \forall k = 0, ..., N, \ \ g_f(y_N^{x,\boldsymbol{u}}) \leq 0 \right\}. \tag{66}$$

Then, for the discrete auxiliary control problem, we define the cost functional J by

$$J(x, z, \boldsymbol{u}) := \Big(\sum_{k=0}^{N-1} L_k(y_k^{x,\boldsymbol{u}}, u_k) + \varphi(y_N^{x,\boldsymbol{u}}) - z \Big) \bigvee \Big(\max_{0 \leq k \leq N} g(y_k^{x,\boldsymbol{u}}) \Big) \bigvee g_f(y_N^{x,\boldsymbol{u}}) \tag{67}$$

(for $(x, z, \boldsymbol{u} = (u_k)) \in \mathbb{R}^d \times \mathbb{R} \times U^N$) and the corresponding auxiliary value is defined, for $(x, z) \in \mathbb{R}^d \times \mathbb{R}$, by

$$W(x, z) := \inf_{\boldsymbol{u}=(u_k)_k \in U^N} J(x, z, \boldsymbol{u}). \tag{68}$$

Notice that $W(x, z)$ converges to $\mathcal{W}(0, x, z)$ as $N \to \infty$ (i.e., $\Delta t = \frac{T}{N} \to 0$), where the continuous value function $\mathcal{W}$ is defined in (41). Under Assumptions $(\mathbf{H_0})$-$(\mathbf{H_4})$, the error estimate of $|W(x, z) - \mathcal{W}(0, x, z)|$ is bounded by $O(\frac{1}{N})$, see [22, Appendix B]. Furthermore, the sequence of discrete-time optimal trajectories (for $N \in \mathbb{N}$) provide convergent approximations of optimal trajectories of the continuous problem (68), see [7].

Remark 5.1. The discrete dynamics F and the discrete cost L can be defined as approximations of the time-continuous function f and ℓ. It is worth mentioning that the algorithms that will be presented in this section can also handle situations where the dynamics F and L are obtained by some statistical models which can be enriched during the computational process.

With similar arguments as in the proof of (42), we have

$$V(x) = \inf\{z \mid W(x, z) \leq 0\}. \tag{69}$$

For the sake of simplicity and without loss of generality, we suppose that the control is of dimension $r = 1$ and we denote by D its maximal diameter ($\forall a, a' \in A, \|a - a'\| \leq D$), although the approach can be generalized to control variables in multiple dimensions.

Planning algorithms are based on the principles of optimistic optimization. In order to minimize the objective function J over the space U^N, we refine, in an iterative way the search space into smaller subsets. A search space, called node and denoted by $\mathbb{U}_i$ with $i \in \mathbb{N}$, is a Cartesian product of sub-intervals of U, i.e., $\mathbb{U}_i := U_{i,0} \times U_{i,1} \times \cdots \times U_{i,N-1} \subseteq U^N$, where $U_{i,k}$ represents the control interval at time step k, for $k = 0, \ldots, N-1$. The collection of nodes will be organized into a tree Υ that will be constructed progressively by expanding the tree nodes. Expanding a node $\mathbb{U}_i$, with $i \in \mathbb{N}$, consists in choosing an interval $U_{i,k}$, for $k = 0, \ldots, N-1$, and splitting it uniformly to M sub-intervals where $M > 1$ is a parameter of the algorithm. The order of expanded nodes and the intervals that have to be split will be chosen in such a way to minimize the cost J. For now, we introduce some useful notations related to the tree Υ:

- We associate, for any node $\mathbb{U}_i \in \Upsilon$, a sample sequence of controls $\boldsymbol{u}_i := (u_{i,k})_{k=0}^{N-1} \in \mathbb{U}_i$ such that $u_{i,k}$ corresponds to the midpoint of the interval $U_{i,k}$ for any $k = 0, \ldots, N-1$.

- Denote $d_{i,k}$, for $k = 0, ..., N-1$, the diameter of the interval $U_{i,k}$ of some node $\mathbb{U}_i \in \Upsilon$. In particular,
$$d_{i,k} = \frac{D}{M^{s_i(k)}},$$
where $s_i(k)$ indicates the number of splits needed to obtain the interval $U_{i,k}$ for $k = 0, \ldots, N-1$.
- The depth p_i of a node $\mathbb{U}_i$ is the total number of splits done to obtain this node:
$$p_i := \sum_{k=0}^{N-1} s_i(k). \tag{70}$$
We denote by $\texttt{Depth}(\Upsilon)$ the maximal depth in the tree Υ.
- A node $\mathbb{U}_i$ is a tree leaf if it has not been expanded. The set of tree leaves is denoted by Λ.
- Finally, we denote by $\Lambda_p := \Big\{ \mathbb{U}_i \in \Upsilon \quad s.t. \quad p_i = p \Big\}$ the set of leaves of Υ of depth $p \in \mathbb{N}$.

By selecting controls at the intervals centers and by taking M odd, we guarantee that after expanding a node $\mathbb{U}_i$ we generate at least one node $\mathbb{U}_j$ with $J(x, z, \boldsymbol{u}_j) \leq J(x, z, \boldsymbol{u}_i)$. Indeed, the middle child $\mathbb{U}_j$ contains the control sequence of $\mathbb{U}_i$.

Proposition 5.1. *By the tree construction, there exists at least a leaf node $\mathbb{U}_i \in \Lambda$ containing an optimal control sequence and satisfying*
$$J(x, z, \boldsymbol{u}_i) - \sigma_i \leq W(x, z) \leq J(x, z, \boldsymbol{u}_i), \tag{71}$$
where $\boldsymbol{u}_i$ is the sample control sequence in $\mathbb{U}_i$ and where
$$\sigma_i := \Big(\sum_{k=0}^{N-1} \beta_k d_{i,k} \Big) \bigvee \Big(\sum_{k=0}^{N-1} \gamma_k d_{i,k} \Big), \tag{72}$$
with β_k and γ_k positive constants only depending on the Lipschitz constants of F, L, Φ and of Ψ.

In the optimistic planning algorithms, at each iteration, one or several optimistic nodes are chosen and split to get from each node M children ($M > 1$ is a fixed parameter of the algorithm). To expand a node $\mathbb{U}_i$, we choose an interval from $U_{i,0} \times U_{i,1} \times \cdots \times U_{i,N-1}$ and we partition it uniformly to M sub-intervals. If we choose to split the interval $U_{i,k}$, for some $k = 0, \ldots, N-1$, then M nodes will be generated and then the new error term $\sigma_i^+(k)$ is
$$\sigma_i^+(k) := \Big(\sum_{j=0, j\neq k}^{N-1} \beta_j d_{i,j} + \beta_k \frac{d_{i,k}}{M} \Big) \bigvee \Big(\sum_{j=0, j\neq k}^{N-1} \gamma_j d_{i,j} + \gamma_k \frac{d_{i,k}}{M} \Big).$$
Henceforth, in order to minimize the error $\sigma_i^+(k)$, the best choice of the interval to split, k_i^*, is given by:
$$k_i^* \in \underset{0 \leq k \leq N-1}{\operatorname{argmin}} \ \sigma_i^+(k). \tag{73}$$

The following result gives an upper bound on the error term σ_i, of any node $\mathbb{U}_i \in \Upsilon$.

Proposition 5.2. *Assume that the number of split $M > L_{F,x} > 1$. Consider a node $\mathbb{U}_i$ at some depth $p_i = p$. For p large enough, the error σ_i (defined in (72)) is bounded as follows:*
$$\sigma_i \leq \delta_p := \ c_1(N)\, \Delta t\, M^{-\frac{p}{N}}, \tag{74}$$
where $c_1(N) > 0$ is bounded independently of N.

We refer to [22] for the proof of this result and for the exact expression of the constant $c_1(N)$.

Now, we will present the rules for refining the search of an optimal control strategy. In the first algorithm, at each iteration, the node $\mathbb{U}_{i^*}$ minimizing the lower bound $(J(x,z,u_i)-\sigma_i)$ will be selected and split to M children. More precisely, we identify an interval $U_{i^*,k^*_{i^*}}$ whose partition in M sub-intervals will produce the lowest error $\sigma_{i^*}(k^*_{i^*})$.

Algorithm Optimistic Planning (OP)

Require: The number of intervals N, the split factor M, the maximal number of expanded nodes $I_{\max}$

1: Initialize Υ with a root $\mathbb{U}_0 := U^N$ and $n=0$ ($n :=$ number of expanded nodes).
2: **while** $n < I_{\max}$ **do**
3: Select an optimistic node to expand: $\mathbb{U}_{i^*} \in \underset{\mathbb{U}_i\in\Lambda}{\operatorname{argmin}}\,(J(x,z,\boldsymbol{u}_i)-\sigma_i)$.
4: Select $k^*_{i^*}$, defined in (73), the interval to split for the node $\mathbb{U}_{i^*}$.
5: Update Υ by expanding $\mathbb{U}_{i^*}$ along $k^*_{i^*}$ and adding its M children.
6: Update $n = n+1$.
7: **end while**
8: **return** Control sequence $\boldsymbol{u}_{i^*} = (u_{i^*,k})_k \in U^N$ of the node $\mathbb{U}_{i^*} \in \underset{\mathbb{U}_i\in\Lambda}{\operatorname{argmin}}\, J(x,z,\boldsymbol{u}_i)$.

Theorem 5.1. *Assume that $M > L_{F,x} > 1$. Let $\boldsymbol{u}_{i^*}$ and $J(x,z,\boldsymbol{u}_{i^*})$ be the output of the OP algorithm, and let $n \geq 1$ be the corresponding number of expanded nodes. We have*

$$0 \leq J(x,z,\boldsymbol{u}_{i^*}) - W(x,z) \longrightarrow 0, \text{ when } n \to +\infty. \tag{75}$$

In Algorithm (OP), the number $I_{\max}$ represents a maximal available computational resource. The number of expanded nodes corresponds to the number of iterations, since at each iteration only one node is expanded. Other optimistic planning methods can be considered. For instance, the *simultaneous optimistic planning* (SOP) algorithm or *simultaneous optimistic planning with multiple steps* (SOPMS) algorithm that expand at each iteration several nodes at every iteration, see [22].

Test 5. To show the relevance of this approach, we consider a variant of Zermelo problem where a boat targets the set $\mathcal{C} := \mathbb{B}(0;0.1)$, at time $T=1$, with minimal fuel consumption. The dynamics is similar to the one considered in (63). We consider also two rectangular obstacles with horizontal and vertical half lengths (r_x, r_y). The first obstacle is centered at $(-2.0, 0.5)$ with $(r_x, r_y) = (0.4, 0.4)$, and the second obstacle is centered at $(-2.5, -1)$ with $(r_x, r_y) = (0.2, 1)$. To take into account the pointwise and final state constraints, we define the functions g and g_f by

$$g(x) := \Big(0.4 - \|x-(-2,0.5)\|_\infty\Big) \bigvee \min\big(0.2-|x_1+2.5|, 1-|x_2+1|\big) \quad \text{and} \quad g_f(x) := \|x\|_\infty - 0.1.$$

For a given N, the discrete control problem becomes:

$$V(x) = \inf\Big\{ \tfrac{1}{N}\textstyle\sum_{k=0}^{N-1} u_{1,k} \quad \text{with} \quad \boldsymbol{u} = ((u_{1,k}, u_{2,k}))_k \in U^N, \\ g(y_k^{x,\boldsymbol{u}}) \leq 0 \quad \text{for } k = 0, \cdots, N, \quad \text{and} \quad g_f(y_N^{x,\boldsymbol{u}}) \leq 0 \Big\},$$

where $(y_k^{x,\boldsymbol{u}})_k$ is the discrete state variable, corresponding to the control policy $\boldsymbol{u} \in U^N$, and starting at the initial position x, while the discrete auxiliary value function is defined as

$$W(x,z) = \inf_{\boldsymbol{u}\in U^N} \Big\{ \Big(\frac{1}{N} \sum_{k=0}^{N-1} u_{1,k} - z \Big) \bigvee \Big(\max_{0\le k\le N} g(y_k^{x,\boldsymbol{u}}) \Big) \bigvee g_f(y_N^{x,\boldsymbol{u}}) \Big\}.$$

Figure 14 displays optimal trajectories obtained from three different initial positions. A simultaneous optimistic planning algorithm is used for this simulation with $I_{\max} = 3200$ and $N = 40$. The optimal controls are displayed on Figure 15. In this example, we can see that the

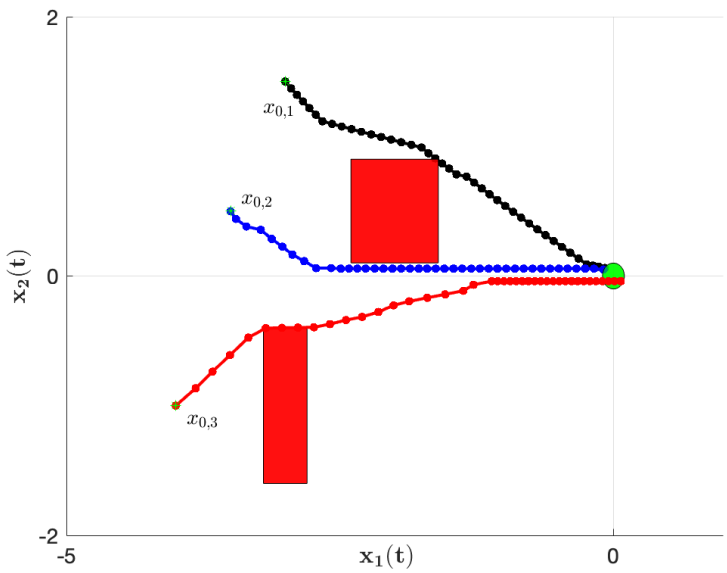

Figure 14: (Test 5) An optimistic planning approach with $N = 40$. Optimal trajectories corresponding to three different initial data. The trajectory in black corresponds to the initial position $x_{0,1} = (-3, 1.5)$, the trajectory in blue corresponds to $x_{0,2} = (-3.5, 0.5)$, and the trajectory in red corresponds to $x_{0,3} = (-4, -1)$

optimal trajectories, computed by an OP approach, reach the target and avoid the obstacles. The main feature of the OP approaches is the fact that they give an approximation of the global solution. For a fixed value of N, the complexity of these (global) approaches depends on the dimension of the control and not on the dimension of the space variable. The complexity increases also when the discretization is refined (i.e., when N increases). As pointed out in the literature [33, 32, 22], the accuracy of the numerical solutions, obtained by OP methods, depends on the available numerical resources $I_{\max}$. The convergence results derived in the literature give some hints on how to choose $I_{\max}$ to obtain a given precision, but this question deserves further analysis.

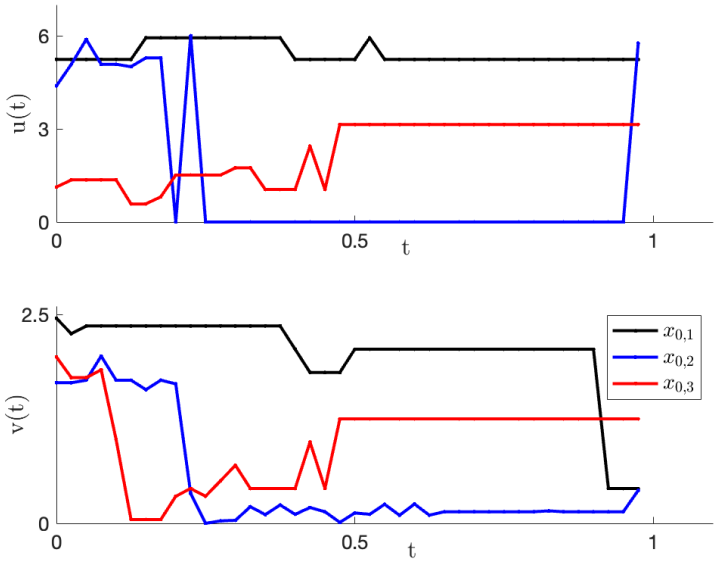

Figure 15: (Test 5) An optimistic planning approach with $N = 40$. Control laws corresponding to three different initial positions. The control law in black corresponds to the initial position $x_{0,1} = (-3, 1.5)$, the control laws in blue corresponds to $x_{0,2} = (-3.5, 0.5)$, and the control laws in red corresponds to $x_{0,3} = (-4, -1)$

References

[1] A. A. Agrachev. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.

[2] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[3] A. Alla and M. Falcone. An adaptive POD approximation method for the control of advection-diffusion equations. *Control and Optimization with PDE Constraints*, pages 1–17, 2013.

[4] A. Alla, M. Falcone, and L. Saluzzi. An efficient DP algorithm on a tree-structure for finite horizon optimal control problems. *SIAM Journal on Scientific Computing*, 41(4):A2384–A2406, 2019.

[5] A. Altarovici, O. Bokanowski, and H. Zidani. A general Hamilton-Jacobi framework for non-linear state-constrained control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(02):337–357, 2013.

[6] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 2018.

[7] M. Assellaou, O. Bokanowski, A. Desilles, and H. Zidani. Value function and optimal trajectories for a maximum running cost control problem with state constraints. application to an abort landing problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(1):305–335, 2018.

[8] J.-P. Aubin and A. Cellina. *Differential inclusions. Set-valued maps and viability theory*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1984.

[9] J. P. Aubin and H. Frankowska. *Set-Valued Analysis. Systems and Control: Foundations and Applications, vol. 2*. Birkhäuser, Boston, 1990.

[10] S. Augoula and R. Abgrall. High order numerical discretization for hamilton–jacobi equations on triangular meshes. *Journal of Scientific Computing*, 15(2):197–229, 2000.

[11] F. Ballarin, G. Rozza, and M. Strazzullo. Space-time pod-galerkin approach for parametric flow control. *Handbook of Numerical Analysis, Volume 1*, (to appear).

[12] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.

[13] G. Barles. Solutions de viscosité des équations de hamilton-jacobi. *Collection SMAI*, 1994.

[14] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.

[15] E. N. Barron. Viscosity solutions and analysis in L_∞. In *Nonlinear analysis, differential equations and control*, volume 528 of *Serie C: Mathematical and Physical Sciences*, pages 1–60. Springer Science+Business Media Dordrecht, 1999.

[16] P. Bettiol, A. Bressan, and R. Vinter. On trajectories satisfying a state constraint: $w^{1,1}$ estimates and counterexamples. *SIAM Journal on Control and Optimization*, 48(7):4664–4679, 2010.

[17] J. T. Betts. *Practical methods for optimal control using nonlinear programming*, volume 36 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, [2020] ©2020. Third edition of [1826768].

[18] O. Bokanowski, A. Briani, and H. Zidani. Minimum time control problems for non autonomous differential equations. *Systems and Control Letters*, 58(10-11):742–746, 2009.

[19] O. Bokanowski, A. Desilles, and H. Zidani. Relationship between maximum principle and dynamic programming in presence of intermediate and final state constraints. *ESAIM: Control, Optimisation and Calculus of Variations*, 27, 2021.

[20] O. Bokanowski, M. Falcone, R. Ferretti, L. Grüne, D. Kalise, and H. Zidani. Value iteration convergence of ϵ-monotone schemes for stationary hamilton-jacobi equations. *Discrete and Continuous Dynamical Systems-Series A*, 35(9):4041–4070, 2015.

[21] O. Bokanowski, M. Falcone, and S. Sahu. An efficient filtered scheme for some first order time-dependent hamilton–jacobi equations. *SIAM Journal on Scientific Computing*, 38(1):A171–A195, 2016.

[22] O. Bokanowski, N. Gammoudi, and H. Zidani. Optimistic planning algorithms for state-constrained optimal control problems. *Computers and Mathematics with Applications*, 109(1):158–179, 2002.

[23] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-lagrangian scheme for first order Hamilton-Jacobi Bellman equations. *Journal of Scientific Computing*, 55(3), 575-605), 2013.

[24] J. F. Bonnans and A. Hermant. Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control. *SIAM J. Control Optim.*, 46(4):1398–1430, 2007.

[25] J. F. Bonnans and A. Hermant. Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(2):561–598, 2009.

[26] J. F. Bonnans and J. Laurent-Varin. Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control. *Numer. Math.*, 103(1):1–10, 2006.

[27] B. Bonnard, J.-B. Caillau, and E. Trélat. Second order optimality conditions in the smooth case and applications in optimal control. *ESAIM Control Optim. Calc. Var.*, 13(2):207–236, 2007.

[28] B. Bonnard and M. Chyba. *Singular trajectories and their role in control theory*, volume 40 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2003.

[29] B. Bonnard, O. Cots, J. Gergaud, and B. Wembé. Abnormal geodesics in 2d-zermelo navigation problems in the case of revolution and the fan shape of the small time balls. *Systems Control Lett.*, 161(105140):1–8, 2022.

[30] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux*, volume 51 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006.

[31] J. B. Buckheit and D. L. Donoho. Wavelab and reproducible research. Technical Report 474, Stanford University, 1995.

[32] L. Buşoniu, E. Páll, and R. Munos. Discounted near-optimal control of general continuous-action nonlinear systems using optimistic planning. In *2016 American Control Conference (ACC)*, pages 203–208. IEEE, 2016.

[33] L. Buşoniu, E. Páll, and R. Munos. Continuous-action planning for discounted infinite-horizon nonlinear optimal control with lipschitz values. *Automatica*, 92:100–108, 2018.

[34] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, volume 58 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA, 2004.

[35] E. Carlini, R. Ferretti, and G. Russo. A weighted essentially nonoscillatory, large time-step scheme for hamilton–jacobi equations. *SIAM Journal on Scientific Computing*, 27(3):1071–1091, 2005.

[36] Y. Chitour, F. Jean, and E. Trélat. Genericity results for singular curves. *J. Differential Geom.*, 73(1):45–73, 2006.

[37] Y. Chitour, F. Jean, and E. Trélat. Singular trajectories of control-affine systems. *SIAM J. Control Optim.*, 47(2):1078–1095, 2008.

[38] F. Clarke, Y. S. Ledyaev, R. Stern, and P. Wolenski. *Nonsmooth analysis and control theory*. Springer, 1998.

[39] F. H. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Springer, 2013.

[40] F. H. Clarke and R. B. Vinter. The relationship between the maximum principle and dynamic programming. *SIAM J. Control Optim.*, 25(5):1291–1311, 1987.

[41] G. Colombo, N. Khalil, and F. Rampazzo. Rotational controls and uniqueness of constrained viscosity solutions of hamilton-jacobi pde. *https://arxiv.org/abs/2110.08530*, 2021.

[42] M. Crandall, L. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc*, 282(2):487–502, 1984.

[43] M. Crandall and P. Lions. Condition d'unicité pour les solutions generalisées des équations de Hamilton-Jacobi du premier ordre. *C. R. Acad. Sci. Paris Sér. I Math.*, 292:183–186, 1981.

[44] M. Crandall and P. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277:1–42, 1983.

[45] M. Crandall and P. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.*, 43:1–19, 1984.

[46] G. N. Elnagar and M. A. Kazemi. Pseudospectral Chebyshev optimal control of constrained nonlinear dynamical systems. *Comput. Optim. Appl.*, 11(2):195–217, 1998.

[47] M. Falcone and R. Ferretti. Discrete time high-order schemes for viscosity solutions of Hamilton-Jacobi-Bellman equations. *Numerische Mathematik*, 67(3):315–344, 1994.

[48] M. Falcone and R. Ferretti. Semi-Lagrangian schemes for Hamilton–Jacobi equations, discrete representation formulae and Godunov methods. *Journal of computational physics*, 175(2):559–575, 2002.

[49] M. Falcone and R. Ferretti. *Semi-Lagrangian approximation schemes for linear and Hamilton—Jacobi equations.* SIAM, 2013.

[50] T. Faulwasser and L. Grüne. Turnpike properties in optimal control: An overview of discrete-time and continuous-time results. In *Numerics of Control*, Handbook of Numerical Analysis. Elsevier, 2022, to appear.

[51] R. Ferretti. Convergence of semi-Lagrangian approximations to convex Hamilton–Jacobi equations under (very) large Courant numbers. *SIAM journal on numerical analysis*, 40(6):2240–2253, 2002.

[52] W. H. Fleming and R. W. Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

[53] R. Fourer, D. Gay, and B. Kernighan. *AMPL: A modeling language for mathematical programming.* Duxbury Press, Second edition, 2002.

[54] H. Frankowska and M. Mazzola. On relations of the adjoint state to the value function for optimal control problems with state constraints. *Nonlinear Differential Equations and Applications NoDEA*, 20:361–383, 2013.

[55] B. D. Froese and A. M. Oberman. Convergent filtered schemes for the Monge–Ampère partial differential equation. *SIAM Journal on Numerical Analysis*, 51(1):423–444, 2013.

[56] J. Garcke and A. Kröner. Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids. *Journal of Scientific Computing*, 70(1):1–28, 2017.

[57] Q. Gong, I. M. Ross, W. Kang, and F. Fahroo. Connections between the covector mapping theorem and convergence of pseudospectral methods for optimal control. *Comput. Optim. Appl.*, 41(3):307–335, 2008.

[58] W. W. Hager. Runge-Kutta methods in optimal control and the transformed adjoint system. *Numer. Math.*, 87(2):247–282, 2000.

[59] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration.* Springer, 2006.

[60] R. F. Hartl, S. Sethi, and R. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.*, 37(2):181–218, 1995.

[61] C. Hermosilla and H. Zidani. Relationship between the maximum principle and dynamic programming for minimax problems. *Applied Mathematics & Optim.*, To appear, 2023.

[62] J. B. Lasserre. *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series*. Imperial College Press, London, 2010.

[63] J. B. Lasserre and D. Henrion. Moments and convex optimization for analysis and control of nonlinear partial differential equations. *Handbook on Numerical Analysis, Volume 1*, (to appear).

[64] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM J. Control Optim.*, 47(4):1643–1666, 2008.

[65] E. B. Lee and L. Markus. *Foundations of optimal control theory*. John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[66] C.-T. Lin and E. Tadmor. l^1-stability and error estimates for approximate Hamilton-Jacobi solutions. *Numerische Mathematik*, 87(4):701–735, 2001.

[67] P. Malisani, F. Chaplais, and N. Petit. An interior penalty method for optimal control problems with state and input constraints of nonlinear systems. *Optimal Control Appl. Methods*, 37(1):3–33, 2014.

[68] R. Munos. From bandits to Monte-Carlo tree search: the optimistic principle applied to optimization and planning. *Foundations and Trends in Machine Learning*, 7(1):1–129, 2014.

[69] J. Noceal and S. J. Wright. *Numerical optimization*. Springer, 2006. Second edition.

[70] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(1):12–49, 1988.

[71] S. Osher and C.-W. Shu. High essentially nonoscillatory schemes for Hamilton-Jacobi equations. *SIAM Journal of Numerical Analysis*, 28(4):907–922, 1991.

[72] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. A Pergamon Press Book. The Macmillan Company, New York, 1964. Translated by D. E. Brown.

[73] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.

[74] L. Rifford and E. Trélat. On the stabilization problem for nonholonomic distributions. *J. Eur. Math. Soc. (JEMS)*, 11(2):223–255, 2009.

[75] I. M. Ross and F. Fahroo. Legendre pseudospectral approximations of optimal control problems. In *New trends in nonlinear dynamics and control, and their applications*, volume 295 of *Lect. Notes Control Inf. Sci.*, pages 327–342. Springer, Berlin, 2003.

[76] J. D. L. Rowland and R. B. Vinter. Construction of optimal feedback controls. *Systems & Control Letters*, 16(5):357–367, 1991.

[77] J. M. Sanz-Serna. Symplectic Runge-Kutta schemes for adjoint equations, automatic differentiation, optimal control, and more. *SIAM Rev.*, 58(1):3–33, 2016.

[78] H. Soner. Optimal control with state-space constraint I. *SIAM Journal on Control and Optimization*, 24(3):552–561, 1986.

[79] G. Stefani. Regularity properties of the minimum-time map. In *Nonlinear synthesis (Sopron, 1989)*, volume 9 of *Progr. Systems Control Theory*, pages 270–282. Birkhäuser Boston, Boston, MA, 1991.

[80] E. Trélat. Optimal control and applications to aerospace: some results and challenges. *J. Optim. Theory Appl.*, 154(3):713–758, 2012.

[81] E. Trélat and E. Zuazua. The turnpike property in finite-dimensional nonlinear optimal control. *J. Differential Equations*, 258(1):81–114, 2015.

[82] P. Tsiotras and H. J. Kelley. Goddard problem with constrained time of flight. *Journal of Guidance, Control and Dynamics*, 15:289–296, 1992.

[83] R. B. Vinter. New results on the relationship between dynamic programming and the maximum principle. *Mathematics of Control, Signals and Systems*, 1(1):97–105, 1988.

[84] R. B. Vinter. *Optimal control*. Birkhaüser, Boston, 2000.

[85] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*, 106(1, Ser. A):25–57, 2006.