Unified Formulation for Element-Based Indirect Trajectory Optimization

By Juan luis GONZALO,[1)] Claudio BOMBARDELLI,[1)] and Francesco TOPPUTO[2)]

[1)]*School of Aerospace Engineering - ETSIAE, Technical University of Madrid - UPM, Madrid, Spain*
[2)]*Dipartimento di Scienze e Tecnologie Aerospaziali, Politecnico di Milano, Milan, Italy*

A general mathematical framework is presented to treat low thrust trajectory optimization problems using the indirect method and employing a generic set of orbital elements (e.g. classical elements, equinoctial, etc.). An algebraic manipulation of the optimality conditions stemming from Pontryagin Maximum Principle reveals the existence of a new quadratic form of the costate, which governs the costate contribution in all the equations of the first order necessary optimality conditions. The quadratic form provides a simple tool for the mathematical development of the optimality conditions for any chosen set of orbital elements and greatly simplifies the computation of a state transition matrix needed in order to improve the convergence of the associated two-point boundary value problem. Objective functions corresponding to minimum-time, minimum-energy and minimum-fuel problems are considered.

Key Words: Trajectory Optimization, Indirect Method, Primer Vector, Element Formulations

Nomenclature

$\mathbf{r}$	:	position vector in Cartesian coordinates
$\mathbf{v}$	:	velocity vector in Cartesian coordinates
$\mathbf{s}$	:	state vector
m	:	mass
$\mathbf{x}$	:	vector of $\mathbf{s}$ and m
$\boldsymbol{\lambda}$	:	vector of costates associated to $\mathbf{s}$
λ_m	:	costate associated to m
$\mathbf{y}$	:	vector of state, mass and costates
$\mathbf{p}$	:	vector of perturbing accelerations
$\mathbf{f}$	:	vector of thrust acceleration
f	:	thruster throttling parameter
$\boldsymbol{\alpha}$	:	thrust pointing unity vector
$\mathbf{u}$	:	control vector $[f\ \boldsymbol{\alpha}]$
$T_{\max}$	:	maximum thrust
c	:	thruster exhaust velocity
J	:	objective (cost) function
$\mathcal{H}$	:	Hamiltonian
$\mathcal{L}$	:	Lagrangian
$\boldsymbol{\psi}$	:	vector of terminal constraints
$\mathbf{g}$	:	vector of path constraints
ε	:	homotopy parameter
S	:	switching function
Subscripts		
i	:	initial
f	:	final
l	:	lower bound
u	:	upper bound
Superscripts		
$*$	:	optimal solution

1. Introduction

One of the greatest challenges when designing optimal trajectories using indirect methods is the treatment of the costate. These variables, introduced by the two point boundary value problem (TPBVP) derived from the first-order necessary optimality conditions and linked to their corresponding state variables, lack a physical interpretation in general. This complicates the task of finding an initial guess for the solution of the TPBVP, required if it is to be solved using iterative numerical algorithms. Furthermore, the introduction of the costate into the state equations through the control term increases their complexity and makes it more difficult to derive variational equations for the accurate integration of the state transition matrix, essential to increase the performance of numerical shooting methods. When dynamics are formulated using Cartesian coordinates, these issues are alleviated by the well known result of Lawden's primer vector theory.[1)] By applying Pontryagin Maximum Principle (PMP), it is possible to verify that the direction of thrust is given by a vector whose components are the costate variables associated to velocity. This leads to a fairly simple expression of the control involving only the costate, while also providing some physical sense for at least part of the costate.

However, there are practical cases where formulations for the dynamics other than Cartesian may be more convenient. For instance, using orbital elements allows one to express constraints in a natural way as fixed initial and final values of the state when moving from one orbit to another without specifying the specific departure and arrival points. Unfortunately, when the system dynamics are parametrized through orbital elements the classical primer vector result is not preserved and the optimal thrust orientation becomes a function of both state and costate. This fact greatly complicates the computation of partial derivatives and the derivation of an STM in analytical form as required in order to improve the convergence of the associated TPBVP.[4,6)]

In this article, we show that by carefully manipulating the equations derived from the first order optimality conditions it is possible to write them in a compact and more convenient way, partially isolating the different contributions of state and costate and leading to a unified, efficient formulation of the indirect equations for trajectory optimization that can be applied to any set of generalized orbital elements or orbital parametrization. In the first section of the article we rewrite the necessary optimality conditions of a generic optimization problem to highlight the existence of a quadratic form of the costate

whose partial derivatives directly provide the contribution of the costate to all the OCP equations (including the variational ones). The computation of the partial derivatives is simplified thanks to the separation of the state and costate contributions in the quadratic form. In the subsequent section we proceed to the derivation of the state transition matrix terms applicable to a generic formulation. Finally, we derive the partial derivatives of the Lagrangian and thrust throttling factor corresponding to minimum-time, minimum-energy and minimum-fuel problems. One application example is cited and conclusions are drawn.

2. Indirect Optimal Control Problem Formulation

A generic formulation for orbital dynamics with an n-dimensional state vector $\mathbf{s}$ and independent variable t (from now on considered a "generalized time") can be expressed in the form

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \mathbf{F}\left(\mathbf{s}, m, \mathbf{u}, t\right) = \mathbf{B}\left(\mathbf{s}, t\right)\ \mathbf{p}\left(\mathbf{s}, m, \mathbf{u}, t\right) + \mathbf{A}\left(\mathbf{s}, t\right)\ , \qquad (1)$$

together with the mass equation

$$\frac{\mathrm{d}m}{\mathrm{d}t} = F_m\left(\mathbf{u}, t\right) = -\frac{T_{\max}}{c} f\ , \qquad (2)$$

where $\mathbf{u} = [\, f\, \boldsymbol{\alpha}\,]$ is the control vector formed by the thruster throttling parameter f and pointing unity vector $\boldsymbol{\alpha}$, $\mathbf{A}$ is an $n\times 1$ vector and $\mathbf{B}$ a $n\times 3$ matrix. Without loss of generality, it is possible to assume that the thrust acceleration is the only perturbing action

$$\mathbf{p} = \mathbf{f} = f\frac{T_{\max}}{m}\boldsymbol{\alpha}$$

leading to

$$\mathbf{F} = f\frac{T_{\max}}{c}\mathbf{B}\cdot\boldsymbol{\alpha} + \mathbf{A}\,. \qquad (3)$$

Note that this expression holds even if there are additional perturbing accelerations, simply by redefining $\mathbf{A}$ as $\mathbf{A} + \mathbf{B}\left(\mathbf{p} - \mathbf{f}\right)$.

In general, an Optimal Control Problem (OCP) can be defined as minimizing (or maximizing) an objective function of the form,[3)]

$$J = \phi\left(\mathbf{x}_f, t_f\right) + \int_{t_i}^{t_f} \mathcal{L}\left(\mathbf{x}, \mathbf{u}, t\right)\ \mathrm{d}t \qquad (4)$$

where $\mathbf{x} = [\mathbf{s}\ m]$, for a given set of constraints of the state and control imposed at the terminal points t_i and t_f

$$\boldsymbol{\psi}_{il} \le \boldsymbol{\psi}_i\left(\mathbf{x}_i, \mathbf{u}_i, t_i\right) \le \boldsymbol{\psi}_{iu}$$
$$\boldsymbol{\psi}_{fl} \le \boldsymbol{\psi}_f\left(\mathbf{x}_f, \mathbf{u}_f, t_f\right) \le \boldsymbol{\psi}_{fu}$$

and along the trajectory (also known as path constraints)

$$\mathbf{g}_l \le \mathbf{g}\left(t, \mathbf{x}\left(t\right), \mathbf{u}\left(t\right)\right) \le \mathbf{g}_u$$
$$\mathbf{x}_l \le \mathbf{x}\left(t\right) \le \mathbf{x}_u$$
$$\mathbf{u}_l \le \mathbf{u}\left(t\right) \le \mathbf{u}_u$$

Indirect optimization methods make use of the calculus of variations to derive the first order optimality conditions corresponding to this OCP,[3)] leading to a TPBVP which is normally solved with numerical algorithms such as shooting methods. The number of equations (both differential and algebraic) and unknowns will depend on the number and nature of the constraints, but in all cases a set of costate variables $\boldsymbol{\lambda}$ and λ_m associated to $\mathbf{s}$ and m are introduced (along with the differential equations describing their evolution). A detailed treatment of the different types of constraints falls beyond the scope of this work. For simplicity, in the following it is assumed that there are no path constraints, and that the terminal constraints can be expressed as fixed of free values of $\mathbf{x}$.

The Hamiltonian associated to the OCP can be written as:

$$\mathcal{H} = \boldsymbol{\lambda}^\top\cdot\mathbf{F} + \lambda_m F_m + \mathcal{L} \qquad (5)$$

where the evolution of the costates is given by the Euler-Lagrange equations derived from the first order optimality conditions using the calculus of variations:[3)]

$$\dot{\boldsymbol{\lambda}} = \mathbf{G} = -\frac{\partial\mathcal{H}}{\partial\mathbf{s}} \qquad (6)$$

$$\dot{\lambda_m} = G_m = -\frac{\partial\mathcal{H}}{\partial m} \qquad (7)$$

An expression for the optimal control $\mathbf{u}^* = [f^*\ \boldsymbol{\alpha}^*]$ as a function of $\mathbf{x}$, $\boldsymbol{\lambda}$, λ_m and t can be obtained by applying the PMP,[2)] which states that the optimal control for a given optimal trajectory is the one that leads to an extreme value of $\mathcal{H}$ over the set of admissible controls. For the minimization case, the PMP can be written in mathematical form as follows

$$\mathbf{u}^* = \underset{\mathbf{u}\in\mathbf{U}}{\arg\min}\ \mathcal{H}\left(\mathbf{x}^*, \boldsymbol{\lambda}^*, \lambda_m^*, \mathbf{u}; t\right) \quad \forall t \in \left[t_i, t_f\right]$$

where

$$\mathbf{U} = \left\{\boldsymbol{\alpha}\in\mathbb{R}^3/\|\boldsymbol{\alpha}\| = 1\right\} \cup \left\{f\in\mathbb{R}/f\in[0,1]\right\}$$

is the allowable set for the control. Substituting Eqs. (2) and (3) into Eq. (5) and rearranging terms one reaches:

$$\mathcal{H} = \frac{T_{\max}}{m} f\left[\boldsymbol{\lambda}^\top\cdot\mathbf{B}\cdot\boldsymbol{\alpha}\right] + \boldsymbol{\lambda}^\top\cdot\mathbf{A} - \lambda_m\frac{T_{\max}}{c} + \mathcal{L}$$

Because $(T_{\max}/m)f$ is semi-positive the term in brackets should be negative and as large (in absolute value) as possible, leading to an expression for $\boldsymbol{\alpha}^*$

$$\boldsymbol{\alpha}^* = -\frac{\boldsymbol{B}^\top\cdot\boldsymbol{\lambda}}{\|\mathbf{B}^\top\cdot\boldsymbol{\lambda}\|}\,. \qquad (8)$$

This expression is a generalization of the well-known result of the primer vector in Cartesian coordinates,[1)] which takes the simple form $\boldsymbol{\alpha}^* = -\boldsymbol{\lambda}_\mathbf{v}/\lambda_\mathbf{v}$. However, for an arbitrary formulation of the dynamics $\boldsymbol{\alpha}^*$ can be a complex expression of both the state and the costate; this is one of the main drawbacks of using element formulations to solve OCPs with indirect methods. On the other hand, the application of the PMP to determine f will depend on the particular expression of $\mathcal{L}$ for each OCP, so a general solution cannot be derived. In the following, it is assumed that f^* is a known function of state, mass and costate.

Plugging the expression for $\boldsymbol{\alpha}^*$, Eq. (8), back into Eq. (3) yields:

$$\mathbf{F} = -f^*\frac{T_{\max}}{m}\frac{\mathcal{B}\cdot\boldsymbol{\lambda}}{\sqrt{\boldsymbol{\lambda}^\top\cdot\mathcal{B}\cdot\boldsymbol{\lambda}}} + \mathbf{A} \qquad (9)$$

where $\mathcal{B} = \mathbf{B}\mathbf{B}^\top$ is a square matrix of dimension $n\times n$. This matrix is, by construction, symmetric and positive (semi-)definite. Consequently, $\boldsymbol{\lambda}^\top\cdot\mathcal{B}\cdot\boldsymbol{\lambda}$ represents a quadratic form of $\boldsymbol{\lambda}$, where

the coefficients of the quadratic form are in turn functions of $\mathbf{s}$. Introducing

$$\Psi = \sqrt{\boldsymbol{\lambda}^\top \cdot \mathcal{B} \cdot \boldsymbol{\lambda}} \tag{10}$$

as the square root of the quadratic form, it is possible to express $\mathbf{F}$ in a more compact way:

$$\mathbf{F} = -f^* \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial \boldsymbol{\lambda}} + \mathbf{A} \tag{11}$$

as well as the Hamiltonian:

$$\mathcal{H} = -f^* \frac{T_{\max}}{m} \Psi + \boldsymbol{\lambda}^\top \cdot \mathbf{A} - \lambda_m f^* \frac{T_{\max}}{c} + \mathcal{L}. \tag{12}$$

Substituting $\mathcal{H}$ into Eqs. (6) and (7) and operating, the ordinary differential equations (ODEs) describing the time evolution of $\boldsymbol{\lambda}$ and λ_m are reached (for clarity, the equations are written using Einstein notation):

$$\begin{aligned} \left(\dot{\boldsymbol{\lambda}}\right)_l = \dot{\lambda}_l = (\mathbf{G})_l = \frac{\partial f^*}{\partial s_l} \frac{T_{\max}}{m} \Psi + f^* \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial s_l} \\ -\lambda_j \frac{\partial A_j}{\partial s_l} + \frac{\partial f^*}{\partial s_l} \lambda_m \frac{T_{\max}}{c} + \frac{\partial \mathcal{L}}{\partial s_l} \end{aligned} \tag{13}$$

$$\dot{\lambda}_m = G_m = \frac{\partial f^*}{\partial m} \frac{T_{\max}}{m} \Psi - f^* \frac{T_{\max}}{m^2} \Psi + \frac{\partial f^*}{\partial m} \lambda_m \frac{T_{\max}}{c} - \frac{\partial \mathcal{L}}{\partial m} \tag{14}$$

with

$$\frac{\partial \Psi}{\partial s_l} = \frac{1}{2} \frac{\lambda_j \frac{\partial B_{jk}}{\partial s_l} \lambda_k}{\Psi} \tag{15}$$

being A_j and B_{jk} the elements of $\mathbf{A}$ and $\mathcal{B}$, respectively.

Gathering the previous results, the ODE system describing the evolution of $\mathbf{s}$, m, $\boldsymbol{\lambda}$ and λ_m can be written in a compact form as:

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \mathbf{s} \\ m \\ \boldsymbol{\lambda} \\ \lambda_m \end{bmatrix} = \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{y} = \mathbf{R} = \begin{bmatrix} \mathbf{F} \\ F_m \\ \mathbf{G} \\ G_m \end{bmatrix} \tag{16}$$

where the expressions for $\mathbf{F}$, F_m, $\mathbf{G}$ and G_m are given in Eqs. (11), (2), (13) and (14), respectively.

2.1. Variational equation and STM

A typical approach for solving the TPBVP associated to the OCP is using shooting methods, which require the Jacobian of the shooting function (or a suitable approximation). Although it can be approximated numerically using finite differences and BFGS updates, more efficiency and robustness can be achieved by constructing it analytically from the State Transition Matrix (STM) of the problem.[4)]

The STM maps small variations of the initial conditions, $\delta \mathbf{y}_i = [\delta \mathbf{s}_i \, \delta m_i \, \delta \boldsymbol{\lambda}_i \, \delta \lambda_{mi}]$, into small variations of $\mathbf{y}$ at a given time t, $\delta \mathbf{y} = [\delta \mathbf{s} \, \delta m \, \delta \boldsymbol{\lambda} \, \delta \lambda_m]$. The STM is calculated through the integration of the variational equation:

$$\dot{\boldsymbol{\Phi}}(t_i, t) = \mathrm{D}_{\mathbf{y}} \mathbf{R} \, \boldsymbol{\Phi}(t_i, t)$$

where $\mathrm{D}_{\mathbf{y}} \mathbf{R}$ is the Jacobian of $\mathbf{R}$, and the initial value for $\boldsymbol{\Phi}$ corresponds to the identity matrix of dimension $2(n+1) \times 2(n+1)$.

Although computing the Jacobian for $\mathbf{R}$ entails no intrinsic challenges from a conceptual point of view, its practical implementation may become rather burdensome. As previously indicated, $\boldsymbol{\alpha}^*$ can be a complex function involving both the state and the costate, and the derivatives of $\mathbf{R}$ may rapidly become too cumbersome to handle even with symbolic manipulators. However, it is relatively simple to derive general expressions for any formulation leveraging the structures previously revealed in the equations. The Jacobian will then be expressed as a set of problem-independent algebraic relations involving $\mathcal{B}$, $\mathbf{A}$, f^*, $\mathcal{L}$ and their derivatives, which will be problem-dependent.

Following this approach, the Jacobian can be written as:

$$\mathrm{D}_{\mathbf{y}} \mathbf{R} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial \mathbf{s}} & \frac{\partial \mathbf{F}}{\partial m} & \frac{\partial \mathbf{F}}{\partial \boldsymbol{\lambda}} & \frac{\partial \mathbf{F}}{\partial \lambda_m} \\ \frac{\partial F_m}{\partial \mathbf{s}} & \frac{\partial F_m}{\partial m} & \frac{\partial F_m}{\partial \boldsymbol{\lambda}} & \frac{\partial F_m}{\partial \lambda_m} \\ \frac{\partial \mathbf{G}}{\partial \mathbf{s}} & \frac{\partial \mathbf{G}}{\partial m} & \frac{\partial \mathbf{G}}{\partial \boldsymbol{\lambda}} & \frac{\partial \mathbf{G}}{\partial \lambda_m} \\ \frac{\partial G_m}{\partial \mathbf{s}} & \frac{\partial G_m}{\partial m} & \frac{\partial G_m}{\partial \boldsymbol{\lambda}} & \frac{\partial G_m}{\partial \lambda_m} \end{bmatrix}$$

with

$$\left(\frac{\partial \mathbf{F}}{\partial \mathbf{s}}\right)_{jk} = \frac{\partial F_j}{\partial s_k} = -\frac{T_{\max}}{m} \frac{\partial \Psi}{\partial \lambda_j} \frac{\partial f^*}{\partial s_k} - f^* \frac{T_{\max}}{m} \frac{\partial^2 \Psi}{\partial s_k \partial \lambda_j} + \frac{\partial A_j}{\partial s_k}$$

$$\left(\frac{\partial \mathbf{F}}{\partial m}\right)_j = \frac{\partial F_j}{\partial m} = -\frac{\partial f^*}{\partial m} \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial \lambda_j} + f^* \frac{T_{\max}}{m^2} \frac{\partial \Psi}{\partial \lambda_j}$$

$$\left(\frac{\partial \mathbf{F}}{\partial \boldsymbol{\lambda}}\right)_{jk} = \frac{\partial F_j}{\partial \lambda_k} = -\frac{T_{\max}}{m} \frac{\partial \Psi}{\partial \lambda_j} \frac{\partial f^*}{\partial \lambda_k} - f^* \frac{T_{\max}}{m} \frac{\partial^2 \Psi}{\partial \lambda_j \partial \lambda_k}$$

$$\left(\frac{\partial \mathbf{F}}{\partial \lambda_m}\right)_j = \frac{\partial F_j}{\partial \lambda_m} = -\frac{\partial f^*}{\partial \lambda_m} \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial \lambda_j}$$

$$\left(\frac{\partial F_m}{\partial \mathbf{s}}\right)_j = \frac{\partial F_m}{\partial s_j} = -\frac{\partial f^*}{\partial s_j} \frac{T_{\max}}{c}$$

$$\frac{\partial F_m}{\partial m} = -\frac{\partial f^*}{\partial m} \frac{T_{\max}}{c}$$

$$\left(\frac{\partial F_m}{\partial \boldsymbol{\lambda}}\right)_j = \frac{\partial F_m}{\partial \lambda_j} = -\frac{\partial f^*}{\partial \lambda_j} \frac{T_{\max}}{c}$$

$$\frac{\partial F_m}{\partial \lambda_m} = -\frac{\partial f^*}{\partial \lambda_m} \frac{T_{\max}}{c}$$

$$\begin{aligned} \left(\frac{\partial \mathbf{G}}{\partial \mathbf{s}}\right)_{jk} = \frac{\partial G_j}{\partial s_k} = \frac{\partial^2 f^*}{\partial s_j \partial s_k} \frac{T_{\max}}{m} \Psi + \frac{T_{\max}}{m} \frac{\partial f^*}{\partial s_j} \frac{\partial \Psi}{\partial s_k} + \\ \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial s_j} \frac{\partial f^*}{\partial s_k} + f^* \frac{T_{\max}}{m} \frac{\partial^2 \Psi}{\partial s_j \partial s_k} - \\ \lambda_l \frac{\partial^2 A_l}{\partial s_j \partial s_k} - \frac{\partial^2 \mathcal{L}}{\partial s_j \partial s_k} + \lambda_m \frac{T_{\max}}{c} \frac{\partial^2 f^*}{\partial s_j \partial s_k} \end{aligned}$$

$$\begin{aligned} \left(\frac{\partial \mathbf{G}}{\partial m}\right)_j = \frac{\partial G_j}{\partial m} = \frac{\partial^2 f^*}{\partial s_j \partial m} \frac{T_{\max}}{m} \Psi - \frac{\partial f^*}{\partial s_j} \frac{T_{\max}}{m^2} \Psi + \\ \frac{\partial f^*}{\partial m} \frac{T_{\max}}{m} \frac{\partial \Psi}{\partial s_j} - f^* \frac{T_{\max}}{m^2} \frac{\partial \Psi}{\partial s_j} - \frac{\partial^2 \mathcal{L}}{\partial s_j \partial m} + \lambda_m \frac{T_{\max}}{c} \frac{\partial^2 f^*}{\partial s_j \partial m} \end{aligned}$$

$$\left(\frac{\partial \mathbf{G}}{\partial \boldsymbol{\lambda}}\right)_{jk} = \frac{\partial G_j}{\partial \lambda_k} = \frac{\partial^2 f^*}{\partial s_j \partial \lambda_k}\frac{T_{\max}}{m}\Psi + \frac{\partial f^*}{\partial s_j}\frac{T_{\max}}{m}\frac{\partial \Psi}{\partial \lambda_k} + \frac{T_{\max}}{m}\frac{\Psi}{\partial s_j}\frac{\partial f}{\partial \lambda_k} + f^*\frac{T_{\max}}{m}\frac{\partial^2 \Psi}{\partial s_j \partial \lambda_k} - \frac{\partial A_k}{\partial s_j} - \frac{\partial^2 \mathcal{L}}{\partial s_j \partial \lambda_k} + \lambda_m \frac{T_{\max}}{c}\frac{\partial^2 f^*}{\partial s_j \partial \lambda_k}$$

$$\left(\frac{\partial \mathbf{G}}{\partial \lambda_m}\right)_j = \frac{\partial G_j}{\partial \lambda_m} = \frac{\partial f^*}{\partial s_j \partial \lambda_m}\frac{T_{\max}}{m}\Psi + \frac{\partial f^*}{\partial \lambda_m}\frac{T_{\max}}{m}\frac{\partial \Psi}{\partial s_j} - \frac{\partial^2 \mathcal{L}}{\partial s_j \partial \lambda_m} + \frac{T_{\max}}{c}\frac{\partial f^*}{\partial s_j} + \lambda_m \frac{T_{\max}}{c}\frac{\partial^2 f^*}{\partial s_j \partial \lambda_m}$$

$$\left(\frac{\partial G_m}{\partial \mathbf{s}}\right)_j = \left(\frac{\partial \mathbf{G}}{\partial m}\right)_j$$

$$\frac{\partial G_m}{\partial m} = \frac{\partial^2 f^*}{\partial m^2}\frac{T_{\max}}{m}\Psi - 2\frac{\partial f^*}{\partial m}\frac{T_{\max}}{m^2}\Psi + f^*\frac{2T_{\max}}{m^3}\Psi - \frac{\partial^2 \mathcal{L}}{\partial m^2} + \lambda_m \frac{T_{\max}}{c}\frac{\partial^2 f^*}{\partial m^2}$$

$$\left(\frac{\partial G_m}{\partial \boldsymbol{\lambda}}\right)_j = \frac{\partial G_m}{\partial \lambda_j} = \frac{\partial^2 f^*}{\partial m \partial \lambda_j}\frac{T_{\max}}{m}\Psi + \frac{\partial f^*}{\partial m}\frac{T_{\max}}{m}\frac{\partial \Psi}{\partial \lambda_j} - \frac{\partial f^*}{\partial \lambda_j}\frac{T_{\max}}{m^2}\Psi - f^*\frac{T_{\max}}{m^2}\frac{\partial \Psi}{\partial \lambda_j} - \frac{\partial^2 \mathcal{L}}{\partial m \partial \lambda_j} + \lambda_m \frac{T_{\max}}{c}\frac{\partial^2 f^*}{\partial m \partial \lambda_j}$$

$$\frac{\partial G_m}{\partial \lambda_m} = \frac{\partial^2 f^*}{\partial m \partial \lambda_m}\frac{T_{\max}}{m}\Psi - \frac{\partial f^*}{\partial \lambda_m}\frac{T_{\max}}{m}\Psi - \frac{\partial^2 \mathcal{L}}{\partial m \partial \lambda_m} + \frac{T_{\max}}{c}\frac{\partial f^*}{\partial m} + \lambda_m \frac{T_{\max}}{c}\frac{\partial^2 f^*}{\partial m \partial \lambda_m}$$

These expressions involve the partial derivatives of Ψ with respect to $\mathbf{s}$ and $\boldsymbol{\lambda}$ up to order two. The first order derivative of Ψ with respect to $\mathbf{s}$ was already given in Eq. (15), and the rest can be calculated as follows:

$$\frac{\partial \Psi}{\partial \lambda_l} = \frac{B_{lj}\lambda_j}{\Psi}$$

$$\frac{\partial^2 \Psi}{\partial s_j \partial s_k} = \frac{1}{2}\frac{\lambda_l \frac{\partial^2 B_{ln}}{\partial s_j \partial s_k}\lambda_n}{\Psi} - \frac{1}{4}\frac{\left(\lambda_l \frac{\partial B_{ln}}{\partial s_j}\lambda_n\right)\left(\lambda_l \frac{\partial B_{ln}}{\partial s_k}\lambda_n\right)}{\Psi^3}$$

$$\frac{\partial^2 \Psi}{\partial \lambda_j \partial \lambda_k} = \frac{B_{jk}}{\Psi} - \frac{\left(B_{jl}\lambda_l\right)\left(B_{kl}\lambda_l\right)}{\Psi^3}$$

$$\frac{\partial^2 \Psi}{\partial s_j \partial \lambda_k} = -\frac{1}{2}\frac{\left(\lambda_l \frac{\partial B_{ln}}{\partial s_j}\lambda_n\right)\left(B_{kl}\lambda_l\right)}{\Psi^3} + \frac{\frac{\partial B_{kl}}{\partial s_j}\lambda_l}{\Psi}$$

2.2. Minimum-time, fuel and energy problems

Three especially interesting OCPs are the minimum-time, minimum-energy and minimum-fuel problems. In this subsection, the values of $\mathcal{L}$, f^* and their derivatives for these families of OCPs are presented.

The objective function for minimum-time (understood as the independent variable, or 'generalized time') problems takes the form:

$$J = \int_{t_i}^{t_f} 1\,\mathrm{d}t = \int_{t_i}^{t_f} \mathcal{L}\,\mathrm{d}t$$

leading to a constant Lagrangian of value 1. Furthermore, it is possible to prove that f^* is also constant and equal to 1. Consequently, all the derivatives of $\mathcal{L}$ and f^* become zero, greatly simplifying the previous expressions for the Jacobian of $\mathbf{R}$.

Minimum-fuel problems have a great practical interest for mission design, but the discontinuities introduced by the characteristic bang-bang structure of f^* greatly difficult their numerical treatment. One way to address this issue is to apply homotopy techniques,[5] solving an easier problem such as minimum-energy first and performing a continuation in the objective function until the original minimum-fuel problem is reached. To this end, a modified J depending on an homotopy parameter ε is introduced as done by Zhang et al.:[4]

$$J = \frac{T_{\max}}{c}\int_{t_i}^{t_f} \left[f - \varepsilon f\,(1-f)\right]\,\mathrm{d}t = \int_{t_i}^{t_f} \mathcal{L}\,\mathrm{d}t, \quad \varepsilon \in [0,1]$$

The two limit cases in ε correspond to minimum-fuel for $\varepsilon = 0$ and minimum-energy for $\varepsilon = 1$. Substituting $\mathcal{L}$ into the Hamiltonian and rearranging terms one reaches:[4]

$$\mathcal{H} = \boldsymbol{\lambda}^\top \cdot \mathbf{A} + f^*\frac{T_{\max}}{c}\left[S - \varepsilon + \varepsilon f\right]$$

where a switching function S has been defined as:

$$S = -\frac{c}{m}\Psi - \lambda_m + 1 \tag{17}$$

This S is a generalization of the expression obtained by Zhang et al.[4] for the CR3BP in Cartesian coordinates. Same as in the Cartesian case, by deriving $\mathcal{H}$ with respect to f and equating to zero a minimum for f is found in the form $f = (\varepsilon - S)/2\varepsilon$. Restricting this solution to the allowable set for the control, $f \in [0,1]$, a piecewise function for f^* is finally obtained:

$$f^* = \begin{cases} 0 & \text{for } S > \varepsilon \\ (\varepsilon - S)/2\varepsilon & \text{for } -\varepsilon \le S \le \varepsilon \\ 1 & \text{for } S < -\varepsilon \end{cases} \tag{18}$$

This expression for f^* coincides with the one given by Zhang et al.[4] using Cartesian coordinates, only changing the definition of S. Furthermore, it is straightforward to check that for the minimum-fuel case ($\varepsilon = 0$) a bang-bang profile is obtained.

The derivatives of f^* will be different from zero only when the thruster is operating in the intermediate regime, that is, for $|S| \le \varepsilon$. In that case, the derivatives of S with respect to $\mathbf{y}$ up to order 2 can be expressed as functions of the known derivatives of Ψ as follows (for brevity, only the derivatives needed to compute $\mathbf{G}$, G_m and the variational equation are reported):

$$\left(\frac{\partial f^*}{\partial \mathbf{s}}\right)_j = \frac{\partial f^*}{\partial s_j} = \frac{1}{2\varepsilon}\frac{c}{m}\frac{\partial \Psi}{\partial s_j}$$

$$\frac{\partial f^*}{\partial m} = -\frac{1}{2\varepsilon}\frac{c}{m^2}\Psi$$

$$\left(\frac{\partial f^*}{\partial \boldsymbol{\lambda}}\right)_j = \frac{\partial f^*}{\partial \lambda_j} = \frac{1}{2\varepsilon}\frac{c}{m}\frac{\partial \Psi}{\partial \lambda_j}$$

$$\frac{\partial f^*}{\partial \lambda_m} = \frac{1}{2\varepsilon}$$

$$\left(\frac{\partial^2 f^*}{\partial \mathbf{s}\partial \mathbf{s}}\right)_{jk} = \frac{\partial^2 f^*}{\partial s_j \partial s_k} = \frac{1}{2\varepsilon}\frac{c}{m}\frac{\partial^2 \Psi}{\partial s_j \partial s_k}$$

$$\left(\frac{\partial^2 f^*}{\partial \mathbf{s}\partial m}\right)_j = \frac{\partial^2 f^*}{\partial s_j \partial m} = -\frac{1}{m}\frac{\partial f^*}{\partial s_j}$$

$$\left(\frac{\partial^2 f^*}{\partial \mathbf{s}\partial \boldsymbol{\lambda}}\right)_{jk} = \frac{\partial^2 f^*}{\partial s_j \partial \lambda_k} = \frac{1}{2\varepsilon}\frac{c}{m}\frac{\partial \Psi}{\partial s_j \partial \lambda_k}$$

$$\left(\frac{\partial^2 f^*}{\partial \mathbf{s}\partial \lambda_m}\right)_j = \frac{\partial^2 f^*}{\partial s_j \partial \lambda_m} = 0$$

$$\frac{\partial^2 f^*}{\partial m^2} = \frac{1}{\varepsilon}\frac{c}{m^3}\Psi$$

$$\left(\frac{\partial^2 f^*}{\partial m\partial \boldsymbol{\lambda}}\right)_j = \frac{\partial^2 f^*}{\partial m \partial \lambda_j} = -\frac{1}{m}\frac{\partial f^*}{\partial \lambda_j}$$

$$\frac{\partial^2 f^*}{\partial m \partial \lambda_m} = 0$$

The derivatives of $\mathcal{L}$ with respect to $\mathbf{y}$ up to order 2 can now be expressed in a compact way in terms of the derivatives of f^* (again, only the terms required to evaluate $\mathbf{G}$, G_m and $\mathrm{D}_\mathbf{y}\mathbf{R}$ are included.):

$$\left(\frac{\partial \mathcal{L}}{\partial \mathbf{s}}\right)_j = \frac{\partial \mathcal{L}}{\partial s_j} = \frac{T_{\max}}{c}\frac{\partial f^*}{\partial s_j}K$$

$$\frac{\partial \mathcal{L}}{\partial m} = \frac{T_{\max}}{c}\frac{\partial f^*}{\partial m}K$$

$$\left(\frac{\partial^2 \mathcal{L}}{\partial \mathbf{s}\partial \mathbf{s}}\right)_{jk} = \frac{\partial^2 \mathcal{L}}{\partial s_j \partial s_k} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial s_j \partial s_k}K + 2\varepsilon\frac{\partial f^*}{\partial s_j}\frac{\partial f^*}{\partial s_k}\right]$$

$$\left(\frac{\partial^2 \mathcal{L}}{\partial \mathbf{s}\partial m}\right)_j = \frac{\partial^2 \mathcal{L}}{\partial s_j \partial m} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial s_j \partial m}K + 2\varepsilon\frac{\partial f^*}{\partial s_j}\frac{\partial f^*}{\partial m}\right]$$

$$\left(\frac{\partial^2 \mathcal{L}}{\partial \mathbf{s}\partial \boldsymbol{\lambda}}\right)_{jk} = \frac{\partial^2 \mathcal{L}}{\partial s_j \partial \lambda_k} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial s_j \partial \lambda_k}K + 2\varepsilon\frac{\partial f^*}{\partial s_j}\frac{\partial f^*}{\partial \lambda_k}\right]$$

$$\left(\frac{\partial^2 \mathcal{L}}{\partial \mathbf{s}\partial \lambda_m}\right)_j = \frac{\partial^2 \mathcal{L}}{\partial s_j \partial \lambda_m} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial s_j \partial \lambda_m}K + 2\varepsilon\frac{\partial f^*}{\partial s_j}\frac{\partial f^*}{\partial \lambda_m}\right]$$

$$\frac{\partial^2 \mathcal{L}}{\partial m^2} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial m^2}K + 2\varepsilon\left(\frac{\partial f^*}{\partial m}\right)^2\right]$$

$$\left(\frac{\partial^2 \mathcal{L}}{\partial m\partial \boldsymbol{\lambda}}\right)_j = \frac{\partial^2 \mathcal{L}}{\partial m \partial \lambda_j} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial m \partial \lambda_j}K + 2\varepsilon\frac{\partial f^*}{\partial m}\frac{\partial f^*}{\partial \lambda_j}\right]$$

$$\frac{\partial^2 \mathcal{L}}{\partial m \partial \lambda_m} = \frac{T_{\max}}{c}\left[\frac{\partial^2 f^*}{\partial m \partial \lambda_m}K + 2\varepsilon\frac{\partial f^*}{\partial m}\frac{\partial f^*}{\partial \lambda_m}\right]$$

with

$$K = (1 - \varepsilon + 2\varepsilon f^*)$$

3. Application

This mathematical framework has been used in a recent work by Gonzalo et al.[6] to study the low thrust end-of-life disposal of Galileo satellites, using modified equinoctial element to formulate the dynamics.[7–10]

4. Conclusion

A mathematical framework for the formulation of OCPs using indirect methods and a generic formulation for the dynamics has been proposed. The mathematical structure of the equations derived from the first order optimality conditions has been leveraged to propose problem-independent formulas involving a set of problem-dependent matrices and their derivatives with respect to the state and mass. In particular, all the contributions of the costate elements associated to the state (excluding mass) have been reduced to a quadratic form of the costate and its derivatives. Relatively simple expression have also been derived for the variational equation needed to compute the STM, key to improve the numerical behavior of solvers for the TPBVP associated to the OCP. Expressions for the minimum-time, minimum-energy and minimum-fuel problems have been provided.

Acknowledgments

This work has been supported by the Spanish Ministry of Education, Culture and Sport through its FPU Program (reference number FPU13/05910), and by the Spanish Ministry of Economy and Competitiveness within the framework of the research project Dynamical Analysis, Advanced Orbital Propagation, and Simulation of Complex Space Systems (ESP2013-41634-P).

References

1) Lawden, D.F., *Optimal Trajectories for Space Navigation*, Butterworths, 1963.
2) Pontryagin, L.: *Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, 1962
3) Bryson, A. E. and Ho, Y. C.: *Applied Optimal Control. Optimization, Estimation and Control*, Taylor & Francis, 1975, ISBN:978-0-89116-228-5.
4) Zhang, C., Topputo, F., Bernelli-Zazzera, F., and Zhao, Y. S., Low-thrust minimum-fuel optimization in the circular restricted three-body problem, Journal of Guidance, Control, and Dynamics, 38(8) (2015),pp. 1501-1510.
5) Bertrand, R. and Epenoy, R.: New Smoothing Techniques for Solving BangBang Optimal Control Problems: Numerical Results and Statistical Interpretation, Optimal Control Applications and Methods, 23(4) (2002), pp. 171197, doi:10.1002/oca.709.
6) Gonzalo, J. L., Topputo, F. and Armellin, R.: Indirect Optimization of End-of-Life Disposal for Galileo Constellation Using Low Thrust Propulsion, 26th International Symposium on Space Flight Dynamics, Matsuyama, Japan, 3-9 June 2017

7) Cefola, P.J.: Equinoctial orbit elements-Application to artificial satellite orbits, AIAA/AAS Astrodynamics Conference, Palo Alto, California, USA, 11-12 Sep. 1972.
8) Broucke, R. A. and Cefola, P. J.: On the equinoctial orbit elements, Celestial Mechanics, 5(3) (1972),pp. 303-310.
9) Walker, M. J. H., Ireland, B., and Owens, J.: A Set of Modified Equinoctial Orbit Elements, Celestial Mechanics, 36(4) (1985), pp. 409419.
10) Walker, M. J. H.: A Set of Modified Equinoctial Orbit Elements, Celestial Mechanics and Dynamical Astronomy, 38(4) (1986), pp. 391-392 (Erratum).