Analytical and Numerical Methods for Optimal Control Problems on Manifolds and Lie Groups

by

Rohit Gupta

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in The University of Michigan
2016

Doctoral Committee:

Professor Anthony M. Bloch, Co-Chair
Professor Ilya V. Kolmanovsky, Co-Chair
Professor Dennis S. Bernstein
Asen L. Dontchev
Professor Ralf J. Spatzier

To my parents.

ACKNOWLEDGEMENTS

I would first like to thank my advisors, Professor Anthony M. Bloch and Professor Ilya V. Kolmanovsky for encouraging, guiding and supporting me throughout my doctoral studies and also giving me the opportunity for doing both applied and theoretical research. I would also like to thank my committee members, Professor Dennis S. Bernstein, D.Sc. Asen L. Dontchev and Professor Ralf J. Spatzier for their valuable comments and suggestions, which helped me in improving my dissertation. I am extremely thankful to Dr. Leonardo J. Colombo and Professor Tomoki Ohsawa, without whose help the last part of my dissertation would not have been possible. I would like to acknowledge my friends from the University of Michigan: Dae Young, Frantisek, Hyeongjun, Khaled, Robert, Shankar, Shardul, Sriram, Uroš, Zhaojian and others. Finally, I would like to thank my parents for their constant encouragement and support during the course of my doctoral studies.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

DP	Dynamic programming
LGVI	Lie group variational integrator
LQR	Linear quadratic regulator
MPC	Model predictive control
NMPC	Nonlinear model predictive control
NEOC	Neighboring extremal optimal control
OCP	Optimal control problem
OCPs	Optimal control problems
PMP	Pontryagin's maximum principle
TPBVP	Two-Point boundary value problem

LIST OF SYMBOLS

$\mathbb{R}^n$	Set of all n-tuples of real numbers
$\mathbb{R}^{n\times m}$	Set of all $n \times m$ matrices of real numbers
$\mathbb{R}_+$	Set of all positive real numbers
$\mathbb{Z}_+$	Set of all positive integers
$2^{\mathcal{X}}$	Power set of the set $\mathcal{X}$
diag	Diagonal matrix
$\mathrm{id}_{\mathcal{X}}$	Identity map on the space $\mathcal{X}$
int	Interior of a set
span	Linear span of a set of matrices/vectors
tr	Trace of a matrix
$X \oplus Y$	Direct sum of the vector spaces X and Y/Whitney sum of the vector bundles X and Y
$L^\infty(\mathcal{I}, \mathcal{X})$	Space of all essentially bounded functions on the interval $\mathcal{I} \subseteq \mathbb{R}$ taking values in the space $\mathcal{X}$
$W^{1,\infty}(\mathcal{I}, \mathcal{X})$	Space of all functions on the interval $\mathcal{I} \subseteq \mathbb{R}$ taking values in the space $\mathcal{X}$ such that the function and its first order weak derivative lies in $L^\infty(\mathcal{I}, \mathcal{X})$ (also known as the Sobolev space $W^{1,\infty}$)
$AC(\mathcal{I}, \mathcal{X})$	Space of all absolutely continuous functions on the interval $\mathcal{I} \subseteq \mathbb{R}$ taking values in the space $\mathcal{X}$
$C^k(\mathcal{I}, \mathcal{X})$	Space of all k-times continuously differentiable functions on the interval $\mathcal{I} \subseteq \mathbb{R}$ taking values in the space $\mathcal{X}$
F_X	First partial derivative of a scalar valued function F with respect to the vector argument X
F_{XY}	Second partial derivative of F with respect to the vector arguments X and Y
$\mathbf{D}_X$	Fréchet/(Gâteaux or the directional) derivative with respect to the argument X
$X \succeq Y$	$X - Y$ is a positive semi-definite matrix
$X \succ Y$	$X - Y$ is a positive definite matrix
$\cdot_A$	Antisymmetric part of a matrix
$\cdot_{ij}$	(i, j) entry of a matrix
$\|.\|_2$	2-norm of a vector
$\|.\|_F$	Frobenius norm of a matrix
$I_{n\times m}$	$n \times m$ identity matrix

$0_{n\times m}$	$n \times m$ zero matrix
$\{u_{k+j}\}_{j=0}^{N-1}$	Control sequence over a prediction horizon of N-steps
$\{x_{k+j}\}_{j=0}^{N}$	State sequence over a prediction horizon of N-steps
$\mathrm{SO}(n)$	Set of all orthogonal matrices with determinant equal to 1, which form a matrix Lie group (also known as the $\frac{n(n-1)}{2}$-dimensional special orthogonal group)
$\mathrm{SE}(n)$	Set of all matrices of the form $\begin{bmatrix} R & x \\ 0 & 1 \end{bmatrix}$, where $R \in \mathrm{SO}(n)$ and $x \in \mathbb{R}^n$, which form a matrix Lie group (also known as the $\frac{n(n+1)}{2}$-dimensional special Euclidean group)
$\mathfrak{so}(n)$	Set of all antisymmetric matrices, which form the Lie algebra of $\mathrm{SO}(n)$
$\mathfrak{se}(n)$	Set of all matrices of the form $\begin{bmatrix} \Omega & v \\ 0 & 0 \end{bmatrix}$, where $\Omega \in \mathfrak{so}(n)$ and $v \in \mathbb{R}^n$, which form the Lie algebra of $\mathrm{SE}(n)$
$\mathfrak{so}(n)^*$	Dual space of $\mathfrak{so}(n)$
$\mathfrak{se}(n)^*$	Dual space of $\mathfrak{se}(n)$
$.^\times$	Lie algebra isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$
$.^\diamond$	Dual Lie algebra isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)^*$
$x \times y$	Cross product of the vectors x and y
S^n	n-sphere
exp	Exponential map in Lie theory/Riemannian geometry
log	Inverse of exp
∇	Levi-Civita connection
$\frac{D}{d.}$	Covariant derivative
$\Gamma_{..}^{.}$	Christoffel symbols of the connection ∇
$R(.,.).$	Curvature tensor of the connection ∇
$[.,.]$	Commutator/Lie bracket of vector fields
$\rtimes$	Semidirect product
$\cong$	Isomorphism

Let $\mathcal{Q}$ be a finite-dimensional smooth manifold

$T_q\mathcal{Q}$	Tangent space of $\mathcal{Q}$ at $q \in \mathcal{Q}$
$T\mathcal{Q}$	Tangent bundle of $\mathcal{Q}$
$T_{v_q}T\mathcal{Q}$	Tangent space of $T\mathcal{Q}$ at $v_q := (q, v) \in T\mathcal{Q}$
$T_q^*\mathcal{Q}$	Cotangent space of $\mathcal{Q}$ at $q \in \mathcal{Q}$
$T^*\mathcal{Q}$	Cotangent bundle of $\mathcal{Q}$
π	Natural projection map from $T\mathcal{Q}$ to $\mathcal{Q}$
$\langle .,. \rangle$	Riemannian metric on $\mathcal{Q}$
$\langle\langle .,. \rangle\rangle$	Riemannian (Sasaki) metric on $T\mathcal{Q}$
$.(.)$	Natural pairing between elements in $T_q\mathcal{Q}$ and $T_q^*\mathcal{Q}$

$.^\sharp$	Canonical isomorphism between $T\mathcal{Q}$ and $T^*\mathcal{Q}$ given by the Riemannian metric (also known as the musical isomorphism and maps elements from $T^*\mathcal{Q}$ to $T\mathcal{Q}$)
$\mathfrak{X}(\mathcal{Q})$	Real vector space of all smooth vector fields on $\mathcal{Q}$
$\mathfrak{X}^*(\mathcal{Q})$	Real vector space of all smooth covector fields on $\mathcal{Q}$

Let G be a finite-dimensional Lie group, with the identity element e

$\mathfrak{g}$	Lie algebra of G
$\mathfrak{g}^*$	Dual space of $\mathfrak{g}$
$\mathrm{Ad}_.$	Adjoint action of G on $\mathfrak{g}$
$\mathrm{Ad}^*_.$	Coadjoint action of G on $\mathfrak{g}^*$
$\mathrm{ad}_.$	Adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$
$\mathrm{ad}^*_.$	Coadjoint action of $\mathfrak{g}$ on $\mathfrak{g}^*$
$\kappa_\mathfrak{g}(.,.)$	Killing form on $\mathfrak{g}$
$[X, Y]$	Span of the set of all $[A, B]$, where $A \in X$ and $B \in Y$, with X and Y being subspaces of $\mathfrak{g}$
L_g	Left translation map on G
$T_h L_g$	Tangent lift of L_g
$T^*_h L_g$	Cotangent lift of L_g
τ	Retraction map, which maps elements from $\mathfrak{g}$ to G (see, e.g., [12], [53], [54], [55])
$\mathbf{d}\,\tau^{-1}$	Right trivialized tangent of τ^{-1} (see, e.g., [12], [53], [54], [55])

CHAPTER I

Introduction

1.1 Motivation and Literature Review

This dissertation addresses four different topics and the motivation and literature review for each of the topics are given below.

1.1.1 Combined Homotopy and Neighboring Extremal Optimal Control

For most OCPs in engineering applications, it is difficult to obtain analytical or closed form solutions using Pontryagin's maximum principle (PMP) or dynamic programming (DP). Consequently, iterative/numerical methods are utilized for solving such OCPs [8], [78]. Two methods, which have been used independently in optimal control theory are homotopy (see, e.g., [10], [18], [51], [79], [91], [100]) and neighboring extremal optimal control (NEOC) (see, e.g., [14]). However, the combination of these two techniques has not been investigated. With this motivation, we combine these two techniques and arrive at a method for obtaining sub-optimal control in OCPs defined on a Euclidean space.

The method exploits the idea of homotopy (see, e.g., [6]) to continuously deform the trajectory from that of a linear system to that of a nonlinear system and it uses NEOC to predict the optimal solution as the homotopy parameter changes. Note that the method presented here is different from [42] as we, additionally, exploit the

idea of NEOC. The main motivation for the approach is that it is easier to solve OCPs for linear systems than for nonlinear systems. Once, we obtain the optimal control for the linear system, the control is iteratively updated using NEOC theory, combined with only a few iterations of a convergent optimizer at each step. We note that while the homotopy method is used in many practical trajectory optimization methods, e.g., in aerospace applications (see [37], [77]), its use is limited to systems with contractible state space, i.e., state space with a trivial fundamental group, such as $\mathbb{R}^n$.

1.1.2 Constrained Spacecraft Attitude Control on SO(3) Using Fast Nonlinear Model Predictive Control

Nonlinear model predictive control (NMPC) is a powerful technique for obtaining sub-optimal control in OCPs (see, e.g., [39]). However, in some cases, the system dynamics might not evolve on a Euclidean space but on a smooth manifold. For such OCPs the use of tools from differential geometry becomes advantageous (see, e.g., [3], [9], [16], [48]). The optimization problem arising in NMPC of spacecraft attitude, where the spacecraft attitude evolves on SO(3) was studied in [49], where it is shown that SO(3) based NMPC feedback laws can accomplish global spacecraft reorientation maneuvers and deal effectively with system nonlinearities and constraints. However, the numerical solution of the optimization problem in [49] is based on a direct method (input parameterization) and standard constrained optimizer in `MATLAB` (`fmincon.m`). This optimizer uses the numerical approximation of the derivatives and does not explicitly take advantage of the underlying Lie group structure. With this motivation, we develop a numerical solver for NMPC problem in [49] exploiting the geometric control formalism.

A nonlinear discrete-time spacecraft dynamics model based on a Lie group variational integrator (LGVI) is exploited in [49]. This model provides higher accuracy

in prediction and unlike continuous-time integrators, preserves the conserved quantities of motion (momentum and energy) to machine precision in absence of external moments (see [61]). As SO(3) is closed under multiplication, the LGVI updates the attitude by multiplying two matrices in SO(3) and hence ensures that the attitude always evolves on SO(3). For a detailed introduction to variational integrators see [69] and for the discrete-time rigid body equations see [73]. The numerical solver uses the solution of the necessary conditions for optimality in a discrete-time OCP defined over a prediction horizon, where the discrete-time dynamics are based on the LGVI model. The inequality constraints (which may represent thrust constraint, inclusion/exclusion zone constraints, etc.) are handled using a exterior penalty function approach. The indirect single shooting method is applied to the nonlinear root finding problem resulting from the necessary conditions for optimality. Our implementation also exploits sensitivity derivative expressions obtained from the necessary conditions for optimality. There is a growing interest in constrained spacecraft attitude control and in exploiting MPC and geometric control formalism to address these and related problems. In particular, MPC of spacecraft attitude based on linearized dynamics is studied in [40], [43], [87], [92]. NMPC problems on SO(3) are addressed in [38], where, however, neither spacecraft attitude control nor LGVI based models are considered. Related literature also includes publications on optimal control and motion planning on Lie groups. Constrained motion planning for multiple vehicles on SE(3) using barrier functions (rather than penalty functions) to handle constraints is considered in [81]. An optimal control technique for control systems evolving on noncompact Lie groups is developed in [82]. The necessary conditions for optimality for a related OCP are derived in [63], where, however, inequality constraints are not considered.

1.1.3 Neighboring Extremal Optimal Control for Mechanical Systems on Riemannian Manifolds

NEOC is well established for OCPs defined on a Euclidean space (see, e.g., [14]). However, the configuration space for most mechanical systems is not a Euclidean space but a smooth manifold. For instance, the configuration space of a spacecraft modeled as a rigid body is $\mathrm{SE}(3) = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$. With this motivation, we extend NEOC to OCPs for mechanical systems evolving on Riemannian manifolds.

1.1.4 Optimal Control Problems on Lie Groups with Symmetry Breaking Cost Functions

Reduction is an indispensable tool in the study of Lagrangian/Hamiltonian systems (which include OCPs), as it allows the dynamics associated with the Lagrangian/ Hamiltonian to be described on a quotient space, e.g., in the case of a Lie group G, the dynamics associated with a G-invariant Lagrangian/Hamiltonian can be described on $\mathfrak{g}/\mathfrak{g}^*$ instead of TG/T^*G.

Consider a G-invariant Lagrangian $L : TG \to \mathbb{R}$, then the dynamics associated with this G-invariant Lagrangian can be described on $\mathfrak{g}$, given by the following Euler-Poincaré equations

$$\frac{d}{dt}\,\mathbf{D}_\xi\,\ell = \mathrm{ad}^*_\xi\,\mathbf{D}_\xi\,\ell,$$

where $\ell : \mathfrak{g} \to \mathbb{R}$ is the reduced Lagrangian and $\ell(\xi) = L(e, \xi)$. For more details see [9], [45], [68]. Similarly, for a G-invariant Hamiltonian $H : T^*G \to \mathbb{R}$, the dynamics associated with this G-invariant Hamiltonian can be described on $\mathfrak{g}^*$, given by the following Lie-Poisson equations

$$\dot{\mu} = \mathrm{ad}^*_{\mathbf{D}_\mu\,h}\,\mu,$$

where $h : \mathfrak{g}^* \to \mathbb{R}$ is the reduced Hamiltonian and $h(\mu) = H(e, \mu)$. For more details see [9], [45], [68].

Reduction is well established for OCPs on Lie groups (see [58]). However, reduction for OCPs on Lie groups with symmetry breaking cost functions has not been investigated much, with an exception of [11]. With this motivation, we investigate the reduction for OCPs on Lie groups with symmetry breaking cost functions.

From the Lagrangian point of view, we obtain the Euler-Poincaré equations and from the Hamiltonian point of view, we obtain the Lie-Poisson equations. We also study the relationship between both formalisms. The theory of reduction for OCPs on Lie groups from a Hamiltonian point of view has been developed in [58]. The general theory of reduction for OCPs from a Hamiltonian point of view has been developed in [75]. However, [58], [75] do not consider symmetry breaking cost functions. Note that the theory for semidirect product reduction for OCPs on Lie groups with symmetry breaking cost functions from a Hamiltonian point of view has been developed in [11]. A variational integrator for OCPs on Lie groups with symmetry breaking cost functions is also developed. Note that variational integrators for OCPs on Lie groups are also developed in [53], [54], [55] but do not consider symmetry breaking cost functions.

1.2 Mathematical Preliminaries

We will now briefly review some of the mathematical tools used in this dissertation but for the most part of this dissertation, we assume that the reader is familiar with the basics of smooth manifold theory, Riemannian geometry, Lie groups and Lie algebras. For an introduction to smooth manifold theory, we refer the unfamiliar reader to [60]. For an introduction to Riemannian geometry, we refer the unfamiliar reader to [29], [44], [72]. For an introduction to Lie groups, we refer the unfamiliar reader to [44], [52], [68], [85]. For an introduction to Lie algebras, we refer the

unfamiliar reader to [44], [46], [85].

Definition I.1 ([9])**.** An n-dimensional smooth manifold $\mathcal{Q}$ is a set of points together with a finite or countably infinite set of subsets $U_\alpha \subset \mathcal{Q}$ and one-to-one mappings $\phi_\alpha : U_\alpha \to V_\alpha \subseteq \mathbb{R}^n$ such that

(a) $\bigcup\limits_{\alpha\in\mathcal{A}} U_\alpha = \mathcal{Q}$,

(b) For each nonempty intersection $U_\alpha \bigcap U_\beta$, the set $\phi_\alpha(U_\alpha \bigcap U_\beta)$ is an open subset of $\mathbb{R}^n$ and the one-to-one and onto mapping $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \bigcap U_\beta) \to \phi_\alpha(U_\alpha \bigcap U_\beta)$ is smooth,

(c) The family $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in\mathcal{A}}$ is maximal with respect to conditions (a) and (b).

Definition I.2 ([9])**.** The tangent space $T_q\mathcal{Q}$ is the set of all tangent vectors of $\mathcal{Q}$ at $q \in \mathcal{Q}$.

Definition I.3 ([9])**.** The tangent bundle $T\mathcal{Q}$ is a smooth manifold, whose underlying set is the disjoint union of the tangent spaces of $\mathcal{Q}$ at all points of $\mathcal{Q}$, i.e., $T\mathcal{Q} = \coprod\limits_{q\in\mathcal{Q}} T_q\mathcal{Q}$.

Definition I.4 ([60])**.** A Riemannian metric on $\mathcal{Q}$ is a smooth symmetric covariant 2-tensor field on $\mathcal{Q}$ that is positive definite at each point.

Definition I.5 ([60])**.** A Riemannian manifold is a pair $(\mathcal{Q}, \langle ., . \rangle)$, where $\mathcal{Q}$ is a smooth manifold and $\langle ., . \rangle$ is a Riemannian metric on $\mathcal{Q}$.

Definition I.6 ([29])**.** An affine connection ∇ on a smooth manifold $\mathcal{Q}$ is a mapping $\nabla : \mathfrak{X}(\mathcal{Q}) \times \mathfrak{X}(\mathcal{Q}) \to \mathfrak{X}(\mathcal{Q})$, which is denoted by $(X, Y) \overset{\nabla}{\to} \nabla_X Y$ and which satisfies the following properties

(a) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$,

(b) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$,

(c) $\nabla_X(fY) = f\nabla_X Y + X(f)Y$,

where X, Y, $Z \in \mathfrak{X}(\mathcal{Q})$, f and g are functions of class C^∞ defined on $\mathcal{Q}$.

Definition I.7 ([29])**.** The curvature R of a Riemannian manifold $\mathcal{Q}$ is a correspondence that associates to every pair X, $Y \in \mathfrak{X}(\mathcal{Q})$ a mapping $R(X,Y) : \mathfrak{X}(\mathcal{Q}) \to \mathfrak{X}(\mathcal{Q})$ given by

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z,$$

where $Z \in \mathfrak{X}(\mathcal{Q})$ and ∇ is the Riemannian connection of $\mathcal{Q}$.

Definition I.8 ([9])**.** A Lie group G is a smooth manifold that is a group and for which the group operations of multiplication $(g,h) \mapsto gh$, for g, $h \in G$ and inversion $g \mapsto g^{-1}$ are smooth.

Definition I.9 ([46])**.** A vector space $\mathfrak{g}$ over a field $\mathbb{F}$, with an operation $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, denoted $(X,Y) \mapsto [X,Y]$ and called the bracket or the commutator of X and Y, is called a Lie algebra over $\mathbb{F}$ if the following axioms are satisfied

(a) The bracket operation is bilinear,

(b) $[X,X] = 0$, for all $X \in \mathfrak{g}$,

(c) $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$,

where X, Y, $Z \in \mathfrak{g}$.

Definition I.10 ([45])**.** The adjoint action of G on $\mathfrak{g}$ is given by

$$\mathrm{Ad}_g\, \xi = T_e(L_g \circ R_{g^{-1}})\xi,$$

for $\xi \in \mathfrak{g}$.

Remark I.11. The coadjoint action of G on $\mathfrak{g}^*$ is given by $\mathrm{Ad}^*_{g^{-1}}$ (see, e.g., [45], [68]).

Definition I.12 ([45])**.** The infinitesimal generator map

$$\xi_{\mathfrak{g}}\eta = \frac{d}{dt}(\mathrm{Ad}_{\exp(t\xi)}\,\eta)\bigg|_{t=0} = \mathrm{ad}_{\xi}\,\eta,$$

where ξ, $\eta \in \mathfrak{g}$, is called the adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$, even though it is not a group action.

Remark I.13. The coadjoint action of $\mathfrak{g}$ on $\mathfrak{g}^*$ is given by ad^* (see, e.g., [45], [68]).

1.3 Contributions

The main contributions of this dissertation are summarized below.

(a) We have developed a method for obtaining sub-optimal control in OCPs defined on a Euclidean space, that is based on the combined use of homotopy and NEOC.

(b) We have developed a numerical solver for NMPC of spacecraft attitude that exploits the underlying Lie group structure of SO(3) and the geometric control formalism. We have also extended the classical penalty convergence theorem to the setting of smooth manifolds and the classical exact penalization theorem to the setting of Riemannian manifolds.

(c) We have extended NEOC, which is well established for OCPs defined on a Euclidean space, to the setting of Riemannian manifolds.

(d) We have extended reduction for OCPs on Lie groups with symmetry breaking cost functions. We have also developed a variational integrator for OCPs on Lie groups with symmetry breaking cost functions.

1.4 Dissertation Outline

The dissertation is organized as follows.

(a) In Chapter II, we will describe a method for obtaining sub-optimal control in OCPs defined on a Euclidean space, that is based on the combined use of homotopy and NEOC. We also present an example along with simulation results.

(b) In Chapter III, we describe the implementation of a numerical solver for NMPC of spacecraft attitude that exploits the underlying Lie group structure of SO(3) and the geometric control formalism. The numerical solver is based on numerically solving the necessary conditions for optimality. The control input/state constraints are handled through the exterior penalty function approach. We also extend the classical penalty convergence theorem to the setting of smooth manifolds and the classical exact penalization theorem to the setting of Riemannian manifolds.

(c) In Chapter IV, we extend NEOC, which is well established for OCPs defined on a Euclidean space, to the setting of Riemannian manifolds. We further specialize the results to the case of Lie groups. We also present an example along with simulation results.

(d) In Chapter V, we investigate the reduction for OCPs on Lie groups with symmetry breaking cost functions. From the Lagrangian point of view, we obtain the Euler-Poincaré equations and from the Hamiltonian point of view, we obtain the Lie-Poisson equations. We also study the relationship between both formalisms and present several examples. A variational integrator for OCPs on Lie groups with symmetry breaking cost functions is also developed.

CHAPTER II

Combined Homotopy and Neighboring Extremal Optimal Control

This chapter presents a new approach to trajectory optimization for nonlinear systems. The method exploits a homotopy between a linear system and a nonlinear system and NEOC, in combination with few iterations of a convergent optimizer at each step, to iteratively update the trajectory as the homotopy parameter changes. To illustrate the proposed method, a numerical example of a three dimensional orbit transfer problem for a spacecraft is presented. We will now briefly discuss homotopy and NEOC. In what follows, we will suppress the explicit dependence of the state, costate and control trajectories on time unless otherwise necessary.

2.1 Homotopy

Homotopy is a topological concept (see, e.g., [41]), which can be used, typically in combination with another optimization method, to solve OCPs. The basic idea is to start out with a simpler problem, whose solution is easy to compute, and then gradually evolve the solution to the solution of the harder problem by changing the homotopy parameter. Consider an OCP, where the objective is to minimize a cost

functional given by

$$\min_{u(.)} J = K(x(T)) + \int_0^T L(x(t), u(t))dt \tag{2.1}$$

subject to

$$\dot{x}(t) = f(x(t), u(t)), \;\; x(0) = x_0, \tag{2.2}$$

where $x(.) \in AC([0, T], \mathbb{R}^n)$, $u(.) \in L^\infty([0, T], \mathbb{R}^m)$, $K : \mathbb{R}^n \to \mathbb{R}$, $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ satisfy appropriate differentiability assumptions. Suppose the OCP (2.1)-(2.2) is difficult to solve with the dynamic constraint given by the model $\dot{x}(t) = f(x(t), u(t))$ but is easier to solve with the dynamic constraint given by the model $\dot{x}(t) = g(x(t), u(t))$ (e.g., $g(x(t), u(t)) = Ax(t) + Bu(t) + d$), where $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ also satisfies appropriate differentiability assumptions. Then by creating a homotopy given by

$$\dot{x}(t) = \lambda f(x(t), u(t)) + (1 - \lambda)g(x(t), u(t)), \tag{2.3}$$

where $\lambda \in [0, 1]$ is the homotopy parameter and under appropriate assumptions, we can solve the original OCP (2.1)-(2.2) by changing λ from 0 to 1 and re-using the solution from the previous homotopy step as an initial guess for the solution at the next homotopy step. For the background on homotopy methods see [4], [42]. The survey paper [91] discusses continuation methods and their application to OCPs. For the use of homotopy method in OCPs see also [10], [18], [51], [79], [100].

2.2 Neighboring Extremal Optimal Control

Consider a parameter dependent OCP, where the objective is to minimize a cost functional given by

$$\min_{u(.)} J = K(x(T), p) + \int_0^T L(x(t), u(t), p)dt \tag{2.4}$$

subject to

$$\dot{x}(t) = f(x(t), u(t), p), \;\; x(0) = x_0, \tag{2.5}$$

where $x(.) \in AC([0, T], \mathbb{R}^n)$, $u(.) \in L^\infty([0, T], \mathbb{R}^m)$, $p \in \mathbb{R}^l$ is a parameter, $K : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$, $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^n$ are functions of class C^2. Let (x_p^*, u_p^*) be a solution for the OCP (2.4)-(2.5), where $u_p^*(t)$ denotes the optimal control, which satisfies the Lagrange multiplier rule in a normal form (see, e.g., [9]). Let Ψ_p^* be the solution corresponding to $(x, u) = (x_p^*, u_p^*)$ of the following costate equation

$$\dot{\Psi} = -H_x(x, u, \Psi, p), \;\; \Psi(T) = K_x(x(T), p),$$

where $\Psi(.) \in AC([0, T], \mathbb{R}^n)$, H is the Hamiltonian and $H(x, u, \Psi, p) := L(x, u) + \Psi^T f(x, u, p)$. Altogether, (x_p^*, u_p^*, Ψ_p^*) satisfy the following necessary conditions for optimality

$$\dot{x}(t) = f(x(t), u(t), p), \;\; x(0) = x_0, \tag{2.6}$$

$$\dot{\Psi}(t) = -H_x(x(t), u(t), \Psi(t), p), \;\; \Psi(T) = K_x(x(T), p), \tag{2.7}$$

$$0 = H_u(x(t), u(t), \Psi(t), p). \tag{2.8}$$

Suppose there is a small variation in the initial condition and/or the parameter, and we would like to update the optimal control. Instead of solving the original OCP again, we employ a first order approximation of the necessary conditions for optimality around the nominal trajectory. This approximation is given by (see, e.g., [14], [30], [31], [32])

$$\delta\dot{x}(t) = \frac{\partial f}{\partial x}\delta x(t) + \frac{\partial f}{\partial u}\delta u(t) + \frac{\partial f}{\partial p}\delta p, \;\; \delta x(0) = \delta x_0, \tag{2.9}$$

$$\delta\dot{\Psi}(t) = -H_{xx}\delta x(t) - H_{xu}\delta u(t) - H_{x\Psi}\delta\Psi(t) - H_{xp}\delta p, \;\; \delta\Psi(T) = K_{xx}\delta x(T) + K_{xp}\delta p, \tag{2.10}$$

$$0 = H_{ux}\delta x(t) + H_{uu}\delta u(t) + H_{u\Psi}\delta\Psi(t) + H_{up}\delta p. \tag{2.11}$$

Under the the second order sufficient optimality condition (see, e.g., [30], [32]), (2.9)-(2.11) represents the optimality condition for the following OCP (see, e.g., [14], [30], [31], [32])

$$\begin{aligned}\min_{\delta u(.)} \delta^2 J = \frac{1}{2}\begin{bmatrix}\delta x(T)\\ \delta p\end{bmatrix}^T \begin{bmatrix}K_{xx}(T) & K_{xp}(T)\\ K_{px}(T) & 0\end{bmatrix}\begin{bmatrix}\delta x(T)\\ \delta p\end{bmatrix} + \\ \frac{1}{2}\int_0^T \left[\begin{bmatrix}\delta x(t)\\ \delta u(t)\\ \delta p\end{bmatrix}^T \begin{bmatrix}H_{xx}(t) & H_{xu}(t) & H_{xp}(t)\\ H_{ux}(t) & H_{uu}(t) & H_{up}(t)\\ H_{px}(t) & H_{pu}(t) & 0\end{bmatrix}\begin{bmatrix}\delta x(t)\\ \delta u(t)\\ \delta p\end{bmatrix}\right] dt\end{aligned} \tag{2.12}$$

subject to the perturbed dynamics

$$\delta\dot{x}(t) = \frac{\partial f}{\partial x}\delta x(t) + \frac{\partial f}{\partial u}\delta u(t) + \frac{\partial f}{\partial p}\delta p, \;\; \delta x(0) = \delta x_0, \tag{2.13}$$

where the matrices in the cost functional (2.12) and the Jacobian matrices in the dynamic constraint (2.13) are evaluated at the nominal trajectories. The optimal

control for the OCP (2.12)-(2.13) is given by

$$\delta u^*(t) = -H_{uu}^{-1}(t)\left[H_{ux}(t)\delta x(t) + f_u^T(t)\delta\Psi(t) + H_{up}(t)\delta p\right], \tag{2.14}$$

where all partial derivative matrices are evaluated at the nominal trajectories and $\delta\Psi(t)$ is a perturbation from $\Psi^*(t)$, ultimately expressible in terms of $\delta x(t)$ and δp.

The updated control is now calculated as the sum of $u^*(t)$ and $\delta u^*(t)$ and can be used directly or to warm start an optimizer for parameter $p+\delta p$. This is the basic idea behind NEOC. For a detailed description of NEOC see [14]. For a mathematically rigorous introduction to NEOC see [86].

Remark II.1. The OCP (2.12)-(2.13) is known as the accessory minimum problem in the calculus of variations (see, e.g., [90]). If there is no variation in the initial condition, i.e., the initial condition remains fixed, then $\delta x(0) = 0$ and similarly, if there is no variation in the parameter, i.e., the parameter remains fixed, then $\delta p = 0$. Note that it is also possible to go back to the conventional NEOC setting (see, e.g., [14]), by adding p as a state, with $\dot{p} = 0$.

For $(x_p^*(t), u_p^*(t))$ to be a strong local minimizer for the OCP (2.4)-(2.5), the second order sufficient condition (strengthened Legendre-Clebsch condition) requires that $H_{uu}(t) \succ 0$, for a.e. $t \in [0, T]$ and conjugate points for the OCP (2.12)-(2.13) must not exist (Jacobi condition) (see, e.g., [86]). An indicator for the existence of conjugate points is that the Riccati equation associated with the OCP (2.12)-(2.13) has a finite escape time (see, e.g., [86]). Existence of a solution of the Riccati equation associated with the OCP (2.12)-(2.13) over the interval $[0, T]$ is enough to rule out the existence of conjugate points. For a modern exposition on conjugate points see [3], [86]. For more on conjugate points for OCPs see [13], [14], [18], [20], [66], [71], [97], [98], [99].

We will now discuss the proposed method that combines the ideas of homotopy

and NEOC.

2.3 Method Description

Consider a linear system and a nonlinear system given below

$$\dot{x} = Ax + Bu + d, \;\; x(0) = x_0, \tag{2.15}$$

$$y = Cx, \tag{2.16}$$

$$\dot{x} = f(x, u), \;\; x(0) = x_0, \tag{2.17}$$

where $x(.) \in AC([0, T], \mathbb{R}^n)$, $u(.) \in L^\infty([0, T], \mathbb{R}^m)$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{q\times n}$, $d \in \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a function of class C^2. Create a homotopy between the linear system and the nonlinear system by

$$\dot{x} = \lambda f(x, u) + (1 - \lambda)(Ax + Bu + d) =: F(x, u, \lambda), \tag{2.18}$$

where $\lambda \in [0, 1]$. Note that the linear system (2.15) can be defined as the linearization of the nonlinear system (2.17) at a selected steady-state operating point (x_{op}, u_{op}), with $d = f(x_{op}, u_{op}) - Ax_{op} - Bu_{op}$. Consider a class of problems with a quadratic type cost defined over a finite horizon given by

$$J = \frac{1}{2}e^T(T)K_f e(T) + \frac{1}{2}\int_0^T [e^T(t)Qe(t) + u^T(t)Ru(t)]dt, \tag{2.19}$$

where K_f, $Q \succeq 0$, $R \succ 0$ and $e(t) = y(t) - y_d(t)$, with $y_d(t)$ being the desired trajectory.

Remark II.2. While we introduce our ideas in the context of a specific OCP with cost functional (2.19), many generalizations are possible. For instance, a minimum time problem can be handled using the given approach by rescaling time and introduc-

ing final time as an additional variable to be optimized. Note that for a minimum time problem, the optimal control is usually discontinuous (at least for control affine systems with a box constraint on u) and for the proposed approach to be used practically, the cost should be "regularized" with a small control-dependent term to make the optimal control continuous (see, e.g., [7], [88]). The case when the homotopy parameter enters the cost or the cost is not quadratic can be handled as well. However, simplifications do occur in the case of quadratic costs as is apparent from the next section.

2.3.1 Algorithm

The proposed algorithm is based on applying neighboring extremal updates to predict the optimal control trajectory as $p = \lambda$ changes. Note the superscripts in the following discussion represent the iteration number.

Step 1: Start with $k = 0$ and set $\lambda^{(0)} = 0$. Solve the OCP with the cost functional (2.19) subject to the dynamic constraint (2.18). The solution to this OCP is given by

$$u^{*(0)} = -R^{-1}B^T P x^{(0)} + R^{-1}B^T r_1, \tag{2.20}$$

where P and r_1 are the solutions of the differential equations

$$-\dot{P} = A^T P + PA - PBR^{-1}B^T P + C^T QC, \;\; P(T) = C^T K_f C, \tag{2.21}$$

$$-\dot{r}_1 = (A - BR^{-1}B^T P)^T r_1 - Pd + C^T Q y_d, \;\; r_1(T) = C^T K_f y_d(T). \tag{2.22}$$

Note that (2.21) is a Riccati differential equation that does not depend on y_d and is solved backwards in time and (2.22) is a linear differential equation which is also solved backwards in time. Obtain $x^*_{\lambda^{(0)}}$ from $\dot{x}^{(0)} = F(x^{(0)}, u^{*(0)}, \lambda^{(0)}) = Ax^{(0)} + Bu^{*(0)}$ and $u^*_{\lambda^{(0)}}$ from (2.20).

Step 2: Set $k = k + 1$ and $\lambda^{(k)} = \lambda^{(k-1)} + \delta\lambda^{(k)}$, where $\delta\lambda^{(k)} > 0$ is small and solve the OCP given below

$$\min_{\delta u^{(k)}(.)} \delta^2 J^{(k)} = \frac{1}{2} \begin{bmatrix} \delta x^{(k)}(T) \\ \delta\lambda^{(k)} \end{bmatrix}^T \begin{bmatrix} C^T K_f C & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta x^{(k)}(T) \\ \delta\lambda^{(k)} \end{bmatrix} +$$
$$\frac{1}{2} \int_0^T \left[\begin{bmatrix} \delta x^{(k)}(t) \\ \delta u^{(k)}(t) \\ \delta\lambda^{(k)} \end{bmatrix}^T \begin{bmatrix} H_{xx}^{(k)}(t) & H_{xu}^{(k)}(t) & H_{x\lambda}^{(k)}(t) \\ H_{ux}^{(k)}(t) & H_{uu}^{(k)}(t) & H_{u\lambda}^{(k)}(t) \\ H_{\lambda x}^{(k)}(t) & H_{\lambda u}^{(k)}(t) & 0 \end{bmatrix} \begin{bmatrix} \delta x^{(k)}(t) \\ \delta u^{(k)}(t) \\ \delta\lambda^{(k)} \end{bmatrix} \right] dt \tag{2.23}$$

subject to the perturbed dynamics

$$\delta\dot{x}^{(k)}(t) = A^{(k)}(t)\delta x^{(k)}(t) + B^{(k)}(t)\delta u^{(k)}(t) + G^{(k)}(t)\delta\lambda^{(k)}, \ \ \delta x^{(k)}(0) = 0, \tag{2.24}$$

where

$$\begin{aligned}
H_{xx}^{(k)}(t) &= \frac{\partial}{\partial x}\frac{\partial H}{\partial x}\Big|_{(x^*_{\lambda^{(k-1)}}(t), u^*_{\lambda^{(k-1)}}(t), \lambda^{(k-1)})}, \\
H_{xu}^{(k)}(t) &= \frac{\partial}{\partial u}\frac{\partial H}{\partial x}\Big|_{(x^*_{\lambda^{(k-1)}}(t), u^*_{\lambda^{(k-1)}}(t), \lambda^{(k-1)})}, \\
&\vdots \\
A^{(k)}(t) &= \frac{\partial F}{\partial x}\Big|_{(x^*_{\lambda^{(k-1)}}(t), u^*_{\lambda^{(k-1)}}(t), \lambda^{(k-1)})}, \\
B^{(k)}(t) &= \frac{\partial F}{\partial u}\Big|_{(x^*_{\lambda^{(k-1)}}(t), u^*_{\lambda^{(k-1)}}(t), \lambda^{(k-1)})}, \\
G^{(k)}(t) &= \frac{\partial F}{\partial \lambda}\Big|_{(x^*_{\lambda^{(k-1)}}(t), u^*_{\lambda^{(k-1)}}(t), \lambda^{(k-1)})},
\end{aligned}$$

with $H(x, u, \Psi, \lambda) := \frac{1}{2}[(Cx - y_d)^T Q(Cx - y_d) + u^T Ru] + \Psi^T F(x, u, \lambda)$. The solution

to the OCP (2.23)-(2.24) is given by (see, e.g., [14])

$$\delta u^{*(k)} = -H_{uu}^{-1(k)}(t)\left[H_{ux}^{(k)}(t)\delta x^{(k)} + B^{T(k)}(t)\delta\Psi^{(k)} + H_{u\lambda}^{(k)}(t)\delta\lambda^{(k)}\right], \tag{2.25}$$

where $\delta\Psi^{(k)} = S^{(k)}\delta x^{(k)} - r_2^{(k)}$, $S^{(k)}$ and $r_2^{(k)}$ are the solutions of the differential equations

$$-\dot{S}^{(k)} = \tilde{A}^{T(k)}(t)S^{(k)} + S^{(k)}\tilde{A}^{(k)}(t) - S^{(k)}\tilde{B}^{(k)}(t)S^{(k)} + \tilde{C}^{(k)}(t),\ S^{(k)}(T) = C^T K_f C, \tag{2.26}$$

$$-\dot{r}_2^{(k)} = (\tilde{A}^{T(k)}(t) - S^{(k)}\tilde{B}^{(k)}(t))r_2^{(k)} - (S^{(k)}\tilde{D}_1^{(k)}(t) + \tilde{D}_2^{(k)}(t))\delta\lambda^{(k)},\ r_2^{(k)}(T) = 0, \tag{2.27}$$

where

$$\begin{aligned}
\tilde{A}^{(k)}(t) &= A^{(k)}(t) - B^{(k)}(t)H_{uu}^{-1(k)}(t)H_{ux}^{(k)}(t),\\
\tilde{B}^{(k)}(t) &= B^{(k)}(t)H_{uu}^{-1(k)}(t)B^{T(k)}(t),\\
\tilde{C}^{(k)}(t) &= H_{xx}^{(k)}(t) - H_{xu}^{(k)}(t)H_{uu}^{-1(k)}(t)H_{ux}^{(k)}(t),\\
\tilde{D}_1^{(k)}(t) &= G^{(k)}(t) - B^{(k)}(t)H_{uu}^{-1(k)}(t)H_{u\lambda}^{(k)}(t),\\
\tilde{D}_2^{(k)}(t) &= H_{x\lambda}^{(k)}(t) - H_{xu}^{(k)}(t)H_{uu}^{-1(k)}(t)H_{u\lambda}^{(k)}(t).
\end{aligned}$$

Obtain $\delta x^*_{\delta\lambda^{(k)}}$ from (2.24), $\delta u^*_{\delta\lambda^{(k)}}$ from (2.25) and $\delta\Psi^*_{\delta\lambda^{(k)}} = S^{(k)}\delta x^*_{\delta\lambda^{(k)}} - r_2^{(k)}$. Calculate $x^*_{\lambda^{(k)}} = x^*_{\lambda^{(k-1)}} + \delta x^*_{\delta\lambda^{(k)}}$, $u^*_{\lambda^{(k)}} = u^*_{\lambda^{(k-1)}} + \delta u^*_{\delta\lambda^{(k)}}$ and $\Psi^*_{\lambda^{(k)}} = \Psi^*_{\lambda^{(k-1)}} + \delta\Psi^*_{\delta\lambda^{(k)}}$.

Step 3: Repeat **Step 2** until $\lambda^{(k)} = 1$.

Following the above steps, we can obtain a sub-optimal control for a nonlinear system with a given cost functional. Note that special methods exist for solving the differential equations (2.26)-(2.27) efficiently (see, e.g., [26], [95]). We consider a numerical example in the next section.

Remark II.3. We would like to clarify that by a sub-optimal control, we mean that we

are close enough to the optimal control, where the closeness of sub-optimal control to the optimal control can be controlled by controlling the rate of change of the homotopy parameter (derivation of the estimates for such an error bound is left to future work). The proposed algorithm can also be extended (under appropriate assumptions see, e.g., [30], [31], [32]) to OCPs with control input/state constraints. An alternative way to extend the proposed algorithm to OCPs with control input/state constraints is by using the penalty function approach. Moreover, the weighting factor multiplying the penalty function could be treated as an additional parameter in applying neighboring extremal predictions, so as to avoid the problem of ill-conditioning caused by starting directly with a very high value of the weighting factor.

Recall that an indicator for the existence of conjugate points is that (2.26) has a finite escape time. We will now give three sufficient conditions for the nonexistence of conjugate points, if the optimal control is obtained at each iteration of the proposed algorithm.

Proposition II.4. *Assume that* $\begin{bmatrix} \tilde{C}^{(k-1)}(t) & \tilde{A}^{T(k-1)}(t) \\ \tilde{A}^{(k-1)}(t) & -\tilde{B}^{(k-1)}(t) \end{bmatrix} \geq \begin{bmatrix} \tilde{C}^{(k)}(t) & \tilde{A}^{T(k)}(t) \\ \tilde{A}^{(k)}(t) & -\tilde{B}^{(k)}(t) \end{bmatrix}$, $H_{uu}^{(k-1)}(t) \geq 0$ *and* $H_{uu}^{(k)}(t) \geq 0$, *for a.e.* $t \in [0,T]$ *and for* $k \in \mathbb{Z}_+$, *then* $S^{(k-1)}(t) \geq S^{(k)}(t)$ *on the interval* $[0,T]$. *Moreover, if there exists a solution* $S^{(k-1)}(t)$ *for (2.26) on the interval* $[0,T]$, *then there exists a solution* $S^{(k)}(t)$ *for (2.26) on the interval* $[0,T]$.

Proof. It is easy to verify that $\tilde{A}^{(k-1)}(t)$, $\tilde{A}^{(k)}(t)$, $\tilde{B}^{(k-1)}(t)$, $\tilde{B}^{(k)}(t)$, $\tilde{C}^{(k-1)}(t)$ and $\tilde{C}^{(k)}(t)$ are integrable on the interval $[0,T]$. It follows from Theorem 4.1.4 of [1] that $S^{(k-1)}(t) \geq S^{(k)}(t)$ on the interval $[0,T]$. It is also easy to verify that $\tilde{B}^{(k-1)}(t) = \tilde{B}^{T(k-1)}(t) \geq 0$, $\tilde{B}^{(k)}(t) \geq 0$, $\tilde{C}^{(k-1)}(t) = \tilde{C}^{T(k-1)}(t)$ and $S^{(k-1)}(t) = S^{T(k-1)}(t)$ on the interval $[0,T]$. It follows from Theorem 5.7 of [35] that there exists a solution $S^{(k)}(t)$ for (2.26) on the interval $[0,T]$. □

Proposition II.5. *Assume that* $\tilde{C}^{(k-1)}(t) \succeq 0$ *and* $H_{uu}^{(k-1)}(t) \succeq 0$, *for a.e.* $t \in [0,T]$ *and for* $k \in \mathbb{Z}_+$, *then there exists a solution* $S^{(k-1)}(t)$ *for (2.26) on the interval* $[0,T]$.

Proof. It is easy to verify that $\tilde{A}^{(k-1)}(t)$, $\tilde{B}^{(k-1)}(t)$ and $\tilde{C}^{(k-1)}(t)$ are integrable on the interval $[0,T]$. It is also easy to verify that $\tilde{B}^{(k-1)}(t) \succeq 0$ on the interval $[0,T]$. It follows from Theorem 4.1.6 of [1] that there exists a solution $S^{(k-1)}(t)$ for (2.26) on the interval $[0,T]$. □

Proposition II.6. *Assume that* $H_{uu}^{(k-1)}(t) \succeq 0$, *for a.e.* $t \in [0,T]$ *and for* $k \in \mathbb{Z}_+$. *In addition, assume that there exists* $\bar{S}^{(k-1)}(.) \in AC([0,T], \mathbb{R}^{n\times n})$ *on the interval* $[0,T]$ *such that*

$$0 \succeq \dot{\bar{S}}^{(k-1)} + \tilde{A}^{T(k-1)}(t)\bar{S}^{(k-1)} + \bar{S}^{(k-1)}\tilde{A}^{(k-1)}(t) - \bar{S}^{(k-1)}\tilde{B}^{(k-1)}(t)\bar{S}^{(k-1)} + \tilde{C}^{(k-1)}(t),$$

for a.e. $t \in [0,T]$ *and* $\bar{S}^{(k-1)}(T) \succeq C^T K_f C$, *then there exists a solution* $S^{(k-1)}(t)$ *for (2.26) on the interval* $[0,T]$ *and* $\bar{S}^{(k-1)}(t) \succeq S^{(k-1)}(t)$ *on the interval* $[0,T]$.

Proof. It is easy to verify that $\tilde{A}^{(k-1)}(t)$, $\tilde{B}^{(k-1)}(t)$ and $\tilde{C}^{(k-1)}(t)$ are integrable on the interval $[0,T]$. It is also easy to verify that $\tilde{B}^{(k-1)}(t) = \tilde{B}^{T(k-1)}(t) \succeq 0$, $\tilde{C}^{(k-1)}(t) = \tilde{C}^{T(k-1)}(t)$ and $S^{(k-1)}(t) = S^{T(k-1)}(t)$ on the interval $[0,T]$. It follows from Theorem 5.8 of [35] that there exists a solution $S^{(k-1)}(t)$ for (2.26) on the interval $[0,T]$ and $\bar{S}^{(k-1)}(t) \succeq S^{(k-1)}(t)$ on the interval $[0,T]$. □

Remark II.7. Note that the proposed algorithm only gives a prediction step and not a correction step. To improve the solution, a prediction step can be augmented by a correction step that can be implemented by a few iterations of a convergent optimizer.

It is possible to obtain a predictor-corrector algorithm in a slightly more general setting and in the spirit of [33], we will now outline the predictor-corrector algorithm.

2.3.2 General Algorithm

Consider the following OCP

$$\min_{u(.)} J = K(x(T)) + \int_0^T L(x(t), u(t))dt \tag{2.28}$$

subject to

$$\dot{x}(t) = F(x(t), u(t), \lambda), \;\; x(0) = x_0, \tag{2.29}$$

where $x(.) \in W^{1,\infty}([0,T], \mathbb{R}^n)$, $u(.) \in L^{\infty}([0,T], \mathcal{U})$, with $\mathcal{U} \subset \mathbb{R}^m$ (nonempty, closed and convex), $K : \mathbb{R}^n \to \mathbb{R}$, $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $F : \mathbb{R}^n \times \mathbb{R}^m \times [0,1] \to \mathbb{R}^n$ are functions of class C^2. Let $(x^*_\lambda, u^*_\lambda)$ be a solution for the OCP (2.28)-(2.29) and Ψ^*_λ be the solution corresponding to $(x,u) = (x^*_\lambda, u^*_\lambda)$ of the following costate equation

$$\dot{\Psi} = -H_x(x, u, \Psi, \lambda), \;\; \Psi(T) = K_x(x(T)),$$

where $\Psi(.) \in W^{1,\infty}([0,T], \mathbb{R}^n)$, H is the Hamiltonian and $H(x,u,\Psi,\lambda) := L(x,u) + \Psi^T F(x,u,\lambda)$. It follows by PMP (see, e.g., [59]) that the following condition holds

$$H_u^T(x^*(t), u^*(t), \Psi^*(t), \lambda)(v - u^*(t))dt \geqslant 0, \;\; \text{for all } v \in \mathcal{U} \text{ and for a.e. } t \in [0,T],$$

where

$$H_u(x^*(t), u^*(t), \Psi^*(t), \lambda) = \left.\frac{\partial H}{\partial u}\right|_{(x^*(t), u^*(t), \Psi^*(t), \lambda)}.$$

Altogether, $(x^*_\lambda, u^*_\lambda, \Psi^*_\lambda)$ satisfy the following necessary conditions for optimality

$$\dot{x}(t) - F(x(t), u(t), \lambda) = 0, \;\; x(0) = x_0, \tag{2.30}$$

$$\dot{\Psi}(t) + H_x(x(t), u(t), \Psi(t), \lambda) = 0, \;\; \Psi(T) = K_x(x(T)), \tag{2.31}$$

$$H_u^T(x(t), u(t), \Psi(t), \lambda)(v - u(t))dt \geqslant 0, \;\; \text{for all } v \in \mathcal{U} \text{ and for a.e. } t \in [0, T]. \tag{2.32}$$

Let the set $N_{\mathbf{U}}(u) := \{\tilde{v}(.) \in L^\infty([0,T], \mathcal{U}) \mid \tilde{v}(t) \in N_{\mathcal{U}}(u(t)), \;\; \text{for all } t \in [0,T]\}$, where the normal cone mapping (set-valued mapping) to the set $\mathcal{U}$ is given by

$$N_{\mathcal{U}}(u(t)) = \begin{cases} \{w \mid \langle w, \tilde{u} - u(t) \rangle \leqslant 0, \;\; \text{for all } \tilde{u} \in \mathcal{U}\}, & \text{for } u(t) \in \mathcal{U}, \\ \varnothing, & \text{for } u(t) \notin \mathcal{U}. \end{cases}$$

Note that there are some subtleties that we have glossed over in the above discussion and we refer the reader to [34] for more details. Let $\omega := (x, u, \Psi)$, $\Omega_1 := \{(x, u, \Psi) \mid (x, u, \Psi) \in W^{1,\infty} \times L^\infty \times W^{1,\infty}, \;\; x(0) = x_0, \;\; \Psi(T) = K_x(x(T)), \;\; u(.) \in L^\infty([0,T], \mathcal{U})\}$ and $\Omega_2 := L^\infty \times L^\infty \times L^\infty$. The necessary conditions for optimality (2.30)-(2.32) can now be re-written as a generalized equation as follows

$$\mathcal{F}(\omega, \lambda) + \mathcal{N}(\omega) \ni 0, \tag{2.33}$$

where $\mathcal{F} : \Omega_1 \times [0,1] \to \Omega_2$ and $\mathcal{N} : \Omega_1 \to 2^{\Omega_2}$, with

$$\mathcal{F}(\omega, \lambda) := \begin{bmatrix} \dot{x} - F(x, u, \lambda) \\ \dot{\Psi} + H_x(x, u, \Psi, \lambda) \\ H_u(x, u, \Psi, \lambda) \end{bmatrix} \;\; \text{and} \;\; \mathcal{N}(\omega) := \begin{bmatrix} 0 \\ 0 \\ N_{\mathbf{U}}(u) \end{bmatrix}.$$

We are now ready to outline the algorithm. Note the superscripts in the following discussion represent the iteration number.

Step 1: Start with $k = 0$ and set $\lambda^{(0)} = 0$. Solve the OCP with the cost functional (2.28) subject to the dynamic constraint (2.29). Obtain $\omega^*_{\lambda^{(0)}} = (x^*_{\lambda^{(0)}}, u^*_{\lambda^{(0)}}, \Psi^*_{\lambda^{(0)}})$.

Step 2: Set $k = k + 1$ and $\lambda^{(k)} = \lambda^{(k-1)} + \delta\lambda^{(k)}$, where $\delta\lambda^{(k)} > 0$ is small. The Euler

predictor and the Newton corrector steps consist of solving the following linearized generalized equations

$$\mathcal{F}(\omega^*_{\lambda^{(k-1)}}, \lambda^{(k-1)}) + \mathbf{D}_\omega \mathcal{F}(\omega_{\lambda^{(k-1)}}, \lambda^{(k-1)})(\bar{\omega} - \omega^*_{\lambda^{(k-1)}}) + \mathbf{D}_\lambda \mathcal{F}(\omega^*_{\lambda^{(k-1)}}, \lambda^{(k-1)})\delta\lambda^{(k)} +$$
$$\mathcal{N}(\bar{\omega}) \ni 0, \tag{2.34}$$
$$\mathcal{F}(\bar{\omega}, \lambda^{(k)}) + \mathbf{D}_\omega \mathcal{F}(\bar{\omega}, \lambda^{(k)})(\omega^*_{\lambda^{(k)}} - \bar{\omega}) + \mathcal{N}(\omega^*_{\lambda^{(k)}}) \ni 0. \tag{2.35}$$

Obtain $\omega^*_{\lambda^{(k)}} = (x^*_{\lambda^{(k)}}, u^*_{\lambda^{(k)}}, \Psi^*_{\lambda^{(k)}})$ from (2.34)-(2.35).

Step 3: Repeat **Step 2** until $\lambda^{(k)} = 1$.

Following the above steps, we can obtain a sub-optimal control for a nonlinear system with a given cost functional, where the space of control parameters is nonempty, closed and convex.

Remark II.8. For computational purposes, (2.34) and (2.35) would result in linear quadratic problems with control input constraints (see, e.g., [30], [31], [32]).

We will now present a numerical example.

2.4 Numerical Example

To illustrate our combined homotopy and NEOC method, we consider a three dimensional orbit transfer problem for a spacecraft from an initial circular orbit of radius R_i (km) to a final circular orbit of radius R_f (km) (see, e.g., [51]). The OCP is given below

$$\min_{u(.)} J = \frac{1}{2}(x(T) - x_d)^T K_f (x(T) - x_d) + \frac{1}{2}\int_0^{14000} u^T(t)u(t)dt \tag{2.36}$$

subject to

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \\ \dot{x}_5(t) \\ \dot{x}_6(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ x_1(t)x_4^2(t)\cos^2(x_5(t)) + x_1(t)x_6^2(t) - \dfrac{\mu}{x_1^2(t)} + u_1(t) \\ x_4(t) \\ -\dfrac{2x_2(t)x_4(t)}{x_1(t)} + 2x_4(t)x_6(t)\tan(x_5(t)) + \dfrac{u_2(t)}{x_1(t)\cos(x_5(t))} \\ x_6(t) \\ -\dfrac{2x_2(t)x_6(t)}{x_1(t)} - x_4^2(t)\sin(x_5(t))\cos(x_5(t)) + \dfrac{u_3(t)}{x_1(t)} \end{bmatrix}, \tag{2.37}$$

$$u^T(t)u(t) \leqslant 10^{-8}, \tag{2.38}$$

where

$$K_f = \mathrm{diag}(10^{-4}, 1, 1, 1, 1, 1),$$

$$x(0) = x_0 = \left[R_e + R_i \;\; 0 \;\; 0 \;\; \sqrt{\frac{\mu}{(R_e + R_i)^3}} \;\; 0 \;\; 0 \right]^T,$$

$$x_d = \left[R_e + R_f \;\; 0 \;\; \frac{17\pi}{4} \;\; \sqrt{\frac{\mu}{(R_e + R_f)^3 \cos^2\left(\frac{5\pi}{180}\right)}} \;\; \frac{5\pi}{180} \;\; 0 \right]^T.$$

In (2.37), $x_1 = r$ (km) (radius of orbit), $x_2 = \dot{r}$ (km/sec), $x_3 = \theta$ (rad) (azimuth angle), $x_4 = \dot{\theta}$ (rad/sec), $x_5 = \phi$ (rad) (elevation angle), $x_6 = \dot{\phi}$ (rad/sec), $u_1 = a_r$ (km/sec^2) (acceleration in the r direction), $u_2 = a_\theta$ (km/sec^2) (acceleration in the θ direction), $u_3 = a_\phi$ (km/sec^2) (acceleration in the ϕ direction), $R_e = 6378$ (km) (radius of earth) and $\mu = 398600.4$ (km^3/sec^2) (gravitational parameter).

We consider a linear system given by $\dot{x} = Ax + Bu + d$, $x(0) = x_0$, which is obtained by the linearization of (2.37) at a selected steady-state operating point $x_{op} = x_0$ and $u_{op} = [0 \;\; 0 \;\; 0]^T$. Instead of solving the OCP (2.36)-(2.38), we use the

penalty function approach and solve the OCP given below

$$\min_{u(.)} J = \frac{1}{2}(x(T) - x_d)^T K_f (x(T) - x_d) + \frac{1}{2}\int_0^{14000} [u^T(t)u(t) + \nu\Phi(h(u(t)))]dt \tag{2.39}$$

subject to

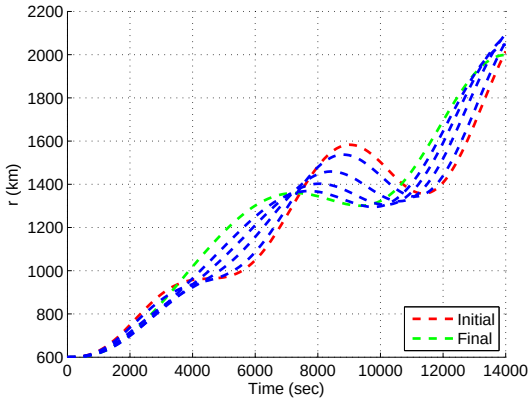
$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \\ \dot{x}_5(t) \\ \dot{x}_6(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ x_1(t)x_4^2(t)\cos^2(x_5(t)) + x_1(t)x_6^2(t) - \dfrac{\mu}{x_1^2(t)} + u_1(t) \\ x_4(t) \\ -\dfrac{2x_2(t)x_4(t)}{x_1(t)} + 2x_4(t)x_6(t)\tan(x_5(t)) + \dfrac{u_2(t)}{x_1(t)\cos(x_5(t))} \\ x_6(t) \\ -\dfrac{2x_2(t)x_6(t)}{x_1(t)} - x_4^2(t)\sin(x_5(t))\cos(x_5(t)) + \dfrac{u_3(t)}{x_1(t)} \end{bmatrix}, \tag{2.40}$$

where $h(u) = u^T u - 10^{-8}$, $(\Phi \circ h)(.) = \max\{0, h(.)\}^4$ is by choice a differentiable penalty function and $\nu \in \mathbb{R}_+$ is the weighting factor. We create a homotopy between the nonlinear system and the linear system and use the indirect single shooting method as a solver for the OCP with the cost functional (2.39) at each homotopy iteration. The indirect single shooting method converts the OCP into a root finding problem and solves for the initial values of the costate variables.

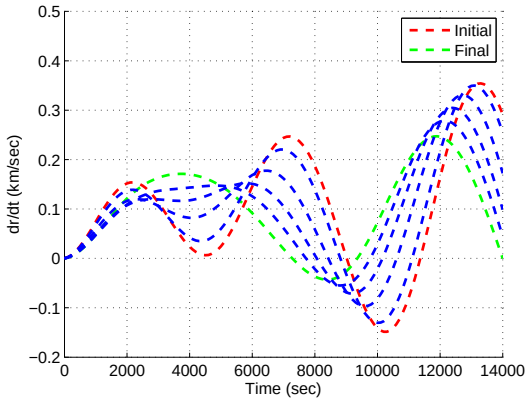
To demonstrate the advantages of the combined homotopy and NEOC method, two cases are considered. In the first case, we set the initial guess for the initial value of the costate variables for the next iteration to be equal to the optimal value of the costate variables obtained from the previous iteration. In the second case, we use the combined homotopy and NEOC method discussed in the previous section to set the initial guess for the initial value of the costate variables for the next iteration. Note that [51] uses (2.3) to solve OCPs but does not use neighboring extremal updates to

predict the change in the initial value of the costate variables. The Matlab function `fsolve.m` has been used to solve the root finding problem, the weighting factor is $\nu = 10^{30}$ and λ has been varied from 0 to 1 in increments of 0.1.

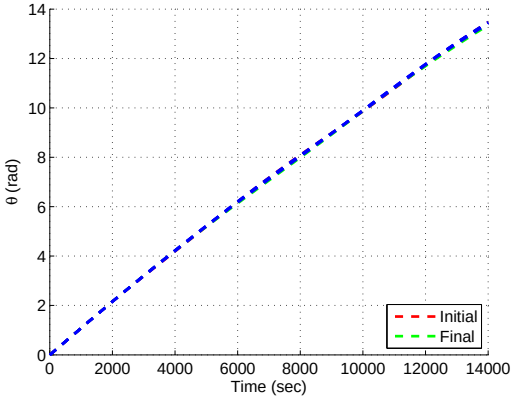
Figures 2.1(a)-(f) show the trajectory for the states of the nonlinear system, along with trajectories for some values of λ, with $R_i = 600$ (km) and $R_f = 2000$ (km). Figures 2.1(g)-(i) show the control inputs to the nonlinear system, along with trajectories for some values of λ. Figure 2.1(j) shows the control input constraint as λ varies from 0 to 1. Figure 2.1(k) shows the total cost for the nonlinear system as λ varies from 0 to 1. Figure 2.1(l) shows the spacecraft maneuver from an initial circular orbit of radius $R_i = 600$ (km) to a final circular orbit of radius $R_f = 2000$ (km). Figure 2.1(m) shows the total number of function evaluations of `fsolve.m` for different spacecraft maneuvers, for the two cases described above. Figure 2.1(n) shows the total number of iterations of `fsolve.m` for different spacecraft maneuvers, for the two cases described above. From Figures 2.1(m)-(n), one can see that the second case described above needs fewer function evaluations and iterations of `fsolve.m`.



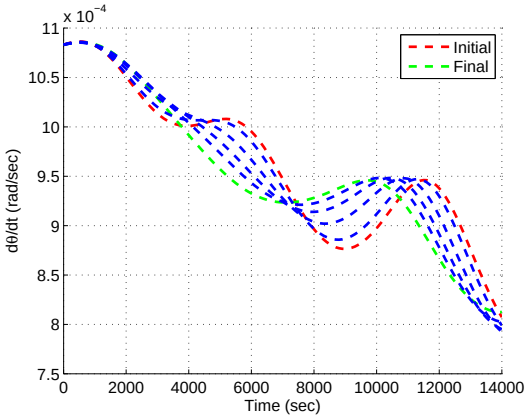
(a) Trajectory of r.



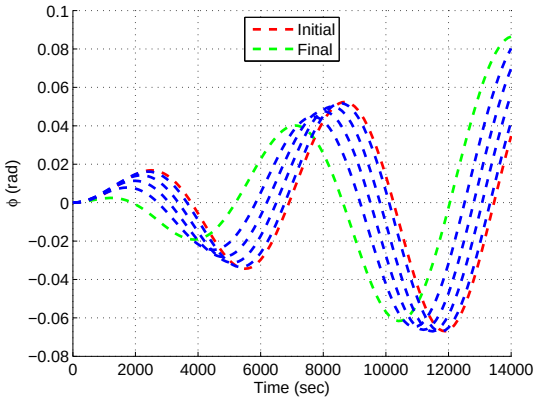
(b) Trajectory of $\dot{r}$.



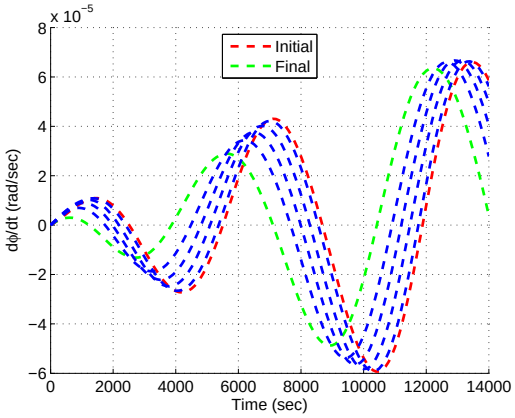
(c) Trajectory of θ.



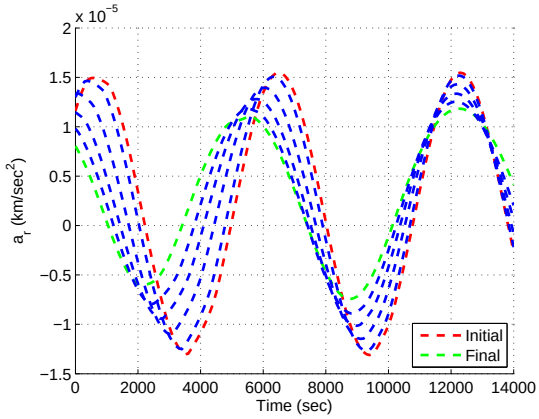
(d) Trajectory of $\dot{\theta}$.



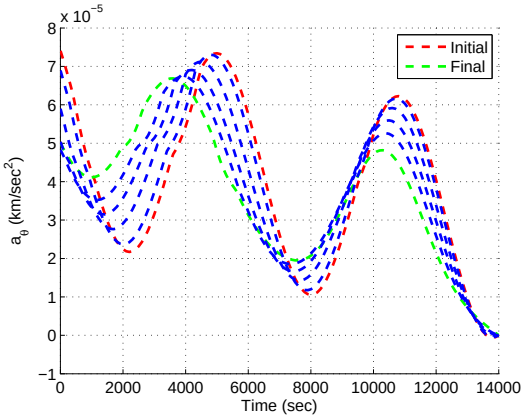
(e) Trajectory of ϕ.



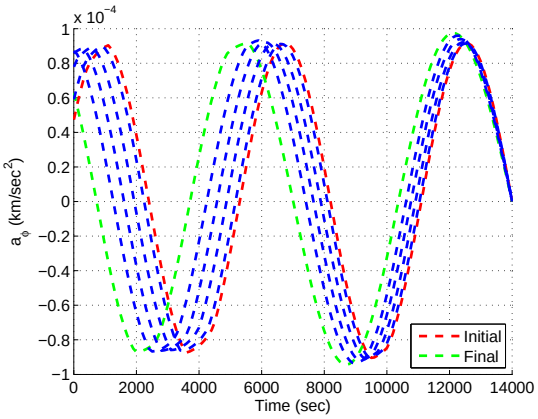
(f) Trajectory of $\dot{\phi}$.



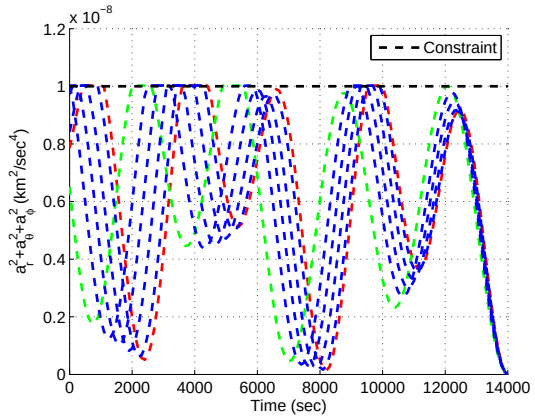
(g) Control Input a_r.



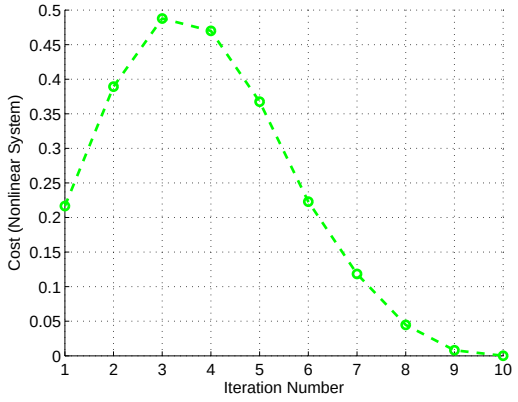
(h) Control Input a_θ.



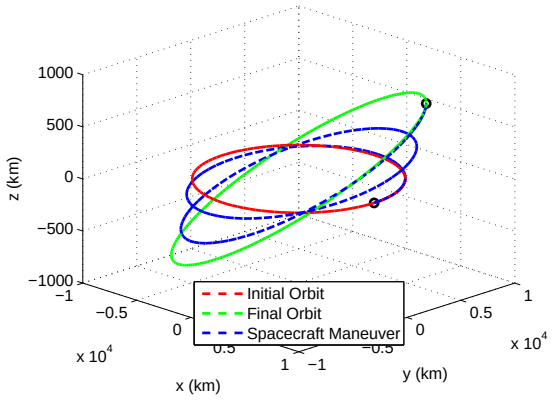
(i) Control Input a_ϕ.



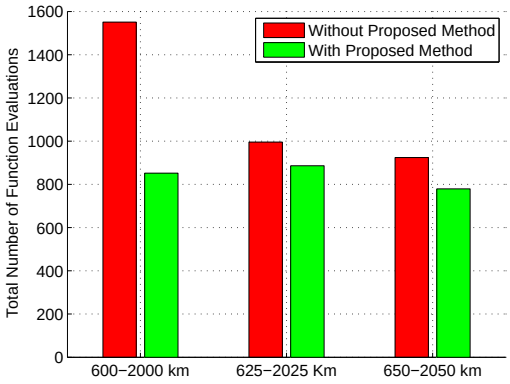
(j) Control Input Constraint.



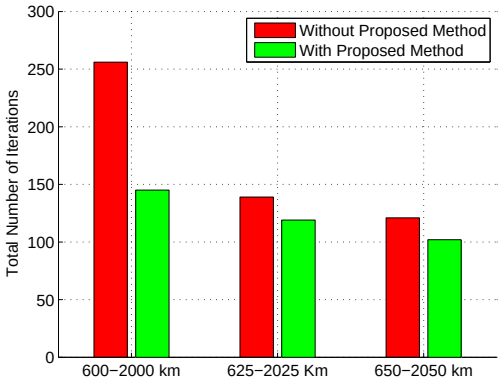
(k) Total Cost for the Nonlinear System.



(l) Spacecraft Maneuver.



(m) Total Number of Function Evaluations.



(n) Total Number of Iterations.

Figure 2.1: Results.

CHAPTER III

Constrained Spacecraft Attitude Control on SO(3) Using Fast Nonlinear Model Predictive Control

In this chapter, a numerical solver for the optimization problem arising in the NMPC of spacecraft attitude is developed and simulation results of its application to constrained spacecraft attitude control are presented. The numerical solver exploits the solution of the necessary conditions for optimality in a discrete-time OCP defined over a prediction horizon, where the discrete-time dynamics are based on the LGVI model. The inequality constraints (thrust constraint, inclusion/exclusion zone constraints, etc.) are handled using a exterior penalty function approach. Our developments exploit the geometric control formalism in deriving the numerical solver for the NMPC problem, which is based on the indirect single shooting method and is faster than the baseline solver (`fmincon.m`), which was used in [49]. In the last section of this chapter, we include some convergence results, which extend the classical penalty convergence theorem to the setting of smooth manifolds and the classical exact penalization theorem to the setting of Riemannian manifolds. We will now discuss the NMPC problem formulation on SO(3) that follows [49].

3.1 Nonlinear Model Predictive Control on SO(3)

Consider the following NMPC problem

$$\min_{\{u^{\times}_{k+j|k}\}_{j=0}^{N-1}} \mathcal{J}_d = K_d(R_{k+N|k}, \Pi^{\times}_{k+N|k}) + \sum_{j=0}^{N-1} C_d(R_{k+j|k}, \Pi^{\times}_{k+j|k}, u^{\times}_{k+j|k}) \tag{3.1}$$

subject to

$$h\Pi^{\times}_{k+j|k} = F_{k+j|k}J_d - J_dF^T_{k+j|k}, \tag{3.2}$$

$$R_{k+1+j|k} = R_{k+j|k}F_{k+j|k}, \tag{3.3}$$

$$\Pi_{k+1+j|k} = F^T_{k+j|k}\Pi_{k+j|k} + hu_{k+j|k}, \tag{3.4}$$

$$H_i(R_{k+j|k}, \Pi^{\times}_{k+j|k}, u^{\times}_{k+j|k}) \leqslant 0, \;\; i = 0, \ldots, m, \tag{3.5}$$

where $R_{k+j|k}$, $F_{k+j|k} \in \mathrm{SO}(3)$, $\Pi_{k+j|k}$, $u_{k+j|k} \in \mathbb{R}^3$ and $h \in \mathbb{R}_+$ is the time step. Note that $R_{k+j|k}$ is the spacecraft attitude, $\Pi_{k+j|k}$ is the spacecraft angular momentum and $u_{k+j|k}$ is the control torque. The terminal cost K_d is a real-valued non-negative function with respect to its arguments such that $K_d(I_{3\times3}, 0_{3\times3}) = 0$. The incremental cost C_d is a real-valued non-negative function with respect to its arguments $R_{k+j|k}$ and $\Pi^{\times}_{k+j|k}$ and a positive function with respect to its argument $u^{\times}_{k+j|k}$ such that $C_d(I_{3\times3}, 0_{3\times3}, 0_{3\times3}) = 0$. It is assumed that the terminal cost K_d, the incremental cost C_d and the inequality constraints H_i satisfy appropriate differentiability assumptions.

Note that $J_d \in \mathbb{R}^{3\times3}$ is the nonstandard moment of inertia matrix and is related to the standard moment of inertia matrix $J \in \mathbb{R}^{3\times3}$ as $J_d = \frac{1}{2}\,\mathrm{tr}(J)I_{3\times3} - J$. To obtain the necessary conditions for optimality, we follow the same variational approach as in [62]. Since, the numerical solver is based on solving the necessary conditions for optimality resulting from a discrete-time OCP over a prediction horizon, where the inequality constraints are incorporated as soft constraints through a penalty function,

we consider the following discrete-time OCP

$$\min_{\{u_k^\times\}_{k=0}^{N-1}} \mathcal{J}_d = K_d(R_N, \Pi_N^\times) + \sum_{k=0}^{N-1} C_d(R_k, \Pi_k^\times, u_k^\times) \tag{3.6}$$

subject to

$$h\Pi_k^\times = F_k J_d - J_d F_k^T, \tag{3.7}$$

$$R_{k+1} = R_k F_k, \tag{3.8}$$

$$\Pi_{k+1} = F_k^T \Pi_k + h u_k, \tag{3.9}$$

$$H_i(R_k, \Pi_k^\times, u_k^\times) \leqslant 0, \;\; i = 0, \ldots, m. \tag{3.10}$$

Define the augmented cost functional as follows

$$\mathcal{J}_d^a = K_d(R_N, \Pi_N^\times) + \sum_{k=0}^{N-1} C_d(R_k, \Pi_k^\times, u_k^\times) + \sum_{k=0}^{N-1} \lambda_k^1(\log(R_k^{-1} R_{k+1}) - \log(F_k)) + \sum_{k=0}^{N-1} \lambda_k^2((\Pi_{k+1} - F_k^T \Pi_k - h u_k)^\diamond) + \sum_{k=0}^{N-1} \sum_{i=0}^{m} \mu_i \Phi_i(H_i(R_k, \Pi_k^\times, u_k^\times)), \tag{3.11}$$

where $\lambda_k^1 \in \mathfrak{so}(3)^*$, $\lambda_k^2 \in \mathfrak{so}(3)$, $\Phi_i(.)$ is a penalty function and $\mu_i \in \mathbb{R}_+$. Note that the exponential map in the case of matrix Lie groups, coincides with the matrix exponential.

Remark III.1. Under the Lie algebra isomorphism $.^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$, given by $x^\times y = x \times y$, for all $x, y \in \mathbb{R}^3$, $\kappa_{\mathfrak{so}(3)}(.,.)$ gets identified with the standard inner product on $\mathbb{R}^3$ (see, e.g., [52]). Specifically, if $\kappa_{\mathfrak{so}(3)}(a^\times, b^\times) := \operatorname{tr}(\operatorname{ad} a^\times \circ \operatorname{ad} b^\times)$, then $\kappa_{\mathfrak{so}(3)}(a^\times, b^\times) = \operatorname{tr}(a^\times b^\times) = -2\langle a, b \rangle$. In fact, as SO(3) is compact and semisimple, $-\kappa_{\mathfrak{so}(3)}(.,.)$ gives a bi-invariant Riemannian metric on SO(3). Using the map $.^\diamond : \mathbb{R}^3 \to \mathfrak{so}(3)^*$ and letting the natural pairing $a^\diamond(b^\times) := \langle a, b \rangle$, it is easily seen that $a^\diamond(b^\times) = -\frac{1}{2}\kappa_{\mathfrak{so}(3)}(a^\times, b^\times) = \frac{1}{2}\operatorname{tr}((a^\times)^T b^\times)$, which also shows that $\mathfrak{so}(3) \cong \mathfrak{so}(3)^*$ (see, e.g., [45]). In this way, the natural pairing between a covector and a vector gets identified with the Killing form

on $\mathfrak{so}(3)$, which further gets identified with the standard inner product on $\mathbb{R}^3$. Using this, it is possible to obtain the necessary conditions for optimality in $\mathbb{R}^3$.

We will now derive the necessary conditions for optimality for the discrete-time OCP (3.6)-(3.10).

3.1.1 Necessary Conditions for Optimality

The variations of R_k, F_k and Π_k are given as follows

$$R_{k,\epsilon} = R_k \exp(\epsilon \eta_k^{\times}), \tag{3.12}$$

$$F_{k,\epsilon} = F_k \exp(\epsilon \xi_k^{\times}), \tag{3.13}$$

$$\Pi_{k,\epsilon} = \Pi_k + \epsilon \delta\Pi_k, \tag{3.14}$$

where η_k, $\xi_k \in \mathbb{R}^3$, with $\eta_0 = 0$, $\xi_0 = 0$ and $\delta\Pi_0 = 0$. The infinitesimal variations of R_k and F_k are given by

$$\begin{aligned} \delta R_k &= \left.\frac{dR_{k,\epsilon}}{d\epsilon}\right|_{\epsilon=0}, \\ &= R_k \eta_k^{\times}, \end{aligned} \tag{3.15}$$

$$\begin{aligned} \delta F_k &= \left.\frac{dF_{k,\epsilon}}{d\epsilon}\right|_{\epsilon=0}, \\ &= F_k \xi_k^{\times}. \end{aligned} \tag{3.16}$$

Before proceeding further, we need a few facts.

Fact 1. ([62]) $\eta_{k+1} = F_k^T \eta_k + \xi_k$.

The variation of (3.7) yields the following fact.

Fact 2. ([62]) $\xi_k = hF_k^T(\mathrm{tr}(F_k J_d) I_{3\times 3} - F_k J_d)^{-1} \delta\Pi_k =: M_k \delta\Pi_k$, where $M_k \in \mathbb{R}^{3\times 3}$.

Fact 3. ([45]) $\frac{1}{2}\mathrm{tr}(B^T a^{\times}) = \langle ((B)_A)^{-\times}, a \rangle$, for all $a \in \mathbb{R}^3$ and $B \in \mathbb{R}^{3\times 3}$.

The above result is used to obtain the following fact.

Fact 4. ([45]) $\mathbf{D}_{R_k}\mathcal{F}(R_k\eta_k^{\times}) = \langle((R_k^T(\mathbf{D}_{R_k}\mathcal{F}))_A)^{-\times}, \eta_k\rangle$.

Using Facts 1-4, the variation of the augmented cost functional is written as follows

$$
\begin{aligned}
\delta\mathcal{J}_d^a &= \langle((R_N^T(\mathbf{D}_{R_N}K_d))_A)^{-\times}, \eta_N\rangle + \langle((\mathbf{D}_{\Pi_N^{\times}}K_d)_A)^{-\times}, \delta\Pi_N\rangle + \\
&\sum_{k=0}^{N-1}[\langle((R_k^T(\mathbf{D}_{R_k}C_d))_A)^{-\times}, \eta_k\rangle + \langle((\mathbf{D}_{\Pi_k^{\times}}C_d)_A)^{-\times}, \delta\Pi_k\rangle + \langle((\mathbf{D}_{u_k^{\times}}C_d)_A)^{-\times}, \delta u_k\rangle] + \\
&\sum_{k=0}^{N-1}\langle\lambda_k^1, \eta_{k+1} - F_k^T\eta_k - \xi_k\rangle + \sum_{k=0}^{N-1}\langle\lambda_k^2, \delta\Pi_{k+1} - (F_k\xi_k^{\times})^T\Pi_k - F_k^T\delta\Pi_k - h\delta u_k\rangle + \\
&\sum_{k=0}^{N-1}\sum_{i=0}^{m}\mu_i[\langle((R_k^T(\mathbf{D}_{R_k}(\Phi_i\circ H_i)))_A)^{-\times}, \eta_k\rangle + \langle((\mathbf{D}_{\Pi_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta\Pi_k\rangle + \\
&\langle((\mathbf{D}_{u_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta u_k\rangle], \\
&= \langle((R_N^T(\mathbf{D}_{R_N}K_d))_A)^{-\times}, \eta_N\rangle + \langle((\mathbf{D}_{\Pi_N^{\times}}K_d)_A)^{-\times}, \delta\Pi_N\rangle + \\
&\sum_{k=0}^{N-1}[\langle((R_k^T(\mathbf{D}_{R_k}C_d))_A)^{-\times}, \eta_k\rangle + \langle((\mathbf{D}_{\Pi_k^{\times}}C_d)_A)^{-\times}, \delta\Pi_k\rangle + \langle((\mathbf{D}_{u_k^{\times}}C_d)_A)^{-\times}, \delta u_k\rangle] + \\
&\sum_{k=0}^{N-1}\langle\lambda_k^1, \eta_{k+1} - F_k^T\eta_k - \xi_k\rangle + \sum_{k=0}^{N-1}\langle\lambda_k^2, \delta\Pi_{k+1} + ((F_k^T\Pi_k)^{\times})^T\xi_k - F_k^T\delta\Pi_k - h\delta u_k\rangle + \\
&\sum_{k=0}^{N-1}\sum_{i=0}^{m}\mu_i[\langle((R_k^T(\mathbf{D}_{R_k}(\Phi_i\circ H_i)))_A)^{-\times}, \eta_k\rangle + \langle((\mathbf{D}_{\Pi_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta\Pi_k\rangle + \\
&\langle((\mathbf{D}_{u_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta u_k\rangle], \\
&= \langle((R_N^T(\mathbf{D}_{R_N}K_d))_A)^{-\times} + \lambda_{N-1}^1, \eta_N\rangle + \sum_{k=0}^{N-1}[\langle -F_k\lambda_k^1 + \lambda_{k-1}^1 + \\
&((R_k^T(\mathbf{D}_{R_k}C_d))_A)^{-\times} + \sum_{i=0}^{m}\mu_i((R_k^T(\mathbf{D}_{R_k}(\Phi_i\circ H_i)))_A)^{-\times}, \eta_k\rangle] + \langle((\mathbf{D}_{\Pi_N^{\times}}K_d)_A)^{-\times} + \\
&\lambda_{N-1}^2, \delta\Pi_N\rangle + \sum_{k=0}^{N-1}[\langle -M_k^T\lambda_k^1 - (F_k - M_k^T(F_k^T\Pi_k)^{\times})\lambda_k^2 + \lambda_{k-1}^2 + ((\mathbf{D}_{\Pi_k^{\times}}C_d)_A)^{-\times} + \\
&\sum_{i=0}^{m}\mu_i((\mathbf{D}_{\Pi_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta\Pi_k\rangle] + \sum_{k=0}^{N-1}[\langle -h\lambda_k^2 + ((\mathbf{D}_{u_k^{\times}}C_d)_A)^{-\times} + \\
&\sum_{i=0}^{m}\mu_i((\mathbf{D}_{u_k^{\times}}(\Phi_i\circ H_i))_A)^{-\times}, \delta u_k\rangle],
\end{aligned}
$$

where the analogue of integration by parts in the discrete-time setting is used along with the fact that the variations η_k and $\delta\Pi_k$ vanish at $k = 0$. Since, $\delta\mathcal{J}_d^a$ should vanish for all variations of η_k, $\delta\Pi_k$ and δu_k, the necessary conditions for optimality are as follows

$$h\Pi_k^{\times} = F_k J_d - J_d F_k^T, \tag{3.17}$$

$$R_{k+1} = R_k F_k, \tag{3.18}$$

$$\Pi_{k+1} = F_k^T \Pi_k + h u_k, \tag{3.19}$$

$$\lambda_{k+1}^1 = F_{k+1}^T[\lambda_k^1 + ((R_{k+1}^T(\mathbf{D}_{R_{k+1}} C_d))_A)^{-\times} + \sum_{i=0}^{m} \mu_i((R_{k+1}^T(\mathbf{D}_{R_{k+1}}(\Phi_i \circ H_i)))_A)^{-\times}], \tag{3.20}$$

$$\lambda_{N-1}^1 = -((R_N^T(\mathbf{D}_{R_N} K_d))_A)^{-\times}, \tag{3.21}$$

$$\lambda_{k+1}^2 = (F_{k+1} - M_{k+1}^T(F_{k+1}^T \Pi_{k+1})^{\times})^{-1}[-M_{k+1}^T \lambda_{k+1}^1 + \lambda_k^2 + ((\mathbf{D}_{\Pi_{k+1}^{\times}} C_d)_A)^{-\times} + \sum_{i=0}^{m} \mu_i((\mathbf{D}_{\Pi_{k+1}^{\times}}(\Phi_i \circ H_i))_A)^{-\times}], \tag{3.22}$$

$$\lambda_{N-1}^2 = -((\mathbf{D}_{\Pi_N^{\times}} K_d)_A)^{-\times}, \tag{3.23}$$

$$h\lambda_k^2 = ((\mathbf{D}_{u_k^{\times}} C_d)_A)^{-\times} + \sum_{i=0}^{m} \mu_i((\mathbf{D}_{u_k^{\times}}(\Phi_i \circ H_i))_A)^{-\times}. \tag{3.24}$$

Remark III.2. Note that we assume that the extremals for the OCP (3.6)-(3.10) are normal (see, e.g., [9]). However, abnormal extremals do occur in practical problems and there might exist abnormal extremals for the OCP (3.6)-(3.10) (a systematic study of the abnormal extremals for the OCP (3.6)-(3.10) is left to future work).

We will now describe the cost and the inequality constraints that are used for the subsequent numerical examples.

3.1.2 Cost and Inequality Constraints

We now consider the terminal cost, the incremental cost and the inequality constraints as given in [49]

$$K_d = \frac{1}{2}\|P_1^{\frac{1}{2}}(R_N - I_{3\times 3})\|_F^2 + \frac{1}{2}\|P_2^{\frac{1}{2}}\Pi_N^{\times}\|_F^2, \tag{3.25}$$

$$C_d = \frac{h}{2}\|Q_1^{\frac{1}{2}}(R_k - I_{3\times 3})\|_F^2 + \frac{h}{2}\|Q_2^{\frac{1}{2}}\Pi_k^{\times}\|_F^2 + \frac{h}{2}\|Q_3^{\frac{1}{2}}u_k^{\times}\|_F^2, \tag{3.26}$$

$$H_0 = \frac{1}{2}\|u_k^{\times}\|_F^2 - \alpha, \tag{3.27}$$

$$H_i = \beta_i - v_i^T R_k^T w_i, \;\; i = 1, \ldots, m, \tag{3.28}$$

where P_1, P_2, Q_1, $Q_2 \geq 0$ and $Q_3 > 0$. Note that (3.27) represents a thrust constraint, where $\alpha \in \mathbb{R}_+$, (3.28) represents inclusion/exclusion zone constraints, where $\beta_i \in \mathbb{R}$, v_i is the spacecraft body-fixed vector and w_i is the inertial direction vector (see, e.g., [94]). Note that $\frac{1}{2}\|B^{1/2}a^{\times}\|_F^2 = \frac{1}{2}a^T\tilde{B}a$, for all $a \in \mathbb{R}^3$ and $B \geq 0/B > 0$, where $\tilde{B} = \mathrm{tr}(B)I_{3\times 3} - B$ (see, e.g., [49]).

Remark III.3. Since, SO(3) is a matrix Lie group, a natural choice is to use the Frobenius norm to define a metric (see, e.g., [22]), which in turn is used to define the terminal and the incremental cost. The specific form of the terminal and the incremental cost in (3.25) and (3.26), respectively, corresponds to a LQR type problem on $\mathrm{SO}(3) \times \mathfrak{so}(3)$.

For the specific form of the terminal and the incremental cost in (3.25) and (3.26), respectively, the necessary conditions for optimality (3.17)-(3.24) take the form

$$h\Pi_k^{\times} = F_k J_d - J_d F_k^T, \tag{3.29}$$

$$R_{k+1} = R_k F_k, \tag{3.30}$$

$$\Pi_{k+1} = F_k^T \Pi_k + h u_k, \tag{3.31}$$

$$\lambda^1_{k+1} = F^T_{k+1}[\lambda^1_k - h((R^T_{k+1}Q_1)_A)^{-\times} - \sum_{i=1}^{m} \mu_i((R^T_{k+1}(\mathbf{D}_{H_i}(\Phi_i \circ H_i)w_i v_i^T))_A)^{-\times}], \tag{3.32}$$

$$\lambda^1_{N-1} = ((R^T_N P_1)_A)^{-\times}, \tag{3.33}$$

$$\lambda^2_{k+1} = (F_{k+1} - M^T_{k+1}(F^T_{k+1}\Pi_{k+1})^{\times})^{-1}[-M^T_{k+1}\lambda^1_{k+1} + \lambda^2_k + h((Q_2\Pi^{\times}_{k+1})_A)^{-\times}], \tag{3.34}$$

$$\lambda^2_{N-1} = -((P_2\Pi^{\times}_N)_A)^{-\times}, \tag{3.35}$$

$$h\lambda^2_k = h((Q_3 u^{\times}_k)_A)^{-\times} + \mu_0((\mathbf{D}_{H_0}(\Phi_0 \circ H_0)u^{\times}_k)_A)^{-\times}, \tag{3.36}$$

where we have chosen the differentiable penalty function, $(\Phi_i \circ H_i)(.) = \dfrac{h}{2}\max\{0, H_i(.)\}^2$.

Remark III.4. To obtain F_k in (3.29), (3.29) is expressed on $\mathbb{R}^3 \cong \mathfrak{so}(3)$ using the exponential map or the Cayley transform, to which a Newton method is applied (see [63], pp. 96–100). Also, if a certain condition is satisfied, then F_k in (3.29) can be obtained by solving a continuous-time algebraic Riccati equation (see [19]). The trajectories for $(R_k, \Pi_k, \lambda^1_k, \lambda^2_k)$ (starting from a given $(R_0, \Pi_0, \lambda^1_0, \lambda^2_0)$), using the necessary conditions for optimality are computed in the same way as in [62], p. 474.

We will now describe the numerical solver.

3.2 Description of the Numerical Solver

The necessary conditions for optimality (3.29)-(3.36) lead to a two-point boundary value problem which is solved using the indirect single shooting method to determine the initial values of the Lagrange multipliers. Sensitivity derivatives obtained from the necessary conditions for optimality are used in the numerical solution. We follow the same procedure as in [62] to characterize these sensitivity derivatives. The sensitivity derivatives for (3.30)-(3.31) are given as follows

$$\begin{bmatrix} \eta_{k+1} \\ \delta\Pi_{k+1} \end{bmatrix} = \begin{bmatrix} F^T_k & M_k \\ 0_{3\times 3} & F^T_k + (F^T_k\Pi_k)^{\times}M_k \end{bmatrix} \begin{bmatrix} \eta_k \\ \delta\Pi_k \end{bmatrix} + \begin{bmatrix} 0_{3\times 3} \\ hI_{3\times 3} \end{bmatrix} \delta u_k. \tag{3.37}$$

The sensitivity derivatives for (3.32) and (3.34) are given as follows

$$\begin{bmatrix} \delta\lambda^1_{k+1} \\ \delta\lambda^2_{k+1} \end{bmatrix} = S_k \begin{bmatrix} \eta_{k+1} \\ \delta\Pi_{k+1} \\ \delta u_k \\ \delta\lambda^1_k \\ \delta\lambda^2_k \end{bmatrix}, \tag{3.38}$$

where $S_k \in \mathbb{R}^{6\times 15}$. Assuming that δu_k is explicitly expressible in terms of u_k and $\delta\lambda^2_k$ (this assumption is not required but helps to present the idea clearly), from (3.37)-(3.38) we obtain

$$\begin{bmatrix} \eta_{k+1} \\ \delta\Pi_{k+1} \\ \delta\lambda^1_{k+1} \\ \delta\lambda^2_{k+1} \end{bmatrix} = T_k \begin{bmatrix} \eta_k \\ \delta\Pi_k \\ \delta\lambda^1_k \\ \delta\lambda^2_k \end{bmatrix}, \tag{3.39}$$

where $T_k \in \mathbb{R}^{12\times 12}$. From (3.39) we obtain

$$\begin{bmatrix} \eta_N \\ \delta\Pi_N \\ \delta\lambda^1_N \\ \delta\lambda^2_N \end{bmatrix} = \left(\prod_{k=0}^{N-1} T_k \right) \begin{bmatrix} \eta_0 \\ \delta\Pi_0 \\ \delta\lambda^1_0 \\ \delta\lambda^2_0 \end{bmatrix}. \tag{3.40}$$

In the indirect single shooting method, the initial values of the Lagrange multipliers are unknowns that are determined by solving a nonlinear root finding problem. To solve this nonlinear root finding problem, we employ a Newton-like method. The

updates have the following form

$$\lambda_0^{(i+1)} = \lambda_0^{(i)} - \gamma \left[\frac{\delta E^{(i)}}{\delta \lambda_0^{(i)}} \right]^{-1} E^{(i)}, \tag{3.41}$$

where the superscripts represent the iteration number, $\gamma \in (0, 1]$ is the step size and $E^{(i)}$ is given as follows

$$E^{(i)} = \begin{bmatrix} \lambda_{N-1}^{1(i)} - ((R_N^{T(i)} P_1)_A)^{-\times} \\ \lambda_{N-1}^{2(i)} + ((P_2 \Pi_N^{\times(i)})_A)^{-\times} \end{bmatrix}. \tag{3.42}$$

Note that $E^{(i)}$ represents the error in satisfaction of the terminal boundary conditions at the *i-th* iteration. The sensitivity derivative for $E^{(i)}$ is computed with the help of the following expression

$$\delta E^{(i)} = \begin{bmatrix} \delta\lambda_{N-1}^{1(i)} + ((\eta_N^{\times(i)} R_N^{T(i)} P_1)_A)^{-\times} \\ \delta\lambda_{N-1}^{2(i)} + ((P_2 \delta\Pi_N^{\times(i)})_A)^{-\times} \end{bmatrix}. \tag{3.43}$$

For a given $(R_0^{(i)}, \Pi_0^{(i)}, \lambda_0^{1(i)}, \lambda_0^{2(i)})$, the trajectories for $(R_k^{(i)}, \Pi_k^{(i)}, \lambda_k^{1(i)}, \lambda_k^{2(i)})$ are obtained using the necessary conditions for optimality (3.29)-(3.36) and $E^{(i)}$ is obtained using (3.42). Letting $\delta\lambda_0^{(i)} = \lambda_0^{(i)} - \lambda_0^{(i-1)}$, $\delta E^{(i)}$ is obtained using (3.44), which in turn is obtained using (3.40) and $(R_k^{(i)}, \Pi_k^{(i)}, \lambda_k^{1(i)}, \lambda_k^{2(i)})$ along with the facts that $\eta_0^{(i)} = 0$ and $\delta\Pi_0^{(i)} = 0$. In this way, we obtain the Jacobian matrix, $\left[\frac{\delta E^{(i)}}{\delta \lambda_0^{(i)}} \right]$. Once the optimal initial values of the Lagrange multipliers are obtained, the optimal trajectories are calculated using the necessary conditions for optimality (3.29)-(3.36) obtained in the previous section.

Remark III.5. Continuation methods (see, e.g., [4]) can be exploited to obtain additional time savings. There are two scenarios where continuation methods can be used. The first scenario occurs for the MPC problem over a fixed prediction horizon, when

the weighting factor multiplying the penalty function is being increased. Generally, starting with a very high value of the weighting factor is not recommended as this might result in numerical ill-conditioning. Continuation with respect to the weighting factor can be used to obtain a desired solution quickly and to avoid the problem of numerical ill-conditioning. The second scenario occurs when going from one prediction horizon to the next, wherein the initial state in the MPC problem changes. If the state changes by a small amount, this can be seen as a small perturbation. Continuation with respect to the state along with the solution computed in the previous prediction horizon can be used to predict a desired solution quickly. The idea of continuation presented here is similar to the one in [27], [28], [36], [76], [96]. While we do not formally take advantage of continuation methods in this chapter, our subsequent numerical examples warm-start the numerical solver with the previous solution.

We will now present numerical examples.

3.3 Numerical Examples

We consider a spacecraft with moment of inertia matrix $J = \text{diag}(1, 0.8, 0.8)$ (kg-m^2), with time step $h = 0.4$ (sec). We take $P_1 = P_2 = Q_1 = Q_2 = 0.01 I_{3\times 3}$ and $Q_3 = I_{3\times 3}$, in (3.25)-(3.26).

In some of the subsequent figures (Figures 3.4 and 3.9), the attitude maneuver is plotted on S^2, where the vectors $[x\ y\ z]^T$ corresponding to the first, second and third column of R_0 are plotted in dashed-red, dashed-green and dashed-blue, respectively. Similarly, the vectors $[x\ y\ z]^T$ corresponding to the first, second and third column of R_N are plotted in red, green and blue, respectively. For all other R_k, $k \neq 0$, N, only the coordinates are shown in the corresponding colors.

3.3.1 Simulation with Thrust Constraint (Case I)

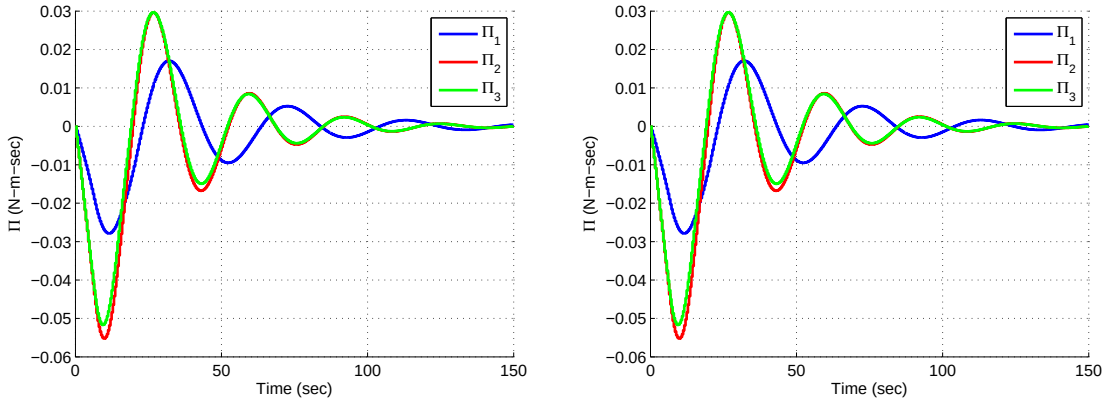
In this simulation we consider only the thrust constraint (3.27), with $\alpha = 10^{-4}$ (N-m). The simulation time is 150 (sec), the prediction horizon is 2 (sec), the weighting factor is $\mu_0 = 10^{10}$ and the step size is $\gamma = 1$. The initial condition for the attitude and angular momentum are given as follows

$$R_0 = \exp(\zeta^{\times}),$$
$$\Pi_0 = [0\ 0\ 0]^T,$$

where $\zeta = [0.25\ 0.5\ 0.5]^T$. Results for the numerical solver are shown in Figures 3.1(a)-3.4(a) and results for the baseline solver are shown in Figures 3.1(b)-3.4(b). It can be seen from Figures 3.1-3.4 that the solution obtained by the numerical solver is close enough to the solution obtained by the baseline solver.

3.3.2 Simulation with Thrust and Exclusion Zone Constraints (Case II)

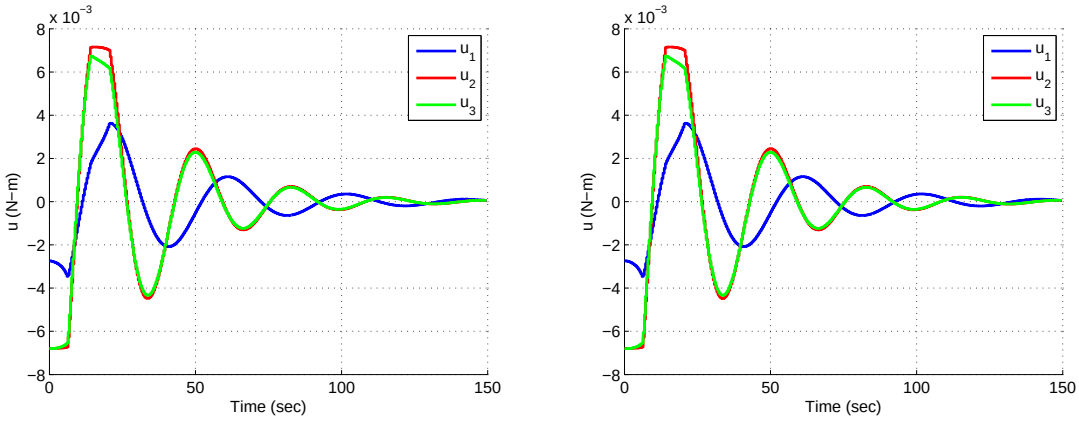
In this simulation we consider the thrust constraint and one exclusion zone constraint (3.27)-(3.28), with $\alpha = 10^{-4}$ (N-m), $\beta_1 = -0.9962$, $v_1 = [1\ 0\ 0]^T$ and



(a) Numerical Solver.

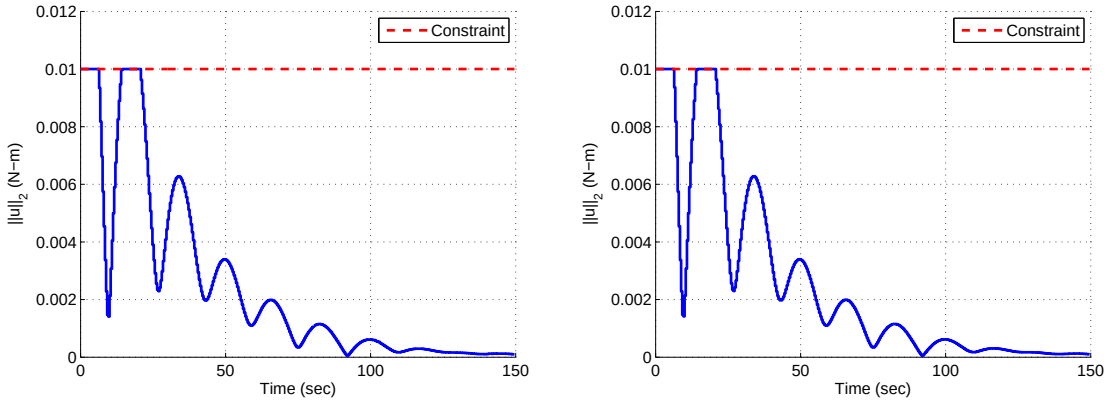
(b) Baseline Solver.

Figure 3.1: Angular Momentum.



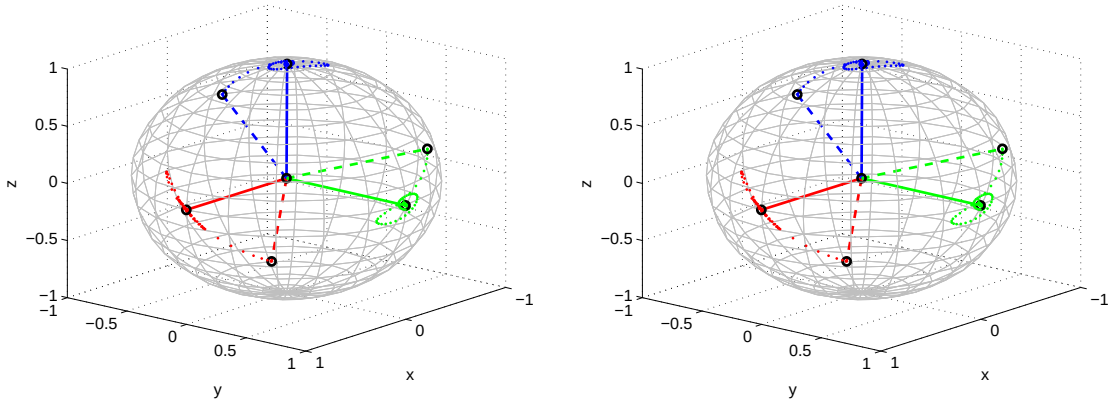
(a) Numerical Solver. (b) Baseline Solver.

Figure 3.2: Control Input.



(a) Numerical Solver. (b) Baseline Solver.

Figure 3.3: Thrust Constraint.



(a) Numerical Solver. (b) Baseline Solver.

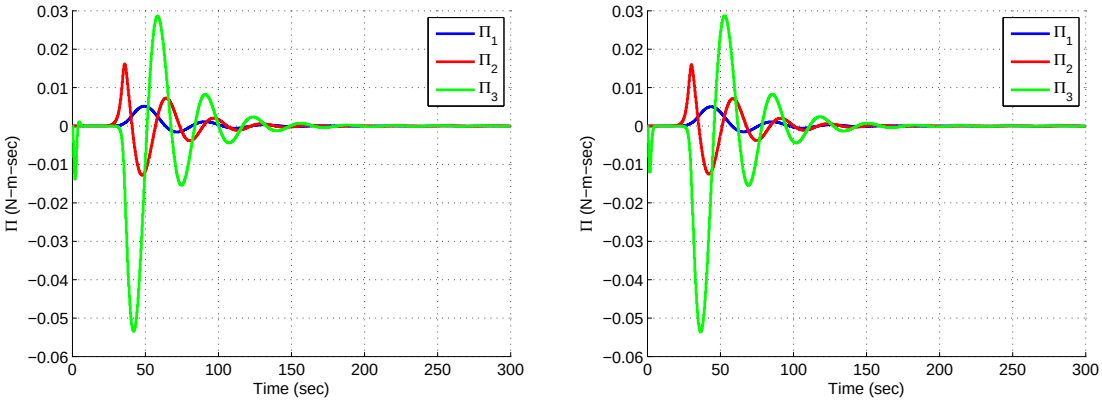
Figure 3.4: Attitude Maneuver.

$w_1 = -[0.9276\ 0.3736\ 0]^T$. The simulation time is 300 (sec), the prediction horizon is 2 (sec), the weighting factors are $\mu_0 = 10^{10}$, $\mu_1 = 10^3$ and a hybrid step size method is used, i.e., we switch from one value of γ to the other during the simulation. We use a hybrid step size method because we observed that $\gamma < 1$ is helpful to ensure convergence of (3.41) when the exclusion zone constraint is active whereas with $\gamma = 1$ (3.41) is convergent, when the exclusion zone constraint is not active. The initial condition for the attitude and angular momentum are given as follows

$$R_0 = \exp(\zeta^\times),$$
$$\Pi_0 = [0\ 0\ 0]^T,$$

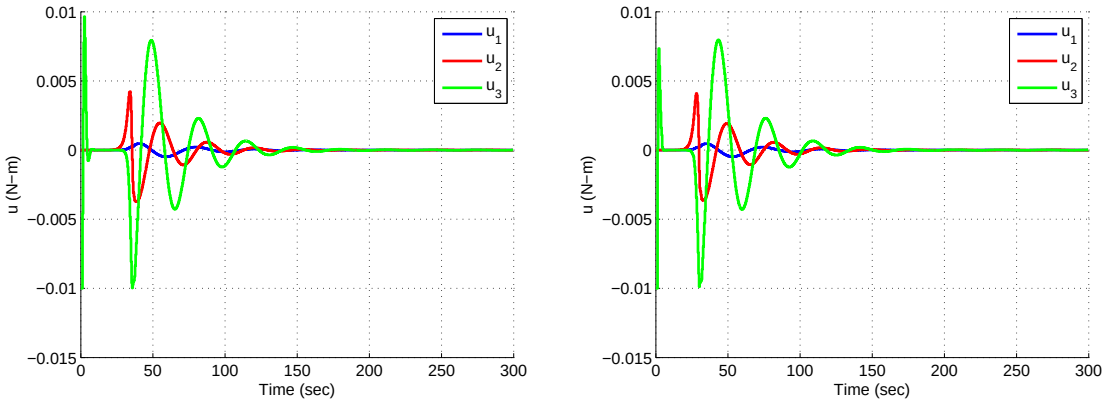
where $\zeta = [0\ 0\ 0.5]^T$. Results for the numerical solver are shown in Figures 3.5(a)-3.9(a) and results for the baseline solver are shown in Figures 3.5(b)-3.9(b). It can be seen from Figures 3.5-3.9 that the solution obtained by the numerical solver is close enough to the solution obtained by the baseline solver.

Table 3.1 compares the total computational time for the numerical solver and the baseline solver, to obtain solutions to the NMPC problems, on a 3.6 GHz Intel Xeon desktop computer with 16 GB of RAM. This comparison demonstrates the time



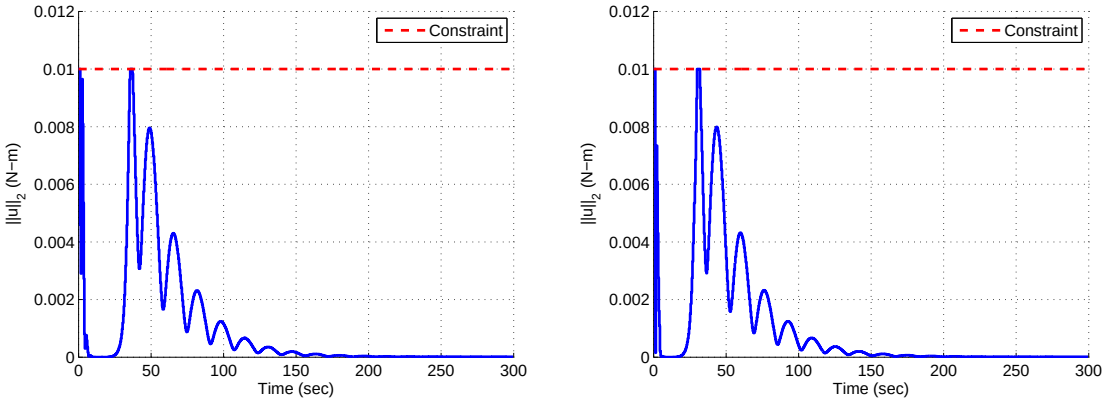
(a) Numerical Solver. (b) Baseline Solver.

Figure 3.5: Angular Momentum.



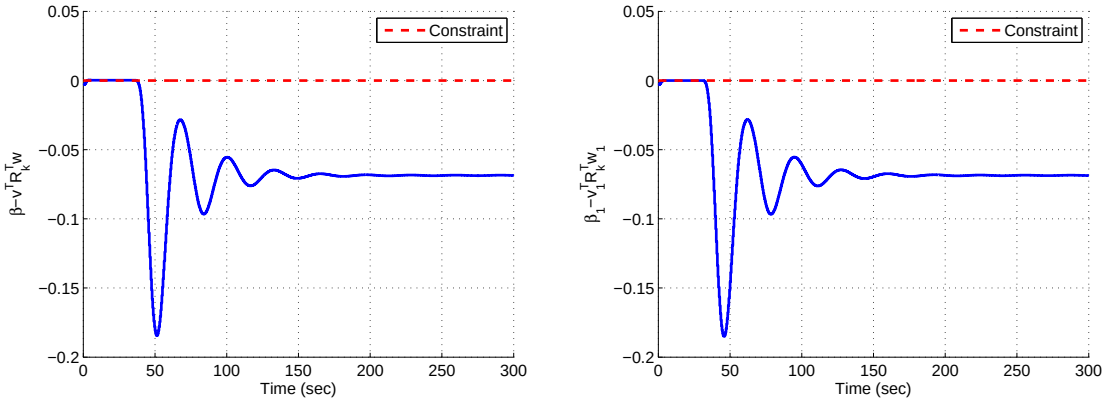
(a) Numerical Solver. (b) Baseline Solver.

Figure 3.6: Control Input.



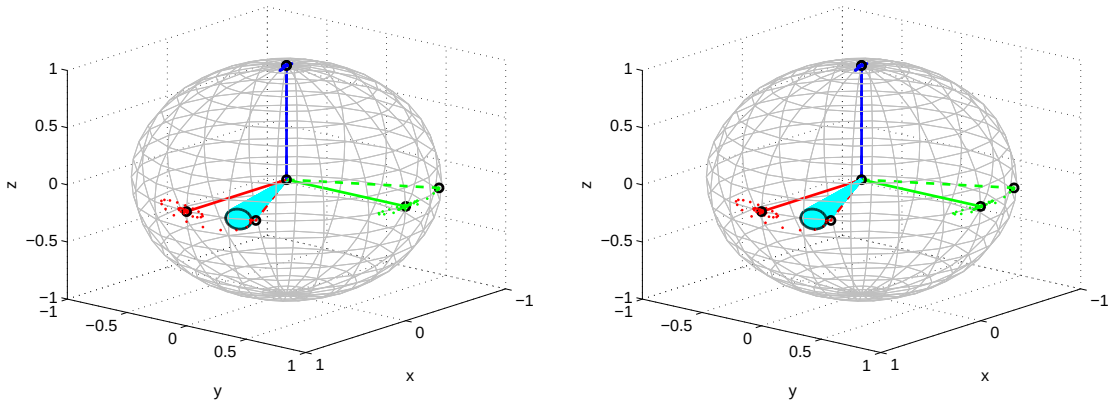
(a) Numerical Solver. (b) Baseline Solver.

Figure 3.7: Thrust Constraint.



(a) Numerical Solver. (b) Baseline Solver.

Figure 3.8: Exclusion Zone Constraint.



(a) Numerical Solver. (b) Baseline Solver.

Figure 3.9: Attitude Maneuver.

savings with the numerical solver versus the baseline solver. For Case I, the maximum time taken by the numerical solver to obtain the optimal solution for one time step is approximately 0.44 (sec). For Case II, the maximum time taken by the numerical solver to obtain the optimal solution for one time step is approximately 3.76 (sec). The code has been implemented using a `MATLAB` `m-file` and the computational time has been assessed using the `tic-toc` function in `MATLAB`.

Case	Numerical Solver	Baseline Solver
I	39.39 (sec) (approx.)	271.01 (sec) (approx.)
II	126.52 (sec) (approx.)	767.97 (sec) (approx.)

Table 3.1: Total Computational Time for both the Solvers.

We will now show that under appropriate assumptions, it is possible to obtain the minimizer for the constrained optimization problem using the exterior penalty function approach. This analysis extends the classical penalty convergence theorem to the setting of smooth manifolds and the classical exact penalization theorem to the setting of Riemannian manifolds.

3.4 Convergence Analysis for the Penalty Function Approach

Let $\mathcal{M}$ be a n-dimensional smooth manifold and $\mathcal{U}$ be a m-dimensional smooth manifold. Consider the following discrete-time OCP

$$\min_{\{u_k\}_{k=0}^{N-1}} \mathcal{J}_d = K_d(q_N) + \sum_{k=0}^{N-1} C_d(q_k, u_k) \tag{3.44}$$

subject to

$$q_{k+1} = F(q_k, u_k), \tag{3.45}$$

$$H_i(q_k, u_k) \leqslant 0, \;\; i = 0, \ldots, L, \tag{3.46}$$

where $K_d : \mathcal{M} \to \mathbb{R}$, $C_d : \mathcal{M} \times \mathcal{U} \to \mathbb{R}$, $F : \mathcal{M} \times \mathcal{U} \to \mathcal{M}$, $H_i : \mathcal{M} \times \mathcal{U} \to \mathbb{R}$, $q_k \in \mathcal{M}$ and $u_k \in \mathcal{U}$. Let $M := \underbrace{\mathcal{M} \times \ldots \times \mathcal{M}}_{N\text{-}copies} \times \underbrace{\mathcal{U} \times \ldots \times \mathcal{U}}_{N\text{-}copies}$. It is easy to verify that M is also a smooth manifold. Let the set $\mathcal{S} \subset M$ be the feasible region for the discrete-time OCP (3.44)-(3.46). The discrete-time OCP (3.44)-(3.46) can be shown to reduce to the following constrained optimization problem **(P)**

$$\min_{m \in \mathcal{S}} \; f(m), \tag{3.47}$$

where $f : M \to \mathbb{R}$. Consider the following unconstrained optimization problem $\mathbf{P}(\boldsymbol{\mu^k})$

$$\min_{m \in M} \; f(m) + \mu^k p(m), \tag{3.48}$$

where $\mu^k \in \mathbb{R}_+$ and $p : M \to \mathbb{R}_+$ is a penalty function, which has the following properties

(a) p is a function of class C^0,

(b) $p(m) \geqslant 0$, for all $m \in M$,

(c) $p(m) = 0$, if and only if $m \in \mathcal{S}$.

In addition, $\{\mu^k\}_{k=1}^{\infty}$ is a strictly increasing sequence such that $\lim_{k\to\infty} \mu^k = \infty$. Let m^* and m^k be the solutions for **(P)** and $\mathbf{P}(\boldsymbol{\mu^k})$, respectively. We are now ready to prove that under appropriate assumptions, the exterior penalty function approach recovers the minimizer for the constrained optimization problem.

Theorem III.6. (Penalty Convergence Theorem) *Assume that f is a function of class C^0 and let $\{m^k\}_{k=1}^{\infty}$ be a sequence of solutions for $\mathbf{P}(\boldsymbol{\mu^k})$. In addition, assume that there exists a chart (U, ϕ) of M such that $m^* \in U$ and with respect to which $\{m^k\}_{k=1}^{\infty}$ converges, then the limit point of $\{m^k\}_{k=1}^{\infty}$ solves* **(P)**.

Proof. The proof follows arguments similar to the one given in [67]. Let $\bar{m}$ be the limit point of $\{m^k\}_{k=1}^{\infty}$. By hypothesis, there exists a $K > 0$ such that $m^k \in U$, for all $k > K$ and $\{\phi(m^k)\}_{k=K}^{\infty}$ converges to $\phi(\bar{m})$. Let $x^k := \phi(m^{k+K-1})$, for all $k \in \mathbb{Z}_+$ and $\bar{x} := \lim_{k\to\infty} x^k = \lim_{k\to\infty} \phi(m^k) = \phi(\bar{m})$. By using the fact that ϕ is a diffeomorphism, it is easy to verify that x^k is the solution for the following unconstrained optimization problem

$$\min_{x\in\phi(U)} \ (f \circ \phi^{-1})(x) + \mu^k (p \circ \phi^{-1})(x).$$

Let $x^* := \phi(m^*)$. Again, by using the fact that ϕ is a diffeomorphism, it is easy to verify that x^* is the solution for the following constrained optimization problem

$$\min_{x\in\phi(\mathcal{S}\bigcap\mathcal{U})} \ (f \circ \phi^{-1})(x).$$

To complete the proof, we require some properties of the exterior penalty function approach.

Lemma III.7. (Penalty Lemma) *The following inequalities hold*

(i) $(f \circ \phi^{-1})(x^k) + \mu^k (p \circ \phi^{-1})(x^k) \leqslant (f \circ \phi^{-1})(x^{k+1}) + \mu^{k+1} (p \circ \phi^{-1})(x^{k+1})$.

(ii) $(p \circ \phi^{-1})(x^k) \geqslant (p \circ \phi^{-1})(x^{k+1})$.

(iii) $(f \circ \phi^{-1})(x^k) \leqslant (f \circ \phi^{-1})(x^{k+1})$.

(iv) $(f \circ \phi^{-1})(x^*) \geqslant (f \circ \phi^{-1})(x^k) + \mu^k(p \circ \phi^{-1})(x^k) \geqslant (f \circ \phi^{-1})(x^k)$.

Proof.

(i) We have the following

$$\begin{aligned}(f \circ \phi^{-1})(x^{k+1}) + \mu^{k+1}(p \circ \phi^{-1})(x^{k+1}) &\geqslant (f \circ \phi^{-1})(x^{k+1}) + \mu^k(p \circ \phi^{-1})(x^{k+1}), \\ &\geqslant (f \circ \phi^{-1})(x^k) + \mu^k(p \circ \phi^{-1})(x^k).\end{aligned}$$

(ii) We have the following

$$\begin{aligned}(f \circ \phi^{-1})(x^k) + \mu^k(p \circ \phi^{-1})(x^k) &\leqslant (f \circ \phi^{-1})(x^{k+1}) + \mu^k(p \circ \phi^{-1})(x^{k+1}), \\ (f \circ \phi^{-1})(x^{k+1}) + \mu^{k+1}(p \circ \phi^{-1})(x^{k+1}) &\leqslant (f \circ \phi^{-1})(x^k) + \mu^{k+1}(p \circ \phi^{-1})(x^k).\end{aligned}$$

Subtracting the above two inequalities gives the following

$$(\mu^{k+1} - \mu^k)(p \circ \phi^{-1})(x^k) \geqslant (\mu^{k+1} - \mu^k)(p \circ \phi^{-1})(x^{k+1}).$$

From the above inequality, it follows that $(p \circ \phi^{-1})(x^k) \geqslant (p \circ \phi^{-1})(x^{k+1})$.

(iii) From (i), we have the following

$$(f \circ \phi^{-1})(x^{k+1}) + \mu^k(p \circ \phi^{-1})(x^{k+1}) \geqslant (f \circ \phi^{-1})(x^k) + \mu^k(p \circ \phi^{-1})(x^k).$$

From (ii), $(p \circ \phi^{-1})(x^k) \geqslant (p \circ \phi^{-1})(x^{k+1})$, which implies that $(f \circ \phi^{-1})(x^k) \leqslant (f \circ \phi^{-1})(x^{k+1})$.

(iv) We have the following

$$\begin{aligned}(f \circ \phi^{-1})(x^*) &= (f \circ \phi^{-1})(x^*) + \mu^k (p \circ \phi^{-1})(x^*), \\ &\geqslant (f \circ \phi^{-1})(x^k) + \mu^k (p \circ \phi^{-1})(x^k), \\ &\geqslant (f \circ \phi^{-1})(x^k).\end{aligned}$$

□

We are now ready to complete the proof of Theorem III.6. From Lemma III.7, it follows that $\{(f \circ \phi^{-1})(x^k) + \mu^k (p \circ \phi^{-1})(x^k)\}_{k=1}^{\infty}$ is a nondecreasing sequence, bounded above by $(f \circ \phi^{-1})(x^*)$, which implies that $\lim_{k \to \infty} [(f \circ \phi^{-1})(x^k) + \mu^k (p \circ \phi^{-1})(x^k)] = r^* \leqslant (f \circ \phi^{-1})(x^*)$. Using the continuity of the function $(f \circ \phi^{-1})$, it follows that $\lim_{k \to \infty} \mu^k (p \circ \phi^{-1})(x^k) = r^* - (f \circ \phi^{-1})(\bar{x})$. Using the facts that $(p \circ \phi^{-1})(x^k) \geqslant 0$ and $\lim_{k \to \infty} \mu^k = \infty$, it follows from the above equality that $\lim_{k \to \infty} (p \circ \phi^{-1})(x^k) = 0$. Using the continuity of the function $(p \circ \phi^{-1})$, it follows that $(p \circ \phi^{-1})(\bar{x}) = 0$. This shows that $\bar{m}$ is a feasible solution for **(P)**. From Lemma III.7, it follows that $(f \circ \phi^{-1})(x^k) \leqslant (f \circ \phi^{-1})(x^*)$, which implies that $(f \circ \phi^{-1})(\bar{x}) \leqslant (f \circ \phi^{-1})(x^*)$, or equivalently, $f(\bar{m}) \leqslant f(m^*)$, which further implies that $f(\bar{m}) = f(m^*)$. This shows that the limit point of $\{m^k\}_{k=1}^{\infty}$ solves **(P)**. □

Corollary III.8. *Let $\mathcal{M} = \mathbb{R}^n$ and $\mathcal{U} = \mathbb{R}^m$. Assume that f is a function of class C^0 and let $\{m^k\}_{k=1}^{\infty}$ be a sequence of solutions for $\mathbf{P}(\boldsymbol{\mu}^k)$. If $\{m^k\}_{k=1}^{\infty}$ converges, then the limit point of $\{m^k\}_{k=1}^{\infty}$ solves* **(P)**.

Proof. By setting $U = M$ and $\phi = \mathrm{id}_M$, it is easy to verify that Theorem III.6 holds. □

The above convergence analysis shows that the exterior penalty function approach recovers the minimizer for the constrained optimization problem only in the limit. This may not be desirable as the solution from the exterior penalty function approach

is not guaranteed to be a feasible solution for the constrained optimization problem and the problem may become numerically ill-conditioned as the penalty factor increases (see Remark III.5). To avoid this situation, under appropriate assumptions, it is possible to come up with an exact penalization approach.

Let $\mathcal{M}$ be a n-dimensional connected Riemannian manifold and $\mathcal{U}$ be a m-dimensional connected Riemannian manifold. It is easy to verify that M is also a connected Riemannian manifold. For p, $q \in M$, the Riemannian distance function is given by

$$d(p,q) := \inf_{\gamma \in \Omega} l(\gamma), \tag{3.49}$$

where Ω denotes the collection of all piecewise C^1 curves joining p, q and the length of γ is given by

$$l(\gamma) = \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)} dt, \tag{3.50}$$

where $\gamma : [a,b] \to M$, with $\gamma(a) = p$, $\gamma(b) = q$ and $\|.\|_m$ denotes the induced norm at the point $m \in M$. The distance between a point $m \in M$ and the set $\mathcal{S}$ is given by

$$d_{\mathcal{S}}(m) := \inf_{m' \in \mathcal{S}} d(m,m'). \tag{3.51}$$

We will now introduce the definition of a Lipschitz function on an open subset of M.

Definition III.9. Let U be an open subset of M. A function $\tilde{f} : U \to \mathbb{R}$ is said to be Lipschitz, with Lipschitz constant $\tilde{K} < \infty$, if

$$|\tilde{f}(m_1) - \tilde{f}(m_2)| \leqslant \tilde{K} d(m_1, m_2), \text{ for all } m_1, m_2 \in U. \tag{3.52}$$

Theorem III.10. (Exact Penalization Theorem) *Assume that f is a Lipschitz function on M, with Lipschitz constant K. Choose any $\bar{K} \geqslant K$, then m^* is also a*

minimizer for the following unconstrained optimization problem

$$\min_{m \in M} \; f(m) + \bar{K} d_{\mathcal{S}}(m). \tag{3.53}$$

If $\bar{K} > K$ and the set $\mathcal{S}$ is closed, then any minimizer $\bar{m}$ for (3.53) is also a minimizer for **(P)** *and so, in particular, $\bar{m} \in \mathcal{S}$.*

Proof. It is well known that with the Riemannian distance function, any connected Riemannian manifold is a metric space whose metric topology is the same as the original manifold topology (see, e.g., [60]). The proof now follows from Theorem 3.2.1 of [93]. □

CHAPTER IV

Neighboring Extremal Optimal Control for Mechanical Systems on Riemannian Manifolds

In this chapter, we extend NEOC, which is well established for OCPs defined on a Euclidean space (see, e.g., [14]), to the setting of Riemannian manifolds. See also Chapter II for the discussion on NEOC for OCPs defined on a Euclidean space. We further specialize the results to the case of Lie groups. An example along with the simulation results is presented. We will now discuss the OCP that will be studied in this chapter. In what follows, we will suppress the explicit dependence of the state, costate and control trajectories on time unless otherwise necessary.

4.1 Optimal Control Problem

Let $\mathcal{Q}$ be a n-dimensional complete connected Riemannian manifold and $\{X_i\}_{i=1}^n$ be smooth vector fields on $\mathcal{Q}$. For a given time interval $[0, T]$, it is assumed that the flow of each vector field in $\{X_i\}_{i=1}^n$ exists, for all $t \in [0, T]$. Additionally, if $\mathcal{Q}$ is compact, then each vector field in $\{X_i\}_{i=1}^n$ is complete (see, e.g., [60]). Note that in this chapter, we only consider the class of fully-actuated controlled mechanical systems for which the Lagrangian $L : T\mathcal{Q} \to \mathbb{R}$ is given by $L(v_q) = \frac{1}{2}\langle v_q, v_q \rangle$, where

$v_q \in T_q\mathcal{Q}$. Consider the following OCP **(P)**

$$\min_{u(.)} J = \frac{1}{2}\int_0^T \langle u(t), u(t)\rangle dt \tag{4.1}$$

subject to

$$\frac{dq}{dt}(t) = v(t),\ \ q(0) = q_0,\ \ q(T) = q_T, \tag{4.2}$$

$$\frac{Dv}{dt}(t) = u(t),\ \ v(0) = v_0,\ \ v(T) = v_T, \tag{4.3}$$

where $q(.) \in C^2([0,T], \mathcal{Q})$, $v(.) \in C^1([0,T], T_{q(.)}\mathcal{Q})$ and the n-tuple of control inputs $[u^1 \dots u^n]^T$ take values in $\mathbb{R}^n$. Note that in general, the n-tuple of control inputs $[u^1 \dots u^n]^T$ are constrained to take values in the set $\mathcal{U} \subset \mathbb{R}^n$ (nonempty, connected, with $0 \in \text{int}(\mathcal{U})$ and also generally assumed to be compact and convex). In a more general setting, e.g., when admissible controls are only assumed to be measurable locally bounded mappings taking values in the set $\mathcal{U}$, more technical assumptions are needed (see, e.g., [3], [16]) but we do not consider such a setting in this chapter.

Remark IV.1. It is possible to generalize the idea presented in this chapter to a cost functional, which has a more general form with a more complicated dynamic constraint (see, e.g., [47]). We choose to work with the cost functional (4.1) as the solution for **(P)** has a nice geometric interpretation thereby helping to present the main idea of the chapter clearly and avoid unnecessary mathematical complications. In fact, **(P)** is equivalent to the well known Riemannian geodesic problem (see, e.g., [9]). The local existence and uniqueness of the solution for **(P)** follows from the theorems on local existence and uniqueness of the solution for ordinary differential equations. The equations of motion for the class of fully-actuated controlled mechan-

ical systems with the Lagrangian defined above are given by

$$\nabla_{\dot{q}}\dot{q} = \sum_{l=1}^{n} u^l X_l(q), \tag{4.4}$$

where $q : [0,T] \to \mathcal{Q}$. The vertical lift of a vector field X on $\mathcal{Q}$ is the vector field X^{vlift} on $T\mathcal{Q}$ given by

$$X^{\text{vlift}}(v_q) = \frac{d}{dt}(v_q + tX(q))\bigg|_{t=0} \in T_{v_q}T\mathcal{Q}, \tag{4.5}$$

where $v_q \in T_q\mathcal{Q}$. In local coordinates, (4.5) has a simple interpretation. Let $(q^1,\ldots,q^n)$ be the local coordinates for $\mathcal{Q}$ and $(q^1,\ldots,q^n,v^1,\ldots,v^n)$ be the corresponding local coordinates for $T\mathcal{Q}$. If $X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial q^i}$, then $X^{\text{vlift}} = \sum_{i=1}^{n} X^i \frac{\partial}{\partial v^i}$, where $(X^1,\ldots,X^n)$ are the component functions of X in some given chart. We can now re-write (4.4) as follows

$$\dot{\gamma} = Z(\gamma) + \sum_{l=1}^{n} u^l X_l^{\text{vlift}}(\gamma), \tag{4.6}$$

where $\gamma : [0,T] \to T\mathcal{Q}$ and Z is the geodesic spray associated with the connection ∇. In local coordinates, $Z = \sum_{i=1}^{n} v^i \frac{\partial}{\partial q^i} - \sum_{i=1}^{n}\sum_{j,k=1}^{n} \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i}$. Note that γ is the canonical lifting of q, i.e., $(\pi \circ \gamma)(t) = q(t)$. It is not difficult to see that (4.4) is equivalent to (4.6). Indeed, in local coordinates, (4.4) has the following form

$$\ddot{q}^i + \sum_{j,k=1}^{n} \Gamma^i_{jk}\dot{q}^j\dot{q}^k = \sum_{l=1}^{n} u^l X_l(q), \;\; i = 1,\ldots,n. \tag{4.7}$$

Observe that (4.7) is system of second order ordinary differential equations on $\mathcal{Q}$, which is equivalent to a system of first order ordinary differential equations on $T\mathcal{Q}$,

which has the following form

$$\dot{q}^i = v^i, \;\; i = 1, \ldots, n, \tag{4.8}$$

$$\dot{v}^i = -\sum_{j,k=1}^{n} \Gamma^i_{jk} v^j v^k + \sum_{l=1}^{n} u^l X_l(q), \;\; i = 1, \ldots, n. \tag{4.9}$$

The connection ∇ induces an Ehresmann connection on $\pi : T\mathcal{Q} \to \mathcal{Q}$ such that, for all $v_q \in T_q\mathcal{Q}$, there is a splitting of $T_{v_q}T\mathcal{Q}$ into a horizontal subspace and a vertical subspace, i.e., $T_{v_q}T\mathcal{Q} \cong H_{v_q}(T\mathcal{Q}) \oplus V_{v_q}(\pi)$, where $H_{v_q}(T\mathcal{Q}) \cong T_q\mathcal{Q}$ and $V_{v_q}(\pi) \cong T_q\mathcal{Q}$. Note that $H_{v_q}(T\mathcal{Q}) = \text{span}\left\{\dfrac{\partial}{\partial q^i} - \displaystyle\sum_{j,k=1}^{n} \Gamma^k_{ij} v^j \frac{\partial}{\partial v^k}\right\}_{i=1}^{n}$ and $V_{v_q}(\pi) = \text{span}\left\{\dfrac{\partial}{\partial v^i}\right\}_{i=1}^{n}$. It is easy to verify that with respect to the above splitting, for all $v_q \in T_q\mathcal{Q}$, $Z(v_q) \in H_{v_q}(T\mathcal{Q})$ and $X^{\text{vlift}}(v_q) \in V_{v_q}(\pi)$. For more details see [2], [5], [9], [17], [64], [65], [89]. In view of the above discussion, we note that (4.2)-(4.3) are equivalent to (4.6). Using the splitting of $T_{v_q}T\mathcal{Q}$ discussed above, for all $r \in T_{v_q}T\mathcal{Q}$, r can be uniquely written as follows

$$r = r^h + r^v,$$

where $r^h \in H_{v_q}(T\mathcal{Q})$ and $r^v \in V_{v_q}(\pi)$. For all pairs r_1, $r_2 \in T_{v_q}T\mathcal{Q}$, the Riemannian metric (Sasaki metric) on $T\mathcal{Q}$ is obtained in terms of the Riemannian metric on $\mathcal{Q}$ as follows

$$\langle\langle r_1, r_2 \rangle\rangle = \langle r_1^h, r_2^h \rangle + \langle r_1^v, r_2^v \rangle.$$

It is easy to verify that (4.1) makes sense as $\dfrac{1}{2}\displaystyle\int_0^T \langle\langle u(t), u(t) \rangle\rangle dt = \frac{1}{2}\int_0^T \langle u(t), u(t) \rangle dt$. For more details see [83], [84], [89].

Before we proceed further, we introduce the concept of a variation (see, e.g., [2], [9], [16], [29], [68], [72]). Let Ω denote the set of all C^2 curves on $\mathcal{Q}$ satisfying the

boundary conditions (4.2)-(4.3). The set Ω is also referred to as the path space of $\mathcal{Q}$ (see, e.g., [72]). For a curve $q(t) \in \Omega$, $T_{q(t)}\Omega$ is a vector space consisting of all C^2 vector fields $w(t)$ along $q(t)$ such that $w(0) = 0$ and $w(T) = 0$.

Definition IV.2 ([72])**.** A one-parameter variation of a curve $q \in \Omega$ is a function $\bar{q} : (-\epsilon, \epsilon) \to \Omega$, for some $\epsilon > 0$ such that

(a) $\bar{q}(0) = q$,

(b) The map $q_\epsilon : [0, T] \times (-\epsilon, \epsilon) \to \mathcal{Q}$ defined by $q_\epsilon(t, \bar{\epsilon}) = \bar{q}(\bar{\epsilon})(t)$ is C^2 on $[0, T] \times (-\epsilon, \epsilon)$.

Note that a one-parameter variation of a curve $q(t) \in \Omega$ defined above is proper (see, e.g., [29]). The vector field $v(t) := \dfrac{\partial q_\epsilon}{\partial t}(t, 0)$ is the velocity vector field along $q(t)$ and the vector field $w(t) := \dfrac{\partial q_\epsilon}{\partial \epsilon}(t, 0)$ is the variation vector field associated with the one-parameter variation q_ϵ (see, e.g., [29], [72]). By setting $q_\epsilon(t, \bar{\epsilon}) := \exp_{q(t)}(\bar{\epsilon} w(t))$, we obtain a one-parameter variation of a curve $q(t) \in \Omega$, where $w(t) \in T_{q(t)}\Omega$ (see, e.g., [29], [72]).

To demonstrate NEOC for **(P)**, we first obtain the nominal trajectory, by solving **(P)** using two methods. The first method is solving **(P)** using Lagrange multipliers and the second method is solving **(P)** as a variational problem.

4.2 Solution Using Lagrange Multipliers

We proceed by following the same procedure as given in [24] and defining the augmented cost functional as follows

$$J^a = \int_0^T \left[\frac{1}{2} \langle u, u \rangle + \lambda_1 \left(\frac{dq}{dt} - v \right) + \lambda_2 \left(\frac{Dv}{dt} - u \right) \right] dt, \tag{4.10}$$

where $\lambda_1(.)$, $\lambda_2(.) \in C^1([0, T], T^*_{q(.)}\mathcal{Q})$. We will now fix some notation.

4.2.1 Notation

For any smooth vector field $y = \sum_{i=1}^{n} y^i(t) X_i(q)$ along the curve q, with velocity vector field v, $\dfrac{Dy}{dt} = \sum_{i=1}^{n} \dot{y}^i(t) X_i(q) + \sum_{i=1}^{n} y^i(t) (\nabla_v X_i)(q)$, or in shorthand is written as $\dfrac{Dy}{dt} = \dot{y} + \nabla_v y$. Using this shorthand, $\left.\dfrac{Dy}{\partial \epsilon}\right|_{\epsilon=0} = \delta y + \nabla_w y$. Similarly, for any smooth covector field $\alpha = \sum_{i=1}^{n} \alpha^i(t) \omega_i(q)$ along the curve q, with velocity vector field v, $\dfrac{D\alpha}{dt} = \sum_{i=1}^{n} \dot{\alpha}^i(t) \omega_i(q) + \sum_{i=1}^{n} \alpha^i(t) (\nabla_v \omega_i)(q)$, or in shorthand is written as $\dfrac{D\alpha}{dt} = \dot{\alpha} + \nabla_v \alpha$. Using this shorthand, $\left.\dfrac{D\alpha}{\partial \epsilon}\right|_{\epsilon=0} = \delta\alpha + \nabla_w \alpha$. For more details see [24]. Before we proceed further, we need a few lemmas.

Lemma IV.3 ([24])**.**

$$\int_0^T \lambda_2 \left(\delta \frac{Dv}{dt} \right) dt = \int_0^T \left[-\frac{D\lambda_2}{dt} (\delta v) + \lambda_2 (\nabla_{\delta v} v) \right] dt.$$

Lemma IV.4 ([29], [72])**.** *If the connection ∇ is symmetric, then*

$$\frac{D}{\partial \epsilon} \frac{\partial q_\epsilon}{\partial t} = \frac{D}{\partial t} \frac{\partial q_\epsilon}{\partial \epsilon}.$$

The necessary conditions for a normal extremal (see, e.g., [9]) for **(P)** are obtained by setting

$$\left. \frac{dJ_\epsilon^a}{d\epsilon} \right|_{\epsilon=0} = 0,$$

where

$$J_\epsilon^a = \int_0^T \left[\frac{1}{2} \langle u_\epsilon, u_\epsilon \rangle + \lambda_1 \left(\frac{\partial q_\epsilon}{\partial t} - v_\epsilon \right) + \lambda_2 \left(\frac{Dv_\epsilon}{\partial t} - u_\epsilon \right) \right] dt.$$

The above condition, with the use of Lemmas IV.3-IV.4, gives the following

$$\begin{aligned}
\left.\frac{dJ_\epsilon^a}{d\epsilon}\right|_{\epsilon=0} &= \int_0^T \Bigg[\langle u, \delta u + \nabla_w u\rangle + \lambda_1\left(\frac{Dw}{dt} - \delta v - \nabla_w v\right) + \\
&\qquad \lambda_2\left(\delta\frac{Dv}{dt} + \nabla_w\frac{Dv}{dt} - \delta u - \nabla_w u\right)\Bigg] dt, \\
&= \int_0^T \Bigg[\langle u, \nabla_w u\rangle + \langle u, \delta u\rangle + \lambda_1\left(\frac{Dw}{dt} - \nabla_w v\right) - \lambda_1(\delta v) + \\
&\qquad \lambda_2\left(\nabla_w\frac{Dv}{dt} - \nabla_w u\right) + \lambda_2\left(\delta\frac{Dv}{dt}\right) - \lambda_2(\delta u)\Bigg] dt, \\
&= \int_0^T \Bigg[-\frac{D\lambda_1}{dt}(w) - \lambda_1(\nabla_w v) + \lambda_2\left(\nabla_w\frac{Dv}{dt} - \nabla_w u\right) + \langle u, \nabla_w u\rangle - \\
&\qquad \frac{D\lambda_2}{dt}(\delta v) - \lambda_1(\delta v) + \lambda_2(\nabla_{\delta v} v) + \langle u, \delta u\rangle - \lambda_2(\delta u)\Bigg] dt,
\end{aligned}$$

where we have used integration by parts along with the fact that the one-parameter variation q_ϵ is proper. We are now ready to state a theorem.

Theorem IV.5 ([24]). *A normal extremal for* **(P)** *satisfies the following equations*

$$\frac{dq}{dt} = v, \tag{4.11}$$

$$\frac{Dv}{dt} = u, \tag{4.12}$$

$$\frac{D\lambda_1}{dt} = -\lambda_1(\nabla v) + \lambda_2(\nabla\lambda_2^\sharp), \tag{4.13}$$

$$\frac{D\lambda_2}{dt} = -\lambda_1 + \lambda_2(\nabla v), \tag{4.14}$$

where $u = \lambda_2^\sharp$.

We assume that the nominal solution has been obtained for a fixed initial condition. Suppose there is a small variation in the initial condition and we would like to update the optimal control for **(P)**. Instead of solving **(P)** from scratch, we employ NEOC as described previously. We will now fix some more notation.

4.2.2 Notation

In what follows, we use superscript n to denote the nominal trajectory and the corresponding vector and covector fields. The one-parameter variation of $q^n(t)$ is denoted by q^n_ϵ. Note that the one-parameter variation of $q^n(t)$ is not proper as there is a small variation in the initial condition. The vector field $v^n(t) := \dfrac{\partial q^n_\epsilon}{\partial t}(t,0)$ is the velocity vector field along $q^n(t)$ and the vector field $w^n(t) := \dfrac{\partial q^n_\epsilon}{\partial \epsilon}(t,0)$ is the variation vector field associated with the one-parameter variation q^n_ϵ.

Employing the NEOC approach described previously, the variational equations for (4.11)-(4.14) are given as follows

$$\left.\frac{D}{\partial\epsilon}\frac{\partial q^n_\epsilon}{\partial t}\right|_{\epsilon=0} = \left.\frac{Dv^n_\epsilon}{\partial\epsilon}\right|_{\epsilon=0}, \tag{4.15}$$

$$\left.\frac{D}{\partial\epsilon}\frac{Dv^n_\epsilon}{\partial t}\right|_{\epsilon=0} = \left.\frac{D\lambda^{n\sharp}_{2,\epsilon}}{\partial\epsilon}\right|_{\epsilon=0}, \tag{4.16}$$

$$\left.\frac{D}{\partial\epsilon}\frac{D\lambda^n_{1,\epsilon}}{\partial t}\right|_{\epsilon=0} = \left.\frac{D}{\partial\epsilon}\left(-\lambda^n_{1,\epsilon}(\nabla v^n_\epsilon) + \lambda^n_{2,\epsilon}(\nabla\lambda^{n\sharp}_{2,\epsilon})\right)\right|_{\epsilon=0}, \tag{4.17}$$

$$\left.\frac{D}{\partial\epsilon}\frac{D\lambda^n_{2,\epsilon}}{\partial t}\right|_{\epsilon=0} = \left.\frac{D}{\partial\epsilon}\left(-\lambda^n_{1,\epsilon} + \lambda^n_{2,\epsilon}(\nabla v^n_\epsilon)\right)\right|_{\epsilon=0}. \tag{4.18}$$

Note that the change in the control trajectory corresponding to the change in the initial condition is given by $\left.\dfrac{D\lambda^{n\sharp}_{2,\epsilon}}{\partial\epsilon}\right|_{\epsilon=0}$. Before we proceed further, we need a few lemmas.

Lemma IV.6 ([29], [72]). *Given any smooth vector field y along q_ϵ, then*

$$\frac{D}{\partial\epsilon}\frac{Dy}{\partial t} - \frac{D}{\partial t}\frac{Dy}{\partial\epsilon} = R\left(\frac{\partial q_\epsilon}{\partial\epsilon}, \frac{\partial q_\epsilon}{\partial t}\right) y.$$

Remark IV.7. Note that the definition of the curvature tensor of the connection ∇ used in this chapter, differs by a negative sign from the one defined in [29], [72].

Lemma IV.8 ([24]). *Given y, $z \in \mathfrak{X}(\mathcal{Q})$ and $\alpha \in \mathfrak{X}^*(\mathcal{Q})$, then*

$$\frac{D}{\partial\epsilon}\alpha(\nabla_z y) = \frac{D\alpha}{\partial\epsilon}(\nabla_z y) + \alpha\left(\frac{D}{\partial\epsilon}(\nabla_z y) - \nabla_{\frac{Dz}{\partial\epsilon}} y\right).$$

Remark IV.9. Note that the expression, $\frac{D}{\partial\epsilon}(\nabla_z y) - \nabla_{\frac{Dz}{\partial\epsilon}} y$ in Lemma IV.8 represents the second covariant derivative.

We are now ready to state two theorems.

Theorem IV.10. *The variational equations (4.15)-(4.18) give the following equations*

$$\dot{w}^n = \delta v^n + [w^n, v^n], \tag{4.19}$$

$$\delta\dot{v}^n + \nabla_{v^n}\delta v^n + \nabla_{\delta v^n} v^n + \nabla_{w^n}\dot{v}^n + \nabla_{w^n}\nabla_{v^n} v^n = \delta\lambda_2^{n\sharp} + \nabla_{w^n}\lambda_2^{n\sharp}, \tag{4.20}$$

$$(\delta\dot{\lambda}_1^n + \nabla_{w^n}\dot{\lambda}_1^n + \nabla_{\delta v^n}\lambda_1^n + \nabla_{v^n}\delta\lambda_1^n + \nabla_{w^n}\nabla_{v^n}\lambda_1^n)(z) = (-\delta\lambda_1^n - \nabla_{w^n}\lambda_1^n)(\nabla_z v^n) -$$
$$\lambda_1^n(\nabla_z\delta v^n + \nabla_{w^n}\nabla_z v^n - \nabla_{\nabla_{w^n} z} v^n) + (\delta\lambda_2^n + \nabla_{w^n}\lambda^2)(\nabla_z\lambda_2^{n\sharp} + \lambda_2^n(\nabla_z\lambda_2^{n\sharp}) +$$
$$\nabla_{w^n}\nabla_z\lambda_2^{n\sharp} - \nabla_{\nabla_{w^n} z}\lambda_2^{n\sharp}), \tag{4.21}$$

$$(\delta\dot{\lambda}_2^n + \nabla_{w^n}\dot{\lambda}_2^n + \nabla_{\delta v^n}\lambda_2^n + \nabla_{v^n}\delta\lambda_2^n + \nabla_{w^n}\nabla_{v^n}\lambda_2^n)(z) = (-\delta\lambda_1^n - \nabla_{w^n}\lambda_1^n)(z) +$$
$$(\delta\lambda_2^n + \nabla_{w^n}\lambda_2^n)(\nabla_z v^n) + \lambda_2^n(\nabla_z\delta v^n + \nabla_{w^n}\nabla_z v^n - \nabla_{\nabla_{w^n} z} v^n), \tag{4.22}$$

where $z \in \mathfrak{X}(\mathcal{Q})$.

Proof. Using Lemma IV.4, (4.15) can be re-written as follows

$$\frac{D}{\partial t}\frac{\partial q_\epsilon^n}{\partial\epsilon}\bigg|_{\epsilon=0} = \frac{Dv_\epsilon^n}{\partial\epsilon}\bigg|_{\epsilon=0}.$$

The above equation gives the following

$$\dot{w}^n = \delta v^n + \nabla_{w^n} v^n - \nabla_{v^n} w^n.$$

Using the symmetry of the connection ∇, the above equation can be re-written as

follows

$$\dot{w}^n = \delta v^n + [w^n, v^n].$$

Using Lemma IV.6, (4.16) can be re-written as follows

$$\frac{D}{\partial t}\frac{Dv^n_\epsilon}{\partial \epsilon}\bigg|_{\epsilon=0} + R(w^n, v^n)v^n = \frac{D\lambda^{n\sharp}_{2,\epsilon}}{\partial \epsilon}\bigg|_{\epsilon=0},$$

where $R(w^n, v^n)v^n := \nabla_{w^n}\nabla_{v^n}v^n - \nabla_{v^n}\nabla_{w^n}v^n - \nabla_{[w^n,v^n]}v^n$. The above equation gives the following

$$\delta\dot{v}^n + \nabla_{\dot{w}^n}v^n + \nabla_{w^n}\dot{v}^n + \nabla_{v^n}\delta v^n + \nabla_{v^n}\nabla_{w^n}v^n + R(w^n, v^n)v^n = \delta\lambda^{n\sharp}_2 + \nabla_{w^n}\lambda^{n\sharp}_2.$$

Substituting $\dot{w}^n = \delta v^n + [w^n, v^n]$ into the above equation, gives the following

$$\delta\dot{v}^n + \nabla_{v^n}\delta v^n + \nabla_{\delta v^n}v^n + \nabla_{w^n}\dot{v}^n + \nabla_{w^n}\nabla_{v^n}v^n = \delta\lambda^{n\sharp}_2 + \nabla_{w^n}\lambda^{n\sharp}_2.$$

Similarly, the other two variational equations can be derived using Lemma IV.8. □

Theorem IV.11. *The variational equations (4.19)-(4.20) give the following Jacobi equation*

$$\ddot{w}^n + 2\nabla_{v^n}\dot{w}^n + \nabla_{\dot{v}^n}w^n + \nabla_{v^n}\nabla_{v^n}w^n + R(w^n, v^n)v^n = \delta\lambda^{n\sharp}_2 + \nabla_{w^n}\lambda^{n\sharp}_2. \tag{4.23}$$

Proof. Substituting (4.19) into (4.20), gives the following

$$\ddot{w}^n - [\dot{w}^n, v^n] - [w^n, \dot{v}^n] + \nabla_{v^n}\dot{w}^n - \nabla_{v^n}[w^n, v^n] + \nabla_{\dot{w}^n}v^n - \nabla_{[w^n,v^n]}v^n + \nabla_{w^n}\dot{v}^n + \nabla_{w^n}\nabla_{v^n}v^n = \delta\lambda^{n\sharp}_2 + \nabla_{w^n}\lambda^{n\sharp}_2.$$

Using the symmetry of the connection ∇, the above equation can be re-written as

follows

$$\ddot{w}^n + \nabla_{v^n}\dot{w}^n - \nabla_{\dot{w}^n}v^n + \nabla_{\dot{v}^n}w^n - \nabla_{w^n}\dot{v}^n + \nabla_{v^n}\dot{w}^n + \nabla_{v^n}\nabla_{v^n}w^n - \nabla_{v^n}\nabla_{w^n}v^n +$$
$$\nabla_{\dot{w}^n}v^n - \nabla_{[w^n,v^n]}v^n + \nabla_{w^n}\dot{v}^n + \nabla_{w^n}\nabla_{v^n}v^n = \delta\lambda_2^{n\sharp} + \nabla_{w^n}\lambda_2^{n\sharp}.$$

Using the definition of the curvature tensor of the connection ∇, the above equation can be re-written as follows

$$\ddot{w}^n + 2\nabla_{v^n}\dot{w}^n + \nabla_{\dot{v}^n}w^n + \nabla_{v^n}\nabla_{v^n}w^n + R(w^n, v^n)v^n = \delta\lambda_2^{n\sharp} + \nabla_{w^n}\lambda_2^{n\sharp},$$

where $R(w^n, v^n)v^n := \nabla_{w^n}\nabla_{v^n}v^n - \nabla_{v^n}\nabla_{w^n}v^n - \nabla_{[w^n,v^n]}v^n$. □

Remark IV.12. It should be noted that (4.23) plays a crucial role in determining conjugate points for **(P)**. It is also worthwhile to note that (4.23) corresponds to (3.3) in Theorem 4 of [15], where the case of a Lie group has been considered but not in a control theoretic setting. For computational purposes, (4.19)-(4.22) would result in a TPBVP and the change in the control trajectory corresponding to the change in the initial condition can then be computed after solving the TPBVP. This point will become more clear, when we consider an example presented later in the chapter.

4.3 Solution as a Variational Problem

We will follow the same procedure as given in [9]. Before we proceed further, we need a lemma.

Lemma IV.13 ([29], [72])**.** *Given* w, x, y, $z \in \mathfrak{X}(\mathcal{Q})$, *then*

$$\langle R(x,y)z, w\rangle = \langle R(w,z)y, x\rangle.$$

The necessary conditions for a normal extremal for **(P)** are obtained by setting

$$\left.\frac{dJ_\epsilon}{d\epsilon}\right|_{\epsilon=0} = 0,$$

where

$$J_\epsilon = \frac{1}{2}\int_0^T \left\langle \frac{D^2 q_\epsilon}{\partial t^2}, \frac{D^2 q_\epsilon}{\partial t^2} \right\rangle dt.$$

The above condition, with the use of Lemmas IV.4, IV.6, IV.13, gives the following

$$\begin{aligned}\left.\frac{dJ_\epsilon}{d\epsilon}\right|_{\epsilon=0} &= \int_0^T \left\langle \frac{Dv}{dt}, \frac{D^2 w}{dt^2} + R(w,v)v \right\rangle dt, \\ &= \int_0^T \left\langle \frac{D^3 v}{dt^3} + R\left(\frac{Dv}{dt}, v\right)v, w \right\rangle dt,\end{aligned}$$

where we have used integration by parts twice along with the fact that the one-parameter variation q_ϵ is proper. We are now ready to state a theorem.

Remark IV.14. It is sometimes appropriate to assume that $\mathcal{Q}$ is parallelizable (see, e.g., [9]). This means that there exist smooth vector fields $\{X_i\}_{i=1}^n$ on $\mathcal{Q}$ such that the vectors $\{X_i(q)\}_{i=1}^n$ form an orthonormal basis for $T_q\mathcal{Q}$, for all $q \in \mathcal{Q}$. Given smooth vector fields $\{X_i\}_{i=1}^n$ on $\mathcal{Q}$, there exist unique smooth covector fields $\{\omega^i\}_{i=1}^n$ on $\mathcal{Q}$ such that the covectors $\{\omega^i(q)\}_{i=1}^n$ are the dual basis for $T_q^*\mathcal{Q}$, for all $q \in \mathcal{Q}$. Equivalently, the assumption that $\mathcal{Q}$ is parallelizable means that $T\mathcal{Q}$ is a trivial bundle. The assumption that $\mathcal{Q}$ is parallelizable is restrictive in some sense but it is satisfied for the case of Lie groups (see, e.g., [60]), which are of special interest.

Theorem IV.15 ([74])**.** *A necessary condition for a curve $q \in C^2([0,T], \mathcal{Q})$ to be a normal extremal for* **(P)** *is that the velocity vector field $v = \dfrac{dq}{dt}$ satisfies the following*

equation

$$\frac{D^3 v}{dt^3} + R\left(\frac{Dv}{dt}, v\right) v = 0. \tag{4.24}$$

Remark IV.16. In [24], it has been shown that (4.11)-(4.14) are equivalent to (4.24). In the case when $\mathcal{Q} = \mathbb{R}^n$, with the standard inner product, the covariant derivative is the usual derivative and $R = 0$. We now see that (4.24) simplifies to the equation $\dddot{q} = 0$, which shows that each coordinate function of a normal extremal q for **(P)** is a cubic spline.

We do not give all the details, as they are similar to the previous section. The variational equation for (4.24) is given as follows

$$\frac{D}{\partial \epsilon}\left(\frac{D^3 v_\epsilon^n}{\partial t^3} + R\left(\frac{D v_\epsilon^n}{\partial t}, v_\epsilon^n\right) v_\epsilon^n\right)\Bigg|_{\epsilon=0} = 0. \tag{4.25}$$

Note that the change in the control trajectory corresponding to the change in the initial condition is given by $\left.\frac{D^2 q_\epsilon}{\partial t^2}\right|_{\epsilon=0}$. We will now specialize the results to the case of Lie groups.

4.4 Application to Lie Groups

We will now present NEOC for OCPs for mechanical systems evolving on Lie groups. Let G be a a finite-dimensional compact semisimple Lie group. Given x, y and z are left invariant vector fields on G and given α is a left invariant one-form on G, then $\nabla_x y = \frac{1}{2}[x, y]$, $\nabla_x \alpha = -\frac{1}{2}\operatorname{ad}_x^* \alpha$ (see, e.g., [24]) and $R(x, y)z = -\frac{1}{4}[x, [y, z]]$ (see, e.g., [23], [89]).

Remark IV.17. Note that the adjoint representation is equivalent to the coadjoint representation for semisimple Lie algebras.

We will still retain the same notation **(P)**, in the case when $\mathcal{Q} = G$. We are now

ready to state a lemma.

Lemma IV.18 ([24]). *A normal extremal for* **(P)** *satisfies the following equations*

$$\dot{g} = T_e L_g(v), \tag{4.26}$$

$$\dot{v} = u, \tag{4.27}$$

$$\dot{\lambda}_1 = \mathrm{ad}_v^* \, \lambda_1, \tag{4.28}$$

$$\dot{\lambda}_2 = -\lambda_1, \tag{4.29}$$

where $u = \lambda_2^\sharp$.

We assume that the nominal solution has been obtained for a fixed initial condition. Suppose there is a small variation in the initial condition and we would like to update the optimal control for **(P)**. Instead of solving **(P)** from scratch, we employ NEOC as described previously. The variational equations for (4.26)-(4.29) are given as follows

$$\dot{w}^n = \delta v^n + [w^n, v^n], \tag{4.30}$$

$$\delta \dot{v}^n = \delta \lambda_2^{n\sharp}, \tag{4.31}$$

$$\delta \dot{\lambda}_1^n = \mathrm{ad}^*_{\delta v^n} \, \lambda_1^n + \mathrm{ad}^*_{v^n} \, \delta \lambda_1^n, \tag{4.32}$$

$$\delta \dot{\lambda}_2^n = -\delta \lambda_1^n. \tag{4.33}$$

To illustrate NEOC for OCPs for mechanical systems evolving on Lie groups, we now consider an example, which is a slightly modified form of the example presented in [24].

4.4.1 Numerical Example

Consider the following OCP

$$\min_{u(.)} J = \frac{1}{2} \int_0^T \|u(t)\|_F^2 dt \tag{4.34}$$

subject to

$$\dot{Q}(t) = Q(t)\Omega_1(t),\ Q(0) = Q_0,\ Q(T) = Q_T, \tag{4.35}$$

$$\dot{\Omega}_1(t) = u(t),\ \Omega_1(0) = \Omega_{10},\ \Omega_1(T) = \Omega_{1T}, \tag{4.36}$$

where $Q(.) \in C^2([0,T], \mathrm{SO}(n))$ and $\Omega_1(.) \in C^1([0,T], \mathfrak{so}(n))$. A normal extremal for the OCP (4.34)-(4.36) satisfies the following equations (see [24])

$$\dot{Q} = Q\Omega_1, \tag{4.37}$$

$$\dot{\Omega}_1 = \lambda_2, \tag{4.38}$$

$$\dot{\lambda}_1 = -\lambda_1\Omega_1^T, \tag{4.39}$$

$$\dot{\lambda}_2 = -\frac{1}{2}(Q^T\lambda_1 - \lambda_1^T Q), \tag{4.40}$$

where $\lambda_1(.),\ \lambda_2(.) \in C^1([0,T], \mathfrak{so}(n))$ and the optimal control $u^* = \lambda_2$. By hypothesizing a solution of the form $\lambda_1 = Q\Omega_2$, with $\Omega_2(.) \in C^1([0,T], \mathfrak{so}(n))$, (4.37)-(4.40) give the following equations

$$\dot{Q} = Q\Omega_1, \tag{4.41}$$

$$\dot{\Omega}_1 = \lambda_2, \tag{4.42}$$

$$\dot{\Omega}_2 = [\Omega_2, \Omega_1], \tag{4.43}$$

$$\dot{\lambda}_2 = -\Omega_2, \tag{4.44}$$

which are in the form of (4.26)-(4.29). For more details see [24]. We assume that the nominal solution has been obtained for a fixed initial condition $[Q(0)\ \Omega_1(0)]^T = [Q_0\ \Omega_{10}]^T$. Suppose there is a small variation in the initial condition, i.e., $[Q(0)\ \Omega_1(0)]^T = [Q_0\bar{Q}_0\ \Omega_{10} + \bar{\Omega}_{10}]^T$, where $\bar{Q}_0 \in \mathrm{SO}(n)$ and $\bar{\Omega}_{10} \in \mathfrak{so}(n)$. We would now like to update the optimal control for the OCP (4.34)-(4.36). Instead of solving the OCP

(4.34)-(4.36) from scratch, we employ NEOC as described previously. The variational equations for (4.41)-(4.44) are given as follows

$$\dot{w}^n = \delta\Omega_1^n + [w^n, \Omega_1^n], \tag{4.45}$$

$$\delta\dot{\Omega}_1^n = \delta\lambda_2^n, \tag{4.46}$$

$$\delta\dot{\Omega}_2^n = [\delta\Omega_2^n, \Omega_1^n] + [\Omega_2^n, \delta\Omega_1^n], \tag{4.47}$$

$$\delta\dot{\lambda}_2^n = -\delta\Omega_2^n, \tag{4.48}$$

with $w^n(0) = \log(\bar{Q}_0)$, $w^n(T) = 0_{n\times n}$, $\delta\Omega_1^n(0) = \bar{\Omega}_{10}$ and $\delta\Omega_1^n(T) = 0_{n\times n}$. Note that the change in the control trajectory corresponding to the change in the initial condition is given by $\delta\lambda_2^n$. We will now present simulation results for the case when $n = 3$, with $T = 10$ (sec) and the following data

$$\begin{aligned}
Q_0 &= \exp(v_1^\times),\\
Q_T &= \exp(v_2^\times),\\
\bar{Q}_0 &= \exp(v_3^\times),\\
\Omega_{10} &= v_4^\times,\\
\Omega_{1T} &= v_5^\times,\\
\bar{\Omega}_{10} &= v_6^\times,
\end{aligned}$$

with

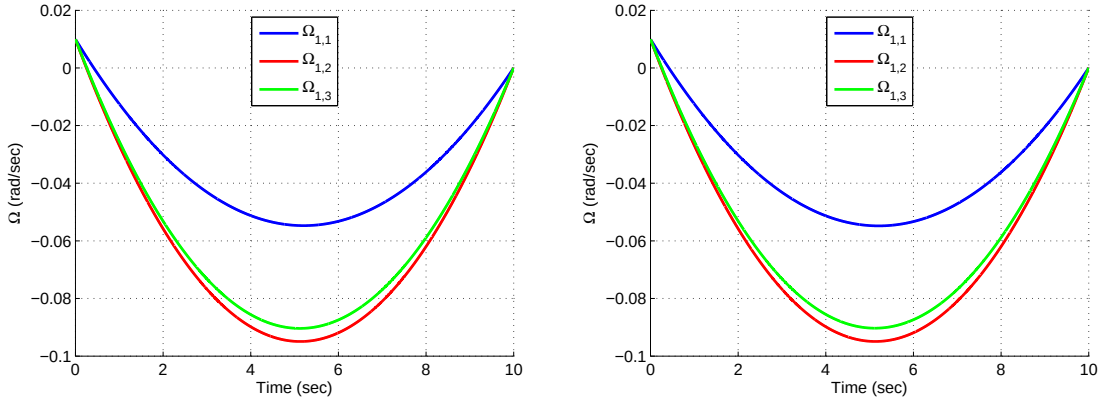
$$\begin{aligned}
v_1 &= [0.25\ \ 0.5\ \ 0.5]^T,\\
v_2 &= [0\ \ 0\ \ 0]^T,\\
v_3 &= [0.1\ \ 0.1\ \ 0.1]^T,\\
v_4 &= [0\ \ 0\ \ 0]^T,
\end{aligned}$$

$$v_5 = [0\ 0\ 0]^T,$$

$$v_6 = [0.01\ 0.01\ 0.01]^T.$$

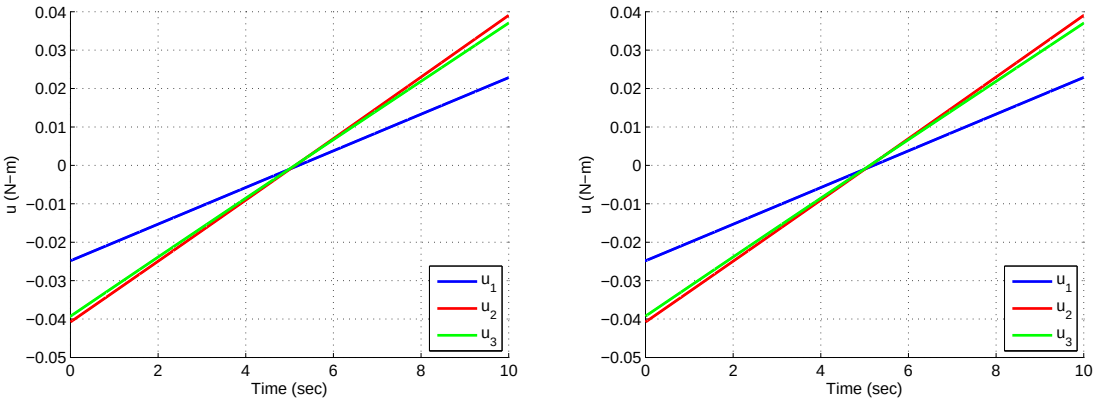
In the subsequent figure (Figure 4.3), the attitude maneuver is plotted on S^2, where the vectors $[x\ y\ z]^T$ corresponding to the first, second and third column of Q_0 are plotted in dashed-red, dashed-green and dashed-blue, respectively. Similarly, the vectors $[x\ y\ z]^T$ corresponding to the first, second and third column of Q_T are plotted in red, green and blue, respectively. For all other $Q(t)$, $t \in (0, T)$, only the coordinates are shown in the corresponding colors.

Figure 4.1 shows the trajectories of Ω_1 obtained from NEOC and by re-solving the OCP (4.34)-(4.36). Figure 4.2 shows the trajectories of u obtained from NEOC and by re-solving the OCP (4.34)-(4.36). Figure 4.3 shows the attitude maneuver obtained from NEOC and by re-solving the OCP (4.34)-(4.36). From Figures 4.1-4.3, one can see that the solution obtained from NEOC is close enough to the solution obtained by re-solving the OCP (4.34)-(4.36).



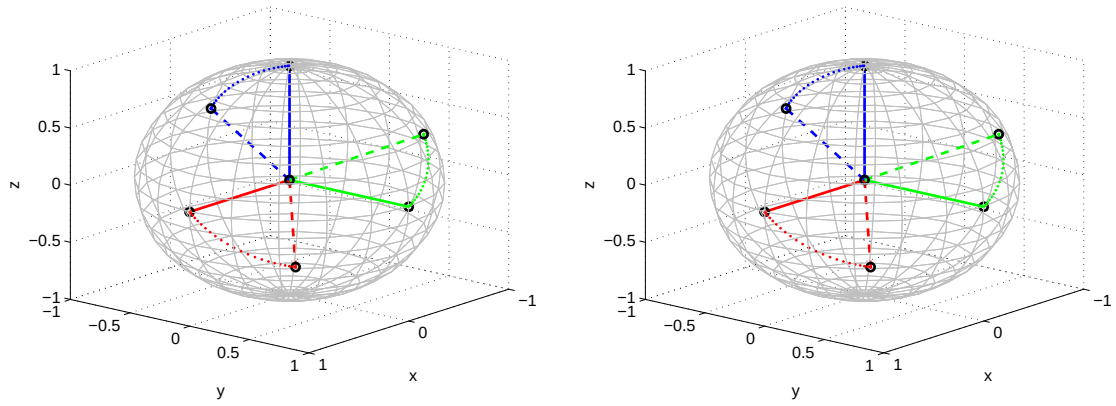
(a) NEOC. (b) Actual.

Figure 4.1: Angular Velocity.



(a) NEOC. (b) Actual.

Figure 4.2: Control Input.



(a) NEOC. (b) Actual.

Figure 4.3: Attitude Maneuver.

CHAPTER V

Optimal Control Problems on Lie Groups with Symmetry Breaking Cost Functions

In this chapter, we investigate the reduction for OCPs on Lie groups with symmetry breaking cost functions. From the Lagrangian point of view, by considering the OCP as a constrained variational problem, we obtain the Euler-Poincaré equations. Furthermore, from the Hamiltonian point of view, we obtain the Lie-Poisson equations. We also study the relation between both formalisms using a reduced Legendre transform.

Several examples are presented, which illustrate the application of the proposed approach. We also develop a variational integrator for OCPs on Lie groups with symmetry breaking cost functions. The resulting variational integrator has the preservation properties of the standard variational integrators. We will now discuss the OCP that will be studied in this chapter. In what follows, we will suppress the explicit dependence of the state, costate and control trajectories on time unless otherwise necessary.

5.1 Optimal Control Problems on Lie Groups

Let G be a n-dimensional Lie group. We will now define a left-invariant control system on G.

Definition V.1. A left-invariant control system on G is given by

$$\dot{g} = T_e L_g(u),$$

where $g(.) \in C^1([0,T],G)$ and u is a curve in the vector space $\mathfrak{g}$. More precisely, if $\mathfrak{g} = \text{span}\{e_1,\ldots,e_m,e_{m+1},\ldots,e_n\}$, then u is given by

$$u(t) = e_0 + \sum_{i=1}^{m} u^i(t)e_i,$$

where the m-tuple of control inputs $[u^1 \ldots u^m]^T$ take values in $\mathbb{R}^m$.

Remark V.2. If $m < n$, then the left-invariant control system is under-actuated otherwise it is fully-actuated.

Consider the following OCP **(P)**

$$\min_{u(.)} J = \int_0^T \left[C(g(t),u(t)) + V(g(t))\right] dt \tag{5.1}$$

subject to

$$\dot{g}(t) = T_e L_{g(t)}(u(t)),\ \ g(0) = g_0,\ \ g(T) = g_T, \tag{5.2}$$

where $C : TG \to \mathbb{R}$ is a G-invariant function, i.e., $C(L_g(h),u) = C(h,u)$, for all $(h,u) \in G \times \mathfrak{g}$ and $V : G \to \mathbb{R}$ (potential function) is not a G-invariant function. We will now study the Euler-Poincaré reduction for **(P)**.

5.1.1 Euler-Poincaré Reduction

We can solve **(P)** as a constrained variational problem using the method of Lagrange multipliers (see, e.g, [9], [57]). The Lagrangian $L : G \times \mathfrak{g} \oplus T^*G \to \mathbb{R}$ for **(P)** is given by

$$L(g, u, \lambda_g) = C(g, u - e_0) + V(g) + \lambda_g(T_e L_g(u - e_0)), \tag{5.3}$$

where $\lambda_g(t) = T^*_g L_{g^{-1}}(\lambda(t)) \in T^*_g G$, with $\lambda(.) \in C^1([0, T], \mathfrak{g}^*)$. Let $\mathfrak{g}^* = \text{span}\{e^1, \ldots, e^m, e^{m+1}, \ldots, e^n\}$, then $\lambda(t) = \lambda_0(t)e^0 + \sum_{i=m+1}^{n} \lambda_i(t)e^i$, where the $(n-m+1)$-tuple of Lagrange multipliers $[\lambda_0 \ \lambda_{m+1} \ldots \lambda_n]^T$ take values in $\mathbb{R}^{n-m+1}$. The reduced Lagrangian $\ell : G \times \mathfrak{g} \oplus \mathfrak{g}^* \to \mathbb{R}$ can now be obtained and is given by

$$\ell(g, u, \lambda) = C(u - e_0) + V(g) + \lambda(u - e_0), \tag{5.4}$$

where with a slight abuse of notation, we write $C(e, u - e_0) = C(u - e_0)$. A normal extremal (see, e.g., [9]) for **(P)** now satisfies the following Euler-Poincaré type equations

$$\frac{d}{dt}(\mathbf{D}_u C + \lambda) = \text{ad}^*_u(\mathbf{D}_u C + \lambda) + T^*_e L_g(\mathbf{D}_g V). \tag{5.5}$$

For more details see [57], [75].

In order to describe the time evolution of u and λ in (5.5), we state the following proposition.

Proposition V.3. *Assume that* $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ *such that*

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{p},$$

$$[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{k},$$

$$[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{p},$$

where $\mathfrak{k} = \text{span}\{e_1, \ldots, e_m\}$ *and* $\mathfrak{p} = \text{span}\{e_0, e_{m+1}, \ldots, e_{n-1}\}$. *The time evolution of* u *and* λ *in (5.5) are given by the following equations*

$$\frac{d}{dt}\mathbf{D}_u C = \text{ad}^*_{e_0}\mathbf{D}_u C + \text{ad}^*_{u_\mathfrak{k}}\lambda + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*},$$
$$\frac{d\lambda}{dt} = \text{ad}^*_{e_0}\lambda + \text{ad}^*_{u_\mathfrak{k}}\mathbf{D}_u C + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*},$$

where $u_\mathfrak{k} = \sum_{i=1}^{m} u^i e_i \in \mathfrak{k}$.

Proof. It is easy to verify that $\mathfrak{g}^* = \mathfrak{k}^* \oplus \mathfrak{p}^*$ such that

$$\text{ad}^*_\mathfrak{k}\,\mathfrak{k}^* \subseteq \mathfrak{p}^*,$$
$$\text{ad}^*_\mathfrak{p}\,\mathfrak{k}^* \subseteq \mathfrak{k}^*,$$
$$\text{ad}^*_\mathfrak{p}\,\mathfrak{p}^* \subseteq \mathfrak{p}^*,$$
$$\text{ad}^*_\mathfrak{k}\,\mathfrak{p}^* \subseteq \mathfrak{k}^*,$$

where $\mathfrak{k}^* = \text{span}\{e^1, \ldots, e^m\}$ and $\mathfrak{p}^* = \text{span}\{e^0, e^{m+1}, \ldots, e^{n-1}\}$. By using the fact that $u = e_0 + u_\mathfrak{k}$, it is also easy to verify that $\mathbf{D}_u C \in \mathfrak{k}^*$. Also, by construction $\lambda \in \mathfrak{p}^*$. We now have a splitting of the left hand side of (5.5) in $\mathfrak{k}^*$ and $\mathfrak{p}^*$. Again, by using the fact that $u = e_0 + u_\mathfrak{k}$, we have the following

$$\text{ad}^*_u\mathbf{D}_u C = \text{ad}^*_{e_0}\mathbf{D}_u C + \text{ad}^*_{u_\mathfrak{k}}\mathbf{D}_u C,$$
$$\text{ad}^*_u\lambda = \text{ad}^*_{e_0}\lambda + \text{ad}^*_{u_\mathfrak{k}}\lambda,$$

where by using the above relations, it is also easy to verify that $\text{ad}^*_{e_0}\mathbf{D}_u C \in \mathfrak{k}^*$, $\text{ad}^*_{u_\mathfrak{k}}\mathbf{D}_u C \in \mathfrak{p}^*$, $\text{ad}^*_{e_0}\lambda \in \mathfrak{p}^*$ and $\text{ad}^*_{u_\mathfrak{k}}\lambda \in \mathfrak{k}^*$. Since, $T^*_e L_g(\mathbf{D}_g V) \in \mathfrak{g}^*$, we define the

following

$$T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*} := \sum_{i=1}^{m} T_e^* L_g(\mathbf{D}_g V) e^i,$$
$$T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*} := T_e^* L_g(\mathbf{D}_g V) e^0 + \sum_{i=m+1}^{n-1} T_e^* L_g(\mathbf{D}_g V) e^i.$$

We now have a splitting of the right hand side of (5.5) in $\mathfrak{k}^*$ and $\mathfrak{p}^*$. So, (5.5) splits into the following equations

$$\frac{d}{dt}\mathbf{D}_u C = \mathrm{ad}^*_{e_0}\mathbf{D}_u C + \mathrm{ad}^*_{u_\mathfrak{k}}\lambda + T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*},$$
$$\frac{d\lambda}{dt} = \mathrm{ad}^*_{e_0}\lambda + \mathrm{ad}^*_{u_\mathfrak{k}}\mathbf{D}_u C + T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*}.$$

□

Remark V.4. Note that semisimple Lie algebras admit a Cartan decomposition, i.e., if $\mathfrak{g}$ is semisimple, then $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ such that

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{p},$$
$$[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{k},$$
$$[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{p},$$

where $\mathfrak{k} = \{x \in \mathfrak{g} \mid \theta(x) = -x\}$ is the -1 eigenspace of the Cartan involution θ and $\mathfrak{p} = \{x \in \mathfrak{g} \mid \theta(x) = x\}$ is the $+1$ eigenspace of the Cartan involution θ. In addition, $\kappa_\mathfrak{g}(.,.)$ is positive definite on $\mathfrak{k}$ and negative definite on $\mathfrak{p}$. So, connected semisimple Lie groups are potential candidates that satisfy the assumption of Proposition V.3. Conversly, a Cartan decomposition (above relations) determines a Cartan involution θ (see, e.g., [52]). For more details see [25], [44], [52]. Also, note that the roles of $\mathfrak{k}$ and $\mathfrak{p}$ can be reversed in Proposition V.3.

We will now present some examples.

5.1.2 Minimum Weighted Input Energy Optimal Control Problem

Consider **(P)**, with e_0 equal to zero and let the cost function be given by

$$C(g,u) = \frac{1}{2}\langle u, I(u)\rangle, \tag{5.6}$$

where $I : \mathfrak{g} \to \mathfrak{g}$ is a linear mapping, with $I \succ 0$. A normal extremal for **(P)**, with e_0 equal to zero and with the cost function (5.7) satisfies the following Euler-Poincaré type equations

$$\begin{aligned} I(\dot{u}) &= \mathrm{ad}_u^* \lambda, \\ \dot{\lambda} &= \mathrm{ad}_u^* I(u). \end{aligned}$$

Note that a similar case is also studied in [57], [58].

5.1.3 Linear Quadratic Regulator Type Problem on SO(3)

Consider **(P)**, with $G = \mathrm{SO}(3)$, e_0 equal to zero and let the cost function be given by

$$C(g,u) = \frac{1}{2}\|R^{\frac{1}{2}}u\|_F^2, \tag{5.7}$$

$$V(g) = \frac{1}{2}\|Q^{\frac{1}{2}}(g - I_{3\times 3})\|_F^2, \tag{5.8}$$

where $Q \succeq 0$ and $R \succ 0$. This is a LQR type problem on SO(3) (see [80]) and it is easy to verify that the potential function V is not invariant under the action of SO(3). Note that $u(t) = \sum_{i=1}^{2} u^i(t) e_i$, where the elements of the basis of $\mathfrak{so}(3)$ are given

by

$$e_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \; e_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Also, note that $\mathfrak{so}(3)$ is semisimple and the elements of the basis of $\mathfrak{so}(3)$ satisfy the following relations

$$[e_1, e_2] = e_3,$$
$$[e_2, e_3] = e_1,$$
$$[e_3, e_1] = e_2.$$

It is also easy to verify that with $\mathfrak{k} = \text{span}\{e_1, e_2\}$ and $\mathfrak{p} = \text{span}\{e_3\}$, $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ such that

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{p},$$
$$[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{k},$$
$$[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{p}.$$

Under the trace pairing, the elements of the basis of $\mathfrak{so}(3)^*$ are given by

$$e^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}, \; e^2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{bmatrix}, \; e^3 = \begin{bmatrix} 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We will now assume that $R = I_{3\times 3}$, for the ease of computations. A normal extremal for **(P)**, with $G = \text{SO}(3)$, e_0 equal to zero and with the cost function (5.7)-(5.8),

satisfies the following Euler-Poincaré type equations

$$\dot{u} = \mathrm{ad}^*_{u_\mathfrak{k}} \lambda + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*}, \tag{5.9}$$

$$\dot{\lambda} = \mathrm{ad}^*_{u_\mathfrak{k}} u + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*}, \tag{5.10}$$

where

$$T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*} = \begin{bmatrix} 0 & 0 & \frac{1}{2}[(g^TQ)_A]_{13} \\ 0 & 0 & -\frac{1}{2}[(g^TQ)_A]_{23} \\ \frac{1}{2}[(g^TQ)_A]_{31} & -\frac{1}{2}[(g^TQ)_A]_{32} & 0 \end{bmatrix},$$

$$T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*} = \begin{bmatrix} 0 & -\frac{1}{2}[(g^TQ)_A]_{12} & 0 \\ -\frac{1}{2}[(g^TQ)_A]_{21} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The coadjoint action of $\mathfrak{so}(3)$ on $\mathfrak{so}(3)^*$ is given by

$$\mathrm{ad}^*_\xi \mu = \mu \times \xi,$$

for $\xi \in \mathbb{R}^3 \cong \mathfrak{so}(3)$ and $\mu \in \mathbb{R}^3 \cong \mathfrak{so}(3)^*$. For more details see [45]. If we write $\lambda = \lambda_3 e^3$, then we have $\mathrm{ad}^*_{[u^1 \ u^2 \ 0]^T}[0 \ 0 \ \lambda_3]^T = [-u^2\lambda_3 \ u^1\lambda_3 \ 0]^T$, which gives

$$\mathrm{ad}^*_{u_\mathfrak{k}} \lambda = \begin{bmatrix} 0 & 0 & \frac{1}{2}u^1\lambda_3 \\ 0 & 0 & \frac{1}{2}u^2\lambda_3 \\ -\frac{1}{2}u^1\lambda_3 & -\frac{1}{2}u^2\lambda_3 & 0 \end{bmatrix}.$$

Similarly, $\mathrm{ad}^*_{[u^1 \ u^2 \ 0]^T}[u^1 \ u^2 \ 0]^T = [0 \ 0 \ 0]^T$, which gives $\mathrm{ad}^*_{u_\mathfrak{k}} u = 0_{3\times 3}$. So, (5.9)-(5.10) give the following equations

$$\dot{u}^1 = -\frac{1}{2}(u^2\lambda_3 - [(g^TQ)_A]_{23}), \tag{5.11}$$

$$\dot{u}^2 = \frac{1}{2}(u^1\lambda_3 + [(g^TQ)_A]_{13}), \tag{5.12}$$

$$\dot{\lambda}_3 = [(g^TQ)_A]_{12}. \tag{5.13}$$

Note that if the potential function is identically equal to zero, i.e., $V(g) = 0$, for all $g \in G$, then (5.11)-(5.13) reduce to

$$\dot{u}^1 = -\frac{1}{2}u^2\lambda_3, \tag{5.14}$$

$$\dot{u}^2 = \frac{1}{2}u^1\lambda_3, \tag{5.15}$$

$$\dot{\lambda}_3 = 0. \tag{5.16}$$

Also, note that the solution for (5.14)-(5.16) is given by

$$\begin{bmatrix} u^1(t) \\ u^2(t) \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{\omega t}{2}\right) & -\sin\left(\frac{\omega t}{2}\right) \\ \sin\left(\frac{\omega t}{2}\right) & \cos\left(\frac{\omega t}{2}\right) \end{bmatrix} \begin{bmatrix} u^1(0) \\ u^2(0) \end{bmatrix},$$

where $\lambda_3 = \omega$ is a constant.

5.1.4 Motion Planning of a Unicycle with Obstacles

We study the OCP for the motion planning of a unicycle with obstacles. To avoid the obstacle, we use the navigation function approach (see, e.g., [50], [56]), which plays the role of the potential function in the cost function of the OCP.

The unicycle is a homogeneous disk on a horizontal plane and it is equivalent to a wheel rolling on a plane. The configuration of the unicycle at any given time is completely determined by the element $g \in \mathrm{SE}(2) \cong \mathbb{R}^2 \times \mathrm{S}^1 \cong \mathbb{R}^2 \times \mathrm{SO}(2)$ (as a set)

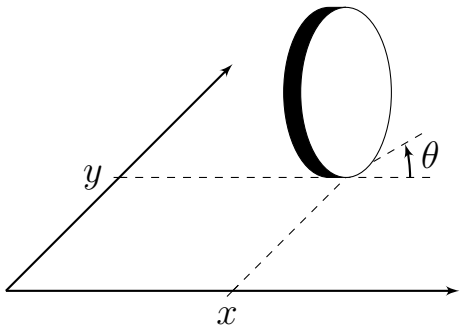


Figure 5.1: The Unicycle.

given by

$$g = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix},$$

where $[x\ y]^T \in \mathbb{R}^2$ represents the point of contact of the wheel with the ground and $\theta \in \mathrm{S}^1$ represents the angular orientation of the overall system (see Figure 5.1). For more details see [9], [58]. The controlled equations for the unicycle are given by

$$\dot{x} = u^2 \cos\theta, \tag{5.17}$$

$$\dot{y} = u^2 \sin\theta, \tag{5.18}$$

$$\dot{\theta} = u^1. \tag{5.19}$$

Note that (5.17)-(5.18) are equivalent to the nonholonomic constraint $\dot{x}\sin\theta - \dot{y}\cos\theta = 0$. Also, note that (5.17)-(5.19) can be viewed as a left-invariant control system on SE(2) (see [58]). A navigation function is a potential field based function used to model an obstacle as a repulsive area or surface. Let the obstacle be circular in shape and be located in the x-y plane, with its center located at the point (x_c, y_c).

Let the potential function $V : \mathbb{R}^2 \backslash \{(x_c, y_c)\} \to \mathbb{R}$ be given by

$$V(x, y) = \frac{1}{2} \frac{k}{(x - x_c)^2 + (y - y_c)^2},$$

where $k \in \mathbb{R}_+$. Equivalently, the potential function $V : \mathrm{SE}(2) \backslash \{g_c\} \to \mathbb{R}$ is given by

$$V(g) = \frac{1}{2} \frac{k}{\|g_c g\|_F^2 - 3},$$

where

$$g_c = \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{SE}(2).$$

It is easy to verify that the potential function V is invariant under the action of $\mathrm{SO}(2)$ but not under the action of $\mathrm{SE}(2)$. With the above motivation, we now consider **(P)**, with $G = \mathrm{SE}(2)$, e_0 equal to zero and with the cost function given by

$$C(g, u) = \frac{1}{2} \|u\|_F^2, \tag{5.20}$$

$$V(g) = \frac{1}{2} \frac{k}{\|g_c g\|_F^2 - 3}. \tag{5.21}$$

Note that $u(t) = \sum_{i=1}^{2} u^i(t) e_i$, where the elements of the basis of $\mathfrak{se}(2)$ are given by

$$e_1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Also, note that $\mathfrak{se}(2)$ is not semisimple and the elements of the basis of $\mathfrak{se}(2)$ satisfy

the following relations

$$[e_1, e_2] = e_3,$$
$$[e_2, e_3] = 0,$$
$$[e_3, e_1] = e_2,$$

It is also easy to verify that with $\mathfrak{k} = \text{span}\{e_1, e_2\}$ and $\mathfrak{p} = \text{span}\{e_3\}$, $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ such that

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{p},$$
$$[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{k},$$
$$[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{p}.$$

Under the trace pairing, the elements of the basis of $\mathfrak{se}(2)^*$ are given by

$$e^1 = \begin{bmatrix} 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; e^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; e^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

A normal extremal for **(P)**, with $G = \text{SE}(2)$, e_0 equal to zero and with the cost function (5.20)-(5.21), satisfies the following Euler-Poincaré type equations

$$\dot{u} = \text{ad}^*_{u_\mathfrak{k}} \lambda + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*}, \tag{5.22}$$
$$\dot{\lambda} = \text{ad}^*_{u_\mathfrak{k}} u + T^*_e L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*}, \tag{5.23}$$

where

$$T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{k}^*} = \begin{bmatrix} 0 & 0 & -\dfrac{k([g^T g_c^T g_c g]_{13})}{(\|g_c g\|_F^2 - 3)^2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$T_e^* L_g(\mathbf{D}_g V)\big|_{\mathfrak{p}^*} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\dfrac{k([g^T g_c^T g_c g]_{23})}{(\|g_c g\|_F^2 - 3)^2} \\ 0 & 0 & 0 \end{bmatrix}.$$

The coadjoint action of $\mathfrak{se}(2)$ on $\mathfrak{se}(2)^*$ is given by

$$\mathrm{ad}^*_{[\xi \ \alpha^T]^T}[\mu \ \beta^T]^T = [\langle \alpha, -\mathbb{J}\beta \rangle \ (\xi \mathbb{J}\beta)^T]^T,$$

where

$$\mathbb{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$[\xi \ \alpha^T]^T \in \mathbb{R}^3 \cong \mathfrak{se}(2)$ and $[\mu \ \beta^T]^T \in \mathbb{R}^3 \cong \mathfrak{se}(2)^*$. For more details see [68]. If we write $\lambda = \lambda_3 e^3$, then we have $\mathrm{ad}^*_{[u^1 \ u^2 \ 0]^T}[0 \ 0 \ \lambda_3]^T = [-u^2\lambda_3 \ u^1\lambda_3 \ 0]^T$, which gives

$$\mathrm{ad}^*_{u_\mathfrak{k}} \lambda = \begin{bmatrix} 0 & \dfrac{1}{2}u^2\lambda_3 & u^1\lambda_3 \\ -\dfrac{1}{2}u^2\lambda_3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Similarly, $\mathrm{ad}^*_{[u^1 \;\; u^2 \;\; 0]^T}[u^1 \;\; u^2 \;\; 0]^T = [0 \;\; 0 \;\; -u^1u^2]^T$, which gives

$$\mathrm{ad}^*_{u_\mathfrak{k}} u = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -u^1u^2 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, (5.22)-(5.23) give the following equations

$$\dot{u}^1 = -\frac{1}{2}u^2\lambda_3, \tag{5.24}$$

$$\dot{u}^2 = u^1\lambda_3 - \frac{k([g^T g_c^T g_c g]_{13})}{(\|g_c g\|_F^2 - 3)^2}, \tag{5.25}$$

$$\dot{\lambda}_3 = -u^1u^2 - \frac{k([g^T g_c^T g_c g]_{23})}{(\|g_c g\|_F^2 - 3)^2}. \tag{5.26}$$

Note that if the potential function is identically equal to zero, i.e., $V(g) = 0$, for all $g \in G$, then (5.24)-(5.26) reduce to

$$\dot{u}^1 = -\frac{1}{2}u^2\lambda_3, \tag{5.27}$$

$$\dot{u}^2 = u^1\lambda_3, \tag{5.28}$$

$$\dot{\lambda}_3 = -u^1u^2. \tag{5.29}$$

Remark V.5. We will now show that (5.24)-(5.26) are equivalently obtained by viewing **(P)**, with $G = \mathrm{SE}(2)$, e_0 equal to zero and with the cost function (5.20)-(5.21), as a constrained variational problem. The above OCP is equivalent to the following constrained variational problem

$$\min_{([x(.) \;\; y(.)]^T, \theta(.))} J = \frac{1}{2}\int_0^T \left(\dot{x}^2(t) + \dot{y}^2(t) + 2\dot{\theta}^2(t) + \frac{k}{(x(t) - x_c)^2 + (y(t) - y_c)^2} \right) dt \tag{5.30}$$

subject to

$$\dot{x}(t)\sin\theta(t) - \dot{y}(t)\cos\theta(t) = 0, \text{ with given boundary conditions } ([x(0)\ y(0)]^T, \theta(0)) \text{ and } ([x(T)\ y(T)]^T, \theta(T)). \tag{5.31}$$

The Lagrangian for the constrained variational problem (5.30)-(5.31) is given by

$$L(\theta, \dot{x}, \dot{y}, \dot{\theta}, \lambda) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + 2\dot{\theta}^2) + \frac{1}{2}\frac{k}{(x-x_c)^2 + (y-y_c)^2} + \lambda(\dot{y}\cos\theta - \dot{x}\sin\theta),$$

where λ is the Lagrange multiplier. A solution for the constrained variational problem (5.30)-(5.31) must satisfy the following Euler-Lagrange equations

$$\ddot{x} - \dot{\lambda}\sin\theta - \lambda\dot{\theta}\cos\theta = -\frac{k(x-x_C)}{((x-x_c)^2 + (y-y_c)^2)^2}, \tag{5.32}$$

$$\ddot{y} + \dot{\lambda}\cos\theta - \lambda\dot{\theta}\sin\theta = -\frac{k(y-y_C)}{((x-x_c)^2 + (y-y_c)^2)^2}, \tag{5.33}$$

$$\ddot{\theta} = -\frac{1}{2}\lambda(\dot{x}\cos\theta + \dot{y}\sin\theta). \tag{5.34}$$

Using the facts that $\dot{x}\sin\theta - \dot{y}\cos\theta = 0$, $u^1 = \dot{\theta}$, $u^2 = \dot{x}\cos\theta + \dot{y}\sin\theta$ and after a few simple calculations, (5.32)-(5.34) give the following equations

$$\dot{u}^1 = -\frac{1}{2}u^2\lambda, \tag{5.35}$$

$$\dot{u}^2 = u^1\lambda - \frac{k([g^T g_c^T g_c g]_{13})}{(\|g_c g\|_F^2 - 3)^2}, \tag{5.36}$$

$$\dot{\lambda} = -u^1 u^2 - \frac{k([g^T g_c^T g_c g]_{23})}{(\|g_c g\|_F^2 - 3)^2}. \tag{5.37}$$

We can now see that (5.24)-(5.26) are the same as (5.35)-(5.37).

We will now use the reduced Legendre transform to derive the Lie-Poisson type equations associated with **(P)**.

5.1.5 Reduced Legendre Transform and Lie-Poisson Type Equations

Consider the reduced Lagrangian $\ell(g, u, \lambda) = C(u - e_0) + V(g) + \lambda(u - e_0)$. If the reduced Lagrangian ℓ is hyper-regular, then we can define the reduced Legendre transform (see, e.g., [45], [68]) to obtain a reduced Hamiltonian $h : G \times \mathfrak{g}^* \oplus \mathfrak{g}^* \to \mathbb{R}$ given by

$$h(g, \mu, \lambda) = \mu(u) - \ell(g, u, \lambda), \tag{5.38}$$

where $\mu = \mathbf{D}_u \, \ell = \mathbf{D}_u \, C + \lambda \in \mathfrak{g}^*$. The Euler-Poincaré type equations for the reduced Lagrangian ℓ can now be written as the Lie-Poisson type equations given below

$$\dot{\mu} = \mathrm{ad}^*_u \, \mu + T^*_e L_g(\mathbf{D}_g \, V). \tag{5.39}$$

We will now study the Lie-Poisson reduction for **(P)** using PMP.

5.1.6 Lie-Poisson Reduction

Define the augmented cost functional as follows

$$J^a = \int_0^T [C(g(t), u(t)) + V(g(t)) + \mu_g(t)(\dot{g}(t) - T_e L_{g(t)}(u(t)))]dt,$$

where $\mu_g(t) = T^*_g L_{g^{-1}}(\mu(t)) \in T^*_g G$, with $\mu(.) \in C^1([0, T], \mathfrak{g}^*)$. We now introduce the Hamiltonian $\bar{H} : G \times \mathfrak{g} \oplus T^*G \to \mathbb{R}$ given by

$$\bar{H}(g, u, \mu_g) = \mu_g(T_e L_g(u)) - C(g, u) - V(g),$$

to rewrite the augmented cost functional as

$$J^a = \int_0^T [\mu_g(t)(\dot{g}(t)) - \bar{H}(g(t), u(t), \mu_g(t))]dt.$$

By PMP, we can obtain the optimal Hamiltonian $H : T^*G \to \mathbb{R}$ given by

$$H(g, \mu_g) = \max_u \bar{H}(g, u, \mu_g) = \bar{H}(g, u^*, \mu_g), \tag{5.40}$$

where u^* denotes the optimal control. The reduced Hamiltonian $h : G \times \mathfrak{g}^* \to \mathbb{R}$ can now be obtained and is given by

$$h(g, \mu) = \mu(u^*) - C(u^*) - V(g). \tag{5.41}$$

A normal extremal for **(P)** now satisfies the following Lie-Poisson type equations

$$\dot{\mu} = \mathrm{ad}^*_{u^*} \mu + T^*_e L_g(\mathbf{D}_g V). \tag{5.42}$$

For more details see [58], [75]. We will now present some examples.

5.1.7 Linear Quadratic Regulator Type Problem on SO(3) Revisited

This example was studied in Section 5.1.3. By PMP, we have the following

$$u^{1*} = \frac{1}{2}\mu_1,$$
$$u^{2*} = \frac{1}{2}\mu_2.$$

A normal extremal for **(P)**, with $G = \mathrm{SO}(3)$, e_0 equal to zero and with the cost function (5.7)-(5.8), satisfies the following Lie-Poisson type equations

$$\dot{\mu} = \mathrm{ad}^*_{u^*} \mu + T^*_e L_g(\mathbf{D}_g V). \tag{5.43}$$

We now have $\mathrm{ad}^*_{[u^{1*}\ u^{2*}\ 0]^T}[\mu_1\ \mu_2\ \mu_3]^T = \left[-\frac{1}{2}\mu_2\mu_3\ \ \frac{1}{2}\mu_1\mu_3\ \ 0\right]^T$ and so, (5.43) gives the following equations

$$\dot{\mu}_1 = -\frac{1}{2}(\mu_2\mu_3 - [(g^TQ)_A]_{23}), \tag{5.44}$$

$$\dot{\mu}_2 = \frac{1}{2}(\mu_1\mu_3 + [(g^TQ)_A]_{13}), \tag{5.45}$$

$$\dot{\mu}_3 = [(g^TQ)_A]_{12}. \tag{5.46}$$

Note that if the potential function is identically equal to zero, i.e., $V(g) = 0$, for all $g \in G$, then (5.44)-(5.46) reduce to

$$\dot{\mu}_1 = -\frac{1}{2}\mu_2\mu_3, \tag{5.47}$$

$$\dot{\mu}_2 = \frac{1}{2}\mu_1\mu_3, \tag{5.48}$$

$$\dot{\mu}_3 = 0. \tag{5.49}$$

Also, note that the solution for (5.47)-(5.49) is given by

$$\begin{bmatrix}\mu^1(t)\\ \mu^2(t)\end{bmatrix} = \begin{bmatrix}\cos\left(\frac{\omega t}{2}\right) & -\sin\left(\frac{\omega t}{2}\right)\\ \sin\left(\frac{\omega t}{2}\right) & \cos\left(\frac{\omega t}{2}\right)\end{bmatrix}\begin{bmatrix}\mu^1(0)\\ \mu^2(0)\end{bmatrix},$$

where $\mu_3 = \omega$ is a constant.

5.1.8 Motion Planning of a Unicycle with Obstacles Revisited

This example was studied in Section 5.1.4. By PMP, we have the following

$$u^{1*} = \frac{1}{2}\mu_1,$$

$$u^{2*} = \mu_2.$$

A normal extremal for **(P)**, with $G = \mathrm{SE}(2)$, e_0 equal to zero and with the cost function (5.20)-(5.21), satisfies the following Lie-Poisson type equations

$$\dot{\mu} = \mathrm{ad}^*_{u^*}\, \mu + T^*_e L_g(\mathbf{D}_g\, V). \tag{5.50}$$

We now have $\mathrm{ad}^*_{[u^{1*}\ u^{2*}\ 0]^T}[\mu_1\ \mu_2\ \mu_3]^T = \left[-\mu_2\mu_3\ \ \frac{1}{2}\mu_1\mu_3\ \ -\frac{1}{2}\mu_1\mu_2\right]^T$ and so, (5.50) gives the following equations

$$\dot{\mu}_1 = -\mu_2\mu_3, \tag{5.51}$$

$$\dot{\mu}_2 = \frac{1}{2}\mu_1\mu_3 - \frac{k([g^T g_c^T g_c g]_{13})}{(\|g_c g\|_F^2 - 3)^2}, \tag{5.52}$$

$$\dot{\mu}_3 = -\frac{1}{2}\mu_1\mu_2 - \frac{k([g^T g_c^T g_c g]_{23})}{(\|g_c g\|_F^2 - 3)^2}. \tag{5.53}$$

Note that if the potential function is identically equal to zero, i.e., $V(g) = 0$, for all $g \in G$, then (5.51)-(5.54) reduce to

$$\dot{\mu}_1 = -\mu_2\mu_3, \tag{5.54}$$

$$\dot{\mu}_2 = \frac{1}{2}\mu_1\mu_3, \tag{5.55}$$

$$\dot{\mu}_3 = -\frac{1}{2}\mu_1\mu_2. \tag{5.56}$$

Also, note that (5.54)-(5.56) are exactly the same as the equations obtained in [58], with the cost function given by

$$C(g, u) = \frac{1}{2}\|u\|_F^2.$$

We will now describe the variational integrator for **(P)**.

5.2 Variational Integrator for Optimal Control Problems on Lie Groups

Recall that the augmented cost functional is given by

$$J^a = \int_0^T [C(g(t), u(t)) + V(g(t)) + \mu_g(t)(\dot{g}(t) - T_e L_{g(t)}(u(t)))]dt.$$

The discrete-time reduced augmented cost functional can now be obtained and is given by

$$J_d^a = \sum_{k=0}^{N-1} h \left[C(u_k) + V(g_k) + \mu_k \left(\frac{1}{h} \tau^{-1}(g_k^{-1} g_{k+1}) - u_k \right) \right],$$

where $h \in \mathbb{R}_+$ is the time step and $Nh = T$. In order to obtain the variational integrator for **(P)**, we will use discrete-time variational calculus. The variation of g_k is given as follows

$$g_{k,\epsilon} = g_k \exp(\epsilon \eta_k), \tag{5.57}$$

where $\eta_k \in \mathfrak{g}$. The infinitesimal variation of g_k is given by

$$\begin{aligned} \delta g_k &= \left. \frac{d g_{k,\epsilon}}{d\epsilon} \right|_{\epsilon=0}, \\ &= g_k \eta_k. \end{aligned} \tag{5.58}$$

Before proceeding further, we need a few facts.

Fact 5. ([12], [53], [55]) $\frac{1}{h} \delta \tau^{-1}(g_k^{-1} g_{k+1}) = \frac{1}{h} \mathbf{d}\, \tau_{hu_k}^{-1}(-\eta_k + \mathrm{Ad}_{\tau(hu_k)}\, \eta_{k+1})$.

Fact 6. ([12]) $\mathbf{d}\, \tau_{\xi_1}^{-1}(\xi_2) = \mathbf{d}\, \tau_{-\xi_1}^{-1}(\mathrm{Ad}_{\tau(-\xi_1)}\, \xi_2)$, for $\xi_1, \xi_2 \in \mathfrak{g}$.

Using Facts 5-6, the variation of the discrete-time reduced augmented cost functional

is written as follows

$$\begin{aligned}
\delta J_d^a &= \sum_{k=0}^{N-1} h\Big[\mathbf{D}_{u_k} C(\delta u_k) + \mathbf{D}_{g_k} V(\delta g_k) + \mu_k \left(\frac{1}{h} \delta \tau^{-1}(g_k^{-1} g_{k+1}) - \delta u_k \right) \Big], \\
&= \sum_{k=0}^{N-1} h\Big[\mathbf{D}_{g_k} V(g_k \eta_k) + \mu_k \left(\frac{1}{h} \mathbf{d}\, \tau_{hu_k}^{-1}(-\eta_k + \mathrm{Ad}_{\tau(hu_k)}\, \eta_{k+1}) \right) + \\
&\qquad (-\mu_k + \mathbf{D}_{u_k} C)(\delta u_k) \Big], \\
&= \sum_{k=0}^{N-1} h\Big[T_e^* L_g\, \mathbf{D}_{g_k} V(\eta_k) + \mu_k \left(\frac{1}{h} \mathbf{d}\, \tau_{hu_k}^{-1}(-\eta_k + \mathrm{Ad}_{\tau(hu_k)}\, \eta_{k+1}) \right) + \\
&\qquad (-\mu_k + \mathbf{D}_{u_k} C)(\delta u_k) \Big], \\
&= \sum_{k=0}^{N-1} \Big[(-(\mathbf{d}\, \tau_{hu_k}^{-1})^* \mu_k + (\mathbf{d}\, \tau_{-hu_{k-1}}^{-1})^* \mu_{k-1} + h T_e^* L_g\, \mathbf{D}_{g_k} V)(\eta_k) + \\
&\qquad h(-\mu_k + \mathbf{D}_{u_k} C)(\delta u_k) \Big],
\end{aligned}$$

where the analogue of integration by parts in the discrete-time setting is used along with the fact that the variation η_k vanishes at $k = 0$, N. Since, δJ_d^a should vanish for all variations of η_k and δu_k, the necessary conditions for optimality are as follows

$$g_{k+1} = g_k \tau(u_k), \tag{5.59}$$

$$(\mathbf{d}\, \tau_{hu_k}^{-1})^* \mu_k = (\mathbf{d}\, \tau_{-hu_{k-1}}^{-1})^* \mu_{k-1} + h T_e^* L_g\, \mathbf{D}_{g_k} V, \tag{5.60}$$

$$\mu_k = \mathbf{D}_{u_k} C. \tag{5.61}$$

Note that if the potential function is identically equal to zero, i.e., $V(g) = 0$, for all $g \in G$, then (5.59)-(5.61) reduce to

$$g_{k+1} = g_k \tau(u_k), \tag{5.62}$$

$$(\mathbf{d}\, \tau_{hu_k}^{-1})^* \mu_k = (\mathbf{d}\, \tau_{-hu_{k-1}}^{-1})^* \mu_{k-1}, \tag{5.63}$$

$$\mu_k = \mathbf{D}_{u_k} C. \tag{5.64}$$

CHAPTER VI

Conclusions and Future Work

This dissertation has focused on extending some of the existing analytical and numerical methods for OCPs on manifolds and Lie groups. The research not only addressed OCPs defined on a Euclidean space but also on Riemannian manifolds. In particular, we considered four different problems. The first problem dealt with obtaining sub-optimal control in OCPs defined on a Euclidean space using the combination of two techniques, homotopy and NEOC. The second problem dealt with constrained spacecraft attitude control on SO(3) using fast NMPC. The third problem dealt with extending NEOC for mechanical systems on Riemannian manifolds. The fourth problem dealt with OCPs on Lie groups with symmetry breaking cost functions.

6.1 Conclusions

The main results of this dissertation are summarized below.

(a) In Chapter II, we described a method for obtaining sub-optimal control in OCPs defined on a Euclidean space, that is based on the combined use of homotopy and NEOC, which to the author's knowledge has not been reported in the previous literature. This approach was illustrated using a numerical example, which

suggested the benefits of the combined use of homotopy and NEOC, in terms of reducing the number of function evaluations and iterations.

(b) In Chapter III, we described the implementation of a numerical solver for NMPC of spacecraft attitude that exploits the underlying Lie group structure of SO(3) and the geometric control formalism. The numerical solver is based on numerically solving the necessary conditions for optimality. The control input/state constraints are handled through the exterior penalty function approach. This work compliments [49] which addressed the NMPC problem formulation and the stability analysis but used a baseline solver for numerical computations which was of direct type and relied on the conventional constrained optimizer in `MATLAB` (`fmincon.m`). The simulation results indicate that the numerical solver we have implemented is faster than the baseline solver and enables the spacecraft to perform a variety of constrained reorientation maneuvers. We also extended the classical penalty convergence theorem to the setting of smooth manifolds and the classical exact penalization theorem to the setting of Riemannian manifolds.

(c) In Chapter IV, we extended NEOC, which is well established for OCPs defined on a Euclidean space, to the setting of Riemannian manifolds. We further specialized the results to the case of Lie groups. We also presented an example along with simulation results.

(d) In Chapter V, we investigated the reduction for OCPs on Lie groups with symmetry breaking cost functions. From the Lagrangian point of view, we obtained the Euler-Poincaré equations and from the Hamiltonian point of view, we obtained the Lie-Poisson equations. We also study the relationship between both formalisms and present several examples. A variational integrator for OCPs on Lie groups with symmetry breaking cost functions is also developed.

6.2 Future Work

The possible future directions are given below.

(a) In the future, we intend to investigate the use of the method described in Chapter II for more complicated control input/state constrained OCPs. We also intend to test numerical examples for the predictor-corrector method described in Chapter II.

(b) The numerical solver implementation in `MATLAB` described in Chapter III is currently slower than real-time but the implementation in `C/C++` is expected to be faster and further computational improvements will be pursued in future research. Extensions of NMPC to mechanical systems evolving on other Lie groups, e.g., $\mathrm{SE}(3) = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$, etc., use of other indirect methods and the integration with continuation methods will also be pursued in future research.

(c) NEOC described in Chapter IV only gives a prediction step and not a correction step. To improve the solution, a prediction step can be augmented by a correction step. In the future, we intend to extend the idea presented in Chapter IV to include a correction step as well along with the generalization to a more general cost function, with a more complicated dynamic constraint.

(d) The idea presented in Chapter V can be taken a step further, if one assumes that the potential function V is invariant under the action of a subgroup of G but not under the action of G (see, e.g., [11], [21], [70] and also Section 5.1.4). Also, in the future, we intend to test numerical examples for the variational integrator developed in Chapter V.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank, *Matrix Riccati equations in control and systems theory*. Birkhäuser, 2012.

[2] R. Abraham and J. E. Marsden, *Foundations of mechanics*. AMS Chelsea Publishing, 2008.

[3] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*. Springer Science & Business Media, 2004.

[4] E. L. Allgower and K. Georg, *Numerical continuation methods: an introduction*. Springer Science & Business Media, 1990.

[5] C. Altafini, "Reduction by group symmetry of second order variational problems on a semidirect product of Lie groups with positive definite Riemannian metric," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 10, no. 4, pp. 526–548, 2004.

[6] J. Z. Ben-Asher, *Optimal control theory with aerospace applications*. American Institute of Aeronautics and Astronautics, 2010.

[7] R. Bertrand and R. Epenoy, "New smoothing techniques for solving bang-bang optimal control problemsnumerical results and statistical interpretation," *Optimal Control Applications and Methods*, vol. 23, no. 4, pp. 171–197, 2002.

[8] J. T. Betts, "Survey of numerical methods for trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 21, no. 2, pp. 193–207, 1998.

[9] A. M. Bloch, *Nonholonomic mechanics and control*. Springer Science & Business Media, 2015.

[10] J. Bonilla, M. Diehl, F. Logist, B. De Moor, and J. F. Van Impe, "A convexity-based homotopy method for nonlinear optimization in model predictive control," *Optimal Control Applications and Methods*, vol. 31, no. 5, pp. 393–414, 2010.

[11] A. D. Borum and T. Bretl, "Geometric optimal control for symmetry breaking cost functions," *Proceedings of IEEE Conference on Decision and Control*, 2014, pp. 5855–5861.

[12] N. Bou-Rabee, *Hamilton-Pontryagin integrators on Lie groups*. PhD Dissertation, California Institute of Technology, 2007.

[13] J. V. Breakwell and H. Yu-Chi, "On the conjugate point condition for the control problem," *International Journal of Engineering Science*, vol. 2, no. 6, pp. 565–579, 1965.

[14] A. E. Bryson, *Applied optimal control: optimization, estimation and control*. CRC Press, 1975.

[15] F. Bullo, *Invariant affine connections and controllability on Lie groups*. Final Project Report for CIT-CDS 141a, California Institute of Technology, 1995.

[16] F. Bullo, *Geometric control of mechanical systems: modeling, analysis, and design for simple mechanical control systems*. Springer Science & Business Media, 2005.

[17] F. Bullo and A. D. Lewis, "Reduction, linearization, and stability of relative equilibria for mechanical systems on Riemannian manifolds," *Acta Applicandae Mathematicae*, vol. 99, no. 1, pp. 53–95, 2007.

[18] J.-B. Caillau, O. Cots, and J. Gergaud, "Differential continuation for regular optimal control problems," *Optimization Methods and Software*, vol. 27, no. 2, pp. 177–196, 2012.

[19] J. R. Cardoso and F. Silva Leite, "The moser-veselov equation," *Linear Algebra and its Applications*, vol. 360, pp. 237–248, 2003.

[20] N. Caroff and H. Frankowska, "Conjugate points and shocks in nonlinear optimal control," *Transactions of the American Mathematical Society*, vol. 348, no. 8, pp. 3133–3153, 1996.

[21] H. Cendra, J. E. Marsden, and T. S. Ratiu, *Lagrangian reduction by stages*. American Mathematical Society, 2001.

[22] G. S. Chirikjian, *Stochastic models, information theory, and Lie groups, volume 2: analytic methods and modern applications*. Springer Science & Business Media, 2012.

[23] P. Crouch and F. Silva Leite, "The dynamic interpolation problem: on Riemannian manifolds, Lie groups, and symmetric spaces," *Journal of Dynamical and Control Systems*, vol. 1, no. 2, pp. 177–202, 1995.

[24] P. Crouch, F. Silva Leite, and M. Camarinha, *A second order Riemannian variational problem from a Hamiltonian perspective*. Pré-publicações do Departamento de Matemática da Universidade de Coimbra, 1998.

[25] D. D'Alessandro, *Introduction to quantum control and dynamics*. CRC press, 2007.

[26] E. J. Davison and M. C. Maki, "The numerical solution of the matrix riccati differential equation," *IEEE Transactions on Automatic Control*, vol. 18, no. 1, pp. 71–73, 1973.

[27] M. Diehl, H. G. Bock, and J. P. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control," *SIAM Journal on Control and Optimization*, vol. 43, no. 5, pp. 1714–1736, 2005.

[28] M. Diehl, R. Findeisen, F. Allgöwer, H. G. Bock, and J. P. Schlöder, "Nominal stability of real-time iteration scheme for nonlinear model predictive control," *IEE Proceedings-Control Theory and Applications*, vol. 152, no. 3, pp. 296–308, 2005.

[29] M. P. do Carmo, *Riemannian geometry*. Birkhäuser, 1992.

[30] A. L. Dontchev and W. W. Hager, "Lipschitzian stability in nonlinear control and optimization," *SIAM Journal on Control and Optimization*, vol. 31, no. 3, pp. 569–603, 1993.

[31] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang, "Optimality, stability, and convergence in nonlinear control," *Applied Mathematics and Optimization*, vol. 31, no. 3, pp. 297–326, 1995.

[32] A. L. Dontchev and W. W. Hager, "Lipschitzian stability for state constrained nonlinear optimal control," *SIAM Journal on Control and Optimization*, vol. 36, no. 2, pp. 698–718, 1998.

[33] A. L. Dontchev, M. I. Krastanov, R. T. Rockafellar, and V. M. Veliov, "An euler-newton continuation method for tracking solution trajectories of parametric variational inequalities," *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 1823–1840, 2013.

[34] A. L. Dontchev and V. M. Veliov, *Regularity Properties of Mappings in Optimal Control*, 2015.

[35] H. Frankowska, *Value function in optimal control*. ICTP Lecture Notes, 115 pages, Summer School of Mathematical Control Theory, 2001.

[36] R. Ghaemi, J. Sun, and I. V. Kolmanovsky, "An integrated perturbation analysis and sequential quadratic programming approach for model predictive control," *Automatica*, vol. 45, no. 10, pp. 2412–2418, 2009.

[37] K. Graichen and N. Petit, "A continuation approach to state and adjoint calculation in optimal control applied to the reentry problem," *Proceedings of IFAC World Congress*, 2008, pp. 14 307–14 312.

[38] S. Gros, M. Zanon, M. Vukov, and M. Diehl, "Nonlinear MPC and MHE for mechanical multi-body systems with application to fast tethered airplanes," *Proceedings of IFAC Nonlinear Model Predictive Control Conference*, 2012, pp. 86–93.

[39] L. Grüne and J. Pannek, *Nonlinear model predictive control: theory and algorithms*. Springer Science & Business Media, 2011.

[40] A. Guiggiani, I. Kolmanovsky, P. Patrinos, and A. Bemporad, "Fixed-point constrained model predictive control of spacecraft attitude," *Proceedings of American Control Conference*, 2015, pp. 2317–2322.

[41] A. Hatcher, "Algebraic topology," *Cambridge University Press*, 2002.

[42] J. Hauser and D. G. Meyer, "Trajectory morphing for nonlinear systems," *Proceedings of American Control Conference*, 1998, pp. 2065–2070.

[43] Ø. Hegrenæs, J. T. Gravdahl, and P. Tøndel, "Spacecraft attitude control using explicit model predictive control," *Automatica*, vol. 41, no. 12, pp. 2107–2114, 2005.

[44] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*. AMS Edition, 2001.

[45] D. D. Holm, T. Schmah, and C. Stoica, *Geometric mechanics and symmetry: from finite to infinite dimensions*. Oxford University Press, 2009.

[46] J. E. Humphreys, *Introduction to Lie algebras and representation theory*. Springer Science & Business Media, 1972.

[47] R. V. Iyer, R. Holsapple, and D. Doman, "Optimal control problems on parallelizable Riemannian manifolds: theory and applications," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 12, no. 1, pp. 1–11, 2006.

[48] V. Jurdjevic, *Geometric control theory*. Cambridge University Press, 1997.

[49] U. V. Kalabić, R. Gupta, S. Di Cairano, A. Bloch, and I. Kolmanovsky, "Constrained spacecraft attitude control on SO(3) using reference governors and nonlinear model predictive control," *Proceedings of American Control Conference*, 2014, pp. 5586–5593.

[50] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The International Journal of Robotics Research*, vol. 5, no. 1, pp. 90–98, 1986.

[51] M. Kim, *Continuous low-thrust trajectory optimization: techniques and applications*. PhD Dissertation, Virginia Polytechnic Institute and State University, 2005.

[52] A. W. Knapp, *Lie groups beyond an introduction*. Springer Science & Business Media, 2002.

[53] M. B. Kobilarov, M. Desbrun, J. E. Marsden, and G. S. Sukhatme, "A discrete geometric optimal control framework for systems with symmetries," *Proceedings of Robotics: Science and Systems*, 2007.

[54] M. B. Kobilarov, K. Crane, and M. Desbrun, "Lie group integrators for animation and control of vehicles," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 2, p. 16, 2009.

[55] M. B. Kobilarov and J. E. Marsden, "Discrete geometric optimal control on Lie groups," *IEEE Transactions on Robotics*, vol. 27, no. 4, pp. 641–655, 2011.

[56] D. E. Koditschek and E. Rimon, "Robot navigation functions on manifolds with boundary," *Advances in Applied Mathematics*, vol. 11, no. 4, pp. 412–442, 1990.

[57] W.-S. Koon and J. E. Marsden, "Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction," *SIAM Journal on Control and Optimization*, vol. 35, no. 3, pp. 901–929, 1997.

[58] P. S. Krishnaprasad, *Optimal control and Poisson reduction*, 1993.

[59] E. B. Lee and L. Markus, *Foundations of optimal control theory*. John Wiley & Sons, 1967.

[60] J. M. Lee, *Introduction to smooth manifolds*. Springer Science & Business Media, 2012.

[61] T. Lee, N. H. McClamroch, and M. Leok, "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum," *Proceedings of Conference on Control Applications*, 2005, pp. 962–967.

[62] T. Lee, M. Leok, and N. H. McClamroch, "Optimal attitude control of a rigid body using geometrically exact computations on SO(3)," *Journal of Dynamical and Control Systems*, vol. 14, no. 4, pp. 465–487, 2008.

[63] T. Lee, *Computational geometric mechanics and control of rigid bodies*. PhD Dissertation, University of Michigan, 2008.

[64] M. B. Liñán, *A geometric study of abnormality in optimal control problems for control and mechanical control systems*. PhD Dissertation, Technical University of Catalonia, 2008.

[65] M. B. Liñán, "Characterization of accessibility for affine connection control systems at some points with nonzero velocity," *Proceedings of IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 6528–6533.

[66] P. D. Loewen and H. Zheng, "Generalized conjugate points for optimal control problems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 22, no. 6, pp. 771–791, 1994.

[67] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming*. Springer Science & Business Media, 2008.

[68] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems.* Springer Science & Business Media, 1999.

[69] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica 2001*, vol. 10, pp. 357–514, 2001.

[70] J. E. Marsden, G. K. Misiolek, J.-P. Ortega, M. Perlmutter, and T. S. Ratiu, *Hamiltonian reduction by stages.* Springer Science & Business Media, 2007.

[71] P. M. Mereau and W. F. Powers, "Conjugate point properties for linear quadratic problems," *Journal of Mathematical Analysis and Applications*, vol. 55, no. 2, pp. 418–433, 1976.

[72] J. W. Milnor, *Morse theory.* Princeton University Press, 1963.

[73] J. Moser and A. P. Veselov, "Discrete versions of some classical integrable systems and factorization of matrix polynomials," *Communications in Mathematical Physics*, vol. 139, no. 2, pp. 217–243, 1991.

[74] L. Noakes, G. Heinzinger, and B. Paden, "Cubic splines on curved spaces," *IMA Journal of Mathematical Control and Information*, vol. 6, no. 4, pp. 465–473, 1989.

[75] T. Ohsawa, "Symmetry reduction of optimal control systems and principal connections," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 96–120, 2013.

[76] T. Ohtsuka, "A continuation/gmres method for fast computation of nonlinear receding horizon control," *Automatica*, vol. 40, no. 4, pp. 563–574, 2004.

[77] J. T. Olympio, "A second-order gradient solver using a homotopy method for space trajectory problems," *Proceedings of AIAA/AAS Astrodynamics Specialist Conference*, 2010.

[78] A. V. Rao, "A survey of numerical methods for optimal control," *Advances in the Astronautical Sciences*, vol. 135, no. 1, pp. 497–528, 2009.

[79] P. Rostalski, I. A. Fotiou, D. J. Bates, A. G. Beccuti, and M. Morari, "Numerical algebraic geometry for optimal control applications," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 417–437, 2011.

[80] A. Saccon, J. Hauser, and A. P. Aguiar, "Exploration of kinematic optimal control on the Lie group SO(3)," *Proceedings of IFAC Symposium on Nonlinear Control Systems*, 2010, pp. 1302–1307.

[81] A. Saccon, A. P. Aguiar, A. J. Hausler, J. Hauser, and A. M. Pascoal, "Constrained motion planning for multiple vehicles on SE(3)," *Proceedings of IEEE Conference on Decision and Control*, 2012, pp. 5637–5642.

[82] A. Saccon, J. Hauser, and A. P. Aguiar, "Optimal control on Lie groups: The projection operator approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2230–2245, 2013.

[83] S. Sasaki, "On the differential geometry of tangent bundles of Riemannian manifolds I," *Tohoku Mathematical Journal, Second Series*, vol. 10, no. 3, pp. 338–354, 1958.

[84] S. Sasaki, "On the differential geometry of tangent bundles of Riemannian manifolds II," *Tohoku Mathematical Journal, Second Series*, vol. 14, no. 2, pp. 146–155, 1962.

[85] D. H. Sattinger and O. L. Weaver, *Lie groups and algebras with applications to physics, geometry, and mechanics*. Springer Science & Business Media, 1986.

[86] H. Schättler and U. Ledzewicz, *Geometric optimal control: theory, methods and examples*. Springer Science & Business Media, 2012.

[87] E. Silani and M. Lovera, "Magnetic spacecraft attitude control: a survey and some new results," *Control Engineering Practice*, vol. 13, no. 3, pp. 357–371, 2005.

[88] C. Silva and E. Trélat, "Smooth regularization of bang-bang optimal control problems," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2488–2499, 2010.

[89] F. Silva Leite, M. Camarinha, and P. Crouch, "Elastic curves as solutions of Riemannian and sub-Riemannian control problems," *Mathematics of Control, Signals, and Systems*, vol. 13, no. 2, pp. 140–155, 2000.

[90] J. L. Speyer and D. H. Jacobson, *Primer on optimal control theory*. SIAM, 2010.

[91] E. Trélat, "Optimal control and applications to aerospace: some results and challenges," *Journal of Optimization Theory and Applications*, vol. 154, no. 3, pp. 713–758, 2012.

[92] J. Vandersteen, M. Diehl, C. Aerts, and J. Swevers, "Spacecraft attitude estimation and sensor calibration using moving horizon estimation," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 3, pp. 734–742, 2013.

[93] R. Vinter, *Optimal control*. Springer Science & Business Media, 2010.

[94] A. Weiss, F. Leve, M. Baldwin, J. R. Forbes, and I. Kolmanovsky, "Spacecraft constrained attitude control using positively invariant constraint admissible sets on $\mathrm{SO}(3)\times\mathbb{R}^3$," *Proceedings of American Control Conference*, 2014, pp. 4955–4960.

[95] H. Yan, F. Fahroo, and I. M. Ross, "Real-time computation of neighboring optimal control laws," *AIAA Paper*, vol. 4657, 2002.

[96] V. M. Zavala and L. T. Biegler, "The advanced-step nmpc controller: Optimality, stability and robustness," *Automatica*, vol. 45, no. 1, pp. 86–93, 2009.

[97] V. Zeidan and P. Zezza, "The conjugate point condition for smooth control sets," *Journal of Mathematical Analysis and Applications*, vol. 132, no. 2, pp. 572–589, 1988.

[98] V. Zeidan and P. Zezza, "Conjugate points and optimal control: counterexamples," *IEEE Transactions on Automatic Control*, vol. 34, no. 2, pp. 254–255, 1989.

[99] V. Zeidan, "The riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency," *SIAM Journal on Control and Optimization*, vol. 32, no. 5, pp. 1297–1321, 1994.

[100] S. S. Zhulin, "Homotopy method for finding extremals in optimal control problems," *Differential Equations*, vol. 43, no. 11, pp. 1495–1504, 2007.