



The minimum time-to-climb and fuel consumption problems and CAS/Mach procedure for aircraft

Olivier Cots, Joseph Gergaud, D Goubinat

► To cite this version:

Olivier Cots, Joseph Gergaud, D Goubinat. The minimum time-to-climb and fuel consumption problems and CAS/Mach procedure for aircraft. 2018. hal-01936193

HAL Id: hal-01936193

<https://hal.science/hal-01936193>

Preprint submitted on 27 Nov 2018

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.

The minimum time-to-climb and fuel consumption problems and CAS/Mach procedure for aircraft

O. Cots¹ and J. Gergaud²

*Toulouse Univ., INP-ENSEEIH-IRIT, UMR CNRS 5505,
2 rue Camichel, 31071 Toulouse, France*

D. Goubinat³

*Thales Avionics SA, 105 av du General Eisenhower,
B.P. 1147, 31047 Toulouse Cedex, France*

In this article, we are interested in optimal aircraft trajectories in climbing phase. We consider the cost index criterion which is a convex combination of the time-to-climb and the fuel consumption. We assume that the thrust is constant and we control the air slope of the aircraft. This optimization problem is modeled as a Mayer optimal control problem with a single-input affine dynamics in the control and with two pure state constraints, limiting the Calibrated AirSpeed (CAS) and the Mach speed. The candidates as minimizers are selected among a set of extremals given by the maximum principle. We first analyze the minimum time-to-climb problem with respect to the bounds of the state constraints, combining small time analysis, indirect multiple shooting and homotopy methods. This investigation emphasizes two strategies: the common CAS/Mach procedure in aeronautics and the classical Bang-Singular-Bang policy in control theory. We then compare these two procedures for the cost index criterion.

I. Introduction

The climbing phase is a normalized phase where most of the civil aircrafts follow the CAS/Mach procedure. This procedure splits the climbing phase in two parts, the first one is an arc at constant *Calibrated AirSpeed* (CAS) and the second one is an arc at constant *Mach* number. According to Refs. [24, 36], the CAS constant is chosen among the ones which maximize the climbing slope, the vertical speed or the climbing rate. The Mach constant depends most of the time on the cruise Mach number but this value is also closed to the one which maximizes the vertical speed. The climbing path resulting from a CAS/Mach couple may be seen as an approximation of an optimal flight, minimizing a cost made by a convex combination of the time-to-climb and the fuel consumption, that we call the *cost index*. Even though historically a lot of studies was done, see for instance Refs. [4, 30, 31], to optimize the climbing flight path, the CAS/Mach procedure is still used mainly because of its simplicity. The CAS and the Mach number are fairly easy quantities to compute since they are defined thanks to the differential pressure which is directly available from the Pitot tubes present all over the fuselage of the aircraft.

In this article, we study first the minimum time-to-climb trajectory submitted to state constraints which are related to the CAS and the Mach number. The dynamics of this climbing phase

¹ Assistant Professor, Institut de Recherche en Informatique de Toulouse, Team APO, olivier.cots@enseeiht.fr

² Professor, Institut de Recherche en Informatique de Toulouse, Team APO, joseph.gergaud@enseeiht.fr

³ FMS system engineer, Thales Avionics Toulouse damien.goubinat@fr.thalesgroup.com

Table 1 Medium-haul aircraft parameters.

Data	Value	Unit	Data	Value	Unit
S	122.6	m ²	$C_{s,1}$	1.055e ⁻⁵	kg.s ⁻¹ .N ⁻¹
g	9.81	m.s ⁻²	$C_{s,2}$	441.54	m.s ⁻¹
$C_{T,1}$	141040	N	R	287.058	J.kg ⁻¹ .K ⁻¹
$C_{T,2}$	14909.9	m	Θ_0	288.15	K
$C_{T,3}$	6.997e ⁻¹⁰	m ⁻²	β	0.0065	K.m ⁻¹
$C_{d,1}$	0.0242		P_0	101325	Pa
$C_{d,2}$	0.0469				

is depicted by the following four-dimensional dynamical system:

$$\frac{dh}{dt} = v \sin \gamma \quad (1)$$

$$\frac{dv}{dt} = \frac{T(h)}{m} - \frac{1}{2} \frac{\rho(h) S v^2}{m} C_d(C_l) - g_0 \sin \gamma \quad (2)$$

$$\frac{dm}{dt} = -C_s(v) T(h) \quad (3)$$

$$\frac{d\gamma}{dt} = \frac{1}{2} \frac{\rho(h) S v}{m} C_l - \frac{g_0}{v} \cos \gamma, \quad (4)$$

where the state variable is composed of the altitude, the true air speed, the mass and the air slope of the aircraft. We refer to Ref. [23] for more details about the dynamics. The altitude h is given in meter (m), the true air speed v in meter per second (m.s⁻¹), the mass m in kilogram (kg) and the air slope γ in radian (rad). In this model, the lift coefficient C_l may be considered as the control variable. The BADA [34] model is chosen to represent the aircraft performance model and provides the following expressions:

$$C_s(v) := C_{s,1} \left(1 + \frac{v}{C_{s,2}} \right), \quad T(h) := C_{T,1} \left(1 - \frac{h}{C_{T,2}} + h^2 C_{T,3} \right), \quad C_d(C_l) := C_{d,1} + C_{d,2} C_l^2,$$

where the constants $C_{T,i}$, $C_{s,i}$ and $C_{d,i}$ depend on the flight phase and on the aircraft. The International Standard Atmospheric (ISA) model is used to represent the atmospheric model from the sea level till the end of the tropopause which here is considered at 11 000 meters. This model provides the evolution of the pressure P , the temperature Θ and the air density ρ with respect to the altitude through the following expressions:

$$P(h) := P_0 \left(\frac{\Theta(h)}{\Theta_0} \right)^{\frac{g_0}{\beta R}}, \quad \Theta(h) := \Theta_0 - \beta h, \quad \rho(h) := \frac{P(h)}{R \Theta(h)}.$$

The remaining data are positive constants: g_0 is the gravitational constant at the sea level, S the wing area, R the specific constant of air, β the thermal gradient and P_0 , Θ_0 represent the pressure and the temperature at the sea level. See Table 1 for the chosen values of the constant parameters for the numerical experiments.

Taking into account the air slope γ in the dynamics introduces numerical instabilities [19, 20] which are known as *singular perturbations* [3, 13, 32] and which come from large time constant differences between the state variables. A comparison of the time constants, see Refs. [3, 23], shows that the dynamics (1)–(4) contains slow (the mass m) and fast (the air slope γ) variables. The altitude h and the true air speed v are fast compare to the mass but slow compare to the air slope. In this article, we consider h and v as slow variables. The singular perturbation phenomenon arise when a system of differential equations contains at least one small parameter multiplying the derivative of one or more state variables. To emphasize the presence of a singular perturbation in the air slope dynamics, we introduce a parameter $\varepsilon > 0$ such that eq. (4) is replaced by:

$$\varepsilon \frac{d\gamma}{dt} = \frac{1}{2} \frac{\rho(h) S v}{m} C_l - \frac{g_0}{v} \cos \gamma. \quad (4')$$

Let $(P_{t_f}^\varepsilon)$ denote the minimum time-to-climb problem with the additional artificial parameter $\varepsilon > 0$. From the control theory, for a fixed value of $\varepsilon > 0$, the candidates as minimizers are selected among a set of *BC-extremals*, solution of a Hamiltonian system given by the *Pontryagin Maximum Principle* (PMP), see Ref. [35]. The application of the PMP leads to define a Boundary Value Problem denoted (BVP^ε) , in terms of state and adjoint variables, which can be solved using indirect multiple shooting methods [11]. It is well known that multiple shooting increases numerical stability and a good alternative would be to use direct multiple shooting [5] to solve the optimal control problem. Difficulties in solving the singularly perturbed boundary value problem (BVP^ε) arise because the solution exhibits narrow regions of very fast variation. In this case, it may be difficult to determine a mesh for the nodes of the multiple shooting method that will give an accurate numerical solution and to give an initial guess that will lead to the convergence of the underlying Newton-like algorithm. In this context, an alternative approach would be to solve the boundary value problem using automatic mesh refinement, see Refs. [15, 16], or to use homotopy techniques, see Ref. [20].

Another point of view to deal with singular perturbations is to approximate the solution using asymptotic expansions. An overview of methods to deal with singular perturbations may be found in Refs. [3, 13, 32]. We consider in this article only the zero-order term of the asymptotic expansion. To uniformly approximate the solution of a singularly perturbed boundary value problem, one has to compute at least two different approximations (the so-called inner and outer solutions) which are accurate only for part of the range of the time variable, and usually they are valid on different length-scales. The Method of Matched Asymptotic Expansion consists in computing the inner and outer solutions and then in matching them to get a uniform approximation which satisfies the boundary conditions. In this article, we consider only the zero-order approximation of the outer solution since it has been shown for the minimum time-to-climb problem with no state constraints that it is a good approximation of the solution, see Ref. [20]. Note that in the aircraft dynamics the zero-order reduction is equivalent to the quasi-steady approximation of the flight which is commonly performed [21, 24, 29, 33, 36].

The reduced-order optimal control problem denoted (P_{t_f}) is obtained putting ε to 0 in eq. (4') and considering γ as the new control variable. The problem (P_{t_f}) has one state variable less than $(P_{t_f}^\varepsilon)$, $\varepsilon > 0$, and can be tackled with the tools from geometric optimal control theory. We refer to Refs. [2, 3, 23] for details about the reduction process. In particular, one important result is that the reduced-order boundary value problem obtained from (BVP^ε) putting ε to 0 is equivalent to the boundary value problem obtained after applying the maximum principle to the reduced-order dynamical system obtained from (1)–(4') putting ε to 0, when we consider at the end γ as the new control variable. Besides, the reduced-order dynamical system with γ as the control variable is affine with respect to the control if we consider the small-angle approximation, that is if we replace $\cos \gamma$ by 1 and $\sin \gamma$ by γ assuming γ is small. In this case, the associated pseudo-Hamiltonian is of the form $H(x, p, u) = H_0(x, p) + u H_1(x, p)$ where $x := (h, v, m)$ is the reduced state, p is the adjoint variable, u the control (*i.e.* the air slope γ), and where H_0 and H_1 are two Hamiltonian lifts. In addition, to complete the definition of the optimal control problem, the control has to satisfy a constraint of the form $u \in [u_{\min}, u_{\max}]$ and the state has to satisfy two constraints denoted $\phi(x) \leq \phi_{\max}$ and $\psi(x) \leq \psi_{\max}$, where ϕ represents the CAS, ψ represents the Mach, and where ϕ_{\max} and ψ_{\max} may be seen as given speed limitations.

The minimum time-to-climb problem with no state constraints, or from another point of view, with state constraints but with ϕ_{\max} and ψ_{\max} big enough, is analyzed in Refs. [18, 20]. In Ref. [18], the influence of the initial mass m_0 and the final true air speed v_f is studied and it is shown in particular that the trajectories are of the form $\sigma_\pm \sigma_s \sigma_\pm$, where σ_- represents a *bang* arc where $u(t) = u_{\min}$ along the arc, where σ_+ represents a *bang* arc with $u(t) = u_{\max}$ and σ_s represents a *singular* arc where $u(t) \in (u_{\min}, u_{\max})$. These trajectories from the state unconstrained case are compliant with the so-called *Maximal Operation Speed* (VMO) and *Maximal Operationnal Mach* (MMO), that is $\phi(x(t)) < \text{VMO}$ and $\psi(x(t)) < \text{MMO}$ along the trajectories. The first goal of this article is to analyze the influence of the bounds ϕ_{\max} and ψ_{\max} on the structure of the trajectories for the minimum time-to-climb problem. The purpose is to classify the different structures with respect to ϕ_{\max} and ψ_{\max} . It is clear that reducing both values of the bounds will increase the final time and it is expected that it will decrease, up to a point, the fuel consumption. The second part of this article is devoted to the comparison between the solutions from this classification and the

minimum fuel consumption solution without speed limitations. Finally, we present a comparison between minimum cost index solutions and trajectories that follow the simple CAS/Mach procedure.

The paper is organized as follows. The physical model with the statements of the optimal control problems are given in Sect. II A while the general necessary conditions of optimality for both the minimum time-to-climb and the fuel consumption problems are given in Sect. II C. A small time analysis which describes the behavior of *hyperbolic* trajectories with state constraints of *order one* is given in Sect. III A. This analysis is applied to the minimum time-to-climb problem in Sect. III B where we give preliminary numerical results about the optimal structures in presence of state constraints on the CAS and Mach speeds. The classification of the optimal structures is then obtained numerically by homotopy methods in Sect. III C. A comparison between minimal cost index solutions and the CAS/Mach procedure is then proposed in Sect. IV. Finally, Sect. V concludes the article.

II. General results about the minimum time-to-climb and fuel consumption problems

A. Definitions of the optimal control problems

We restrict the dynamics to the vertical motion of the aircraft. A complete description of the motion can be found in Ref. [23]. The aircraft is subjected to four forces, the Drag \vec{D} , the Lift \vec{L} , the Thrust \vec{T} and its own weight \vec{P} . A non linear point mass representation is used and we consider that all the forces apply on the center of gravity of the aircraft. We assume that the thrust is colinear to the velocity vector \vec{V} , which means that the angle of attack is neglected here, and that the aircraft evolves in an horizontal constant wind field. The application of the first dynamics principle assuming that the earth is galilean, provides the four eqs. (1)–(4).

During the flight, the lift coefficient C_l depends on the variation of the angle of attack. Since this quantity is not taken into account here, it is quite natural to consider the lift coefficient C_l as the control variable. However, according to the work presented in Refs. [3, 13] this four-dimensional dynamics contains slow (m) and fast (γ) variables. The time scale separation between the slow and fast variables is handled by a singular perturbation analysis which consists here in computing a zero-order approximation of the solution. The reduction process has two main features. In a first step, we substitute eq. (4) by the quasi-steady approximation:

$$0 = \frac{1}{2} \frac{\rho S v}{m} C_l - \frac{g_0}{v} \cos \gamma,$$

which gives

$$C_l = \frac{2mg_0}{\rho S v^2} \cos \gamma.$$

In a second time, the lift coefficient is replaced by the previous expression in eqs. (1)–(3) and the air slope γ is taken as the new control variable. We also consider that the air slope remains small (γ varies from $u_{\min} := 0$ to $u_{\max} := 0.262$ rad) during the climbing so we set $\cos \gamma \approx 1$ and $\sin \gamma \approx \gamma$. These considerations lead to the new affine control system:

$$\dot{x}(t) = F_0(x(t)) + u(t) F_1(x(t)), \quad (5)$$

where $x := (h, v, m)$, $u := \gamma$,

$$F_0(x) := \begin{pmatrix} 0 \\ \frac{T(h)}{m} - \frac{1}{2} \frac{\rho(h) S v^2}{m} C_{d,1} - 2 \frac{mg_0^2}{\rho(h) S v^2} C_{d,2} \\ -C_s(v)T(h) \end{pmatrix} \quad \text{and} \quad F_1(x) := \begin{pmatrix} v \\ -g_0 \\ 0 \end{pmatrix}.$$

According to Ref. [20], this reduced dynamics is a sufficient approximation of the initial dynamics considering the state unconstrained minimum time-to-climb problem.

The true air speed of an aircraft is quite difficult to measure during the flight, that is why historically, the concept of CAS was introduced. This speed can be computed using the differential ratio of pressure and can be expressed, through Bernoulli's equations, as a function of the true air

speed. The Mach speed, for its part, is defined by the ratio between the true air speed and the speed of the sound. Here, the CAS is denoted by ϕ and the Mach by ψ and the corresponding expressions are given by:

$$\phi(x) := \sqrt{2 \frac{R \Theta_0}{\kappa} \left(\left(\frac{P(h)}{P_0} \left(\frac{\kappa v^2}{2R \Theta(h)} + 1 \right)^{\frac{1}{\kappa}} + 1 \right)^{\kappa} - 1 \right)}, \quad \psi(x) := \frac{v}{\sqrt{\gamma_{\text{air}} R \Theta(h)}},$$

where $\gamma_{\text{air}} := 1.4 \text{ J.K}^{-1}$ is the heat capacity of the air and $\kappa := \gamma_{\text{air}}/(1-\gamma_{\text{air}})$. From these expressions, we define the state constraints c_1 and c_2 by:

$$c_1(x) := \phi(x) - \phi_{\max}, \quad (6)$$

$$c_2(x) := \psi(x) - \psi_{\max}, \quad (7)$$

where ϕ_{\max} (resp. ψ_{\max}) can be fixed to VMO (resp. MMO) or to smaller specified values.

In this article we are interested in the minimization of different criteria: the final time t_f , the fuel consumption $\Delta_m := m_0 - m_f$, where m_0 is the initial mass and m_f the final mass, and a convex combination of the final time and the fuel consumption:

$$g_\alpha(t_f, m_f) := \alpha t_f + (1 - \alpha)(m_0 - m_f),$$

where the initial mass m_0 will be fixed and given. The criterion g_α , for $\alpha \in [0, 1]$, is called the cost index. The initial state will be fixed to the realistic values $x_0 := (h_0, v_0, m_0) := (3480, 128.6, 69000)$. We define the set \mathcal{U}_t of *admissible controls*, that is the subset of

$$\{u \mid u: [0, t] \rightarrow [u_{\min}, u_{\max}] \text{ measurable}\}$$

such that the corresponding trajectory solution of eq. (5) with $x(0) = x_0$ is well defined over $[0, t]$. The *minimum cost index problem* can be summarized by:

$$(P_\alpha) \quad \begin{cases} \min_{(t_f, u) \in D} g_\alpha(t_f, m(t_f)), \\ \dot{x}(t) = F_0(x(t)) + u(t) F_1(x(t)), \quad u(t) \in U, \quad t \in [0, t_f] \text{ a.e.}, \quad x(0) = x_0, \\ c_1(x(t)) \leq 0, \quad t \in [0, t_f], \\ c_2(x(t)) \leq 0, \quad t \in [0, t_f], \\ b(x(t_f)) = 0, \end{cases}$$

where $U := [u_{\min}, u_{\max}]$, $D := \{(t_f, u) \mid t_f \geq 0, u \in \mathcal{U}_{t_f}\}$ and

$$b(x) := \begin{pmatrix} h - h_f \\ v - v_f \end{pmatrix}$$

with $h_f := 9144$ and $v_f := 191$. Thus, the final altitude and true air speed are fixed while the final mass is free. The *minimum time-to-climb problem*, denoted (P_{t_f}) , is defined as $(P_\alpha|_{\alpha=1})$ while the *minimum fuel consumption problem*, denoted (P_{Δ_m}) , is defined as $(P_\alpha|_{\alpha=0})$.

B. Notation

Let F_0, F_1 be two smooth vector fields and c a scalar smooth function on \mathbb{R}^n . The *Lie derivative* of c along F_0 denoted $F_0 \cdot c$ is simply the directional derivative of c at x along $F_0(x)$, given by $(F_0 \cdot c)(x) := c'(x) F_0(x)$. The *Lie bracket* between F_0 and F_1 is given by $[F_0, F_1] := F_0 \cdot F_1 - F_1 \cdot F_0$, with $(F_0 \cdot F_1)(x) := F_1'(x) F_0(x)$. Denoting p the adjoint variable and denoting $H_0(x, p) := \langle p, F_0(x) \rangle$ and $H_1(x, p) := \langle p, F_1(x) \rangle$ the Hamiltonian lifts of F_0 and F_1 , then the *Poisson bracket* of H_0 and H_1 is given by $\{H_0, H_1\} := \vec{H}_0 \cdot H_1$, where $\vec{H}_0 := (\frac{\partial H_0}{\partial x}, -\frac{\partial H_0}{\partial p})$ is the *Hamiltonian system* associated to H_0 . We also use the notation H_{01} (resp. F_{01}) to write the bracket $\{H_0, H_1\}$ (resp. $[F_0, F_1]$) and so forth. Besides, since H_0 and H_1 are two Hamiltonian lifts, we have $\{H_0, H_1\} = \langle p, [F_0, F_1] \rangle$.

Remark 1. We refer to Ref. [23, Chapter 2] for the exact and quite long expressions of F_{01} , F_{001} and F_{101} associated to eq. (5).

C. Necessary conditions of optimality

We recall the necessary conditions due to Refs. [25, 28] and follow the presentation of Ref. [10], which exhibits the role of the Lie and Poisson brackets.

1. Statement of the optimal control problem

We consider in this section the problem (P_α) but with only one state constraint, since, according to the numerical results from the next sections, the constraints c_1 and c_2 are active at the same time only at isolated times. In other words, we never have $c_1(x(t)) = 0$ and $c_2(x(t)) = 0$ along a time interval of non-empty interior. Due to these considerations, we present the necessary conditions of optimality considering only one scalar state constraint of the form $c(x(t)) \leq 0$, for every $t \in [0, t_f]$. The optimal control problem may be written in the abstract form:

$$\min \{g_\alpha(t_f, m(t_f, x_0, u)) \mid (t_f, u) \in D, b(x(t_f, x_0, u)) = 0 \text{ and } c(x(t, x_0, u)) \leq 0, \forall t \in [0, t_f]\},$$

where $x(\cdot, x_0, u)$ is the solution of eq. (5) with the initial condition $x(0, x_0, u) = x_0$.

2. Boundary arcs: definitions and assumptions

We call a *boundary arc*, labeled σ_c , an arc defined on an interval $I := [a, b]$ (not reduced to a singleton), such that $c(\sigma_c(t)) = 0$, for every $t \in I$. The times a and b are called the *entry-* and *exit-time* of the boundary arc; a and b are also termed *junction times*. An arc σ is said to have a *contact point* with the boundary at $\bar{t} \in [0, t_f]$ if $c(\sigma(\bar{t})) = 0$ and $c(\sigma(t)) < 0$ for $t \neq \bar{t}$ in a neighborhood of \bar{t} . A subarc σ on which $c(\sigma(t)) < 0$ is called an *interior arc*.

The *generic order* of the constraint c is the integer m such that $F_1 \cdot c = F_1 \cdot (F_0 \cdot c) = \dots = F_1 \cdot (F_0^{m-2} \cdot c) = 0$ and $F_1 \cdot (F_0^{m-1} \cdot c) \neq 0$. If the order of a boundary arc σ_c is m , then its associated feedback control can be generically computed by differentiating m times the mapping $t \mapsto c(\sigma_c(t))$ and solving with respect to u the linear equation:

$$c^{(m)} = F_0^m \cdot c + u F_1 \cdot (F_0^{m-1} \cdot c) = 0.$$

The boundary feedback control denoted u_c is given by

$$u_c := -\frac{F_0^m \cdot c}{F_1 \cdot (F_0^{m-1} \cdot c)}.$$

Let $t \mapsto \sigma_c(t)$, $t \in [t_1, t_2] \subset [0, t_f]$, be a boundary arc associated to $u_c(\cdot)$. We introduce the following assumptions:

- (A₁) $(F_1 \hat{A} \cdot (F_0^{m-1} \cdot c))(\sigma_c(t)) \neq 0$ for every $t \in [t_1, t_2]$, with m the order of the constraint.
- (A₂) $u_c(t) \in [u_{\min}, u_{\max}]$ for $t \in [t_1, t_2]$, i.e. the boundary control is admissible.
- (A₃) $u_c(t) \in (u_{\min}, u_{\max})$ for $t \in (t_1, t_2)$, i.e. u_c is not saturating on (t_1, t_2) .

Remark 2. These assumptions are numerically checked *a posteriori*.

3. Maximum principle with a single scalar state constraint

Define the pseudo-Hamiltonian:

$$\begin{aligned} H: R^n \times (R^n)^* \times \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (x, p, u, \eta) &\longmapsto H(x, p, u, \eta) := \langle p, F_0(x) + u F_1(x) \rangle + \eta c(x), \end{aligned}$$

where η is the *Lagrange multiplier* of the constraint c and $n := 3$ is the dimension of the state. Consider $(\bar{t}_f, \bar{u}(\cdot)) \in D$ an optimal solution with associated trajectory $\bar{x}(\cdot)$. Assume that the set of contact and junction times with the boundary, denoted \mathcal{T} , is finite. Assume also that the optimal control is piecewise smooth and that along each boundary arc, assumptions **A₁** and **A₂** are satisfied. Then we have the following necessary optimality conditions:

1. There exists a function $\bar{\eta}(\cdot) \leq 0$, a real number $p^0 \leq 0$ and a function of bounded variation $\bar{p}(\cdot) \in BV([0, \bar{t}_f], (\mathbb{R}^n)^*)$ such that:

$$\dot{\bar{x}}(t) = \frac{\partial H}{\partial p}(\bar{x}(t), \bar{p}(t), \bar{u}(t), \bar{\eta}(t)), \quad \dot{\bar{p}}(t) = -\frac{\partial H}{\partial x}(\bar{x}(t), \bar{p}(t), \bar{u}(t), \bar{\eta}(t)), \quad t \in [0, \bar{t}_f] \text{ a.e.}$$

2. The *maximization condition* holds for almost every $t \in [0, \bar{t}_f]$:

$$H(\bar{x}(t), \bar{p}(t), \bar{u}(t), \bar{\eta}(t)) = \max_{u \in U} H(\bar{x}(t), \bar{p}(t), u, \bar{\eta}(t)). \quad (8)$$

3. The *boundary conditions* $b(\bar{x}(\bar{t}_f)) = 0$ are satisfied and we have the *transversality conditions*:

$$\bar{p}_m(\bar{t}_f) = -p^0(1 - \alpha)$$

and since t_f is free, if $\bar{u}(\cdot)$ is continuous at time \bar{t}_f , then

$$H(\bar{x}(\bar{t}_f), \bar{p}(\bar{t}_f), \bar{u}(\bar{t}_f), \bar{\eta}(\bar{t}_f)) = -p^0\alpha.$$

4. The function $\bar{\eta}(\cdot)$ is continuous on the interior of the boundary arcs and we have the *complementarity condition*:

$$\bar{\eta}(t) c(\bar{x}(t)) = 0, \quad \forall t \in [0, \bar{t}_f].$$

5. For any $\tau \in \mathcal{T}$ we have

$$\begin{aligned} H[\tau^+] &= H[\tau^-], \quad \text{where } [\tau] \text{ stands for } (\bar{x}(\tau), \bar{p}(\tau), \bar{u}(\tau), \bar{\eta}(\tau)), \\ \bar{p}(\tau^+) &= \bar{p}(\tau^-) - \nu_\tau c'(\bar{x}(\tau)), \quad \text{where } \nu_\tau \leq 0 \text{ is called a } \textit{jump}. \end{aligned}$$

Remark 3. Either $p^0 = 0$ (abnormal case), or p^0 can be set to -1 by homogeneity (normal case). We restrict our study to the normal case.

Definition 1.

- We call an *extremal* a quadruple $(x(\cdot), p(\cdot), u(\cdot), \eta(\cdot))$ satisfying items 1, 2, 4 and 5. It is called a *BC-extremal* if it satisfies also item 3.
- Let $H_0(x, p) := \langle p, F_0(x) \rangle$ and $H_1(x, p) := \langle p, F_1(x) \rangle$ denote the Hamiltonian lifts of F_0 and F_1 . Then, along any extremal, we call $\Phi(t) := H_1(x(t), p(t))$ the *switching function*.
- It follows from (8) that along any extremal, we have $u(t) = u_{\min}$ if $\Phi(t) < 0$ and $u(t) = u_{\max}$ if $\Phi(t) > 0$. We say that a trajectory $x(\cdot)$ restricted to a subinterval $I \subset [0, t_f]$, not reduced to a singleton, is a *bang arc* if $u(\cdot)$ is constant on I , taking values in $\{u_{\min}, u_{\max}\}$. The trajectory is called *bang-bang* if it is the concatenation of a finite number of bang arcs.
- We say that a trajectory $x(\cdot)$ restricted to a subinterval $I \subset [0, t_f]$, not reduced to a singleton, is a *singular arc* if it is an interior arc and if the associated extremal lift satisfies $\Phi(t) = 0$, for every $t \in I$.
- We say that an extremal is a *bang*, *singular*, *boundary* or *interior* extremal if the associated trajectory is respectively a bang, singular, boundary or interior arc.

Definition 2. A bang arc such that $u(\cdot) \equiv u_{\min}$ (resp. u_{\max}) is labeled σ_- (resp. σ_+). A singular arc is labeled σ_s while a boundary arc associated to the constraint c is labeled σ_c . Besides, we denote by $\sigma_1\sigma_2$ an arc σ_1 followed by an arc σ_2 .

Remark 4. Along a boundary arc, the maximization condition (8) with assumption **A₃** imply $\Phi = 0$ on the interior of the boundary arc. Besides, the adjoint vector may be discontinuous at $\tau \in \mathcal{T}$.

4. Parameterization of the singular extremals

Relaxing the control bounds, singular trajectories are parameterized by the constrained Hamiltonian system:

$$\dot{x} = \partial_p H, \quad \dot{p} = -\partial_x H, \quad 0 = \partial_u H = H_1,$$

with $H_1(x, p) = \langle p, F_1(x) \rangle$ the Hamiltonian lift of F_1 . The constraint $H_1 = 0$ has to be differentiated at least twice along a singular extremal to compute the control. This gives:

$$H_1 = H_{01} = H_{001} + u H_{101} = 0,$$

along any singular extremal. A singular extremal along which $H_{101} \neq 0$ is called of *minimal order* and the corresponding singular control is given by:

$$u_s(z) := -\frac{H_{001}(z)}{H_{101}(z)},$$

with $z := (x, p)$. In this case, we have the following additional necessary condition of optimality deduced from the high-order maximum principle, see Ref. [26]. If the singular control is not saturating along the singular extremal, then the *generalized Legendre-Clebsch condition* must hold along the singular extremal, that is:

$$\frac{\partial}{\partial u} \frac{\partial^2}{\partial t^2} \frac{\partial H}{\partial u} = H_{101} \geq 0. \quad (9)$$

Besides, we have the following well-known result (see Ref. [7, Prop. 21]) that we use to define the numerical methods, see Sect. III B 2.

Proposition II.1. *Assume the open subset $\Omega := \{z \mid H_{101}(z) \neq 0\}$ is not empty and let us define on Ω the Hamiltonian $H_s(z) := H_0(z) + u_s(z) H_1(z)$. Then, the singular extremals with minimal order are the solutions of $\dot{z}(t) = \vec{H}_s(z(t))$, starting from the set $\{z \mid H_1(z) = H_{01}(z) = 0\}$.*

For three-dimensional systems, introducing

$$D_\xi(x) := \det(F_1(x), F_{01}(x), F_\xi(x)),$$

the singular control is given in feedback form by

$$u_s(x) = -\frac{D_{001}(x)}{D_{101}(x)},$$

whenever $D_{101}(x) \neq 0$ and $p \neq 0$, since, along a singular extremal

$$\langle p, F_1(x) \rangle = \langle p, F_{01}(x) \rangle = \langle p, F_{001}(x) + u F_{101}(x) \rangle = 0.$$

Prop. II.1 is replaced by:

Proposition II.2. *Assume the open subset $\Omega := \{z \mid H_{101}(z) \neq 0\}$ is not empty and let us define on Ω the Hamiltonian $H_s(z) := H_0(z) + u_s(x) H_1(z)$, $z = (x, p)$. Then, the singular extremals with minimal order are the solutions of $\dot{z}(t) = \vec{H}_s(z(t))$, starting from the set $\{z \mid H_1(z) = H_{01}(z) = 0\}$.*

Remark 5. Assuming $D_0(x) \neq 0$, then the generalized Legendre-Clebsch condition (9) becomes $D_0 D_{101} \geq 0$ (when $\alpha \neq 0$) and on the set $H_1 = H_{01} = 0$ we have $D_{101} = 0 \Rightarrow H_{101} = 0$ and $(H_{101} = 0 \text{ and } p \neq 0) \Rightarrow D_{101} = 0$.

5. Parameterization of the boundary extremals

We may find in Refs. [9, 28] the determination of the multiplier η and the jump ν_τ together with the analysis of the junction conditions, which is based on the concept of order and related to the classification of extremals. We give next some results only for $m = 1$ since the constraints c_1 and c_2 are of order 1, according to the following lemmas:

Lemma II.3. *Partial derivatives of the function ϕ is given by:*

$$\begin{aligned}\frac{\partial \phi}{\partial h}(x) &= \frac{1}{\phi(x)} \frac{\Theta_0 P(h)}{P_0 \Theta(h)} \left(\frac{P(h)}{P_0} \left(\left(\frac{\kappa v^2}{2R\Theta(h)} + 1 \right)^{1/\kappa} - 1 \right) + 1 \right)^{\kappa-1} \\ &\quad \left(-g_0 \left(\left(\frac{\kappa v^2}{2R\Theta(h)} + 1 \right)^{1/\kappa} - 1 \right) + \frac{v^2 \beta}{2\Theta(h)} \left(\frac{\kappa v^2}{2R\Theta(h)} + 1 \right)^{1/\kappa-1} \right), \\ \frac{\partial \phi}{\partial v}(x) &= \frac{v}{\phi(x)} \frac{\Theta_0 P(h)}{P_0 \Theta(h)} \left(\frac{P(h)}{P_0} \left(\left(\frac{\kappa v^2}{2R\Theta(h)} + 1 \right)^{1/\kappa} - 1 \right) + 1 \right)^{\kappa-1} \left(\frac{\kappa v^2}{2R\Theta(h)} + 1 \right)^{1/\kappa-1}.\end{aligned}$$

Lemma II.4. *Partial derivatives of the function ψ is given by:*

$$\frac{\partial \psi}{\partial h}(x) = \frac{\beta}{2\Theta(h)} \psi(x), \quad \frac{\partial \psi}{\partial v}(x) = \frac{\psi(x)}{v}.$$

Lemma II.5. *The two state constraints c_1 and c_2 are of order 1.*

Proof. The differentiation of $c_1(x(t)) = \phi(x(t)) - \phi_{\max}$ with respect to the time t leads to

$$\dot{c}_1(x(t)) = \frac{\partial \phi}{\partial h}(x(t)) \dot{h}(t) + \frac{\partial \phi}{\partial v}(x(t)) \dot{v}(t) \quad \text{and} \quad \frac{\partial \dot{c}_1}{\partial u}(x) = v \frac{\partial \phi}{\partial h}(x) - g_0 \frac{\partial \phi}{\partial v}(x).$$

We check numerically thanks to the software Sage, see Ref. [14], that the last quantity does not vanish for any values of h and v which can appear in the study and then, by definition, c_1 is a state constraint of order 1. The same reasoning holds for c_2 and gives the result. \square

For a first-order constraint, assuming \mathbf{A}_1 and \mathbf{A}_3 , we have the following result from Ref. [9].

Proposition II.6. *Let $m = 1$. Then:*

1. *along the boundary, the control and the multiplier are given by*

$$u_c(x) = -\frac{(F_0 \cdot c)(x)}{(F_1 \cdot c)(x)} \quad \text{and} \quad \eta_c(z) = \frac{H_{01}(z)}{(F_1 \cdot c)(x)}, \quad z = (x, p);$$

2. *if the control is discontinuous at a contact or a junction time τ between a bang arc and the boundary then the jump $\nu_\tau = 0$;*

3. *we have*

$$\nu_\tau = \frac{\Phi(\tau^-)}{(F_1 \cdot c)(x(\tau))} \text{ at an entry point and } \nu_\tau = -\frac{\Phi(\tau^+)}{(F_1 \cdot c)(x(\tau))} \text{ at an exit point.}$$

Remark 6. We refer to Ref. [23, Chapter 3] for the exact and quite long expressions of $F_0 \cdot c$, $F_1 \cdot c$ and H_{01} , which give u_c , η_c and ν_τ , for both constraints c_1 and c_2 .

Likewise the singular case, in the state constrained case, we have the following result excerpt from [17, Prop. 4.5] which is useful to define the numerical shooting method, see Sect. III B 2.

Proposition II.7. *Let $c(x) \leq 0$ be a smooth scalar state constraint of order 1 and assume the open subset $\Omega := \{x \mid (F_1 \cdot c)(x) \neq 0\}$ is not empty. On Ω we define the Hamiltonian*

$$H_c(z) := H_0(z) + u_c(x) H_1(z) + \eta_c(z) c(x), \quad z = (x, p),$$

where u_c and η_c are defined in Prop. II.6. Then, the boundary extremals of order 1 contained in $H_1 = 0$ are the solutions of $\dot{z}(t) = \vec{H}_c(z(t))$, starting from $\{z = (x, p) \mid c(x) = H_1(z) = 0\}$.

III. Analysis of the minimum time-to-climb problem with state constraints

The minimum time-to-climb problem with no state constraints is analyzed in Refs. [18, 20]. In Ref. [18], the influence of the initial mass m_0 and the final true air speed v_f is studied and it is shown in particular that the trajectories are of the form $\sigma_{\pm}\sigma_s\sigma_{\pm}$, where σ_s is an hyperbolic singular extremal. We thus restrict the small time analysis, presented in Sect. III A 2, to hyperbolic singular extremals, that we introduce in Sect. III A 1. A first numerical result is then presented in Sect. III B while the influence of ϕ_{\max} and ψ_{\max} on the state constraints is studied in Sect. III C.

A. Small time synthesis for planar hyperbolic trajectories with an order one state constraint

We first recall some results from Refs. [7, 27] in Sect. III A 1. This leads to introduce hyperbolic singular extremals that we analyze in Sect. III A 2.

1. Generic classification of bang-bang extremals

Along a singular extremal we have $\partial_u H \equiv 0$. In our particular case of a single-input affine control system this condition becomes $\Phi(t) = H_1(z(t)) = 0$. Let us define the switching surface

$$\Sigma := \{z := (x, p) \in \mathbb{R}^n \times (\mathbb{R}^n)^* \mid H_1(z) = 0\},$$

and the set

$$\Sigma_s := \{z := (x, p) \in \mathbb{R}^n \times (\mathbb{R}^n)^* \mid H_1(z) = H_{01}(z) = 0\}$$

containing all the singular extremals. Let us introduce the notation Φ_+ (resp. Φ_-) if the control along a bang extremal is u_{\max} (resp. u_{\min}). The first and second derivatives of Φ_{\pm} are given by:

$$\begin{aligned}\dot{\Phi}_{\pm}(t) &= H_{01}(z(t)), \\ \ddot{\Phi}_{+}(t) &= H_{001}(z(t)) + u_{\max} H_{101}(z(t)), \\ \ddot{\Phi}_{-}(t) &= H_{001}(z(t)) + u_{\min} H_{101}(z(t)).\end{aligned}$$

A crucial point to analyze the minimum time-to-climb problem, is then to apply the results from Ref. [27] (see also Ref. [7]) to classify the extremal curves near the switching surface Σ . We have the following:

1. **Ordinary switching time.** It is a time t such that two bang arcs switch with $\Phi(t) = 0$ and $\dot{\Phi}(t) = H_{01}(z(t)) \neq 0$. According to the maximum principle, near Σ , the extremal is of the form $\sigma_- \sigma_+$ if $\dot{\Phi}(t) > 0$ and $\sigma_+ \sigma_-$ if $\dot{\Phi}(t) < 0$.
2. **Fold point.** It is a time where a bang arc has a contact of order 2 with Σ . We have three cases (if $\ddot{\Phi}_{\pm} \neq 0$) depending on $\ddot{\Phi}_{\pm}$ at the switching time:
 - *Hyperbolic case:* $\ddot{\Phi}_{+} > 0$ and $\ddot{\Phi}_{-} < 0$. A connection with a singular extremal is possible at Σ_s and locally each extremal is of the form $\sigma_{\pm}\sigma_s\sigma_{\pm}$ (by convention each arc of the sequence can be empty).
 - *Parabolic case:* $\ddot{\Phi}_{+}\ddot{\Phi}_{-} > 0$. The singular extremal at the switching point is not admissible and every extremal curve is locally bang-bang with at most two switchings, *i.e.* $\sigma_+\sigma_-\sigma_+$ or $\sigma_-\sigma_+\sigma_-$.
 - *Elliptic case:* $\ddot{\Phi}_{+} < 0$ and $\ddot{\Phi}_{-} > 0$. A connection with a singular arc is not possible and locally each extremal is bang-bang but with no uniform bound on the number of switchings.

When dealing with three-dimensional systems we can give conditions depending only on the state x to classify the three cases around fold points. Assuming $D_0(x) \neq 0$, then the family $(F_0(x), F_1(x), F_{01}(x))$ forms a basis of \mathbb{R}^3 and there exists $(\alpha_0, \alpha_1, \alpha_{01}) \in \mathbb{R}^3$ and $(\beta_0, \beta_1, \beta_{01}) \in \mathbb{R}^3$ such that

$$\begin{aligned}F_{001}(x) + u_{\min} F_{101}(x) &= \alpha_0 F_0(x) + \alpha_1 F_1(x) + \alpha_{01} F_{01}(x), \\ F_{001}(x) + u_{\max} F_{101}(x) &= \beta_0 F_0(x) + \beta_1 F_1(x) + \beta_{01} F_{01}(x).\end{aligned}$$

By linearity of the determinant, we have

$$\begin{aligned} D_{001}(x) + u_{\min} D_{101}(x) &= \alpha_0 D_0(x), \\ D_{001}(x) + u_{\max} D_{101}(x) &= \beta_0 D_0(x), \end{aligned}$$

and since $D_0(x) \neq 0$ then we can compute α_0 and β_0 along the trajectory. Besides, along any singular extremal, we have:

$$\begin{aligned} \ddot{\Phi}_-(t) &= \langle p(t), F_{001}(x(t)) + u_{\min} F_{101}(x(t)) \rangle = \alpha_0(t) H_0(z(t)), \\ \ddot{\Phi}_+(t) &= \langle p(t), F_{001}(x(t)) + u_{\max} F_{101}(x(t)) \rangle = \beta_0(t) H_0(z(t)), \end{aligned}$$

with $H_0(z(t)) > 0$ in the normal case with the convention of the maximum principle. Denoting \bar{t} the time when the extremal has a contact of order two with the switching surface Σ , then the singular extremal is hyperbolic if $\alpha_0(\bar{t}) < 0$ and $\beta_0(\bar{t}) > 0$, is elliptic if $\alpha_0(\bar{t}) > 0$ and $\beta_0(\bar{t}) < 0$, and parabolic if $\alpha_0(\bar{t}) \beta_0(\bar{t}) > 0$.

2. Small time analysis

The time-optimal trajectories are concatenations of bang, singular and boundary arcs. The main difficulty is then to determine the number and the sequence of these arcs. We can see these concatenations as a sequence of patterns, each pattern being made with only few concatenations. These patterns depend on the classification presented in Sect. III A 1 but also on the order of the state constraint. Since we are dealing with order one state constraints and hyperbolic singular extremals, the number of possible patterns is drastically reduced. Besides, a (long) time-optimal trajectory may be seen as a concatenation of small time-optimal trajectories since it is necessarily optimal on every sub-interval of time. Finally, the pattern we are looking for are thus given by small time-optimal trajectories that we analyze in this section.

Remark 7. In Ref. [18], it is shown that the state-unconstrained time-optimal trajectories are of the form $\sigma_{\pm} \sigma_s \sigma_{\pm}$, *i.e.* the optimal sequences are made with only one single pattern.

The aircraft dynamics is a multi-scale system and since the time constant of the mass is a hundred, resp. a thousand, times smaller than the time constant of the altitude, resp. the speed, we can neglect the evolution of the mass when considering short time intervals and thus, in this case, we can reduce our three dimensional dynamics into a planar dynamics. We present hereinafter some results from Refs. [9, 10] about small time analysis considering only a planar single-input control system, in the hyperbolic case and with an order one scalar state constraint $c(q(t)) \leq 0$, of the form:

$$\dot{q}(t) = F_0(q(t)) + u(t) F_1(q(t)),$$

where $|u(t)| \leq 1$ and $q(t) := (x(t), y(t)) \in \mathbb{R}^2$. Let us take q_0 such that $c(q_0) = 0$ and let us identify q_0 to the origin. Let us assume that $F_0(q_0)$ and $F_1(q_0)$ are linearly independent, and that the constraint c is of order one, *i.e.* $(F_1 \cdot c)(q_0) \neq 0$. Then replacing, if necessary, u by $-u$, we can find a local diffeomorphism preserving $q_0 = 0$ and transforming the constrained system into the following generic model:

$$\dot{x}(t) = 1 + y(t) a(q(t)), \quad \dot{y}(t) = b(q(t)) + u(t), \quad y(t) \leq 0.$$

To describe the hyperbolic case, we consider also that $\det(F_1(q_0), F_{01}(q_0)) = a(q_0) = 0$ and we assume that the set

$$S := \{q \in \mathbb{R}^2 \mid \det(F_1(q), F_{01}(q)) = 0\}$$

containing the singular trajectories is a simple curve that we approximate by a straight line in our small time model. The equations of the system become:

$$\dot{x}(t) = 1 + y(t) (a y(t) + b x(t)), \quad \dot{y}(t) = c + u(t), \quad y(t) \leq 0,$$

where S is identified to $\{q \in \mathbb{R}^2 \mid 2a y + b x = 0\}$ and we assume that $a \neq 0$ and $b \neq 0$. Note that the generalized Legendre-Clebsch condition is related to the sign of a . The singular control at q_0 is

given by $u_s(q_0) = -c - b/2a$ and the boundary control is simply $u_c(q) = -c$. Taking the constraint on the control into account, we see that the condition of admissibility implies $|c + b/2a| \leq 1$ and $|c| \leq 1$. Finally, the *hyperbolic case* corresponds to:

$$a < 0, \quad |c + b/2a| < 1, \quad |c| < 1, \quad \text{and} \quad b \neq 0,$$

that is the strict generalized Legendre-Clebsch condition is satisfied, the singular and boundary controls are strictly admissible, and the singular set S is not identified with the boundary set $y = 0$.

For the state unconstrained problem, the singular arc is optimal, each optimal trajectory has at most two switchings and the local optimal synthesis is of the form $\sigma_{\pm}\sigma_s\sigma_{\pm}$. For the state constrained problem, the so-called clock form and Stokes theorem (see Refs. [9, 10] for details) can be used to conclude on the optimality of the boundary arc when $x \geq 0$ and its nonoptimality when $x < 0$, for the case $b > 0$. For $b < 0$, the boundary arc is optimal when $x \leq 0$ and nonoptimal when $x > 0$. See Fig. 1 for a representation of the bang, singular and boundary arcs in the hyperbolic case with $b > 0$, and for a comparison between boundary and bang-bang arcs.

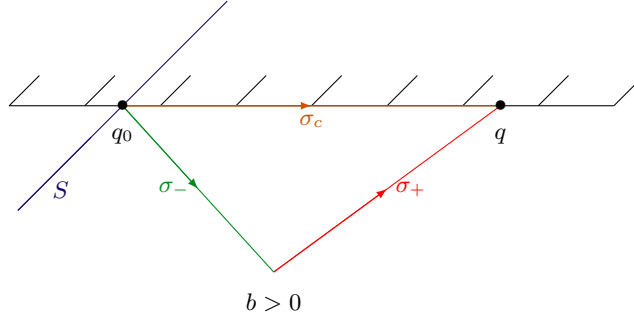


Fig. 1 Small time representation of different trajectories in the neighborhood of a point q_0 belonging to the singular set S and such that $c(q_0) = 0$, in the case $b > 0$. Two trajectories joining q_0 and q are represented: one boundary arc σ_c and a bang-bang trajectory $\sigma_-\sigma_+$. Using the clock form and the Stokes theorem, we can show that the trajectory σ_c is time-optimal.

When $b > 0$, the small time optimal synthesis joining two points q_1 and q_2 on the boundary $y = 0$, with q_1 and q_2 on each side of S and q_1 on the left, is then of the form $\sigma_-\sigma_s\sigma_c$. In this case, each optimal curve in a neighborhood of q_0 has at most three switchings and the local optimal synthesis is of the form $\sigma_{\pm}\sigma_s\sigma_c\sigma_{\pm}$. For $b < 0$, the local optimal synthesis is of the form $\sigma_{\pm}\sigma_c\sigma_s\sigma_{\pm}$. We summarize this analysis in the following proposition excerpted from Ref. [9].

Proposition III.1. *Under our assumptions, in the hyperbolic case each small time optimal trajectory has at most three switchings. Moreover,*

1. For $b < 0$, a boundary arc is optimal iff $x \leq 0$ and each optimal arc has the form $\sigma_{\pm}\sigma_c\sigma_s\sigma_{\pm}$.
2. For $b > 0$, a boundary arc is optimal iff $x \geq 0$ and each optimal arc has the form $\sigma_{\pm}\sigma_s\sigma_c\sigma_{\pm}$.

Remark 8. The Prop. III.1 gives two different patterns $\sigma_{\pm}\sigma_s\sigma_c\sigma_{\pm}$ and $\sigma_{\pm}\sigma_c\sigma_s\sigma_{\pm}$, depending on the local model, for small time optimal trajectories around the boundary. These two patterns generalize the pattern $\sigma_{\pm}\sigma_s\sigma_{\pm}$, presented in Sect. III A 1, in the hyperbolic state unconstrained case.

B. Application to the minimum time-to-climb problem and multiple shooting

1. Small time analysis and the CAS/Mach state constraints

In this section, we explain on an example why the CAS/Mach procedure makes sense. Let us take from Ref. [18, Fig. 1] the state unconstrained trajectory of the form $\sigma_-\sigma_s\sigma_+$, for the minimum time-to-climb problem, with data given in Sect. I and II A. Note that in Ref. [18], $u_{\min} = -0.262$ instead of 0 but this has a small influence and it does not change the structure $\sigma_-\sigma_s\sigma_+$. From this excerpted trajectory, denoted $x(\cdot)$, we compute and display on Fig. 2, the state constraints $c_1(x(t)) = \phi(x(t)) - \phi_{\max}$ and $c_2(x(t)) = \psi(x(t)) - \psi_{\max}$, for $\phi_{\max} := 160$ and $\psi_{\max} = 0.7$. These

realistic values of ϕ_{\max} and ψ_{\max} are chosen to emphasize the following comment: the boundary $c_1 = 0$ is reached during the first bang arc, before the singular arc, while the boundary $c_2 = 0$ is reached during the singular arc, after reaching $c_1 = 0$ and before the last bang arc.

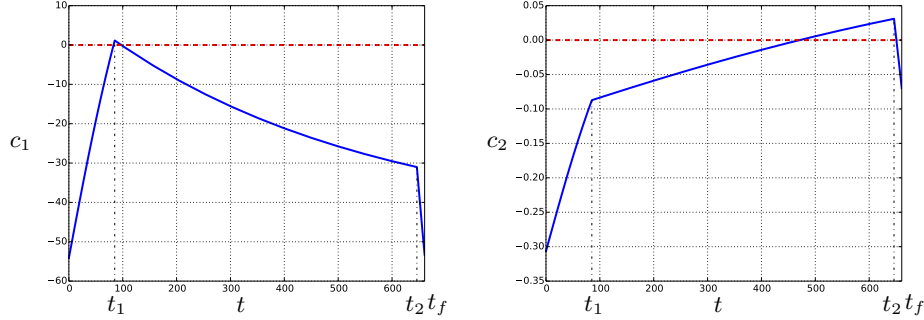


Fig. 2 The plain lines depict the evolution of the state constraints c_1 and c_2 along the state unconstrained hyperbolic trajectory of the form $\sigma_- \sigma_s \sigma_+$, with $\phi_{\max} = 160$ and $\psi_{\max} = 0.7$. The time t_1 (resp. t_2) represents the switching time between the negative bang (resp. the singular) arc and the singular (resp. the positive bang) arc.

From Fig. 2, one can also notice that $\phi(x(\cdot))$ and $\psi(x(\cdot))$ reach their maxima respectively at the switching times t_1 and t_2 . Let us consider now other values for ϕ_{\max} and ψ_{\max} . Assume we fix $\phi_{\max} = \phi(x(t_1)) - \varepsilon$, $\varepsilon > 0$ small, and $\psi_{\max} > \psi(x(t_2))$. Considering these values, the state unconstrained trajectory is not admissible anymore for the state constrained case since in a small neighborhood of t_1 , we have $c_1(x(t)) > 0$. The trajectory is thus slightly modified and the updated optimal sequence becomes $\sigma_- \sigma_{c_1} \sigma_s \sigma_+$ corresponding to the first case of Prop. III.1. On the contrary, if $\phi_{\max} > \phi(x(t_1))$ and $\psi_{\max} = \psi(x(t_2)) - \varepsilon$, $\varepsilon > 0$ small, then the updated optimal trajectory is of the form $\sigma_- \sigma_s \sigma_{c_2} \sigma_+$ corresponding to the second case of Prop. III.1. Finally, if we set $\phi_{\max} = \phi(x(t_1)) - \varepsilon_1$ and $\psi_{\max} = \psi(x(t_2)) - \varepsilon_2$, both $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ sufficiently small, then the state constrained trajectory is of the form $\sigma_- \sigma_{c_1} \sigma_s \sigma_{c_2} \sigma_+$, that is the boundary arc σ_{c_1} appears before the boundary arc σ_{c_2} and this sequence is made of two patterns. In other words, the trajectory contains one arc at constant CAS which is before the arc at constant Mach, and this is consistent with the actual CAS/Mach procedure. Let us recall that this procedure splits the climbing phase in two parts, the first one is an arc at constant CAS and the second one is an arc at constant Mach number. So, comparing to the CAS/Mach procedure there is an additional singular arc joining the boundary arcs, but, according to the results presented in Sect. III C, for smaller values of ϕ_{\max} and ψ_{\max} , the singular arc vanishes and we get the CAS/Mach procedure.

Remark 9. In Sect. III C, we analyze the influence of ϕ_{\max} and ψ_{\max} on the structure of the trajectories by deforming the BC-extremal associated to the state unconstrained trajectory, taken here as example. Just note that for all the numerical case studies presented in this article, whatever the values of ϕ_{\max} and ψ_{\max} , the boundary arc σ_{c_1} is always before σ_{c_2} .

2. A preliminary numerical result by indirect multiple shooting

We present in this section the *indirect multiple shooting method* [11] that we use to solve the necessary conditions of optimality given by the maximum principle presented in Sect. II C. We describe the method on only one example and we refer to Ref. [23] for more details. Let us consider the same example as the one presented in the previous section and let us fix $\phi_{\max} > \phi(x(t_1))$ and $\psi_{\max} = 0.7$. In this case, the resulting trajectory is of the form $\sigma_- \sigma_s \sigma_{c_2} \sigma_+$.

The unknowns of the shooting method are the junction times t_1 , t_2 and t_3 , the final time t_f , the jumps ν_{t_2} and ν_{t_3} at the junction times t_2 and t_3 and the initial costate denoted p_0 . With p_0 , t_1 , t_2 , t_3 , ν_{t_2} and ν_{t_3} , since we know the structure of the trajectory, we can retrieve the states-costates at the junction times t_1 , t_2 and t_3 simply by integration, starting from $x(0) = x_0$ and applying the controls u_{\min} , u_s , u_{c_2} and u_{\max} , respectively on $[0, t_1]$, $[t_1, t_2]$, $[t_2, t_3]$ and $[t_3, t_f]$. However, to improve numerical stability, we add the states-costates z_1 , z_2 and z_3 at the junction times to the

unknowns of the shooting method. The first thing to notice now is that the jumps are zero. Indeed, $\Phi(t_2^-) = 0$ and so $\nu_{t_2} = 0$ (from Prop. II.6) since the switching function $\Phi(\cdot) = 0$ along the singular arc σ_s . Now, combining Prop. II.6 and item 5 of the maximum principle, we have:

$$\nu_{t_3} = -\frac{\Phi(t_3^+)}{(F_1 \cdot c_2)(x(t_3))} \leq 0.$$

But, $\Phi(t_3^+) \geq 0$ since the last bang arc is positive. So, if we have $(F_1 \cdot c_2)(x(t_3)) < 0$, then necessarily $\nu_{t_3} = 0$. Numerical experiments from Ref. [23] suggest that $F_1 \cdot c_2 < 0$ all along the trajectory and so in particular $(F_1 \cdot c_2)(x(t_3)) < 0$. Finally, grouping all together, we write

$$y := (p_0, t_1, t_2, t_3, t_f, z_1, z_2, z_3) \in \mathbb{R}^{25}$$

the unknown of the shooting method.

Remark 10. For the jump ν_{t_3} to be not zero, it is necessary that $u_{c_2}(x(t_3)) = u_{\max}$, which may happen in only very particular cases that we do not encounter in the numerical experiments all through this article.

Now, we need to describe the shooting equations and define the shooting function. First of all, we define the following Hamiltonians:

$$\begin{aligned} H_+(z) &:= H_0(z) + u_{\max} H_1(z), \\ H_-(z) &:= H_0(z) + u_{\min} H_1(z), \\ H_s(z) &:= H_0(z) + u_s(x) H_1(z), \\ H_{c_2}(z) &:= H_0(z) + u_{c_2}(x) H_1(z) + \eta_{c_2}(z) c_2(x), \end{aligned}$$

where $z = (x, p)$ and where u_s , u_{c_2} and η_{c_2} are given in Sect. II C. Note that we can replace $u_s(x) = -D_{001}(x)/D_{101}(x)$ by $u_s(z) = -H_{001}(z)/H_{101}(z)$. Then, we define the *exponential mapping* $e^{t\varphi}(q_0)$ as the solution at time t of the Cauchy problem $\dot{q}(s) = \varphi(q(s))$, $q(0) = q_0$, where φ is any dynamical system and where the state here is denoted q . We define also the canonical projections $\pi_x(z) = x$ and $\pi_{p_m}(z) = p_m$, recalling that $z = (x, p)$, $x = (h, v, m)$ and writting $p := (p_h, p_v, p_m)$. The *multiple shooting function*, denoted S_{c_2} , associated to the structure $\sigma_- \sigma_s \sigma_{c_2} \sigma_+$ is then defined by:

$$S_{c_2}(y) := \begin{pmatrix} H_1(e^{t_1 \vec{H}^-}(x_0, p_0)) \\ H_{01}(e^{t_1 \vec{H}^-}(x_0, p_0)) \\ c_2(\pi_x(e^{(t_2-t_1) \vec{H}_s}(z_1))) \\ \Psi(e^{(t_f-t_3) \vec{H}_+}(z_3)) \\ z_1 - e^{t_1 \vec{H}^-}(x_0, p_0) \\ z_2 - e^{(t_2-t_1) \vec{H}_s}(z_1) \\ z_3 - e^{(t_3-t_2) \vec{H}_{c_2}}(z_2) \end{pmatrix} \in \mathbb{R}^{25},$$

where

$$\Psi(z) := (b(\pi_x(z)), \pi_{p_m}(z), H_+(z) + p^0), \quad p^0 = -1$$

and the *multiple shooting method* consists in finding a zero of the multiple shooting function S_{c_2} , i.e. in solving $S_{c_2}(y) = 0$.

Remark 11. We can replace the first three equations by $H_1(z_1) = H_{01}(z_1) = c_2(\pi_x(z_2)) = 0$.

Let $\bar{y} := (\bar{p}_0, \bar{t}_1, \bar{t}_2, \bar{t}_3, \bar{t}_f, \bar{z}_1, \bar{z}_2, \bar{z}_3)$ be a zero of the shooting function S_{c_2} . Then, to \bar{y} is associated a unique BC-extremal $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u}(\cdot), \bar{\eta}(\cdot))$ if $0 \leq \bar{t}_1 \leq \bar{t}_2 \leq \bar{t}_3 \leq \bar{t}_f$ and if $H_1(\bar{z}(t)) < 0$ for almost every $t \in [0, \bar{t}_1]$ and $H_1(\bar{z}(t)) > 0$ for almost every $t \in [\bar{t}_3, \bar{t}_f]$. These two conditions are not included in the shooting equations but they are easily verified *a posteriori* and should be ok if the structure is the right one. Let us explain now why these shooting equations define a BC-extremal of the form $\sigma_- \sigma_s \sigma_{c_2} \sigma_+$. The first two equations in $S_{c_2}(y) = 0$ impose the extremal to be singular

on $[\bar{t}_1, \bar{t}_2]$ according to Prop. II.2, the third equation impose a boundary arc on $[\bar{t}_2, \bar{t}_3]$ according to Prop. II.7, noticing that $H_1(e^{(\bar{t}_2 - \bar{t}_1)H_s}(\bar{z}_1)) = 0$, while the fourth equation $\Psi = 0$ contains the limit and transversality conditions. The last equations are the so-called *matching conditions*.

We use the **HamPath** software, see Refs. [12, 17], to obtain a (numerical) zero, denoted \bar{y} , of the shooting function with high accuracy: $\|S_{c_2}(\bar{y})\| \approx 1.511 \times 10^{-11}$. The switching times are $\bar{t}_1 \approx 88.61\text{s}$, $\bar{t}_2 \approx 455.7\text{s}$ and $\bar{t}_3 \approx 651.46\text{s}$. The final time is $\bar{t}_f \approx 661.37\text{s}$ and the initial adjoint vector is $\bar{p}_0 \approx (3.180 \times 10^{-2}, 6.594 \times 10^{-1}, -2.300 \times 10^{-1})$. Note that the **HamPath** software is based on a Newton-like algorithm which is initialized thanks to the state unconstrained trajectory. Figure 3 depicts the control and the constraint c_2 along the resulting state constrained trajectory.

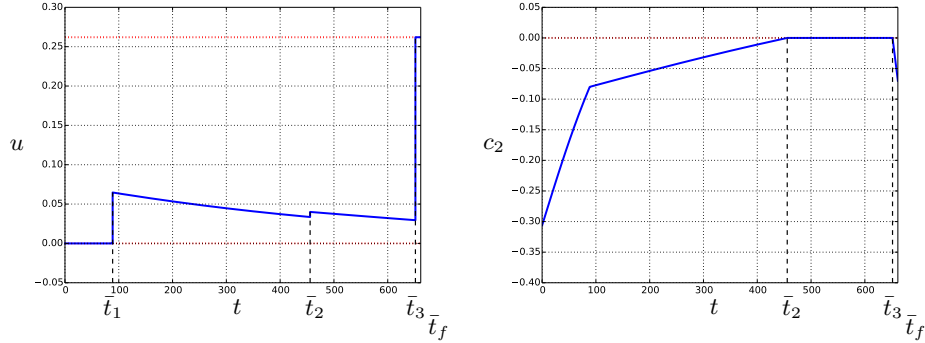


Fig. 3 Evolution of the control and the constraint c_2 along the trajectory associated to the zero \bar{y} of the shooting function S_{c_2} .

Remark 12. Since all the controls depend only on the state x , we can replace the resolution of the shooting equations by the resolution of the following optimization problem:

$$\min_{X \in \mathbb{R}^{13}} f(X), \quad \text{subjected to } g(X) = 0_{\mathbb{R}^{11}},$$

where $X := (t_1, t_2, t_3, t_f, x_1, x_2, x_3)$, $f(X) := t_f$ and

$$g(X) := \begin{pmatrix} b(e^{(t_f - t_3)F_+}(x_3)) \\ x_1 - e^{t_1 F_-}(x_0) \\ x_2 - e^{(t_2 - t_1)F_s}(x_1) \\ x_3 - e^{(t_3 - t_2)F_{c_2}}(x_2) \end{pmatrix} \in \mathbb{R}^{11},$$

where the vector fields are defined by $F_+ := F_0 + u_{\max} F_1$, $F_- := F_0 + u_{\min} F_1$, $F_s := F_0 + u_s F_1$ and $F_{c_2} := F_0 + u_{c_2} F_1$. We have in $g(X) = 0$ only the limit and matching (to improve numerical stability) conditions.

C. Influence of ϕ_{\max} and ψ_{\max} on the BC-extremals

1. Introduction to differential path following methods

The shooting method is used to solve a single optimal control problem. To solve a one-parameter family of optimal control problems, *e.g.* for different values of ψ_{\max} , we use *differential path following* techniques with arclength parameterization (or homotopy), see Refs. [1, 22]. Let $h: \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^N$, $h(y, \lambda)$, denote an *homotopic function* with λ the *homotopic parameter*. For example, one can consider the homotopic function defined by S_{c_2} , with $y := (p_0, t_1, t_2, t_3, t_f, z_1, z_2, z_3)$, $N = 25$ and $\lambda := \psi_{\max}$. That is to say, in this example $h(y, \lambda) = S_{c_2}(y)$ with $\psi_{\max} = \lambda$. We are interested in solving $h = 0$ for $\lambda \in [\lambda_0, \lambda_1]$, $(\lambda_0, \lambda_1) \in \mathbb{R}^2$ given. The classical difficulties about homotopic methods consist in assuring that a curve in $h^{-1}(0)$ exists, is sufficiently smooth and will intersect the fixed target λ_1 in finite length starting from $\lambda = \lambda_0$. As a first result we have the following. Suppose h is continuously differentiable and that we know y_0 such that $h(y_0, \lambda_0) = 0$ and assume

$$\text{rank} \left(\frac{\partial h}{\partial y}(y_0, \lambda_0) \right) = N.$$

Suppose also that 0 is a regular value of h . Then, a continuously differentiable curve crossing (y_0, λ_0) and transverse to $\lambda = \lambda_0$ exists and is either diffeomorphic to a circle or the real line. The curves in $h^{-1}(0)$ are disjoint, and we call each branch of $h^{-1}(0)$ a path of zeros. To compute numerically these paths of zeros, we use the homotopic method from the **HamPath** software, which is based on a Predictor-Corrector algorithm with a high order and variable step size Runge-Kutta scheme for the prediction and with a classical simplified Newton method for the correction.

2. Homotopy and monitoring for the minimum time-to-climb problem

We present in this section how we combine homotopy and *monitoring* (that we define hereinafter) to obtain, for a fixed climbing scenario, a cartography of possible time-optimal structures with respect to the values of the bounds of the state constraints c_1 and c_2 , *i.e.* on the values of ϕ_{\max} and ψ_{\max} . We do not give all the details and we refer to Ref. [17] for a more detailed description of the methodology to obtain such a cartography. We start from the state unconstrained trajectory, denoted $x(\cdot)$, presented in Sect. IIIB1 of the form $\sigma_- \sigma_s \sigma_+$. The maximal values of ϕ and ψ along the trajectory are reached respectively at the switching times t_1 and t_2 . One can notice that $\phi(x(t_1)) < VMO$ and $\psi(x(t_2)) < MMO$, so for this trajectory, the state constraints c_1 and c_2 are not violated if we fix $\phi_{\max} = VMO$ and $\psi_{\max} = MMO$. The idea is then to deform by homotopy on ϕ_{\max} or ψ_{\max} the BC-extremal associated to this trajectory and then to detect by monitoring when a change in the structure occurs. The monitoring consists in checking some conditions after each step of prediction and correction. Technically, we only have to code the conditions and then the **HamPath** software automatically check them after each prediction-correction step and stop the homotopy process if at least one condition is violated. Here are the three different monitorings we do:

M1 check if c_1 is violated at the entry point of the singular arc. If yes, then we add a σ_{c_1} arc.

M2 check if c_2 is violated at the exit point of the singular arc. If yes, then we add a σ_{c_2} arc.

M3 check if the singular arc has a positive length. If not, then the arc is removed.

When a change is detected, we update the homotopic function accordingly to the new structure and repeat the process. We thus have to limit the range of values for ϕ_{\max} and ψ_{\max} . We choose to build the cartography for $(\phi_{\max}, \psi_{\max}) \in [\phi_0, VMO] \times [\psi_0, MMO]$, where $VMO \approx 180$, $MMO \approx 0.82$, $\phi_0 := \max(\phi(x(0)), \phi(x(t_f))) \approx 107$ and $\psi_0 := \max(\psi(x(0)), \psi(x(t_f))) \approx 0.63$, with t_f the final time. With these chosen minimal values ψ_0 and ϕ_0 , whatever ϕ_{\max} and ψ_{\max} , the trajectory will always start with an arc σ_- and end with an arc σ_+ because the state constraints depend only on h and v which are fixed at the initial and final times to some given values which do not depend on ϕ_{\max} and ψ_{\max} . This explains why we say that we study the deformation of the BC-extremal for a fixed climbing scenario.

3. Numerical results and classification

Let simply recall that we start from the BC-extremal associated to the state unconstrained trajectory of the form $\sigma_- \sigma_s \sigma_+$, denoted $x(\cdot)$, with t_1 and t_2 the switching times and t_f the final time.

1. Let $\phi_{c_1} := \phi(x(t_1))$, $\psi_{c_2} := \psi(x(t_2))$ and let us label δ_{c_1} , δ_{c_2} and δ_s , a touch point respectively with the state constraint $c_1 = 0$, $c_2 = 0$ and the singular set Σ_s . Then, we have the following straightforward first result:

ϕ_{\max}	ψ_{\max}	structure
ϕ_{c_1}	ψ_{c_2}	$\sigma_- \delta_{c_1} \sigma_s \delta_{c_2} \sigma_+$
ϕ_{c_1}	$(\psi_{c_2}, MMO]$	$\sigma_- \delta_{c_1} \sigma_s \sigma_+$
$(\phi_{c_1}, VMO]$	ψ_{c_2}	$\sigma_- \sigma_s \delta_{c_2} \sigma_+$
$(\phi_{c_1}, VMO]$	$(\psi_{c_2}, MMO]$	$\sigma_- \sigma_s \sigma_+$

2. Let fix first $\psi_{\max} = MMO$. For $\phi_{\max} = \phi_{c_1} - \varepsilon$, $\varepsilon > 0$ small enough, the BC-extremal has a structure of the form $\sigma_- \sigma_{c_1} \sigma_s \sigma_+$. Then, we perform an homotopy on $\lambda = \phi_{\max}$ from $\lambda_0 = \phi_{c_1}$ to $\lambda_1 = \phi_0$ which stops because of monitoring **M3** around $\lambda \approx 129.8 =: \phi_s$. Hence, at $\lambda = \phi_s$ the structure is $\sigma_- \sigma_{c_1} \delta_s \sigma_+$. The second homotopy from $\lambda_0 = \phi_s$ to $\lambda_1 = \phi_0$ with a structure $\sigma_- \sigma_{c_1} \sigma_+$ is not stopped by the monitoring, thus, no change in the structure is detected anymore. Let fix now $(\phi_{\max}, \psi_{\max}) = (\phi_s, MMO)$. The structure is of the form $\sigma_- \sigma_{c_1} \delta_s \sigma_+$. Then, we define $\psi_s := \max \psi(\cdot) \approx 0.734$ along the corresponding trajectory, and we have the following:

ϕ_{\max}	ψ_{\max}	structure
$[\phi_0, \phi_s)$	MMO	$\sigma_- \sigma_{c_1} \sigma_+$
ϕ_s	MMO	$\sigma_- \sigma_{c_1} \delta_s \sigma_+$
(ϕ_s, ϕ_{c_1})	MMO	$\sigma_- \sigma_{c_1} \sigma_s \sigma_+$
ϕ_s	ψ_s	$\sigma_- \sigma_{c_1} \delta_s \delta_{c_2} \sigma_+$
ϕ_s	$(\psi_s, MMO]$	$\sigma_- \sigma_{c_1} \delta_s \sigma_+$

Fig. 4 depicts the previous results.

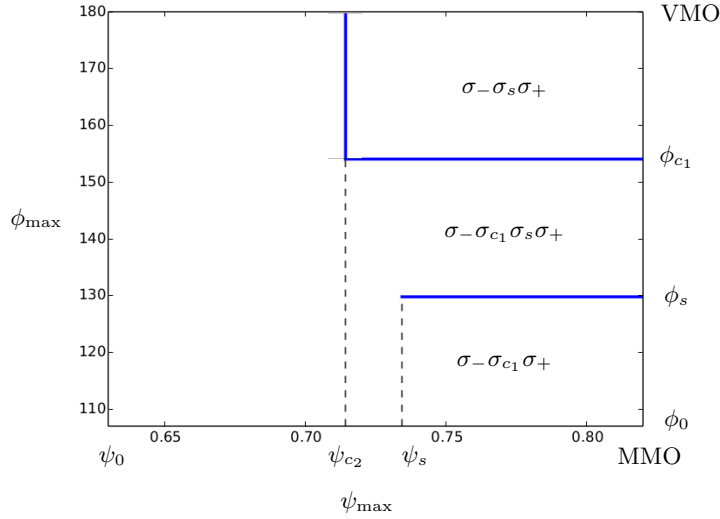


Fig. 4 Sketch of the classification of BC-extremals for the minimum time-to-climb problem.

3. The same methodology is used each times the homotopy is stopped by a monitoring. The new structure is then guessed thanks to the results from section III A. Fig. 5 presents the resulting cartography of the BC-extremal structures for $(\phi_{\max}, \psi_{\max}) \in [\phi_0, VMO] \times [\psi_0, MMO]$, for the minimum time-to-climb problem. Fig. 6 gives the time-to-climb related to the trajectories from Fig. 5. Of course, the state unconstrained trajectory of the form $\sigma_- \sigma_s \sigma_+$ minimizes the time-to-climb and the minimal value is $t_{f, \min} := 658s$. Note that we retrieve the CAS/Mach procedure for some values of ϕ_{\max} and ψ_{\max} . Up to this point, we can say that the CAS/Mach procedure is not the optimal structure for the minimum time-to-climb problem with the Maximal Operation Speed (VMO) and the Maximal Operationnal Mach (MMO) speeds as constraints. However, when decreasing these constraints, the CAS/Mach procedure may be optimal since the classification from Fig. 5 gives in this case BC-extremals of the form $\sigma_- \sigma_{c_1} \sigma_{c_2} \sigma_+$.

Remark 13. In the previous paragraph, we say “may be optimal” since we only check necessary conditions of optimality given by the PMP, that is we only compute BC-extremals. The questions of local and global optimality are more intricate in this context and go beyond the scope of this article.

Remark 14. According to the Sect. IV B, for the solutions from Fig. 5, the fuel consumption is minimal for $(\phi_{\max}, \psi_{\max}) \approx (128.9, 0.661)$ where the structure follows the CAS/Mach procedure.

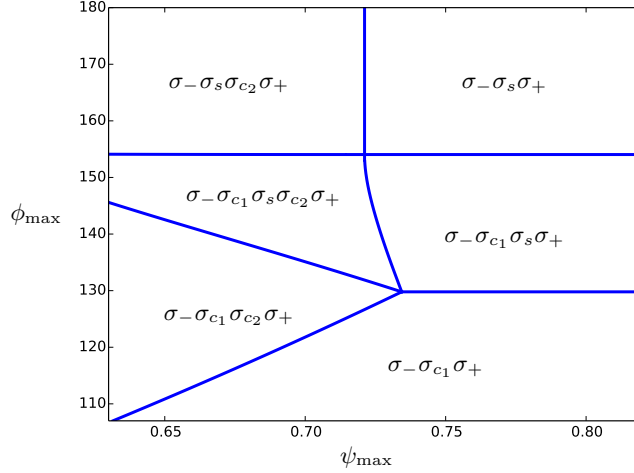


Fig. 5 Classification of the BC-extremal structures with respect to ϕ_{\max} and ψ_{\max} . To the separating blue lines is associated structures with touch points. These blue lines are computed also by homotopy and monitoring.

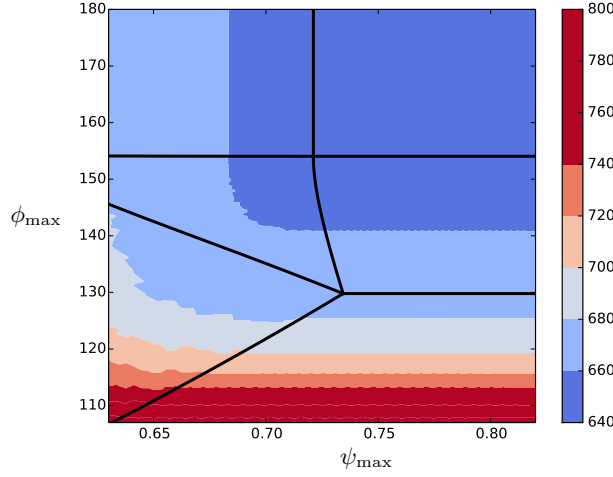


Fig. 6 The final time with respect to ϕ_{\max} and ψ_{\max} .

IV. The minimal cost index: the CAS/Mach and Singular Arc procedures

A. Description of the CAS/Mach (CM) and Singular Arc (SA) procedures

1. The CAS/Mach (or CM) procedure usually splits the climbing phase in two parts, the first one is an arc σ_{c1} at constant CAS and the second one is an arc σ_{c2} at constant Mach number. In our case, we have seen that we need a first negative bang arc σ_- and we need a final positive

bang arc σ_+ to reach the limit conditions ($b(x(t_f)) = 0$) which define the cruise phase. We consider the cost index g_α (see Sect. II A) as the objective function and we want to optimize the CAS and Mach numbers, that is ϕ_{\max} and ψ_{\max} , followed respectively along the arcs σ_{c_1} and σ_{c_2} . The NLP problem may be summarize this way:

$$(CM_\alpha) \quad \min_{X \in \mathbb{R}^{15}} f_\alpha^{\text{CM}}(X), \quad \text{subjected to} \quad g_\alpha^{\text{CM}}(X) = 0,$$

where the decision variable is $X = (t_1, t_2, t_3, t_f, x_1, x_2, x_3, \phi_{\max}, \psi_{\max})$, where the cost function is given by

$$f_\alpha^{\text{CM}}(X) := \alpha t_f + (1 - \alpha) \left[m_0 - \pi_m \left(e^{(t_f - t_3)F_+}(x_3) \right) \right] = g_\alpha \left(t_f, \pi_m \left(e^{(t_f - t_3)F_+}(x_3) \right) \right),$$

with $\pi_m(x) = m$, recalling that $x = (h, v, m)$, and where m_0 is the initial mass, which is fixed. The equality constraints are defined by:

$$g_\alpha^{\text{CM}}(X) := \begin{pmatrix} b(e^{(t_f - t_3)F_+}(x_3)) \\ x_1 - e^{t_1 F_-}(x_0) \\ x_2 - e^{(t_2 - t_1)F_{c_1}}(x_1) \\ x_3 - e^{(t_3 - t_2)F_{c_2}}(x_2) \\ \phi_{\max} - \phi(x_1) \\ \psi_{\max} - \psi(x_2) \end{pmatrix} \in \mathbb{R}^{13},$$

recalling that $x_0 = (h_0, v_0, m_0)$ is fixed (see Sect. II A for the numerical values).

Remark 15. One can check *a posteriori* that $0 \leq t_1 \leq t_2 \leq t_3 \leq t_f$ or include these inequality constraints.

2. We define now what we call the Singular Arc (or SA) procedure, which is, as we can see in the next sections, better than the CM procedure. The trajectories following a SA procedure have a simple structure of the form $\sigma_- \sigma_s \sigma_+$. The cost function is still the cost index and the equality constraints are simply given by the limit and matching conditions. This gives the following NLP problem:

$$(SA_\alpha) \quad \min_{X \in \mathbb{R}^9} f_\alpha^{\text{SA}}(X), \quad \text{subjected to} \quad g_\alpha^{\text{SA}}(X) = 0,$$

where the decision variable is here $X = (t_1, t_2, t_f, x_1, x_2)$, where the cost function is given by

$$f_\alpha^{\text{SA}}(X) := \alpha t_f + (1 - \alpha) \left[m_0 - \pi_m \left(e^{(t_f - t_2)F_+}(x_2) \right) \right] = g_\alpha \left(t_f, \pi_m \left(e^{(t_f - t_2)F_+}(x_2) \right) \right),$$

and where the equality constraints are defined by:

$$g_\alpha^{\text{SA}}(X) := \begin{pmatrix} b(e^{(t_f - t_2)F_+}(x_2)) \\ x_1 - e^{t_1 F_-}(x_0) \\ x_2 - e^{(t_2 - t_1)F_s}(x_1) \end{pmatrix} \in \mathbb{R}^8.$$

Remark 16. One can also check *a posteriori* that $0 \leq t_1 \leq t_2 \leq t_f$ or include these inequality constraints.

Remark 17. Let us denote by (SA_{t_f}) the problem (SA_α) with $\alpha = 1$ and (CM_{t_f}) the problem (CM_α) with $\alpha = 1$. According to the results from Sect. III B, we have that (SA_{t_f}) is better than (CM_{t_f}) .

B. The minimal fuel consumption

We have already compared (SA_α) and (CM_α) when $\alpha = 1$. Before comparing them for all possible values of α , we present a brief analysis when $\alpha = 0$, that is for the minimum fuel consumption problem. We denote by (SA_{Δ_m}) the problem (SA_α) with $\alpha = 0$ and (CM_{Δ_m}) the problem (CM_α) with $\alpha = 0$.

1. For the minimum time-to-climb problem, the classification of the different structures is given on Fig. 5. Associated to these BC-extremals, we have the final time t_f presented on Fig. 6. One can notice that the minimal value ($t_{f,\min} = 658\text{s}$) is of course for the structure following the SA procedure, and the associated fuel consumption is $\Delta_m \approx 882\text{kg}$. For the BC-extremals from Fig. 5, one can also compute the fuel consumption, which is given on Fig. 7. Even if it is not clear from the figure, the minimal value of fuel consumption is obtained when $(\phi_{\max}, \psi_{\max}) \approx (128.9, 0.661)$, that is for a structure following the CM procedure. To obtain these optimal values we solve the problem (CM_{Δ_m}) with the *fmincon* function from the *Matlab* software and we use the *ode45* function to compute the numerical integrations. The numerical solution we obtain satisfies the first-order optimality conditions with a tolerance of 10^{-7} . The switching times are $t_1 \approx 32\text{s}$, $t_2 \approx 450\text{s}$ and $t_3 \approx 676\text{s}$ while we have a fuel consumption $\Delta_m \approx 863\text{kg}$, a final time $t_f \approx 677\text{s}$ and the CAS and Mach optimal values are $(\phi_{\max}, \psi_{\max}) \approx (128.9, 0.661)$. Hence, there is a gap of 19s on the final time and 19kg on the fuel consumption between the solution of (SA_{t_f}) and (CM_{Δ_m}) . That is, the decrease of the bounds ϕ_{\max} and ψ_{\max} of the state constraints c_1 and c_2 allows to save up 19 kg of fuel with a 19 seconds longer flight.

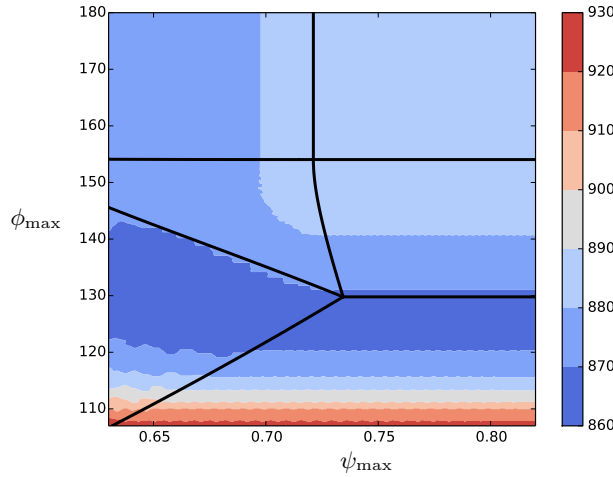


Fig. 7 The fuel consumption with respect to ϕ_{\max} and ψ_{\max} for the minimum time-to-climb problem. See the classification on Fig. 5.

2. In the previous item, we have compared (SA_{t_f}) with (CM_{Δ_m}) . As expected, solving (SA_{t_f}) gives a better final time and solving (CM_{Δ_m}) gives a better fuel consumption. But we can do better. Let us come back to the original problem of minimizing the fuel consumption, that is to the optimal control problem (P_{Δ_m}) defined in Sect. II A. To obtain the best possible fuel consumption, we set $\phi_{\max} = VMO$ and $\psi_{\max} = MMO$. In this case, thanks to the direct collocation method from the *Bocop* software [6], we determine that the structure $\sigma_- \sigma_s \sigma_+$ is still relevant, that is the solution we obtain follows the SA procedure. This means that the problem (SA_{Δ_m}) gives a better solution than the problem (CM_{Δ_m}) . To get a BC-extremal solution of (P_{Δ_m}) with $\phi_{\max} = VMO$ and $\psi_{\max} = MMO$, we use the solution from the *Bocop* software to initialize the shooting method. This technique is classical, see Ref. [8, 23] for details and examples. We then obtain a trajectory for which the switching times are $t_1 \approx 47\text{s}$ and $t_2 \approx 668\text{s}$, the final time is $t_f \approx 675\text{s}$ and the fuel consumption is $\Delta_m \approx 860\text{kg}$. There is a gap of 3kg on the fuel consumption with the solution from (CM_{Δ_m}) and a difference about 2s on the final time. Hence, even if (SA_{Δ_m}) is better than (CM_{Δ_m}) , the difference is very small.

Remark 18. The *Bocop* software transforms an infinite dimensional optimal control problem into a finite dimensional optimization problem, applying a full time discretization of the state variables, the control variables, the dynamics and the constraints.

Let us recap on the following table, the fuel consumptions and the final times for the different problems we have encountered so far:

	\parallel (SA $_{t_f}$)	\parallel (CM $_{t_f}$)	\parallel (SA $_{\Delta_m}$)	\parallel (CM $_{\Delta_m}$)
t_f	658s	660s	675s	677s
Δ_m	882kg	884kg	860kg	863kg

From this table, it is clear that the SA procedure is slightly better than the CM procedure in terms of time-to-climb and fuel consumption. In the next section we analyze the case when $\alpha \in [0, 1]$. This will complete the comparison between (CM $_{\alpha}$) and (SA $_{\alpha}$).

C. The minimal cost index

We are interested now in the so-called cost index criterion which is a convex combination of the time-to-climb and the fuel consumption. Instead of comparing directly (CM $_{\alpha}$) and (SA $_{\alpha}$) for $\alpha \in [0, 1]$, we solve (P_{α}) with $\phi_{\max} = VMO$, $\psi_{\max} = MMO$, and we show that the trajectories we get follow the SA procedure. Doing this, we conclude at the end of this section, that not only the SA procedure is better than the CM procedure but it is also the best solution we have. The problem (CM $_{\alpha}$) is solved with the *fmincon* function from the **Matlab** software and we use the *ode45* function to compute the numerical integrations. The optimal control problem (P_{α}) is solved with the **Bocop** software.

We present in Table 2, for different values of α , the optimized CAS/Mach couple (that is ϕ_{\max} and ψ_{\max}), the fuel consumption and the time-to-climb, associated to the solutions of problems (CM $_{\alpha}$) and (P_{α}). Since all the trajectories from problem (P_{α}) with $\phi_{\max} = VMO$, $\psi_{\max} = MMO$ follow the SA procedure, it is clear from Table 2 that the SA procedure is better than the CM procedure. Indeed, for any α , the final time and the fuel consumption are better, so is the cost index. It is interesting to notice that not only the objective function is better (that is the cost index) but also each part of the criterion (that is t_f and Δ_m). This fact is visible on Fig. 8, where the left sub-graph depicts the evolution of the fuel consumption with respect to α while the right sub-graph presents the time-to-climb.

To conclude, this numerical investigation indicates that the SA procedure of the form $\sigma_- \sigma_s \sigma_+$ is the best policy but the actual Cas/Mach procedure of the form $\sigma_- \sigma_{c_1} \sigma_{c_2} \sigma_+$ provides good sub-optimal trajectories regarding the cost index criterion g_{α} .

α	(CM $_{\alpha}$)					(P $_{\alpha}$)		α	(CM $_{\alpha}$)					(P $_{\alpha}$)	
	CAS	Mach	Δ_m	t_f	Δ_m	t_f	CAS		Mach	Δ_m	t_f	Δ_m	t_f		
0	128.9	0.6611	862.7	677.1	860.0	675.4		0.526	134.9	0.6808	867.3	665.1	864.7	663.2	
0.056	129.4	0.6631	862.8	675.7	860.1	674.0		0.579	135.6	0.6825	868.4	664.2	865.8	662.3	
0.105	129.9	0.6651	862.9	674.3	860.2	672.6		0.631	136.4	0.6844	869.7	663.4	867.1	661.4	
0.158	130.4	0.6670	863.1	673.0	860.4	671.2		0.684	137.2	0.6863	871.1	662.7	868.5	660.7	
0.210	131.0	0.6690	863.4	671.7	860.7	669.9		0.737	138.1	0.6882	872.7	662.0	870.1	660.0	
0.263	131.6	0.6709	863.8	670.5	861.1	668.6		0.790	139.0	0.6901	874.5	661.5	872.0	659.5	
0.316	132.2	0.6729	864.2	669.3	861.6	667.4		0.842	139.9	0.6920	876.6	661.0	874.0	659.0	
0.368	132.8	0.6748	864.8	668.2	862.2	666.3		0.895	140.9	0.6939	878.8	660.7	876.3	658.7	
0.421	133.5	0.6767	865.5	667.1	862.9	665.2		0.947	142.0	0.6959	881.4	660.4	878.8	658.5	
0.474	134.2	0.6787	866.4	666.1	863.7	664.1		1.000	143.2	0.6978	884.3	660.4	881.6	658.4	

Table 2 Fuel consumption and time-to-climb for different values of α associated to the solutions of problems (CM $_{\alpha}$) and (P_{α}). For problem (P_{α}), we set $\phi_{\max} = VMO$ and $\psi_{\max} = MMO$.

V. Conclusion

This article is about aircraft trajectory optimization during the climbing phase considering the cost index criterion. The cost index criterion is a convex combination of the time-to-climb and the fuel consumption. The problem is first modeled as an optimal control problem in Mayer form

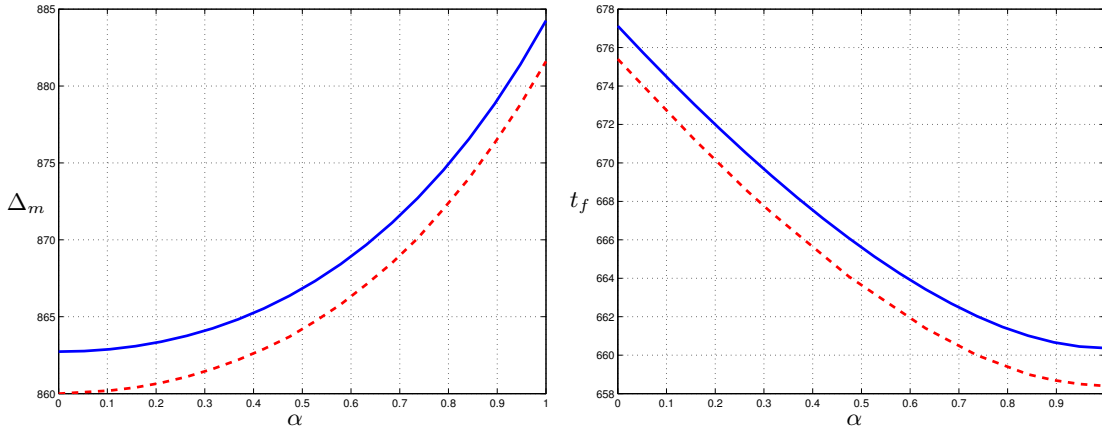


Fig. 8 Evolution of the fuel consumption (on the left) and of the time-to-climb (on the right) with respect to α . The plain lines are for problem (CM_α) while the dotted lines are for problem (P_α) with $\phi_{\max} = VMO$, $\psi_{\max} = MMO$. The two curves are fairly similar, the relative gap between the plain and the dotted curves is about 0.3%.

with a single-input affine control system and with pure state constraints. The two state constraints give CAS and Mach speeds limitations while the control variable is the air slope of the aircraft. In Sect. III, for the minimum time-to-climb problem, we classify the BC-extremals structures with respect to the bounds of the state constraints, that is ϕ_{\max} and ψ_{\max} , using small time analysis, indirect multiple shooting and homotopy with monitoring methods. This classification emphasizes the role of the SA and CM procedures that we define and compare in Sect. IV. The CM procedure is the actual CAS/Mach procedure of the form $\sigma_- \sigma_{c_1} \sigma_{c_2} \sigma_+$ while the SA procedure has the simple form $\sigma_- \sigma_s \sigma_+$. Fixing the structure, that is considering the SA or CM procedure, the optimal control problem can be reduced to a simpler optimization problem in finite and small dimension, where the adjoint vector p is not needed. This is possible in particular because of the parameterization of the control in the feedback form $u_s(x)$, $u_{c_1}(x)$ and $u_{c_2}(x)$ respectively along the arcs σ_s , σ_{c_1} and σ_{c_2} . The numerical investigation shows that the SA procedure is better than the actual CM procedure, in terms of time-to-climb and fuel consumption when considering the cost index criterion, but the difference is small. Besides, thanks to the numerical results obtained with the direct collocation method, considering the optimal control problem (P_α) with $\phi_{\max} = VMO$ and $\psi_{\max} = MMO$, we have also that the SA procedure is possibly the best policy.

It is worth to mention that in this article we consider only the air slope as control variable. This means that the thrust is assumed to be constant. It would be interesting to consider the thrust as a control variable and in this case, we would have a bi-input control system still affine in the control and more complex structures may appear. A preliminary study has been done in Ref. [19] but the singular perturbation phenomenon has not been taken into account, which makes the numerical resolution much more difficult.

References

- [1] E. Allgower and K. Georg. *Introduction to numerical continuation methods*, volume 45 of *Classics in Applied Mathematics*. Soc. for Industrial and Applied Math., Philadelphia, PA, USA, 2003.
- [2] M. D. Ardema. Solution of the Minimum Time-to-Climb Problem by Matched Asymptotic Expansions. *AIAA Journal*, 14(7):pp. 843–850, July 1976.
- [3] M. D. Ardema. *Singular perturbations in flight mechanics*. PhD thesis, 1977.
- [4] J. F. Barman, H. Erzenberg, and J. D. McLean. Fixed-range optimum trajectories for short-haul aircraft. Technical note NASA TN D-8115, NASA Ames Research Center, 1975.
- [5] H. G. Bock and K. J. Plitt. A multiple shooting algorithm for direct solution of optimal control problems. *IFAC Proceedings Volumes*, 17(2):1603 – 1608, 1984. 9th IFAC World Congress: A Bridge Between Control Science and Technology, Budapest, Hungary, 2-6 July 1984.
- [6] J. F. Bonnans, D. Giorgi, V. Grelard, B. Heymann, S. Maindault, P. Martinon, and O. Tissot. Bocop - a collection of examples. Technical report, INRIA, 2016.

- [7] B. Bonnard and M. Chyba. *Singular Trajectories and their Role in Control Theory*. Springer-Verlag, Berlin, Heidelberg, Paris, 2003.
- [8] B. Bonnard, M. Claeys, O. Cots, and P. Martinon. Geometric and numerical methods in the contrast imaging problem in nuclear magnetic resonance. *Acta Applicandae Mathematicae*, 135(1):5–45, Feb 2015.
- [9] B. Bonnard, L. Faubourg, G. Launay, and E. Trélat. Optimal control with state constraints and the space shuttle re-entry problem. *J. Dyn. Control Syst.*, 9(2):155–199, 2003.
- [10] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux*. Springer-Verlag, Berlin, Heidelberg, New-York, 2006.
- [11] R. Bulirsch and J. Stoer. *Introduction to numerical analysis*. Springer, New-York, 2002.
- [12] J.-B. Caillau, O. Cots, and J. Gergaud. Differential continuation for regular optimal control problems. *Optim. Methods Softw.*, 27(2):177–196, 2012.
- [13] A. J. Calise and D. S. Naidu. Singular Perturbations and Time Scales in Guidance and Control of Aerospace Systems: A Survey. *J. Guidance, Control and Dynamics*, 24(6):pp. 1057–1078, November-December 2001.
- [14] A. Casamayou, N. Cohen, G. Connan, T. Dumont, L. Fousse, and al. Calcul mathématique avec Sage. Technical report, 2013.
- [15] J. R. Cash. The numerical solution of nonlinear two-point boundary value problems using iterated deferred correction - a survey. *Opuscula Math.*, 26(2), 2006.
- [16] J. R. Cash, G. Moore, and R. W. Wright. An automatic continuation strategy for the solution of singularly perturbed nonlinear boundary value problems. *ACM Trans. Math. Softw.*, 27(2):245–266, June 2001.
- [17] O. Cots. Geometric and numerical methods for a state constrained minimum time control problem of an electric vehicle. *ESAIM: COCV*, 23(4):1715–1749, 2017.
- [18] O. Cots, P. Delpy, J. Gergaud, and D. Goubinat. On the minimum time optimal control problem of an aircraft in its climbing phase. In *EUropean Conference for Aeronautics and Aerospace sciences (EUCASS), Milano, 03/07/2017-06/07/2017*, 2017.
- [19] O. Cots, J. Gergaud, and D. Goubinat. Direct and indirect methods in optimal control with state constraints and the climbing trajectory of an aircraft. *Optim Control Appl and Meth.*, pages 1–22, 2017.
- [20] O. Cots, J. Gergaud, and D. Goubinat. Time-optimal aircraft trajectories in climbing phase and singular perturbations (regular paper). In *IFAC World Congress, Toulouse, 09/07/2017-14/07/2017*, 2017.
- [21] A. F. Espin. *Aircraft Trajectory Optimization Using Singular Optimal Control Theory*. PhD thesis, Universidad de Sevilla, 2014.
- [22] J. Gergaud. *Sur la résolution numérique de problèmes de contrôle optimal à solution bang-bang via les méthodes homotopiques*. Habilitation à diriger des recherches, INP-ENSEEIH-IRIT, 2008.
- [23] D. Goubinat. *Contrôle optimal géométrique et méthodes numériques : application à un problème de montée d’un avion*. PhD thesis, INP-ENSEEIH-IRIT, juin 2017.
- [24] D. G. Hull. *Fundamentals of Airplane Flight Mechanics*. Springer, Heidelberg, 2007.
- [25] D. H. Jacobson, M. M. Lele, and J. L. Speyer. New necessary conditions of optimality for control problems with state-variable inequality constraints. *J. Math. Anal. Appl.*, 35:255–284, 1971.
- [26] A. J. Krener. The high order maximal principle and its application to singular extremals. *SIAM J. Control Optim.*, 1(299):256–293, 1977.
- [27] I. Kupka. Geometric theory of extremals in optimal control problems. i. the fold and maxwell case. *Trans. Amer. Math. Soc.*, 299(1):225–243, 1987.
- [28] H. Maurer. On optimal control problems with bounded state variables and control appearing linearly. *SIAM J. Cont. Optim.*, 15(3):345–362, 1977.
- [29] M. Le Merrer. *Optimisation de trajectoire d’avion pour la prise en compte du bruit dans la gestion du vol*. PhD thesis, ISAE-SUPAERO, 2012.
- [30] A. Miele. General solutions of optimum problems in nonstationnary flight. Technical memorandum 1388, National Advisory Committee for Aeronautics (NACA), 1955.
- [31] A. Miele. Optimum flight paths of turbojet aircraft. Technical memorandum 1389, National Advisory Committee for Aeronautics (NACA), 1955.
- [32] N. Moiseev. *Problèmes mathématiques d’analyse des systèmes*. Mir Moscou, 1985.
- [33] N. Nguyen. Singular arc time-optimal climb trajectory of aircraft in a two-dimensional wind field. *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2006.
- [34] D. Poles. Base of Aircraft DATA (BADA) aircraft performance modelling report. Technical Report 2009-09, Eurocontrol, september 2009.
- [35] L. Pontryagin, V. Boltianskii, R. Gamkrelidze, and E. Mishchenko. *The mathematical theory of optimal processes*. Translated from the Russian by K. Trilogoff; edited by L. W. Neustadt Interscience Publishers John Wiley & Sons, New-York-London, 1962.
- [36] J. Verriere. Mécanique du vol, performances. Notes de cours, Tome 2, ISAE-Formation ENSICA, 1997.