# Higher-Order Euler-Poincaré Equations and their Applications to Optimal Control 

Leonardo Colombo<br>Instituto de Ciencias Matemáticas,(CSIC-UAM-UCM-UC3M)



Trabajo de fin de Master

Advisor: Dr. David Martín de Diego-Instituto de Ciencias Matemáticas (CSIC-UAM-UC3M-UCM)

Madrid
Septiembre de 2012

## Agradecimientos

Este trabajo ha resultado para mí la culminación de un largo camino, el final de una de las etapas de formación como alumno de doctorado del Instituto de Ciencias Matemáticas (CSIC-UAM-UC3M-UCM). Estoy seguro que esta etapa, este tiempo como alumno de un master, quedará muy marcada en mí, ya que me llevo muchísimos aprendizajes, tanto en sus comienzos en la universidad Carlos Tercero de Madrid, como en su segunda etapa en la Universidad Autónoma de Madrid me han aportado conocimientos muy importantes en mi formacion profesional.

Este transcurso no puedo haber sido posible sin aquellas personas que me soportaron este tiempo. Gente que me dió muchísimas alegrias y hace que disfrute diariamente vivir en Madrid.

Tengo el privilegio de haber tenido (y seguir teniendo) como director, a David Martín de Diego. Muchas gracias David por introducirme en ese mundo, incentivarme constantemente con nuevos tópicos para mi investigación, por estar siempre que te necesité como director y como persona. Involucrarme en tus projectos, invitarme a congresos, workshops, cursos y corregir las charlas siempre entre muchas otras qué yo quedo alucinado por tener un director asi!. El ha realizado un trabajo mas que responsable conmigo y nunca tendré tantas palabras como para agradecerle todo lo que hace día a día por mi bienestar profesional y personal (iy ni hablar de mi salud mental!).

A toda mi familia, especialmente a mi papá y mamá. Sin su contención y apoyo durante este tiempo tampoco esto hubiese sido posible.

A Ramiro y Joaquín por amistad y paciencia durante estos años a distancia.
A mis hermanos en la matemática Fer y Cédric. Aunque no lo crean y me quieran matar muchas veces, me han abierto un gran camino en todo esto sin darse cuenta. Gracias por sus consejos siempre y por tantas alegrias compartidas juntos. ¡Les quiero mucho! Fer, muchísimas gracias por leer este trabajo y corregir mi inglés siempre.

A todos mis compañeros del master y del ICMAT, especialmente a Miguel, Jeza, Javi, Alberto, Luis, Emilio y Ana.Z

Gloria y Sandra por su constante ayuda todo el tiempo y brindarme sus oidos muchas veces, para lo bueno y lo no tan bueno y sus cafés de media mañana.

Este trabajo tambien está dedicado a Rosa; su constante ayuda me solucinó muchas veces mi vida laboral infinitamente. Además de ser una super secretaria del instituto es una gran persona de la que aprendo diariamente con su charla, almuerzos y viajes en tren. Gracias por ser siempre tan positiva y enseñarme (o intentarlo) a que yo lo sea a la hora de enfrentar los retos.

A Manuel, Edith y Juan Carlos por todo lo que me enseñaron, mostrandome, junto a David como ir construyendo este largo camino para ser el día de mañana un buen y
responsable profesional. Siempre han estado por cualquier cosa que necesité en este master ayudandome e incentivandome. ¡Muchas gracias!

No puedo no nombrar a esas tres personas que han hecho para mi del ICMAT un lugar donde me guste y disfrute pasar el día trabajando. Gracias a mi gran amigo Rafa Granero, por todo el apoyo incondicional. Haberte encontrado acá fué una de las cosas que mas valoro de esto. Vivi y María, ya son como mis hermanas mayores. Su compañia y amistad han sido uno de los pilares principales en este tiempo para poder levantarme todos los dias e ir por mas. ¡Gracias chicas por aguantarme tanto!

No quiero olvidarme de nadie, pero seguramente lo este haciendo de muchos. Muchas gracias a todos mis companeros de la red de Geometría, Mecánica y Control. Por tantos buenos momentos compartidos juntos. Junto a ustedes ;si que da gusto dedicarse a la investigación!

A todo el Instituto de Ciencias Matemática por hacerme sentir como en mi propia casa especialemnte a Eduardo Frechilla y Rafael Orive. Al Consejo Superior de Investigaciones Científicas por sustentar estos estudios y porsupuesto ¡a Madrid! Gracias por ser ese bello refugio elegido.

Leonardo J. Colombo
Londres. 05 de septiembre de 2012

## Contents

1 Introduction ..... 5
2 Basic Tools in Geometric Mechanics ..... 9
2.1 Lie Groups and Group Actions ..... 10
2.2 Riemannian Geometry ..... 12
2.3 Hamiltonian Mechanics and Symplectic Geometry ..... 13
2.4 The Gotay-Nester-Hinds Algorithm ..... 15
2.5 Skinner and Rusk Formalism: An unifying framework ..... 17
3 Mechanical Systems on Lie Groups ..... 21
3.1 Trivialization of a cotangent bundle of a Lie group ..... 23
3.2 Euler-Arnold Equations ..... 25
3.2.1 Euler-Arnold equations for the rigid body ..... 26
3.3 Euler-Lagrange Equations on Lie Groups ..... 29
3.3.1 Legendre Transformation ..... 32
3.3.2 Simplecticity and Momentum Preservation ..... 32
3.3.3 Attitude Dynamic of a Rigid Body on $S O(3)$ ..... 34
3.4 Euler-Poincaré Equations ..... 38
3.5 Euler-Poincaré Equations for the Motion of a Rigid Body ..... 40
3.5.1 The Lagrangian of the Rigid Body ..... 41
3.5.2 Euler-Poincaré Reduction Theorem for the Motion of a Rigid Body ..... 43
4 Higher-Order Mechanical Systems ..... 47
4.1 Higher-Order Tangent Bundles ..... 47
4.1.1 The Case of Lie Groups ..... 49
4.2 Hamilton's Principle and Euler-Lagrange Equations ..... 49
4.3 Higher-order Mechanical Systems with Constraints ..... 54
4.3.1 Lagrange-D'Alembert Principle ..... 55
4.3.2 Higher-Order Variational Calculus with Constraints ..... 57
4.3.3 Geometric Formulation for Higher-Order Constrained Mechanics. ..... 59
5 Higher-Order Mechanical Systems on Lie Groups ..... 67
5.1 Higher-Order Euler-Poincaré equations ..... 68
5.2 Higher-order Euler-Arnold's equations on $T^{*}\left(T^{(k-1)} G\right)$ ..... 70
5.3 Higher-Order Unified Mechanics on Lie Groups ..... 73
5.3.1 Unconstrained problem ..... 73
5.3.2 Constrained problem ..... 77
6 Optimal Control of Mechanical Systems ..... 81
6.1 Optimal Control ..... 81
6.1.1 Optimal Control and Maximum Principle ..... 81
6.2 Variational Problems and Optimal Control ..... 83
6.3 Lagrangian and Hamiltonian Control Systems ..... 85
6.4 Optimal Control of Mechanical Systems on Lie Groups ..... 86
6.4.1 Left-invariant control systems ..... 86
6.4.2 Accessibility and Controllability ..... 86
6.4.3 Properties of Reachable Sets ..... 87
6.4.4 Optimal Control of the Position of a Rigid Body ..... 89
7 Optimal Control of Underactuated Mechanical Systems ..... 93
7.1 Optimal Control of underactuated mechanical systems ..... 93
7.2 Quasivelocities and Optimal Control of Underactuated Systems ..... 101
7.2.1 Quasivelocities ..... 102
7.2.2 Optimal Control for Underactuated Mechanical Systems ..... 102
7.3 Underactuated Mechanical Control Systems on Lie Groups ..... 108
7.3.1 Optimal Control of an Underactuated Rigid Body ..... 112
7.3.2 Optimal control of a Cosserat rod ..... 115

## Chapter 1

## Introduction

The mathematical activity in the last century in dynamical systems, mechanics and related areas has been extraordinary. The number of applications has grown exponentially and both basic science as well as several engineering technologies are profiting from this development. In the 1960s more sophisticated and powerful techniques coming from modern differential geometry and topology have been introduced in their study, experiencing a spectacular growth in the last 50 years. Control and optimal control of mechanical systems have not ignored these developments, becoming now a principal research focus of nonlinear control theory. In particular, there are an increasing interest in the control of underactuated mechanical systems (see [17, 31]). These type of mechanical systems are characterized by the fact that there are more degrees of freedom than actuators. This type of system is quite different from a mathematical and engineering perspective than fully actuated control systems where all the degrees of freedom are actuated.

In control system engineering, the underlying geometric features of a dynamic system are often not considered carefully. For example, many control systems are developed for the standard form of ordinary differential equations, namely $\dot{x}=f(x, u)$, where the state and control input are denoted by $x$ and $u$, respectively. It is assumed that the state and control input lie in Euclidean spaces. However, for many interesting mechanical systems the configuration space cannot be expressed globally as a Euclidean space.

In this work, dynamics and optimal control problems for mechanical systems, and in particular, for rigid bodies are studied, incorporating careful consideration, of their geometric features.

Optimal control problems on Lie group are described by Euler-Poincaré systems. An Euler-Poincaré system is a mechanical system whose configuration manifold is a Lie group, $G$, and whose Lagrangian $L: T G \rightarrow \mathbb{R}$ is left or right invariant under the action of that group. To be specific, in this work assumes that the Lagrangian in left-invariant. Let the tangent bundle and the cotangent bunlde of $G$ be denoted by $T G$ and $T^{*} G$ respectively, and its Lie algebra and dual are denoted by $\mathfrak{g}$ and $\mathfrak{g}^{*}$ respectively. The quotient space $T G / G$ is called the reduced space and by a left-trivialization of $T G$ is diffeomorphic to $\mathfrak{g}=T_{e} G$. The restriction of the Lagrangian to the reduced space is called the reduced Lagrangian $l: \mathfrak{g} \rightarrow \mathbb{R}$. A Lagrangian $L: T G \rightarrow \mathbb{R}$ is left-invariant if its left-trivialization $l: G \times \mathfrak{g} \rightarrow \mathbb{R}$ does not deppend of the first entry.

Given an initial condition $\left(g_{0}, \dot{g}_{0}\right) \in T G$ the Euler-Lagrange equations for $L$ on $T G$
describe an initial value problem (IVP). This IVP can be left-trivialized to $G \times \mathfrak{g}$ to give

$$
\begin{align*}
\dot{g} & =\xi g, \quad g(a)=g_{0},  \tag{1.0.1}\\
\frac{d}{d t} \frac{\delta l}{\delta \xi} & =a d_{\xi}^{*} \frac{\delta l}{\delta \xi}+\mathcal{L}_{g}^{*} \frac{\delta l}{\delta g}, \quad \xi(a)=g_{0}^{-1} \dot{g}_{0}, \tag{1.0.2}
\end{align*}
$$

where $\mathcal{L}$ denotes the left translation on the Lie group $G$.
If the Lagrangian is left-invariant, these equations becomes

$$
\begin{align*}
\dot{g} & =\xi g, \quad g(a)=g_{0}  \tag{1.0.3}\\
\frac{d}{d t} \frac{\delta l}{\delta \xi} & =a d_{\xi}^{*} \frac{\delta l}{\delta \xi}, \quad \xi(a)=g_{0}^{-1} \dot{g}_{0} \tag{1.0.4}
\end{align*}
$$

This equations define an IVP in the body angular velocity $\xi(t) \in \mathfrak{g}$ and the configuration $g(t) \in G$ over the interval $[a, b]$. However, due to the invariance of the Lagrangian with respect to the action of the Lie group, (1.0.2) is decoupled from (1.0.1). (1.0.2) is the EP equation and describes the dynamics reduced to $\mathfrak{g}$ to recover the configuration dynamics on $G$, one solves the EP equations to obtain a curve $\xi(t)$ for $t \in[a, b]$, substitutes the solution into (1.0.1) and then solves the IVP for $g(t)$ in the interval $[a, b]$, in a procedure called reconstruction. Consequently, (1.0.1) is called reconstruction equation.

So far, the first order case of mechanics on Lie groups is pretty well understand. But in many optimization problems in mechanics the Lagrangians which appear are of higherorder (as for instance in optimal control problems, interpolation problems, etc), therefore it is interesting to find a full geometric setting for these theories, that is when one considers a Lagrangian $L: T^{(k)} Q \rightarrow \mathbb{R}$ where $T^{(k)} Q$ is the $k$ th-order tangent bundle.

Specifically, this work presents a new global geometric and intrinsic schemes to obtain the Euler-Lagrange equations when the configuration space is the higher-order space associated with a Lie group $G, T^{(k)} G$. During the last decades of the past century, there have been studies and attempts to define the higher-order variational calculus. The main objectives are to describe the Euler-Lagrange equations for Lagrangians defines on these higher-order bundles. The standard framework of higher-order mechanics on Lie groups starts by looking for the extremals of the functional

$$
\mathcal{A}=\int_{a}^{b} L\left(\xi, \dot{\xi}, \ldots, \xi^{k}\right) d t
$$

for $L: k \mathfrak{g} \rightarrow \mathbb{R}$ where $k \mathfrak{g}$ denotes $k$-copies of the Lie algebra $\mathfrak{g}$ and $\xi^{j}, j=1, \ldots, k$ are the $j$ time derivatives of $\xi \in \mathfrak{g}$. Variational calculus states that the extremizers of this integral action must satisfy the higher-order Euler-Poincaré equations

$$
\left(\frac{d}{d t}-\operatorname{ad}_{\xi}^{*}\right) \sum_{l=0}^{k-1}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial \xi^{(l)}}\right)=0 .
$$

Instead of reducing the higher-order space $T^{(k)} G$ one can trivialize this bundle as $G \times$ $k \mathfrak{g}$ and work with the higher-order Euler-Lagrange equations on $G \times k \mathfrak{g}$ and when the

Lagrangian is left-invariant one can obtain the higher-order Euler-Poincaré equations. In this work, we propose an alternative way, avoiding the use of additional structures, working only with intrinsic objects from the Lagrangian and Hamiltonian sides. The main results are published here ([26], [30], [29], [27]). This formalism is strongly based on the one developed by Skinner and Rusk [65]. In order to deal with singular Lagrangian systems, Skinner and Rusk construct a Hamiltonian system on the Whitney sum $T Q \oplus T^{*} Q$ of the tangent and cotangent bundles of the configuration manifold $Q$. The advantages of their approach lies on the fact the second-order condition of the dynamics is automatically satisfied. This does not happened in the Lagrangian side of the Gotay and Nester formulation, where the second-order condition problem has to be considered after the implementation of the constraint algorithm (see [37]).

For higher-order mechanics on Lie groups we start with a Lagrangian function defined on the left trivialized tangent bundle of $T^{(k)} G, G \times k \mathfrak{g}$. We consider the fibration, after a lefttivialization, $\pi_{W, G \times(k-1) \mathfrak{g}}: W \rightarrow G \times(k-1) \mathfrak{g}$, where $W=G \times k \mathfrak{g} \times k \mathfrak{g}^{*}$ is a fiber product. In $W$ we construct the presymplectic 2-form by pulling back the canonical symplectic 2-form of $G \times k \mathfrak{g}^{*}$ and we define a convenient Hamiltonian from a natural canonical pairing and the given Lagrangian function. In $W$ we obtain a global, intrinsic and a unique expression for the Euler-Lagrange equations for higher-order mechanics on Lie groups. Additionally, we obtain a resultant constraint algorithm.

Apart from the lack of ambiguity inherent in our construction, it is worth to emphasize that this formalism is easily extended to the case of higher-order mechanics on Lie groups with higher-order constraints and optimal control problems. Therefore, we introduce constraints in our picture, which are geometrically defined as a submanifold $\mathcal{M}$ of $T^{(k)} G \simeq G \times k \mathfrak{g}$.

The outline of this monograph is the following: First, in Chapter 2 we give the notation used along this work and the basic mathematical background needed: Lie groups and group actions, Riemannian and symplectic geometry, Hamiltonian and Lagrangian mechanics, etc. There is also a sketch of Gotay-Nester-Hinds algorithm and the Skinner-Rusk formalism for the mechanics.

Chapter 3, is a brief introduction to mechanics on Lie groups, variational principles on Lie groups, Euler-Lagrange equations on Lie groups, Euler-Poincaré equations and EulerArnold equations. After this we develop the Euler-Lagrange, Euler-Arnold and EulerPoincaré equations for the motion of a rigid body on $S O(3)$.

Chapter 4, is devoted to the study of Higher-Order Mechanics. The reader will may find first a variational approximation for Higher-Order Mechanics and Higher-Order Mechanics with constraints. The chapter focuses on the geometric derivation of the equations of motions using and extension to higher-order theories by the Skinner-Rusk formalism for the mechanics. It is also introduce constraints in the picture.

Chapter 5, the Euler-Arnold's equations for a hamiltonian system defined on a higherorder cotangent bundle of a Lie group are developed. After this, we define the Pontryaguin bundle $G \times k \mathfrak{g} \times k \mathfrak{g}^{*}$ where we introduce the dynamics using a presymplectic hamiltonian formalism. We deduce the $k$-order Euler-Lagrange equations and, as a particular example, the $k$-order Euler-Poincaré equations. Since the dynamics is presymplectic it is necessary to analyze the consistency of the dynamics using a constraint algorithm. Finally we introduce the case of constrained dynamics. We show that our techniques are easily adapted to this
particular case.
Chapter 6, gives a general background in optimal control theory including basic definitions of controllability, accessibility and the relationship of optimal controls problems and second-order variational problems with constraints. The normal and abnormal solutions to the optimal control problem of the position of a rigid body are also studding.

Finally, in Chapter 7, as an illustration of the applicability of our setting, we analyze the case of underactuated control of mechanical systems. Interesting examples are including in this Chapter: The optimal control problem of a cart with a pendulum, a planar rigid body, elastic rods, and a family of underactuated problems for the rigid body on $S O(3)$.

## Chapter 2

## Basic Tools in Geometric Mechanics

In this chapter we summarize some basic concepts on the geometric mechanics of unconstrained Lagrangian and Hamiltonian systems, assuming familiarity with basic differential geometry [16], [44, [51]. We begin with the traditional variational formulation of mechanical systems in terms of Hamilton's Principle on the Lagrangian side. Through the Legendre transform we then summarize the geometric structure behind the Hamiltonian side. After this, we introduce basic concepts on Lie groups and Lie groups actions that will be used through this work.

In some sense, symplectic geometry is complementary to Riemannian geometry [9, [22], [54]. While Riemannian geometry is based on the study of smooth manifolds that carry a nondegenerate symmetric tensor, symplectic geometry covers the study of smooth manifolds that are equipped with a non-degenerate skew-symmetric tensor. Although both have several similarities, by their nature, they also show to have strong differences. In this section we remember some features of Riemannian geometry and symplectic geometry. In this latter context we derive the Hamiltonian side of mechanics.

After introducing the Gotay-Nester-Hinds [37] algorithm for singular systems, we shall give a brief summary on mechanical systems using an unifying framework: the SkinnerRusk formalism for the mechanics [65]. Our exposition here is largely based on that found in [1], [55], [39] and [51], and we wish to remark that the Einstein summation convention is enforced throughout this thesis unless otherwise is noted.

Let $Q$ be a manifold and $T Q$ its tangent bundle. Denote by $\left(q^{i}\right)$ the local coordinates on $Q$ and by $\left(q^{i} ; \dot{q}^{i}\right)$ the induced coordinates on $T Q$. Define the mechanical Lagrangian $L: T Q \rightarrow \mathbb{R}$ given by $L=T-V$, where $K(v)=\mathcal{G}(v, v)$; is the kinetic energy associated with the Riemannian metric $\mathcal{G}$ and where $V: Q \rightarrow \mathbb{R}$ is the potential energy. The trajectories of an unconstrained mechanical system are then given by Hamilton's Principle, which states that among the set of possible motions $q(t)$ of our mechanical system in any time interval $[a ; b]$, the actual trajectories are such that

$$
\begin{equation*}
\delta \int_{a}^{b} L(q(t), \dot{q}(t)) d t=0 . \tag{2.0.1}
\end{equation*}
$$

We say that a mechanical system (unconstrained or constrained) is variational if its equations of motion can be derived from Hamilton's principle. Basic results in the calculus of variations (see [10, (39] or [55]) show that the condition (2.0.1) is equivalent to the
requirement that $q(t)$ satisfies the Euler-Lagrange equations:

$$
\begin{equation*}
\frac{d}{d t} \frac{\partial L}{\partial \dot{q}^{i}}-\frac{\partial L}{\partial q^{i}}=0 . \tag{2.0.2}
\end{equation*}
$$

Now, if we define the fiber derivative $\mathbb{F} L: T Q \rightarrow T^{*} Q$ in coordinates by the map $\left(q^{i} ; \dot{q}^{i}\right) \mapsto\left(q^{i} ; p_{i}\right)$, where $p_{i}=\frac{\partial L}{\partial \dot{q}^{i}}$ is called the momentum conjugate to $q^{i}$, then assuming that $L$ is hyperregular ${ }^{1]}$ we can define the Hamiltonian $H$ by $H(q ; p)=p_{i} \dot{q}^{i}-L$. The coordinates $\left(q^{i} ; p_{i}\right)$ on the cotangent bundle $T^{*} Q$ are called the canonical cotangent coordinates and $\mathbb{F} L$ is called the Legendre transform.

As we shall see below, the Hamiltonian $H$ is related to the total energy of the mechanical system, and since the cotangent space $T^{*} Q$ carries a natural symplectic structure, we will summarize the rich geometry of Hamiltonian mechanics below and present the analogue of (2.0.2), the Hamiltonian equations of motion.

### 2.1 Lie Groups and Group Actions

First, we give the basic definitions and properties of Lie groups. A Lie group is a differentiable manifold that has a group structure such that the group operation is smooth. A Lie algebra is the tangent space of the Lie group $G$ at the identity of the group, $e \in G$, with the bracket $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ that is bilinear, skew symmetric and satisfies the Jacobi identity [31], 39], [55].

A Lie group $H$ is said to be a Lie subgroup of a Lie group $G$ if it is a submanifold of $G$ and the inclusion mapping $i: H \hookrightarrow G$ is a group homomorphism.

Example 2.1.1. Basic examples of Lie groups which will appear in this work include the unit circle $S^{1}$, the group of $n \times n$ invertible matrices $G L(n, R)$ with the matrix multiplication, and several of its Lie subgroups: the group of rigid motions in 3-dimensional Euclidean space, $S E(3)$; the group of rigid motions in the plane, $S E(2)$; and the group of rotations in $\mathbb{R}^{3}, S O(3)$.

For $g, h \in G$, the left-translation map is defined as $L_{h}: G \rightarrow G$, by $L_{h} g=h g$. Similarly, the right-translation $R_{h}: G \rightarrow G$ is defined as $R_{h} g=g h$. Given $\xi \in \mathfrak{g}$ define a vector field $X_{\xi}: G \rightarrow T G$ such that $X_{\xi}(g)=T_{e} L_{g} \cdot \xi$, and let the corresponding unique integral curve passing through the identity $e$ at $t=0$ be denoted by $\gamma_{\xi}(t)$. The exponential map $\exp : \mathfrak{g} \rightarrow G$ is defined by $\exp \xi=\gamma_{\xi}(1)$. The application $\exp$ is a local diffeomorphism from a neighborhood of zero in $\mathfrak{g}$ onto a neighborhood of $e \in G$.

Define the inner automorphism $I_{g}: G \rightarrow G$ as $I_{g}(h)=g h g^{-1}$. The adjoint operator $A d_{g}: \mathfrak{g} \rightarrow \mathfrak{g}$ is the differential of $I_{g}(h)$ with respect to $h$ at $h=e$ along the direction $\eta \in \mathfrak{g}$, that is $A d_{g} \eta=T_{e} I_{g} \cdot h$. Roughly speaking, the adjoint action measures the noncommutativity of the multiplication of the Lie group: if $G$ is Abelian, then the adjoint action $A d_{g}$ is simply the identity mapping on $G$. In addition, when considering motion along non-Abelian Lie groups, a choice must be made as to whether to represent translation by left or right multiplication. The adjoint action provides the transition between these two possibilities.

[^0]The ad operator $a d_{\xi}: \mathfrak{g} \rightarrow \mathfrak{g}$ is obtained by differentiating $A d_{g} \eta$ with respect to $g$ at $e$ in the direction $\xi$, that is $a d_{\xi} \eta=T_{e}\left(A d_{g} \xi\right) \cdot \eta$. This corresponds to Lie bracket (i,e; $a d_{\xi} \eta=[\xi, \eta]$ ).

We define the coadjoint operator $A d^{*}: G \times \mathfrak{g}^{*} \rightarrow \mathfrak{g}^{*}$ from the paring between vectors and covectors by $\left\langle A d_{g}^{*} \alpha, \xi\right\rangle=\left\langle\alpha, A d_{g} \xi\right\rangle$ for $\alpha \in \mathfrak{g}^{*}$. The co-ad operator ad ${ }^{*}: \mathfrak{g} \times \mathfrak{g}^{*} \rightarrow \mathfrak{g}^{*}$ is defined by $\left\langle a d_{\eta}^{*} \alpha, \eta\right\rangle=\left\langle\alpha, a d_{\eta} \xi\right\rangle$ for $\alpha \in \mathfrak{g}^{*}$.

The notion of symmetry or invariance of the system is formally expressed through the concept of action.

Definition 2.1.2. A (left) action of a Lie group $G$ on a manifold $Q$ is a smooth mapping $\Phi: G \times Q \rightarrow Q$ such that, $\Phi(e, q)=q$ for all $q \in Q$ and $\Phi(g, \Phi(h, q))=\Phi(g h, q)$ for all $g, h \in G, q \in Q$.

The same definition can be stated for right actions, but we consider here left actions, which is the usual convention in mechanics. Normally, we will only be interested in the action as a mapping from $Q$ to $Q$, and so we will write the action as $\Phi_{g}: Q \rightarrow Q$, where $\Phi_{g}(q)=\Phi(g, q)$, for all $g \in G$. In some cases, we shall make a slight abuse of notation and write $g q$ instead of $\Phi_{g}(q)$. The orbit of the $G$-action through a point $q$ is $\operatorname{Orb}_{G}(q)=$ $\{g q \mid g \in G\}$. An action is said to be free if all its isotropy groups are trivial, that is, the relation $\Phi_{g}(q)=q$ implies $g=e$, for any $q \in Q$ (note that, in particular, this implies that there are no fixed points). An action is said to be proper if $\tilde{\Phi}: G \times Q \rightarrow Q \times Q$ defined by $\tilde{\Phi}(g, q)=(q, \Phi(g, q))$ is a proper mapping, i.e., if $K \subset Q \times Q$ is compact, then $\tilde{\Phi}^{-1}(K)$ is compact. Finally, an action is said to be simple or regular if the set $Q / G$ of orbits has a differentiable manifold structure such that the canonical projection of $Q$ onto $Q / G$ is a submersion.

If $\Phi$ is a free and proper action, then $\Phi$ is simple, and therefore $Q / G$ is a smooth manifold and $\pi: Q \rightarrow Q / G$ is a submersion.

Let $\xi$ be an element of the Lie algebra $\mathfrak{g}$. Consider the $\mathbb{R}$-action on $Q$ defined by

$$
\Phi^{\xi}(t, q)=\Phi(\exp (t \xi), q) \in Q
$$

We can interpret $\Phi^{\xi}$ as a flow on the manifold $Q$. Consequently, it determines a vector field on $Q$, given by

$$
\xi_{Q}(q)=\left.\frac{d}{d t}\right|_{t=0}(\Phi(\exp (t \xi), q))
$$

which is called the fundamental vector field or infinitesimal generator of the action corresponding to $\xi$. Given a Lie group $G$, we can consider the natural action of $G$ on itself by left multiplication $\Phi: G \times G \rightarrow G,(g, h) \mapsto g h$. For any $\xi \in \mathfrak{g}$, the corresponding fundamental vector field of the action is given by

$$
\xi_{G}(h)=\left.\frac{d}{d t}\right|_{t=0}(\exp (t \xi) \cdot h)=T_{e} R_{h} \xi
$$

that is, the right-invariant vector field defined by $\xi$.
An action $\Phi$ of $G$ on a manifold $Q$ induces an action of the Lie group on the tangent bundle of $Q, \hat{\Phi}: G \times T Q \rightarrow T Q$ defined by $\hat{\Phi}\left(g, v_{q}\right)=T \Phi_{g}\left(v_{q}\right)\left(=\Phi_{g *}\left(v_{q}\right)\right)$ for any $g \in G$ and $v_{q} \in T_{q} Q . \hat{\Phi}$ is called the lifted action of $\Phi$.

### 2.2 Riemannian Geometry

A Riemannian metric $\mathcal{G}$ is a $(0,2)$-tensor on a manifold $Q$ which is symmetric and positivedefinite. This means that $\mathcal{G}\left(v_{q}, w_{q}\right)=\mathcal{G}\left(w_{q}, v_{q}\right)$ for all $v_{q}, w_{q} \in T_{q} Q, \mathcal{G}\left(v_{q}, v_{q}\right) \geq 0$, and $\mathcal{G}\left(v_{q}, v_{q}\right)=0$, if and only if $v_{q}=0$.

A Riemannian manifold is a pair $(Q, \mathcal{G})$, where $Q$ is a differentiable manifold and $\mathcal{G}$ is a Riemannian metric. The metric is locally determined by the matrix $\left(g_{i j}\right)_{1 \leq i, j \leq n}$ where $\left(g_{i j}\right)=\mathcal{G}\left(\frac{\partial}{\partial q^{i}}, \frac{\partial}{\partial q^{j}}\right)$.

Given a Riemannian manifold, we may consider the "musical" isomorphisms by $\mathrm{b}_{\mathrm{g}}$ : $T Q \rightarrow T^{*} Q$ the induced vector bundle isomorphism and by $\# \mathrm{~g}: T^{*} Q \rightarrow T Q$ the inverse isomorphism defined as $b_{\mathcal{G}}(v)=\mathcal{G}(v, \cdot)$ and $\#_{g}=b_{g}^{-1}$, If $f \in \mathcal{C}^{\infty}(Q)$, we define its gradient as the vector field $\operatorname{grad}(f)=\# g(d f)$.

Every Riemannian manifold is endowed with a canonical affine connection, called the Levi-Civita connection. In general, an affine connection is defined as an assignment $\nabla$ : $\mathfrak{X}(Q) \times \mathfrak{X}(Q) \rightarrow \mathfrak{X}(Q),(X, Y) \mapsto \nabla_{X} Y$ which is $\mathbb{R}$-bilinear and satisfies, for any $X, Y, Z \in$ $\mathfrak{X}(Q), f \in \mathcal{C}^{\infty}(Q), \nabla_{f X+g Y} Z=f \nabla_{X} Z+g \nabla_{Y} Z$ and $\nabla_{X}(f Y)=f \nabla_{X} Y+(d f \cdot X) Y$, where $d f \cdot X$ is the directional derivative of $f$ along $X$, or Lie derivative [31].

We shall call $\nabla_{X} Y$ the covariant derivative of $Y$ with respect to $X$. In local coordinates $\left(q^{A}\right)$ on $Q$, we have that

$$
\nabla_{X} Y=\left(\frac{\partial Y^{i}}{\partial q^{j}} X^{j}+\Gamma_{j k}^{i} X^{j} Y^{k}\right) \frac{\partial}{\partial q^{i}},
$$

where in local coordinates, the $n^{3}$ functions $\Gamma_{i j}^{k}$ (Christoffel symbols for $\nabla$ ) are given by

$$
\nabla_{\frac{\partial}{\partial q^{i}}} \frac{\partial}{\partial q^{j}}=\Gamma_{i j}^{k} \frac{\partial}{\partial q^{k}},
$$

A curve $c:[a, b] \rightarrow Q$ is a geodesic for $\nabla$ if $\nabla_{\dot{c}(t)} \dot{c}(t)=0$. Locally, the condition for a curve $t \mapsto\left(q^{1}(t), \ldots, q^{n}(t)\right)$ to be a geodesic can be expressed as

$$
\ddot{q}^{i}+\Gamma_{j k}^{i} \dot{q}^{j} \dot{q}^{k}=0, \quad 1 \leq i \leq n
$$

Other important objects related to an affine connection are the torsion tensor and the curvature tensor, which are defined, respectively, by

$$
\begin{array}{ll}
T: & \mathfrak{X}(Q) \times \mathfrak{X}(Q) \rightarrow \mathfrak{X}(Q) \\
& (X, Y) \mapsto \nabla_{X} Y-\nabla_{Y} X-[X, Y] \\
R: & \mathfrak{X}(Q) \times \mathfrak{X}(Q) \times \mathfrak{X}(Q) \rightarrow \mathfrak{X}(Q) \\
& (X, Y, Z) \mapsto \nabla_{X} \nabla_{Y} Z-\nabla_{Y} \nabla_{X} Z-\nabla_{[X, Y]} Z .
\end{array}
$$

Locally, these objects are expressed as

$$
T_{i j}^{k}=\Gamma_{i j}^{k}-\Gamma_{j i}^{k}, \quad R_{i j k}^{w}=\frac{\partial \Gamma_{j k}^{w}}{\partial q^{i}}-\frac{\partial \Gamma_{i k}^{w}}{\partial q^{j}}+\Gamma_{j k}^{z} \Gamma_{i z}^{w}-\Gamma_{i k}^{z} \Gamma_{j z}^{w} .
$$

Now we consider a Riemannian metric $\mathcal{G}$ specifying the kinetic energy of the mechanical system. Consider the Levi-Civita connection $\nabla^{\mathcal{G}}$ on $Q$ as the unique affine connection which is torsion-less and metric with respect to $\mathcal{G}$. It is determined by the standard formula

$$
\begin{aligned}
2 \mathcal{G}\left(\nabla_{X}^{\mathcal{G}} Y, Z\right)= & X(\mathcal{G}(Y, Z))+Y(\mathcal{G}(X, Z))-Z(\mathcal{G}(X, Y)) \\
& +\mathcal{G}(X,[Z, Y])+\mathcal{G}(Y,[Z, X])-\mathcal{G}(Z,[Y, X])
\end{aligned}
$$

for all $X, Y, Z \in \mathfrak{X}(Q)$.
Fixed a potential function $V: Q \rightarrow \mathbb{R}$, the mechanical system is defined by the mechanical Lagrangian $L: T Q \longrightarrow \mathbb{R}$,

$$
\begin{equation*}
L\left(v_{q}\right)=\frac{1}{2} \mathcal{G}\left(v_{q}, v_{q}\right)-V(q), \quad \text { where } v_{q} \in T_{q} Q \tag{2.2.1}
\end{equation*}
$$

In this Riemannian context, we may write the equations of motion of the unconstrained mechanical system as

$$
\begin{equation*}
\nabla_{\dot{c}(t)}^{\mathcal{G}} \dot{c}(t)+\operatorname{grad}_{g} V(c(t))=0 \tag{2.2.2}
\end{equation*}
$$

Here, $\operatorname{grad}_{g} V$ is the vector field on $Q$ characterized by

$$
\mathcal{G}\left(\operatorname{grad}_{\mathcal{G}} V, X\right)=X(V), \quad \text { for every } X \in \mathfrak{X}(Q)
$$

In local coordinates, $\operatorname{grad}_{\mathcal{G}} V(c(t))=g^{i j} \frac{\partial V}{\partial q^{j}}$, where $\left(g^{i j}\right)$ is the inverse matrix of $\left(g_{i j}\right)$ an therefore may rewrite the equations of motion as

$$
\ddot{q}^{k}(t)+\Gamma_{i j}^{k}(c(t)) \dot{q}^{i}(t) \dot{q}^{j}(t)+g^{k i}(c(t)) \frac{\partial V}{\partial q^{i}}=0
$$

where $t \mapsto\left(q^{1}(t), \ldots, q^{n}\right)$ is the local representative of $c$.

### 2.3 Hamiltonian Mechanics and Symplectic Geometry

We begin our discussion by recalling some basic definitions in symplectic geometry [57, [54, [51]. Along this section, $V$ and $M$ respectively denote a real vector space and a smooth manifold. They do not necessarily have finite dimension.

Definition 2.3.1. Let $\omega: V \times V \longrightarrow \mathbb{R}$ be a bilinear map and define the morphism $\omega^{b}: V \longrightarrow V^{*}$ by

$$
\left\langle\omega^{b}(v) \mid w\right\rangle=\omega(v, w)
$$

We say that $\omega$ is weakly (resp. strongly) non-degenerate whenever $\omega^{b}$ is a monomorphism (resp. an isomorphism).

It turns out that, if $V$ is finite-dimensional, weak and strong non-degeneracy coincide. Thus, in this case, we simply use the term non-degenerate. We denote by $\Lambda^{p} V$ the set of $p$-sections on $V$, then we have that,

Proposition 2.3.2. Let $V$ be a finite-dimensional real vector space and let $\omega \in \Lambda^{2} V^{*}$ be a skew-symmetric bilinear map. The following holds,

1. $\omega$ is non-degenerate if and only if $V$ is even-dimensional ( $\operatorname{dim} V=2 n$ ) and the exterior nth-power $\omega^{n}$ is a volume form on $V$;
2. if $\omega$ is non-degenerate, then there exists a basis $\left(\varepsilon^{i}\right)_{i=1}^{2 n}$ in $V^{*}$ such that

$$
\left(\omega_{i j}\right)=\left(\begin{array}{cc}
0 & I \\
-I & 0
\end{array}\right)
$$

where $\omega=\omega_{i j} \varepsilon^{i} \otimes \varepsilon^{j}, 0$ is the $n$-by-n null matrix and $I$ is the $n$-dimensional identity matrix. Equivalently, $\omega=\sum_{i=1}^{n} \varepsilon^{i} \wedge \varepsilon^{n+i}$.

Definition 2.3.3. $A$ weak (resp. strong) symplectic form on a real vector space $V$ is a weakly (resp. strongly) non-degenerate 2-form $\omega$ on $V$. The pair $(V, \omega)$ is called a weak (resp. strong) symplectic vector space.

As before, we avoid the use of the terms weak and strong in the case of finite-dimensional vector spaces.

Example 2.3.4. Let $V$ be a real vector space of dimension n. Let $\left(e_{i}\right)_{i=1}^{n}$ be a basis of $V$ and let $\left(\varepsilon^{i}\right)_{i=1}^{n}$ be its dual counterpart (i.e; $\varepsilon^{i}\left(e_{j}\right)=\delta_{j}^{i}$ ). Then, with some abuse of notation, $\omega=\sum_{i=1}^{n} \varepsilon^{i} \wedge e_{i}$ is a non-degenerate 2-form in $V \times V^{*}$. Note that $\omega$ does not depend on the chosen basis $\left(e_{i}\right)_{i=1}^{n}$ of $V$. In fact, $\omega$ may be defined intrinsically by the following expression,

$$
\omega\left(\left(v_{1}, \alpha_{1}\right),\left(v_{2}, \alpha_{2}\right)\right)=\alpha_{2}\left(v_{1}\right)-\alpha_{1}\left(v_{2}\right) .
$$

In the following we denote by $\Omega^{p}(M)$ the set of $p$-forms on $M$.
Definition 2.3.5. Let $M$ be a smooth manifold, a tensor field $\omega \in \Omega^{2}(M)$ is weakly (resp. strongly) non-degenerate if the bilinear map $\omega_{x}: T_{x} M \times T_{x} M \longrightarrow \mathbb{R}$ is weakly (resp. strongly) non-degenerate, for each $x \in M$.

Proposition 2.3.6. Given a tensor field $\omega$ over $M$ of type $(0,2)$, let $\omega^{b}: \mathfrak{X}(M) \rightarrow \Omega(M)$ be the mapping defined by the contraction $\omega^{b}(X)=i_{X} \omega$. We have that $\omega^{b}$ is $\mathcal{C}^{\infty}(M)$ linear. Moreover, if $\omega$ is weakly (resp. strongly) non-degenerate, then $\omega^{b}$ is injective (resp. bijective).

Definition 2.3.7. Let $M$ be a smooth manifold, a weak (resp. strong) symplectic form is a weakly (resp. strongly) non-degenerate 2-form $\omega \in \Omega^{2}(M)$ which is in addition closed. The pair $(M, \omega)$ is called $a$ weak (resp. strong) symplectic manifold.

Theorem 2.3.8 (Darboux). Let $\omega$ be a 2-form over a finite-dimensional smooth manifold $M$. Then, $(M, \omega)$ is a symplectic manifold if and only if $M$ has even dimension $(\operatorname{dim} M=$ $2 n$ ) and there exist local coordinates $\left(q^{1}, \ldots, q^{n}, p_{1}, \ldots, p_{n}\right)$ such that $\omega$ has locally the form

$$
\omega=\mathrm{d} q^{i} \wedge \mathrm{~d} p_{i}
$$

Such coordinates are called Darboux or canonical coordinates.

Let $M_{1}$ and $M_{2}$ be finite dimensional differentiable manifolds, if ( $M_{1}, \omega_{1}$ ) and ( $M_{2}, \omega_{2}$ ) are symplectic manifolds then a $\mathcal{C}^{\infty}$ mapping $\varphi: M_{1} \rightarrow M_{2}$ is called symplectic if $\varphi^{*} \omega_{2}=\omega_{1}$. Using this, we can define a Hamiltonian system in general.

Definition 2.3.9. Let $(M, \omega)$ be a symplectic manifold and $H \in \mathcal{C}^{\infty}(M, \mathbb{R})$ a smooth real valued function on $M$. The vector field $X_{H}$ determined by the condition

$$
\begin{equation*}
i_{X_{H}} \omega=d H \tag{2.3.1}
\end{equation*}
$$

is called the Hamiltonian vector field with energy function $H$. We call $(M, \omega, H)$ a Hamiltonian mechanical system.

Let us now take the case when the finite dimensional manifold is $M=T^{*} Q$. In this case there exists a unique 1 -form $\theta$ on $T^{*} Q$ such that in any choice of canonical cotangent coordinates, $\theta=p_{i} d q^{i}$. Using this we can then define the canonical 2-form $\omega$ by $\omega=-d \theta=$ $d q^{i} \wedge d p_{i}$. It is then clear that $\left(T^{*} Q, \omega\right)$ is a symplectic manifold. A simple computation then shows that $(q(t), p(t))$ is an integral curve of $X_{H}$ if and only if Hamilton's equations holds:

$$
\begin{equation*}
\dot{q}^{i}=\frac{\partial H}{\partial p_{i}}, \quad \dot{p}_{i}=-\frac{\partial H}{\partial q^{i}} . \tag{2.3.2}
\end{equation*}
$$

Moreover, (2.0.2) and (2.3.2) are equivalent via the Legendre transform.
Embedded in the definition of Hamiltonian system above are the following facts.
Proposition 2.3.10. (Conservation of the Energy) Let $(M, \omega, H)$ be a Hamiltonian mechanical system and let $c(t)$ be an integral curve for $X_{H}$. Then $H(c(t))$ is a constant in $t$. Moreover, if $\phi_{t}$ is the flow of $X_{H}$ then $H \circ \phi_{t}=H$ for each $t$.

Proposition 2.3.11. (Volumen Preservation) Let $(M, \omega, H)$ be a Hamiltonian mechanical system and let $\phi_{t}$ be the flow of $X_{H}$. Then for each $t, \phi_{t}^{*} \omega=\omega$, that is, $\phi_{t}$ is a symplectic and volumen preserving.

Proposition 2.3.12. Let $F_{t} \in \operatorname{Diff}(M)$ be the flow of $X_{H}$, then $F_{t}^{*} \omega=\omega$ for each $t$, i.e, $\left\{F_{t}\right\}$ is a family of symplectomorphisms.

### 2.4 The Gotay-Nester-Hinds Algorithm

By definition, if $(M, \omega)$ is a symplectic manifold then the equation

$$
\begin{equation*}
i_{X} \omega=\alpha \tag{2.4.1}
\end{equation*}
$$

has always a unique solution $X \in \mathfrak{X}(M)$, whatever the 1 -form $\alpha \in T^{*} M$ is (Proposition (2.3.6). We suppose that $\operatorname{dim} M=2 n$; then the solution vector field $X \in X(M)$ is locally given by

$$
\begin{equation*}
X=\omega^{\sharp}(\alpha)=\left(\omega^{b}\right)^{-1}(\alpha)=\sum_{i, j=1}^{2 n} \omega^{i j} \alpha_{j} \frac{\partial}{\partial x^{i}}, \tag{2.4.2}
\end{equation*}
$$

where $\left(x^{1}, \ldots, x^{2 n}\right)$ are arbitrary local coordinates on $M, \omega^{i j}$ is the inverse coeficient matrix of $\omega$, with $\omega=\sum_{1 \leq i<j \leq 2 n} \omega_{i j} d x^{i} \wedge d x^{j}$, and $\alpha=\sum_{j=1}^{2 n} \alpha_{j} d x^{j}$. If we instead choose Darboux coordinates $\left(q^{1}, \ldots, q^{n}, p_{1}, \ldots, p_{n}\right)$ for $M$ and write

$$
X=X^{i} \frac{\partial}{\partial q^{i}}+X_{i} \frac{\partial}{\partial p_{i}} \quad \text { and } \quad \alpha=\alpha_{i} \mathrm{~d} q^{i}+\alpha^{i} \mathrm{~d} p_{i}
$$

then

$$
\begin{equation*}
X^{i}=\alpha^{i} \quad \text { and } \quad X_{i}=-\alpha_{i} . \tag{2.4.3}
\end{equation*}
$$

This equations will appear again in later sections in slightly different ways.
The aim of the Gotay-Nester-Hinds (GNH) algorithm (see reference [37]) is to study the equation 2.4.1 whenever the closed 2 -form $\omega$ is symplectic or degenerate, that is, when $\omega$ is presymplectic. It manages to circumvent the degeneracy problems that often appear in mechanics, even though it is totally geometric and may be studied appart of any physical meaning. Equation (2.4.1) could not be solvable for a presymplectic form $\omega$ over the whole manifold $M$, but it could be at some points of $M$. The objective of the GNH algorithm is to find a submanifold $N$ of $M$ such that equation 2.4.1 has solutions in $N$, more precisely, to find the biggest submanifold $N$ of $M$ such that there exists a vector field $X \in \mathfrak{X}(N)$ that satisfies

$$
\begin{equation*}
\left.i_{j_{*} X} \omega\right|_{N}=\left.\alpha\right|_{N} \tag{2.4.4}
\end{equation*}
$$

for a prescribed 1-form $\alpha \in \Omega(M)$, where $j$ is the inclusion $j: N \hookrightarrow M$. The manifold $N$ will, of course, depend on the 1 -form $\alpha$.

Remark 2.4.1. Even though they are quite similar, Equation 2.4.4) should not be confused with the following one

$$
i_{X}\left(j^{*} \omega\right)=j^{*} \alpha
$$

While the latter must be satisfied for any vector field $Y$ "over" $N$, that is

$$
\left(j^{*} \omega\right)(X, Y)=\left(j^{*} \alpha\right)(Y), \quad \forall Y \in \mathfrak{X}(N),
$$

the former is more restrictive and must be satisfied for any vector field $Y$ "along" $N$, that is

$$
\omega\left(j_{*} X, Y\right)=\alpha(Y), \quad \forall Y \in \mathfrak{X}(j) .
$$

Given a presymplectic 2-form $\omega$ over a manifold $M$, let $\alpha \in \Omega(M)$ be any 1-form. We start defining the subset $M_{1}$ of points $x$ of $M$ such that $\alpha(x)$ is in the range of $\omega^{b}(x)$, that is,

$$
M_{1}:=\left\{x \in M: \alpha(x) \in \omega^{b}(T M)\right\} .
$$

The subset $M_{1}$ does not need to be a manifold, fact that is imposed, being $j_{1}: M_{1} \hookrightarrow M$ the inclusion. The equation (2.4.1) restricted to $M_{1}$,

$$
\left.i_{X} \omega\right|_{M_{1}}=\left.\alpha\right|_{M_{1}},
$$

is solvable, but this does not imply that $X$ is a solution in the sense of equation (2.4.4). It could be possible that, at some point $x \in M_{1}$, the vector $X(x)$ is not tangent to $M_{1}$, what will happen when $\alpha(x)$ is not in the range of $\omega^{b}(x)$ restricted to $j_{1 *}\left(T M_{1}\right)$. We are then forced to define a new "submanifold" $j_{2}: M_{2} \hookrightarrow M_{1}$ by

$$
M_{2}:=\left\{x \in M_{1}: \alpha(x) \in \omega^{b}\left(j_{1 *}\left(T M_{1}\right)\right)\right\} .
$$

As before, the solutions of the equation (2.4.1) restricted to $M_{2}$,

$$
\left.i_{X} \omega\right|_{M_{2}}=\left.\alpha\right|_{M_{2}},
$$

may not be tangent to $M_{2}$, therefore we require $\left.\alpha\right|_{M_{2}}$ to be in the range of $\omega^{b}$ restricted to $\left(j_{2} \circ j_{1}\right)_{*}\left(T M_{2}\right)$.

We thus continue this process, defining a chain of further constraint submanifolds

$$
\ldots \hookrightarrow M_{l} \stackrel{j_{l}}{\hookrightarrow} M_{l-1} \hookrightarrow \ldots \hookrightarrow M_{1} \stackrel{j_{1}}{\hookrightarrow} M
$$

as follows

$$
\begin{equation*}
M_{l+1}:=\left\{x \in M_{l}: \alpha(x) \in \omega^{b}\left(\left(j_{1} \circ \cdots \circ j_{l}\right)_{*}\left(T M_{l}\right)\right)\right\} . \tag{2.4.5}
\end{equation*}
$$

At each step, we must assume that the constraint set $M_{l}$ is a smooth manifold (an alternate algorithm for the case when the constraint sets are not smooth submanifolds may be found in [46]). In the end, the algorithm will stop when, for some $k \geq 0, M_{k+1}=M_{k}$. We then take $N:=M_{k}$ and $j:=j_{k} \circ \cdots \circ j_{1}$. Mainly, two different cases may happen:
$-\operatorname{dim} N=0$ : The Hamiltonian system $(M, \omega, \alpha)$ has no dynamics. Furthermore, if $N=\emptyset$, there are no solutions at all and $(M, \omega, \alpha)$ does not accurately describe the dynamics of any system. On the contrary, if $N \neq \emptyset$, then $N$ consists of (steady) isolated points.

- $\operatorname{dim} N \neq 0:(M, \omega, \alpha)$ describes a dynamical system restricted to $N$ and we have completely consistent equations at motion on $N$ of the form

$$
\left.\left(i_{X} \omega-\alpha\right)\right|_{N}=0
$$

### 2.5 Skinner and Rusk Formalism: An unifying framework

In this section we describe the unifying formalism of the Lagrangian-Hamiltonian mechanics introduced by R. Skinner and R. Rusk in [65]. This formalism includes the Lagrangian and Hamiltonian formalism of first order autonomous systems, and allows us to recover the dynamic equations in one or another, just as we shall see below.

We consider a dynamical system of $n$ degrees of freedom modeled by a configuration space $Q$ of dimension $n$. The behavior of this is described by the Lagrangian $L \in \mathcal{C}^{\infty}(T Q)$ which contain the information of the dynamics associated with the system.

Consider the following phase space,

$$
T Q \times_{Q} T^{*} Q \simeq T Q \oplus_{Q} T^{*} Q
$$

that is, the Whitney sum of the velocities space and phase space. This space is endowed with two canonical projections, $p r_{1}: T Q \times_{Q} T^{*} Q \rightarrow T Q$ and $p r_{2}: T Q \times_{Q} T^{*} Q \rightarrow T^{*} Q$. In what follows we denote by $W$ the Whitney sum. Using the canonical projections of the fiber bundle and cotangent bundle over the manifold $Q$ we have the following commutative diagram,


Figure 2.1: Skinner and Rusk formalism
If $(U, \varphi)$ is a local chart of $Q$, and $\varphi=\left(q^{i}\right), i=1, \ldots, n$; we can obtain natural coordinates of $T Q$ and $T^{*} Q$ by $\tau_{Q}^{-1}(U)$ and $\pi_{Q}^{-1}(U)$ respectively. These coordinates are $\left(q^{i}, v^{i}\right)$ and $\left(q^{i}, p_{i}\right)$ respectively. Therefore, $\left(q^{i}, v^{i}, p_{i}\right)$ are natural coordinates in $W$. Observe that $\operatorname{dim}(W)=3 n$.

Let $\theta \in \Omega\left(T^{*} Q\right)$ be the Liouville one-form of the cotangent bundle and $\omega=-d \theta \in$ $\Omega^{2}\left(T^{*} Q\right)$ the canonical symplectic two form on $T^{*} Q$. From this we can define the 2-form $\Omega$ on $W$ as

$$
\Omega=p r_{2}^{*}(\omega) \in \Omega^{2}(W)
$$

Observe that $\Omega$ is a closed 2 -form, because

$$
\Omega=p r_{2}^{*}(\omega)=p r_{2}^{*}(-d \theta)=-d p r_{2}^{*}(\theta)
$$

Nevertheless, this form is degenerate and therefore is a presymplectic form. Using the expression on local coordinates, $\omega=d q^{i} \wedge d p_{i}$ and $p r_{2}\left(q^{i}, v^{i}, p_{i}\right)=\left(q^{i}, p_{i}\right)$ we have that

$$
\begin{aligned}
\Omega & =p r_{2}^{*}(\omega)=p r_{2}^{*}\left(d q^{i} \wedge d p_{i}\right)=p r_{2}^{*}\left(d q^{i}\right) \wedge p r_{2}^{*}\left(d p_{i}\right) \\
& =d p r_{2}^{*}\left(q^{i}\right) \wedge d p r_{2}^{*}\left(p_{i}\right)=d q^{i} \wedge d p_{i}
\end{aligned}
$$

From this local expression, is clear that $\left\{\frac{\partial}{\partial v^{i}}\right\}$ is a local basis of the kernel for $\Omega$, that is,

$$
\operatorname{Ker} \Omega=\left\langle\frac{\partial}{\partial v^{i}}\right\rangle
$$

and therefore the 2 -form $\Omega$ is degenerate.
Then, we have a presymplectic manifold $(W, \Omega)$ and our objective is to obtain a presymplectic Hamiltonian system in order to deduce the dynamic equations following the procedure given in [61, [18, [5]. Nevertheless, in this formalism we suppose that the information of the dynamics is specified by a Lagrangian $L \in \mathcal{C}^{\infty}(T Q)$; which is not enough to directly define a Hamiltonian system.

To define a Hamiltonian function first consider the function $C \in \mathcal{C}^{\infty}(W)$, defined canonically in the following way: if $p \in Q, v_{p} \in T_{p} Q$ is a tangent vector of $Q$ at $p$ and $\alpha_{p} \in T^{*} Q$ is a covector on $p$, we define $C$ as

$$
\begin{array}{ll}
C: & T Q \times_{Q} T^{*} Q \rightarrow \mathbb{R} \\
& \left(p, v_{p}, \alpha_{p}\right) \mapsto\left\langle\alpha_{p}, v_{p}\right\rangle,
\end{array}
$$

where $\left\langle\alpha_{p}, v_{p}\right\rangle \equiv \alpha_{p}\left(v_{p}\right)$ is the canonical pairing between elements of $T Q$ and $T^{*} Q$. In local coordinates, if we consider a local chart on $p \in Q, \alpha_{p}=\left.p_{i} d q^{i}\right|_{p}, v_{p}=\left.v^{i} \frac{\partial}{\partial q^{i}}\right|_{p}$; the local expression of $C$ is

$$
C\left(p, v_{p}, \alpha_{p}\right)=\left\langle\alpha_{p}, v_{p}\right\rangle=\left\langle\left. p_{i} d q^{i}\right|_{p},\left.v^{i} \frac{\partial}{\partial q^{i}}\right|_{p}\right\rangle=\left.p_{i} v^{i}\right|_{p}
$$

Then, we define the Hamiltonian $H \in \mathcal{C}^{\infty}(W)$ by

$$
H=C-p r_{1}^{*}(L)=p_{i} v^{i}-L\left(q^{j}, v^{j}\right)
$$

and therefore we have a presymplectic Hamiltonian system $(W, \Omega, H)$. Since we have a presymplectic Hamiltonian system, the GNH algorithm given in Section 2.4 can be applied and the equations of motion are given by the solutions of

$$
i_{X} \Omega=d H
$$

where $X \in \mathfrak{X}(W)$ is the Hamiltonian vector field of the system.

## Chapter 3

## Mechanical Systems on Lie Groups

Poincaré contributions are relevant in the development of mechanical systems involving Lie groups. First is his work on gravitating fluid problem [63], continuing the line of investigation begun by MacLaurin, Jacobi and Riemann. Following the historical review given by D. Holm [39], [40], Poincare's work is summarized in Chandrasekhar [1967,1977] (see Poincaré [62] for original treatments). This background led to his famous paper, Poincaré [62] in which he laid out the basic equations of Euler type, including the rigid body, heavy top and fluids as special cases.

In this chapter, we give an abstract version of these equations which are determined by a Lagrangian on a Lie algebra. In the paper by Poincaré 62] these equations have the name of Euler-Poincaré equations. This is why now a days the equations of motion for this kind of mechanical systems are known as the Euler-Poincaré equations.

These equations are characterized by Euler-Poincaré theorem. This theorem will be presented in this chapter. Briefly, it stat that if we suppose that the configuration space for our mechanical systems is a Lie group $G$ and let $L: T G \rightarrow \mathbb{R}$ be a left-invariant Lagrangian; denoting by $l: \mathfrak{g} \rightarrow \mathbb{R}$ the restriction of $L$ to the tangent space of $G$ at the identity and for a curve $g(t) \in G$ let $\xi(t)=g^{-1}(t) \dot{g}(t)$; then the following are equivalent (see [8], [39, [55]):
(i) $g(t)$ satisfies the Euler-Lagrange equations for $L$ on $G$,
(ii) The Euler-Poincaré equations hold:

$$
\begin{equation*}
\frac{d}{d t} \frac{\partial l}{\partial \xi}=a d_{\xi}^{*} \frac{\partial l}{\partial \xi}, \tag{3.0.1}
\end{equation*}
$$

where $a d_{\xi}: \mathfrak{g} \rightarrow \mathfrak{g}$ is defined by $a d_{\xi} \eta=[\xi, \eta]$ and $a d_{\xi}^{*}$ is its dual.
These equations are valid for either finite or infinite dimensional Lie algebras and for this reason sometimes we use the notation $\delta$ instead of $\partial$ to denote the partial derivative. For fluids, Poincaré was aware that one needs to use infinite dimensional Lie algebras (see [62]). He was also aware that one has to be careful with the signs in the equations; for example, for rigid body dynamics one uses the equations as they stand, but for fluids, one needs to be careful about the conventions for the Lie algebra operation $a d_{\xi}$. (see [40] and reference therein).

Moreover, in the finite dimensional case, by making the following Legendre transformation from $\mathfrak{g}$ to $\mathfrak{g}^{*}$ :

$$
\begin{equation*}
\mu=\frac{\partial l}{\partial \xi}, \quad h(\mu)=\mu_{i} \xi^{i}-l(\xi) \tag{3.0.2}
\end{equation*}
$$

it follows that the Euler-Poincaré equations are equivalent to the Lie-Poisson equations:

$$
\begin{equation*}
\frac{d \mu}{d t}=a d_{\frac{\partial b}{\partial \mu}}^{*} \mu . \tag{3.0.3}
\end{equation*}
$$

For example, consider the Lie algebra $\mathbb{R}^{3}$ with the usual vector cross product. For $l: \mathbb{R}^{3} \rightarrow \mathbb{R}$, the Euler-Poincaré equations become

$$
\frac{d}{d t} \frac{\partial l}{\partial \Omega}=\frac{\partial l}{\partial \Omega} \times \Omega
$$

which are the Euler equations for the rigid body motion. Here $\Omega=\left(\Omega_{1}, \Omega_{2}, \Omega_{3}\right)$ is the body angular velocity and $l(\Omega)=\frac{1}{2}\left(I_{1} \Omega_{1}^{2}+I_{2} \Omega_{2}^{2}+I_{3} \Omega_{3}^{2}\right)$ the Lagrangian which is the (rotational) kinetic energy of the rigid body.

These equations were written for a certain class of Lagrangians $l$ by Lagrange 48, while Poincaré 62] generalized them to an arbitrary Lie algebra [40].

The rigid body equations are usually written as

$$
\begin{aligned}
I_{1} \dot{\Omega}_{1} & =\left(I_{2}-I_{3}\right) \Omega_{2} \Omega_{3}, \\
I_{2} \dot{\Omega}_{2} & =\left(I_{3}-I_{1}\right) \Omega_{3} \Omega_{1}, \\
I_{3} \dot{\Omega}_{3} & =\left(I_{1}-I_{2}\right) \Omega_{1} \Omega_{2},
\end{aligned}
$$

where $I_{1}, I_{2}, I_{3}$ are the moments of inertia of the rigid body.
Now, we introduce the angular momenta

$$
\Pi_{i}=I_{i} \Omega_{i}=\frac{\partial l}{\partial \Omega_{i}} \quad \quad i=1,2,3
$$

so that the Euler equations become

$$
\begin{aligned}
& \dot{\Pi}_{1}=\frac{I_{2}-I_{3}}{I_{2} I_{3}} \Pi_{2} \Pi_{3} \\
& \dot{\Pi}_{2}=\frac{I_{2}-I_{1}}{I_{3} I_{1}} \Pi_{3} \Pi_{1}, \\
& \dot{\Pi}_{3}=\frac{I_{1}-I_{2}}{I_{1} I_{2}}
\end{aligned}
$$

that is,

$$
\dot{\Pi}=\Pi \times \Omega
$$

which are the Lie-Poisson equations for the rigid body motion. These equations are the Hamiltonian version of Euler'e equations for the Hamiltonian

$$
H=\frac{1}{2}\left(\frac{\Pi_{1}^{2}}{I_{1}}+\frac{\Pi_{2}^{2}}{I_{2}}+\frac{\Pi_{3}^{2}}{I_{3}}\right) .
$$

### 3.1 Trivialization of a cotangent bundle of a Lie group

Let $G$ be a Lie group and $\mathfrak{g}$ its Lie algebra. Given $\xi \in \mathfrak{g}$ the differential equation

$$
\dot{g}=g \xi, \quad g(0)=g_{0}, \quad g \in G
$$

has an unique solution $g(t)=g_{0} \exp (t \xi)$.
Proposition 3.1.1. Given $g \in G$,

$$
T_{g} G=\{g \xi, \mid \xi \in \mathfrak{g}\}=g \mathfrak{g}
$$

Proof: The tangent space at $g$ of $G$ is:

$$
T_{g} G=\{\dot{g}(0) \mid g(t) \in G, g(0)=g\}
$$

We can translate this curve starting at $g$ to another curve starting at the identity $e$ :

$$
c(t)=g^{-1}(0) g(t)
$$

Then

$$
\dot{c}(0)=g^{-1}(0) \dot{g}(0) \in \mathfrak{g}
$$

Denoting by $\xi=\dot{c}(0)$ we have to

$$
\dot{g}(0)=g \xi
$$

Thus, we deduce that $T_{g} G \subseteq g \mathfrak{g}$. Since both have the same dimension, it follows that $T_{g} G=g \mathfrak{g}$.

Definition 3.1.2. Let $E$ a vector space where $\operatorname{dim} E=\operatorname{dim} Q$. A trivialization of a cotangent bundle $T^{*} Q$ of a differentiable manifold $M$ is a diffeomorphism

$$
\Psi: Q \times E \rightarrow T^{*} Q
$$

such that

- $\Psi(q, e) \in T_{q}^{*} Q$, con $(q, e) \in Q \times E$,
- $\Psi(q, \cdot): E \rightarrow T_{q}^{*} Q$ is an isomorphism between vector spaces for each $q \in Q$.

For all points ( $q, e$ ) we have the following identifications [56]

$$
\begin{aligned}
& T_{(q, e)}(E \times Q) \simeq T_{e} E \oplus T_{q} Q \simeq E \times T_{q} Q \\
& T_{(q, e)}^{*}(E \times Q) \simeq T_{e}^{*} E \oplus T_{q}^{*} Q \simeq E^{*} \times T_{q}^{*} Q
\end{aligned}
$$

Now, let $G$ be a Lie group and consider the left-multiplication on itself

$$
G \times G \longrightarrow G, \quad(g, h) \rightarrow £_{g}(h)=g h .
$$

Obviously $£_{g}$ is a diffeomorphism. (The same is valid for the right-translation, but in the sequel we only work with the left-translation, for sake of simplicity).

This left multiplication allows us to trivialize the tangent bundle $T G$ and the cotangent bundle $T^{*} G$ as follows

$$
\begin{aligned}
T G & \rightarrow G \times \mathfrak{g}, \quad(g, \dot{g}) \longmapsto\left(g, g^{-1} \dot{g}\right)=\left(g, T_{g} £_{g^{-1}} \dot{g}\right)=(g, \xi), \\
T^{*} G & \rightarrow G \times \mathfrak{g}^{*}, \quad\left(g, \alpha_{g}\right) \longmapsto\left(g, T_{e}^{*} £_{g}\left(\alpha_{g}\right)\right)=(g, \alpha),
\end{aligned}
$$

where $\mathfrak{g}=T_{e} G$ is the Lie algebra of $G$ and $e$ is the neutral element of $G$. In the same way, we have the following identifications: $T T G \equiv G \times 3 \mathfrak{g}, T^{*} T G=G \times \mathfrak{g} \times 2 \mathfrak{g}^{*}, T T^{*} G=$ $G \times \mathfrak{g}^{*} \times \mathfrak{g} \times \mathfrak{g}^{*}$ and $T^{*} T^{*} G=G \times 3 \mathfrak{g}^{*}$.

Now, we will write the as is write the Liouville 1 -form $\Theta \in \Lambda^{1}\left(T^{*} G\right)$ as a 1-form $\widehat{\Theta}=\mathcal{L}^{*} \Theta \in \Lambda^{1}\left(G \times \mathfrak{g}^{*}\right)$. In a similar way, we will obtain the 2-form $\widetilde{\omega}=-d \widehat{\Theta}=\mathcal{L}^{*} \omega$, where $\omega$ is the canonical symplectic two form on $T^{*} G$.

To compute these forms we need first to find out what is the tangent application of $\tau \circ \mathcal{L}$. Given $\xi \in \mathfrak{g}$ and $\beta \in \mathfrak{g}^{*}$ consider the application

$$
\begin{aligned}
\varphi_{t}^{(\xi, \beta)}: G \times \mathfrak{g}^{*} & \longrightarrow G \times \mathfrak{g}^{*} \\
(g, \alpha) & \longrightarrow(g \exp t \xi, \alpha+t \beta)
\end{aligned}
$$

$\varphi_{t}^{(\xi, \beta)}$ is the flow of the vector fields $X^{(A, \beta)}$ on $G \times \mathfrak{g}^{*}$ defined by

$$
X^{(\xi, \beta)}(g, \alpha)=(g \xi, \beta)
$$

Therefore, the tangent application of $\tau \circ \mathcal{L}$ is:

$$
\begin{aligned}
T_{(g, \alpha)}(\tau \circ \mathcal{L})(g \xi, \beta) & =\left.\frac{d}{d t}\right|_{t=0} \tau \circ \mathcal{L}\left(\varphi_{t}^{(\xi, \beta)}(g, \alpha)\right) \\
& =\left.\frac{d}{d t}\right|_{t=0} g \exp t \xi=g \xi
\end{aligned}
$$

To this we can compute $\widehat{\Theta}$ :

$$
\begin{aligned}
\left\langle\widehat{\Theta}_{(g, \alpha)},(g \xi, \beta)\right\rangle & =\left\langle\Theta_{\mathcal{L}(g, \alpha)}, T_{(g, \alpha)} \mathcal{L}(g \xi, \beta)\right\rangle \\
& =\left\langle\alpha_{g}, T_{(g, \alpha)}(\tau \circ \mathcal{L})(g \xi, \beta)\right\rangle \\
& =\left\langle\alpha_{g}, g \xi\right\rangle=\langle\alpha, \xi\rangle=\alpha(\xi) .
\end{aligned}
$$

To compute the 2 -form $\widehat{\omega}$ we use the formula

$$
\begin{aligned}
& -d \widehat{\Theta}\left(X^{(\xi, \beta)}, X^{(\eta, \gamma)}\right)= \\
& \left.\left.\left.\left.\left.-X^{(\xi, \beta)}\right\lrcorner d\left(X^{(\eta, \gamma)}\right\lrcorner \widehat{\Theta}\right)+X^{(\eta, \gamma)}\right\lrcorner d\left(X^{(\xi, \beta)}\right\lrcorner \widehat{\Theta}\right)+\left[X^{(\xi, \beta)}, X^{(\eta, \gamma)}\right]\right\lrcorner \widehat{\Theta}
\end{aligned}
$$

where $\lrcorner$ denotes the contraction operator.
After some computations, we have

$$
\begin{aligned}
\left.\left.X^{(\xi, \beta)}\right\lrcorner d\left(X^{(\eta, \gamma)}\right\lrcorner \widehat{\Theta}\right)(g, \alpha) & \left.=L_{X(\xi, \beta)}\left(X^{(\eta, \gamma)}\right\lrcorner \widehat{\Theta}\right)(g, \alpha) \\
& \left.=\left.\frac{d}{d t}\right|_{t=0}\left(\varphi_{t}^{(\xi, \beta)}\right)^{*}\left(X^{(\eta, \gamma)}\right\lrcorner \widehat{\Theta}\right)(g, \alpha) \\
& =\left.\frac{d}{d t}\right|_{t=0}\left\langle\widehat{\Theta}(g \exp t \xi, \alpha+t \beta), X^{(\eta, \gamma)}(g \exp t \xi, \alpha+t \beta)\right\rangle \\
& =\left.\frac{d}{d t}\right|_{t=0}\langle\widetilde{\Theta}(g \exp t \xi, \alpha+t \beta),(g \exp (t \xi) \eta, \gamma)\rangle \\
& =\left.\frac{d}{d t}\right|_{t=0}(\alpha+t \beta)(\eta)=\beta(\eta)
\end{aligned}
$$

The second term is computed in a similar form being this $\alpha(\xi)$.
For the third factor we observe that

$$
\begin{aligned}
{\left[X^{(\xi, \beta)}, X^{(\eta, \gamma)}\right](g, \alpha) } & =\left.\frac{d}{d t}\right|_{t=0}\left(\varphi_{-\sqrt{t}}^{(\eta, \gamma)} \circ \varphi_{-\sqrt{t}}^{(\xi, \beta)} \circ \varphi_{\sqrt{t}}^{(\eta, \gamma)} \circ \varphi_{\sqrt{t}}^{(\xi, \beta)}\right)(g, \alpha) \\
& =\left.\frac{d}{d t}\right|_{t=0}(g \exp \sqrt{t} \xi \exp \sqrt{t} \eta \exp -\sqrt{t} \xi \exp -\sqrt{t} \eta, \alpha) \\
& =(g[\xi, \eta], 0)
\end{aligned}
$$

Therefore,

$$
\widehat{\Theta}\left(\left[X^{(\xi, \beta)}, X^{(\eta, \gamma)}\right]\right)(g, \alpha)=\widehat{\Theta}(g, \alpha)(g[\xi, \eta], 0)=\alpha([\xi, \eta])
$$

Finally,

$$
\widetilde{\omega}(g, \alpha)((g \xi, \beta),(g \eta, \gamma))=-\beta(\eta)+\alpha(\xi)+\alpha([\xi, \eta])
$$

### 3.2 Euler-Arnold Equations

In this section we find the Hamiltonian equation on the cotangent bundle of a Lie group, which has been trivialized using a left translation.

We suppose that $\mathcal{H}: T^{*} G \longrightarrow \mathbb{R}$ is a differentiable function. The Hamiltonian vector field $X_{\mathscr{H}}$ over the symplctic manifold $\left(T^{*} G, \Omega\right)$ is defined by the equations

$$
i_{X_{\mathfrak{H}}} \omega=d \mathcal{H}
$$

Consider the Hamiltonian function $H: G \times \mathfrak{g}^{*} \longrightarrow \mathbb{R}$ defined by $H(g, \alpha)=\mathcal{H}(\mathcal{L}(g, \alpha))=$ $\mathcal{H}\left(\alpha_{g}\right)$. In this way we want to search the vector field

$$
i_{X_{H}} \widetilde{\omega}=d H
$$

We will find the Hamiltonian equations; that is, the equations which characterize the integral curves of the Hamiltonian vector field $X_{H}$ on $G \times \mathfrak{g}^{*}$. In this context the equations are known as Euler-Arnold equations.

For each function $f: G \times \mathfrak{g}^{*} \longrightarrow \mathbb{R}$ denote by $\frac{\partial f}{\partial g}$ and $\frac{\partial f}{\partial \alpha}$ the partial derivatives of $f$; that is

- $\frac{\partial f}{\partial g}$ is the differential of the restriction of $f$ with values $\alpha=$ constant,
- $\frac{\partial f}{\partial \alpha}$ is the differential of the restriction of $f$ with values $g=$ constant.

Then $\frac{\partial f}{\partial g}(g, \alpha) \in T_{g}^{*} G$ and $\frac{\partial f}{\partial \alpha}(g, \alpha) \in\left(\mathfrak{g}^{*}\right)^{*}=\mathfrak{g}$. Therefore, for $\eta \in \mathfrak{g}, \beta \in \mathfrak{g}^{*}$.

$$
\begin{aligned}
d H(g, \alpha)(g \eta, \beta) & =\frac{\partial f}{\partial g}(g, \alpha) g \eta+\frac{\partial f}{\partial \alpha}(g, \alpha) \beta \\
& =\frac{\partial f}{\partial g}(g, \alpha) T_{e} L_{g} \eta+\beta\left(\frac{\partial f}{\partial \alpha}(g, \alpha)\right)
\end{aligned}
$$

As $X_{H}(g, \alpha) \in T_{g} G \times \mathfrak{g}^{*}$ the, we can write

$$
X_{H}(g, \alpha)=(g X(g, \alpha), \Lambda(g, \alpha))=\left(T_{g} L_{g}(X(g, \alpha)), \Lambda(g, \alpha)\right)
$$

But, by definition of $X_{H}$ :

$$
d H(g, \alpha)(g \eta, \beta)=\widetilde{\omega}(g, \alpha)(g X(g, \alpha), \Lambda(g, \alpha)),(g \eta, \beta))
$$

for all $\eta \in \mathfrak{g}$ y $\beta \in \mathfrak{g}^{*}$. In other form,

$$
d H(g, \alpha)(g \eta, \beta)=-\Lambda(g, \alpha)(\eta)+\beta(X(g, \alpha))+\alpha([X(g, \alpha), \eta])
$$

Taking $\eta=0$ we obtain that

$$
\beta\left(\frac{\partial H}{\partial \alpha}(g, \alpha)\right)=\beta(X(g, \alpha))
$$

for all $\beta \in \mathfrak{g}^{*}$. Therefore,

$$
X(g, \alpha)=\frac{\partial H}{\partial \alpha}(g, \alpha)
$$

Now, if we take $\beta=0$, then,

$$
\frac{\partial H}{\partial g}(g, \alpha) g \eta=-\Lambda(g, \alpha) \eta+\alpha([X(g, \alpha), \eta])
$$

or alternatively

$$
\frac{\partial H}{\partial g}(g, \alpha) T_{e} L_{g} \eta=-\Lambda(g, \alpha) \eta+\left(a d_{X}^{*} \alpha\right)(\eta)
$$

for all $\eta \in \mathfrak{g}$. Therefore,

$$
\Lambda(g, \alpha)=-\left(T_{e} L_{g}\right)^{*} \frac{\partial H}{\partial g}(g, \alpha)+a d_{\partial H / \partial \alpha}^{*} \alpha
$$

Then, the Euler-Arnold equations are

$$
\begin{aligned}
\dot{g} & =T_{I} L_{g} \frac{\partial H}{\partial \alpha}(g, \alpha) \\
\dot{\alpha} & =-\left(T_{e} L_{g}\right)^{*} \frac{\partial H}{\partial g}(g, \alpha)+a d_{\partial H / \partial \alpha}^{*} \alpha
\end{aligned}
$$

When the Hamiltonian $\mathcal{H}$ is left-invariant, $\frac{\partial H}{\partial g}=0$ and the Euler-Arnold equations are written as the Lie-Poisson equations

$$
\begin{aligned}
\dot{g} & =T_{e} L_{g} \frac{\partial H}{\partial \alpha}(g, \alpha) \\
\dot{\alpha} & =a d_{\partial H / \partial \alpha}^{*} \alpha
\end{aligned}
$$

### 3.2.1 Euler-Arnold equations for the rigid body

In this subsection we compute the Euler-Arnold equations of the rigid body. We refer to [56], [55], 39] and [50] on the results given in this subsection.

First recall that the Lie algebra $\mathfrak{s o}(3)$ is equipped with the Lie bracket [, ] defined by the relations

$$
\left[E_{1}, E_{2}\right]=E_{1} E_{2}-E_{2} E_{1}=E_{3}, \quad\left[E_{2}, E_{3}\right]=E_{1}, \quad\left[E_{3}, E_{1}\right]=E_{2}
$$

where $\left\{E_{1}, E_{2}, E_{3}\right\}$ is the standard basis of $\mathfrak{s o}(3)$ :

$$
E_{1}=\left(\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & 1 & 0
\end{array}\right), \quad E_{2}=\left(\begin{array}{lll}
0 & 0 & 1 \\
0 & 0 & 0 \\
-1 & 0 & 0
\end{array}\right), \quad E_{3}=\left(\begin{array}{ccc}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right) .
$$

Shortly,

$$
\left[E_{i}, E_{j}\right]=\sum_{k=1}^{3} \epsilon_{i j k} E_{k}
$$

where $\epsilon_{i j k}=0$ if some subindex coincides; $\epsilon_{i j k}=1$ if $(i j k)$ is an even permutation of (123); and $\epsilon_{i j k}=-1$ if $(i j k)$ is an odd permutation.

Now, consider the linear application $i: \mathfrak{s o}(3) \longrightarrow \mathbb{R}^{3}$ defined by

$$
i\left(\begin{array}{lll}
0 & -a_{3} & a_{2} \\
a_{3} & 0 & -a_{1} \\
-a_{2} & a_{1} & 0
\end{array}\right)=\left(a_{1}, a_{2}, a_{3}\right)=a
$$

Observe that $i\left(E_{1}\right)=(1,0,0), i\left(E_{2}\right)=(0,1,0)$ and $i\left(E_{3}\right)=(0,0,1)$.
By the identification

$$
i([\xi, \eta])=i\left(\sum_{i, j, k=1}^{3} \epsilon_{i j k} \xi_{i} \eta_{j} E_{k}\right)=\sum_{i, j, k=1}^{3} \epsilon_{i j k} \xi_{i} \eta_{j}=\xi \times \eta
$$

we have that,
Proposition 3.2.1. $i$ is an isomorphism between the Lie algebras $(\mathfrak{s o}(3),[]$,$) and \left(\mathbb{R}^{3}, \times\right)$, where $\times$ denotes the cross product in $\mathbb{R}^{3}$.

Sometimes we denote $i^{-1}(\mathbf{a})=\hat{\mathbf{a}}$. We shall revisit the last property using this notation. As $\mathbf{a} \times \mathbf{b}=\hat{\mathbf{a}} \mathbf{b}$ then, for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^{3}$ it follows that

$$
\begin{aligned}
{[\hat{\mathbf{a}}, \hat{\mathbf{b}}] \mathbf{c} } & =(\hat{\mathbf{a}} \hat{\mathbf{b}}) \mathbf{c}-(\hat{\mathbf{b}} \hat{\mathbf{a}}) \mathbf{c} \\
& =\hat{\mathbf{a}}(\mathbf{b} \times \mathbf{c})-\hat{\mathbf{b}}(\mathbf{a} \times \mathbf{c}) \\
& =\mathbf{a} \times(\mathbf{b} \times \mathbf{c})-\mathbf{b} \times(\mathbf{a} \times \mathbf{c}) \\
& =(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}=\hat{\mathbf{a} \times \mathbf{b}} \mathbf{c}
\end{aligned}
$$

Lemma 3.2.2. Let us consider an element $g \in S O(3)$. Then for all vectors $\boldsymbol{u}, \boldsymbol{v}$ in $\mathbb{R}^{3}$,

$$
g(\boldsymbol{u} \times \boldsymbol{v})=(g \boldsymbol{u}) \times(g \boldsymbol{v}) .
$$

Proof. Fix an orthogonal basis oriented to the right in $\mathbb{R}^{3}$. Let $\left\{u_{1}, u_{2}, u_{3}\right\}$ and $\left\{v_{1}, v_{2}, v_{3}\right\}$ be the components of $\mathbf{u}$ and $\mathbf{v}$ is this basis. In the same basis denote the components of $g$ by $g^{i j}$. As $g$ is a rotation then, $g^{i j} g^{k j}=\delta^{i k}$ where $\delta^{i k}$ is the Kronecker delta. The $i$-th coordinate of $\mathbf{u} \times \mathbf{v}$ will be equal to $(\mathbf{u} \times \mathbf{v})_{i}=\sum_{i, j=1}^{3} \epsilon_{i j k} u_{j} v_{k}$. For the $k$-th coordinate of $(g \cdot \mathbf{u}) \times(g \cdot \mathbf{v})$ it holds that

$$
\begin{aligned}
((g \cdot \mathbf{u}) \times(g \cdot \mathbf{v}))_{k} & =\epsilon_{i m k} g^{i j} g^{m n} u_{j} v_{n} \\
& =\epsilon_{i m l} \delta^{k l} g^{i j} g^{m n} u_{j} v_{n} \\
& =\epsilon_{i m l} g^{k r} g^{l r} g^{i j} g^{m n} u_{j} v_{n} \\
& =\epsilon_{j n r} \operatorname{det} g g^{k r} u_{j} v_{n} .
\end{aligned}
$$

understanding the sum over the same indices. In the last line we use that $\epsilon_{i j k} g^{i m} g^{j n} g^{k r}=$ $\epsilon_{m n r} \epsilon_{i j k} g^{i 1} g^{j 2} g^{k 3}=\epsilon_{m n r} \operatorname{det} g$. As $\operatorname{det} g=1$, it follows that

$$
\begin{aligned}
((g \cdot \mathbf{u}) \times(g \cdot \mathbf{v}))_{k} & =g^{k r} \epsilon_{j n r} u^{j} v^{n} \\
& =g^{k r}(\mathbf{u} \times \mathbf{v})_{r} \\
& =(g \cdot \mathbf{u} \times \mathbf{v})_{k}
\end{aligned}
$$

Proposition 3.2.3. $i$ changes the adjoint action of $S O(3)$ on $\mathfrak{s o ( 3 )}$ with the usual action of $S O(3)$ in $\mathbb{R}^{3}$.

Proof: Let $g \in S O(3)$ and define the application $\Psi_{g}: \mathbb{R}^{3} \longrightarrow \mathbb{R}^{3}$ by $x \longrightarrow i\left(A d_{g} i^{-1}(x)\right)=$ $i\left(A d_{g} \hat{x}\right)$. For all $y \in \mathbb{R}^{3}$ it can be verified that

$$
\begin{aligned}
\left(\Psi_{g}(x), y\right) & =\left(g \hat{x} g^{-1}\right) y=g \hat{x}\left(g^{-1} y\right) \\
& =g\left(x \times g^{-1} y\right) \\
& =(g x) \times y=(g x, y)
\end{aligned}
$$

applying (3.2.2).
Also, we can identify $\mathfrak{s o}(3)^{*}$ with $\mathbb{R}^{3}$ by the standard scalar product in $\mathbb{R}^{3}$, therefore $\alpha \in \mathfrak{s o}(3)^{*}$ is identified with an element $\tilde{\alpha} \in \mathbb{R}^{3}$ such as

$$
\alpha(\hat{a})=:(\tilde{\alpha}, a), \text { for all } a \in \mathbb{R}^{3}
$$

Consider the Hamiltonian $H: S O(3) \times \mathfrak{s o}(3)^{*} \longrightarrow \mathbb{R}$ defined by

$$
H(g, \alpha)=\frac{1}{2}\left(\frac{\left(\alpha\left(E_{1}\right)\right)^{2}}{I_{1}}+\frac{\left(\alpha\left(E_{2}\right)\right)^{2}}{I_{2}}+\frac{\left(\alpha\left(E_{3}\right)\right)^{2}}{I_{3}}\right)
$$

which express as a Hamiltonian $H: S O(3) \times \mathbb{R}^{3} \longrightarrow \mathbb{R}$ as: if $\tilde{\alpha}=\left(\Pi_{1}, \Pi_{2}, \Pi_{3}\right) \in \mathbb{R}^{3}$ then

$$
H(g, \tilde{\alpha})=\frac{1}{2}\left(\frac{\Pi_{1}^{2}}{I_{1}}+\frac{\Pi_{2}^{2}}{I_{2}}+\frac{\Pi_{3}^{2}}{I_{3}}\right)
$$

the classical Hamiltonian for the rigid body motion.

As $\frac{\partial H}{\partial \alpha} \in \mathfrak{s o}(3)$ then, for all $\xi \in \mathfrak{s o}(3)$ it follows that

$$
\begin{aligned}
\left\langle a d\left(\frac{\partial H}{\partial \alpha}\right)^{*} \alpha, \xi\right\rangle & =\left\langle\alpha, a d\left(\frac{\partial H}{\partial \alpha}\right) \xi\right\rangle \\
& =\left\langle\alpha,\left[\frac{\partial H}{\partial \alpha}, \xi\right]\right\rangle \\
& =\alpha\left(i\left(\frac{\partial \bar{H}}{\partial \alpha}\right) \times i(\xi)\right) \\
& =\left(\tilde{\alpha}, i\left(\frac{\partial H}{\partial \alpha}\right) \times i(\xi)\right) \\
& =\left(\tilde{\alpha} \times i\left(\frac{\partial H}{\partial \alpha}\right), i(\xi)\right)
\end{aligned}
$$

By (3.2.3), since $i\left(\frac{\partial H}{\partial \alpha}\right)=\left(\frac{\Pi_{1}}{I_{1}}, \frac{\Pi_{2}}{I_{2}}, \frac{\Pi_{3}}{I_{3}}\right)$ the Euler-Arnold equations (in this case, LiePoisson) are:

$$
\begin{aligned}
\dot{g} & =g\left(\begin{array}{ccc}
0 & -\frac{\Pi_{3}}{I_{3}} & \frac{\Pi_{2}}{I_{2}} \\
\frac{\Pi_{3}}{I_{3}} & 0 & \frac{\Pi_{1}}{I_{1}} \\
-\frac{\Pi_{2}}{I_{2}} & \frac{\Pi_{1}}{I_{1}} & 0
\end{array}\right) \\
\left(\dot{\Pi}_{1}, \dot{\Pi}_{2}, \dot{\Pi}_{3}\right) & =\left(\Pi_{1}, \Pi_{2}, \Pi_{3}\right) \times\left(\frac{\Pi_{1}}{I_{1}}, \frac{\Pi_{2}}{I_{2}}, \frac{\Pi_{3}}{I_{3}}\right)
\end{aligned}
$$

The last equation can be written as:

$$
\begin{aligned}
\dot{\Pi}_{1} & =\frac{I_{2}-I_{3}}{I_{2} I_{3}} \Pi_{2} \Pi_{3} \\
\dot{\Pi}_{2} & =\frac{I_{3}-I_{1}}{I_{1} I_{3}} \Pi_{1} \Pi_{3} \\
\dot{\Pi}_{3} & =\frac{I_{1}-I_{2}}{I_{1} I_{2}} \Pi_{1} \Pi_{2}
\end{aligned}
$$

### 3.3 Euler-Lagrange Equations on Lie Groups

The configuration of a satellite can be described by the location of its mass center and the orientation of this in a three-dimensional space. The location can be expressed in Euclidean space, but the attitude evolves in a nonlinear space that has a certain geometry. The attitude of a rigid body (for example, a satellite) is defined as the direction of a bodyfixed frame with respect to a reference frame, considered as a linear transformation on the vector space $\mathbb{R}^{3}$; the attitude of a rigid body can be represented mathematically by a $3 \times 3$ orthonormal matrix (in general will be an element of the Lie group configuration). We require that its determinant is positive in order to preserve the ordering of the orthonormal axes according to the right-hand rule. These are some of our motivations to study in what follows the variational derivation of the Euler-Lagrange equations for Lagrangians defined in the left trivialized space $G \times \mathfrak{g}$. These equations generalize the the equations of motion for the attitude dynamics of a rigid body when the Lie group is $S O(3)$ and they are of interest in aeronautics and spacial engineering.

Consider a mechanical system evolving on a Lie group $G$. We derive the corresponding Euler-Lagrange equations from a variational principle. We trivialize (by a left trivialization) the tangent space $T G$ as $G \times \mathfrak{g}$. A tangent vector $(g, \dot{g}) \in T_{g} G$ is expressed as $\dot{g}=T_{e} L_{g} \cdot \xi=g \xi$.

In the sequel, we assume that the Lagrangian of the mechanical system is given by $L(g, \xi): G \times \mathfrak{g} \rightarrow \mathbb{R}$ and we define the action map as

$$
\mathcal{A}=\int_{t_{0}}^{t_{f}} L(g, \xi) d t, \quad t_{0}, t_{f} \in[0, T] \subset \mathbb{R}
$$

Hamilton's principle states that the variation of the action integral is equal to zero,

$$
\delta \mathcal{A}=\delta \int_{t_{0}}^{t_{f}} L(g, \xi) d t=0
$$

Let $g(t)$ be a differential curve in $G$ defined for $t \in\left[t_{0}, t_{f}\right]$. The variation is a differentiable mapping $g_{\epsilon}(t):(-c, c) \times\left[t_{0}, t_{f}\right] \rightarrow 0$ for $c>0$ such that $g_{0}(t)=g(t), \forall t \in\left[t_{0}, t_{f}\right]$ and $g_{\epsilon}\left(t_{0}\right)=g\left(t_{0}\right), g_{\epsilon}\left(t_{f}\right)=g\left(t_{f}\right) \forall \epsilon \in(-c, c)$. We express the variation using the exponential map (see [2], [39] and [55] for other approaches), $g_{\epsilon}(t)=g \exp \epsilon \eta(t)$, for any arbitrary curve $\eta(t) \in \mathfrak{g}$. These variations are well defined for some constant $c$ because the exponential map is a local diffeomorphism between $\mathfrak{g}$ and $G$, and it satisfies the properties of the fixed points $\eta\left(t_{0}\right)=\eta\left(t_{f}\right)=0$. Since this is obtained by a group operation, it is also guaranteed that the variation lies on $G$ for any $\eta(t)$.

The infinitesimal variation of $g$ is given by,

$$
\begin{equation*}
\delta g(t)=\left.\frac{d}{d t}\right|_{\epsilon=0} g_{\epsilon}(t)=\left.T_{e} L_{g(t)} \frac{d}{d t}\right|_{\epsilon=0} \exp \epsilon \eta(t)=g(t) \eta(t) . \tag{3.3.1}
\end{equation*}
$$

For each $t \in\left[t_{0}, f_{f}\right]$, the infinitesimal variation $\delta g(t)$ lies in the tangent space $T_{g(t)} G$. Using this expression and $\dot{g}=g \xi$, the infinitesimal variation of $\xi(t)$ is obtained as follows (see [10], [50] for example).

$$
\begin{equation*}
\delta \xi(t)=\dot{\eta}+a d_{\xi(t)} \eta(t) . \tag{3.3.2}
\end{equation*}
$$

The equations (3.3.1) and (3.3.2) are infinitesimal variations of $(g(t), \xi(t)):\left[t_{0}, t_{f}\right] \rightarrow$ $G \times \mathfrak{g}$, respectively.

The variation of the Lagrangian is written as

$$
\delta L(g, \xi)=\frac{\partial L}{\partial g} \delta g+\frac{\partial L}{\partial \xi} \delta \xi,
$$

where $\frac{\partial L}{\partial g} \in T^{*} G$ denotes the derivative of $L$ with respect to $g$, given by

$$
\left.\frac{d}{d \epsilon}\right|_{\epsilon=0} L\left(g_{\epsilon}, \xi\right)=\frac{\partial L}{\partial g} \delta g
$$

and $\frac{\partial L}{\partial \xi}(g, \xi) \in \mathfrak{g}^{*}$ is defined similarly.

Therefore,

$$
\begin{aligned}
& \delta L(g, \xi)=\left\langle\frac{\partial L}{\partial g}(g, \xi), \delta g\right\rangle+\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \delta \xi\right\rangle= \\
& \left\langle\frac{\partial L}{\partial g}(g, \xi),\left(T_{e} L_{g} \circ T_{g} L_{g^{-1}}\right) \delta g\right\rangle+\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \delta \xi\right\rangle
\end{aligned}
$$

because $T\left(L_{g} \circ L_{g^{-1}}\right)=T L_{g} \circ T L_{g^{-1}}$ is equal to the identity on $T G$. Substituting (3.3.1) and 3 3.3.2 we have that

$$
\begin{align*}
\delta L(g, \xi) & =\left\langle\frac{\partial L}{\partial g}(g, \xi), T_{e} L_{g} \cdot \eta\right\rangle+\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \dot{\eta}+a d_{\xi} \eta\right\rangle  \tag{3.3.3}\\
& =\left\langle T_{e}^{*} L_{g} \cdot \frac{\partial L}{\partial g}(g, \xi)+a d_{\xi}^{*} \cdot \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle+\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \dot{\eta}\right\rangle
\end{align*}
$$

Therefore, the variation of the action integral is given by

$$
\delta \mathcal{A}=\int_{t_{0}}^{t_{f}} \delta L(g, \xi) d t
$$

Substituting (3.3.3) and using integration by parts, the variation of the action integral is given by

$$
\begin{aligned}
& \delta \mathcal{A}=\int_{t_{0}}^{t_{f}}\left(\left\langle T_{e}^{*} L_{g} \cdot \frac{\partial L}{\partial g}(g, \xi)+a d_{\xi}^{*} \cdot \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle+\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \dot{\eta}\right\rangle\right) d t \\
& =\left.\left\langle\frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle\right|_{t_{0}} ^{t_{f}}-\int_{t_{0}}^{t_{f}}\left\langle T_{e}^{*} L_{g} \cdot \frac{\partial L}{\partial g}(g, \xi)+a d_{\xi}^{*} \cdot \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle d t \\
& -\int_{t_{0}}^{t_{f}}\left\langle\frac{d}{d t} \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle d t .
\end{aligned}
$$

Since $\eta(t)=0$ at $t=t_{0}$ and $t=t_{f}$, the first term of the above equation vanishes, thus we obtain

$$
\begin{equation*}
\delta \mathcal{A}=\int_{t_{0}}^{t_{f}}\left(\left\langle T_{e}^{*} L_{g} \cdot \frac{\partial L}{\partial g}(g, \xi)+a d_{\xi}^{*} \cdot \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle-\left\langle\frac{d}{d t} \frac{\partial L}{\partial \xi}(g, \xi), \eta\right\rangle\right) d t \tag{3.3.4}
\end{equation*}
$$

From Hamilton's principle $\delta \mathcal{A}=0 \forall \eta \in \mathfrak{g}$. Then, the corresponding Euler-Lagrange equations for $L: G \times \mathfrak{g} \rightarrow \mathbb{R}$ are given by

$$
\begin{align*}
0 & =\frac{d}{d t} \frac{\partial L}{\partial \xi}(g, \xi)-a d_{\xi}^{*} \frac{\partial L}{\partial \xi}(g, \xi)-\left(T_{e}^{*} L_{g}\right) \cdot \frac{\partial L}{\partial g}(g, \xi),  \tag{3.3.5}\\
\dot{g} & =g \xi . \tag{3.3.6}
\end{align*}
$$

If the Lagrangian is left-invariant the resulting equation is equivalent to the EulerPoincaré eqs. and (3.3.6) is the reconstruction equation (see [55]). Therefore, both (3.3.5) and (3.3.6) can be considered as a generalization of the Euler-Poincaré equations.

Remark 3.3.1. If we consider the identification of the tangent bundle $T G$ with $G \times \mathfrak{g}$ by a right-trivialization, the corresponding Euler-Lagrange equations for $L: G \times \mathfrak{g} \rightarrow \mathbb{R}$ are given by

$$
\begin{align*}
0 & =\frac{d}{d t} \frac{\partial L}{\partial \xi}(g, \xi)+a d_{\xi}^{*} \frac{\partial L}{\partial \xi}(g, \xi)-\left(T_{e}^{*} R_{g}\right) \cdot \frac{\partial L}{\partial g}(g, \xi),  \tag{3.3.7}\\
\dot{g} & =\xi g \tag{3.3.8}
\end{align*}
$$

### 3.3.1 Legendre Transformation

As before, we identify the tangent space $T G$ with $G \times \mathfrak{g}$ using a left-trivialization. In the same way, we can identify the cotangent bundle $T^{*} G$ with $G \times \mathfrak{g}^{*}$. For a given Lagrangian, the Legendre transformation $\mathbb{F} L: G \times \mathfrak{g} \rightarrow G \times \mathfrak{g}^{*}$ is defined as

$$
\mathbb{F} L(g, \xi)=(g, \mu),
$$

where $\mu \in \mathfrak{g}^{*}$ is given by $\mu=\frac{\partial L}{\partial \xi}(g, \xi)$.
If the Lagrangian is hyperregular, the induced Hamiltonian system on $G \times \mathfrak{g}^{*}$, via the Legendre transformation yields Hamilton's equation that are equivalent to Euler-Lagrange equations

$$
\begin{aligned}
0 & =\frac{d}{d t} \mu-a d_{\xi}^{*} \mu-\left(T_{e}^{*} L_{g}\right) \cdot \frac{\partial L}{\partial g}(g, \xi), \\
\dot{g} & =g \xi
\end{aligned}
$$

### 3.3.2 Simplecticity and Momentum Preservation

In this subsection we show two properties, the simplecticity and momentum preservation of the Lagrangian flow.

Simplecticity: Let $\Theta_{L}$ be the Lagrangian one-form on $G \times \mathfrak{g}$ defined by

$$
\Theta_{L}(g, \xi) \cdot(\delta g, \delta \xi)=\left\langle\frac{\partial L}{\partial \xi}(g, \xi), g^{-1} \delta g\right\rangle .
$$

The canonical symplectic 2-form is given by $\Omega_{L}=-d \Theta_{L}$ and the flow map $\mathcal{F}_{L}:(G \times$ $\mathfrak{g}) \times\left[0, t_{f}-t_{0}\right] \rightarrow G \times \mathfrak{g}$ as the flow of the Euler-Lagrange equations for $L: G \times \mathfrak{g} \rightarrow \mathbb{R}$

Proposition 3.3.2 ([50], [39], [55]). The lagrangian flow preserves the Lagrangian symplectic 2-form,

$$
\left(\mathcal{F}_{L}^{T}\right)^{*} \Omega_{L}=\Omega_{L}
$$

for $T=t_{f}-t_{0}$.
Proof: Define the solution space $C_{L}$ to be the set of solutions $g(t):\left[t_{0} ; t_{f}\right] \rightarrow G$ of the Euler-Lagrange equations for $L$ over $G \times \mathfrak{g}$. Since an element of $C_{L}$ is uniquely determined
by the initial condition $(g(0) ; \xi(0)) \in G \times \mathfrak{g}$, we can identify $C_{L}$ with the space of initial conditions $G \times \mathfrak{g}$. Define the restricted action map $\widehat{\mathcal{A}}: G \times \mathfrak{g} \rightarrow \mathbb{R}$ by $\widehat{\mathcal{A}}\left(g_{0}, \xi_{0}\right)=\mathcal{A}(g(t))$, where $g(t) \in C_{L}$ with $\left(g(0) ; g^{-1}(0) \dot{g}_{( }(0)\right)=\left(g_{0} ; \xi_{0}\right)$. Since the curve $g(t)$ satisfies the EulerLagrange equations we have that

$$
d \widehat{\mathcal{A}} \cdot w=\left\langle\left(\left(\mathcal{F}_{L}^{T}\right)^{*} \Theta_{L}-\Theta_{L}\right) \cdot ; w\right\rangle=0
$$

where $w=\left(\delta g_{0}, \delta \xi_{0}\right)$. We take the exterior derivative in previous equality since exterior derivatives and pull back commute, we obtain

$$
d^{2} \widehat{\mathcal{A}} \cdot w=\left\langle\left(\left(\mathcal{F}_{L}^{T}\right)^{*} d \Theta_{L}-d \Theta_{L}\right) ; w\right\rangle
$$

Using that $d^{2} \widehat{\mathcal{A}}=0$ we proof that the Lagrnagian flow preserves the canonical symplectic 2-form.

Noether's Theorem Suppose that a Lie group $H$ with Lie algebra $\mathfrak{h}$ acts on $G$. We consider the left action $\Phi: H \times G \rightarrow G$ such that $\Phi(e, g)=g$ and $\Phi\left(h, \Phi\left(h^{\prime}, g\right)\right)=\Phi\left(h h^{\prime}, g\right)$ for any $g \in G$ and $h, h^{\prime} \in H$. The left trivialization is given by $\phi_{L}: T G \rightarrow G \times \mathfrak{g}$ as $\phi_{L}(g, \dot{g})=\left(g, g^{-1} \dot{g}\right)$. The infinitesimal generators $\zeta_{G}: G \rightarrow G \times \mathfrak{g}$ and $\zeta_{G \times \mathfrak{g}}: G \times \mathfrak{g} \rightarrow$ $T(G \times \mathfrak{g}) \simeq G \times \mathfrak{g}$ for the action where $\zeta \in \mathfrak{h}$, are given by

$$
\begin{gathered}
\zeta_{G}(g)=\left.\phi_{L} \circ \frac{d}{d t}\right|_{\epsilon=0} \Phi_{\exp _{H} \epsilon \zeta}(g) \\
\zeta_{G \times \mathfrak{g}}(g, \xi)=\left.\frac{d}{d t}\right|_{\epsilon=0} \phi_{L} \circ T_{g} \Phi_{\exp _{H} \epsilon \zeta}(g) \cdot\left(\phi_{L}^{-1}(g, \xi)\right)
\end{gathered}
$$

We define the momentum map $J_{L}: G \times \mathfrak{g} \rightarrow \mathfrak{h}^{*}$ as

$$
J(g, \xi) \cdot \zeta=\Theta_{L} \cdot \zeta_{G \times \mathfrak{g}}(g, \xi)
$$

Proposition 3.3.3 ([39, [50], [55]). Suppose that the Lagrangian is infinitesimal invariant under the lifted action for any $\zeta \in \mathfrak{h}$. Then, the Lagrangian flow preserves the momentum map

$$
J_{L}\left(\mathcal{F}_{L}^{T}(g, \xi)\right)=J_{L}(g, \xi)
$$

Proof: Since $d L(g ; \xi) \cdot \zeta_{G \times \mathfrak{g}}=0$ implies that $d \mathcal{A} \cdot \zeta_{G \times \mathfrak{g}}=0$, where we consider that the group action $\Phi_{h}$ is applied to each point of a curve. The invariance of the action integral implies that the action maps a solution curve to another solution curve. Thus, we can restrict $d \mathcal{A} \cdot \zeta_{G \times \mathfrak{g}}=0$ to the solution space to obtain $d \widehat{\mathcal{A}} \cdot \zeta_{G \times \mathfrak{g}}=0$, and using that

$$
d \widehat{\mathcal{A}} \cdot \zeta_{G \times \mathfrak{g}}=\left(\left(\mathcal{F}_{L}^{T}\right)^{*} \Theta_{L}-\Theta_{L}\right) \cdot \zeta_{G \times \mathfrak{g}}=0
$$

and substituting the definition of the momentum map into this, we have the preservation of the flow under the momentum map.

### 3.3.3 Attitude Dynamic of a Rigid Body on $S O(3)$

In this subsection we develop following the ideas given in [[50, [39, [55], [49] the continuous Euler-Lagrange equations for the attitude dynamics of the rigid body on the special orthogonal group $S O(3)$ according to the Hamilton's variational principle.

Taking the control of a satellite as motivation, consider a rigid body that can freely rotate around pivot point fixed in an element frame. The pivot point may not be located at the mass center of the rigid body, and it is assumed that there exists a potential field that depends on the attitude. We consider a body fixed frame whose origin is located at the pivot point.

The configuration manifold for the attitude dynamics of a rigid body is the special orthogonal group $S O(3)$. A rotation matrix $R \in S O(3)$ is a linear transformation from a representation of a vector in the body fixed frame into a representation of the vector in the inertial frame. The attitude kinematics equations are given by

$$
\begin{equation*}
\dot{R}=R \hat{\Omega}, \tag{3.3.9}
\end{equation*}
$$

where the angular velocity represented in the body fixed frame is denoted by $\Omega \in \mathbb{R}^{3}$, and the hat map $\hat{:}: \mathbb{R}^{3} \rightarrow \mathfrak{s o}(3)$ is an isomorphism between $\mathbb{R}^{3}$ and the set of skewsymmetric matrices, the Lie algebra $\mathfrak{s o}(3)$, for $\Omega=\left[\Omega_{1}, \Omega_{2}, \Omega_{3}\right] \in \mathbb{R}^{3}$. The Lie bracket on $\mathfrak{s o}(3)$ corresponds to cross product on $\mathbb{R}^{3}$, that is, $\left[\hat{\Omega}, \hat{\Omega}^{\prime}\right]=\Omega \times \Omega^{\prime}$ for $\Omega, \Omega^{\prime} \in \mathbb{R}^{3}$.

Using these kinematic equations, the tangent bundle $T S O(3)$ can be identified with $S O(3) \times \mathfrak{s o}(3)$ after a left trivialization. Then we can define the Lagrangian function $L$ over $S O(3) \times \mathfrak{s o}(3)$.

The Lagrangian $L: S O(3) \times \mathfrak{s o}(3) \rightarrow \mathbb{R}$ is the difference between the kinetic energy $T: S O(3) \times \mathfrak{s o}(3) \rightarrow \mathbb{R}$ and the attitude dependent potential $U: S O(3) \rightarrow \mathbb{R}$.

$$
L(R, \Omega)=T(R, \Omega)-U(R)
$$

Let $\rho \in \mathbb{R}^{3}$ be the vector form the pivot to a mass element represented in the body fixed frame. The mass element has a velocity $\Omega \times \rho$. Thus, the kinematic energy is given by

$$
\begin{equation*}
\frac{1}{2} \int_{\mathcal{B}}\|\hat{\Omega} \rho\|^{2} d m(\rho) \tag{3.3.10}
\end{equation*}
$$

where the region of the body is denoted by $\mathcal{B}$. Since $\hat{\Omega} \rho=-\hat{\rho} \Omega$, the equation 3.3.10 can be written as

$$
\begin{aligned}
T(\Omega) & =\frac{1}{2} \int_{\mathcal{B}}\|\hat{\rho} \Omega\|^{2} d m(\rho)=\int_{\mathcal{B}}(\hat{\rho} \Omega)^{T}(\hat{\rho} \Omega) d m(\rho) \\
& =\frac{1}{2} \int_{\mathcal{B}} \Omega^{T} \hat{\rho}^{T} \hat{\rho} \Omega d m(\rho)=\frac{1}{2} \Omega^{T} J \Omega,
\end{aligned}
$$

where the moment of inertia matrix $J \in \mathbb{R}^{3}$ is defined as $J=\int_{\mathcal{B}} \hat{\rho}^{T} \hat{\rho} d m$.
Alternatively, using the property $\|x\|^{2}=x^{T} x=\operatorname{tr}\left(x x^{T}\right)$ for any $x \in \mathbb{R}^{3}$, equation (3.3.10) can be written as

$$
\begin{align*}
T(\Omega) & =\frac{1}{2} \int_{\mathcal{B}} \operatorname{tr}\left(\hat{\Omega} \rho \rho^{T} \hat{\Omega}^{T}\right) d m(\rho)  \tag{3.3.11}\\
& =\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d} \hat{\Omega}^{T}\right), \tag{3.3.12}
\end{align*}
$$

where a nonstandard moment of inertia matrix is defined as $J_{d}=\int_{\mathcal{B}} \rho \rho^{T} d m$.
In summary, the kinetic energy can be written in the standard form (3.3.11) or in a non-standard form (3.3.12). In (3.3.11), the kinetic energy is expressed as a function of the angular moment of inertia matrix, and in (3.3.12), it is expressed as a function of the Lie algebra with the non-standard momenta of inertia matrix. In this thesis we use the non-standard form. The Lagrangian function of the attitude dynamics of the rigid body is given by

$$
\begin{equation*}
L(R, \Omega)=\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d} \hat{\Omega}^{T}\right)-U(R) \tag{3.3.13}
\end{equation*}
$$

Before proceeding to the next step, we are going to study the relationship between the moment of inertia matrix $J$ and the non-standard moment of inertial matrix $J_{d}$. If we express $\rho$ in coordinates as $\rho=[x, y, z]$, the inertia momenta are given by

$$
\begin{aligned}
& J=\int_{\mathcal{B}}\left(\begin{array}{ccc}
y^{2}+z^{2} & -x y & -z x \\
-x y & z^{2}+x^{2} & -y z \\
-z x x & -y z & x^{2}+y^{2}
\end{array}\right) d m, \\
& J_{d}=\int_{\mathcal{B}}\left(\begin{array}{ccc}
x^{2} & x y & z x \\
x y & y^{2} & y z \\
z x & y z & z^{2}
\end{array}\right) d m .
\end{aligned}
$$

Using the property $\hat{\rho}^{T} \hat{\rho}=\left(\rho^{T} \rho\right) I_{3 \times 3}-\rho \rho^{T}$, it can be shown that

$$
\begin{equation*}
J_{d}=\frac{1}{2} \operatorname{tr}(J) I_{3 \times 3}-J . \tag{3.3.14}
\end{equation*}
$$

Furthermore, the following equation is satisfied for any $\Omega \in \mathbb{R}^{3}$.

$$
\begin{equation*}
\widehat{J \Omega}=\hat{\Omega} J_{d}+J_{d} \hat{\Omega} \tag{3.3.15}
\end{equation*}
$$

Using the expression of the Lagrangian function, the action integral is defined as,

$$
\begin{align*}
\mathcal{A} & =\int_{t_{0}}^{t_{f}} L(R, \Omega) d t  \tag{3.3.16}\\
& =\int_{t_{0}}^{t_{f}}\left(\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d} \hat{\Omega}^{T}\right)-U(R)\right) d t . \tag{3.3.17}
\end{align*}
$$

Hamilton's principle states that this action integral does not vary to the first order for all possible variations of a curve in $S O(3)$.

$$
\begin{equation*}
\delta \mathcal{A}=\delta \int_{t_{0}}^{t_{f}}\left(\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d} \hat{\Omega}^{T}\right)-U(R)\right) d t=0 \tag{3.3.18}
\end{equation*}
$$

Let $R(t)$ be a differentiable curve in $S O(3)$ defined for $t \in\left[t_{0}, t_{f}\right]$. The variation is a differentiable mapping $R_{\epsilon}(t):(-c, c) \times\left[t_{0}, t_{f}\right] \rightarrow S O(3)$ for $c>0$ such that $R_{0}(t)=$ $R(t), R_{\epsilon}\left(t_{0}\right)=R\left(t_{0}\right), R_{\epsilon}\left(t_{f}\right)=R\left(t_{f}\right)$ for any $\epsilon \in(-c, c)$. The infinitesimal variation is given by

$$
\delta R(t)=\left.\frac{d}{d t}\right|_{\epsilon=0} R_{\epsilon}(t) \in T_{R(t)} S O(3)
$$

The variation determines a family of neighboring curves for $R(t)$ that have the same end points parameterized by a single variable $\epsilon$. The infinitesimal variation of the rotation matrix using the exponential map as

$$
R_{\epsilon}(t)=R(t) \exp \epsilon \hat{\eta}(t),
$$

where $\eta(t)$ is defined as a differentiable curve in $\mathbb{R}^{3}$ so that $\hat{\eta}$ is a differentiable curve in $\mathfrak{s o}(3)$. This is well defined since the exponential map ia a local diffeomorphism between $\mathfrak{s o}(3)$ and $S O(3)$. Thus for any $\eta(t)$, there exists a constant $c>0$ such that this variation is defined for any $\epsilon \in(-c, c)$. The corresponding infinitesimal variation is given by

$$
\begin{align*}
\delta R(t) & =\left.\frac{d}{d t}\right|_{\epsilon=0} R_{\epsilon}(t)=\left.R(t) \sum_{i=0}^{\infty} \frac{d}{d t} \frac{1}{i} \frac{\epsilon^{i}}{} \hat{\eta}^{i}\right|_{\epsilon=0}  \tag{3.3.19}\\
& =R(t) \hat{\eta}(t) \in T_{R(t)} S O(3) .
\end{align*}
$$

Since differentiation and the variation commute, we obtain

$$
\delta \dot{R}(t)=\frac{d}{d t}(\delta R(t))=\dot{R}(t) \hat{\eta}(t)+R(t) \hat{\dot{\eta}}(t)
$$

The infinitesimal variation of the angular velocity can be obtained from the kinematic equation (3.3.9) as

$$
\begin{align*}
\delta \hat{\Omega}(t) & =\delta\left(R^{T}(t) \dot{R}(t)\right)=\delta R^{T}(t) \dot{R}(t)+R^{T}(t) \delta \dot{R}(t)  \tag{3.3.20}\\
& =-\hat{\eta}(t) \hat{\Omega}(t)+\hat{\Omega}(t) \hat{\eta}(t)+\dot{\dot{\eta}}(t) .
\end{align*}
$$

Since $\hat{x} \hat{y}-\hat{y} \hat{x}=\widehat{x \times y}$ for any $x, y \in \mathbb{R}^{3}$, this can be written as

$$
\begin{equation*}
\delta \Omega(t)=\dot{\eta}(t)+\Omega(t) \times \eta(t) . \tag{3.3.21}
\end{equation*}
$$

Euler-Lagrange equations: Now, we find the infinitesimal variation of the action integral using (3.3.19) and (3.3.20) as follows,

$$
\begin{aligned}
& \delta \mathcal{A}=\int_{t_{0}}^{t_{f}} \frac{1}{2} \operatorname{tr}\left(\delta \hat{\Omega} J_{d} \hat{\Omega}^{T}\right)+\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d} \delta \hat{\Omega}^{T}\right)-\delta U(R) d t \\
& =\int_{t_{0}}^{t_{f}}\left(-\frac{1}{2} \operatorname{tr}\left((\hat{\dot{\eta}}+\hat{\Omega} \hat{\eta}-\hat{\eta} \hat{\Omega}) J_{d} \hat{\Omega}\right)+\frac{1}{2} \operatorname{tr}\left(\hat{\Omega} J_{d}(-\hat{\dot{\eta}}+\hat{\eta} \hat{\Omega}-\hat{\Omega} \hat{\eta})\right)-\delta U(R)\right) d t \\
& =\int_{t_{0}}^{t_{f}}\left(-\frac{1}{2} \operatorname{tr}\left(\hat{\dot{\eta}}\left(J_{d} \hat{\Omega}+\Omega \hat{J}_{d}\right)\right)+\frac{1}{2} \operatorname{tr}\left(\hat{\eta} \hat{\Omega}\left(J_{d} \hat{\Omega}+\hat{\Omega} J_{d}\right)-\hat{\eta}\left(J_{d} \hat{\Omega}+\hat{\Omega} J_{d}\right) \hat{\Omega}\right)\right) d t \\
& -\int_{t_{0}}^{t_{f}} \delta U(R) d t,
\end{aligned}
$$

where we use the property $\operatorname{tr}(A B)=\operatorname{tr}(B A)$ for any matrices $A, B \in \mathbb{R}^{n \times n}$ repeatedly. Substituting (3.3.15), we obtain

$$
\begin{align*}
\delta \mathcal{A} & =\int_{t_{0}}^{t_{f}}\left(-\frac{1}{2} \operatorname{tr}(\hat{\dot{\eta}} \widehat{J \Omega})+\frac{1}{2} \operatorname{tr}(\hat{\eta}(\hat{\Omega} \widehat{J \Omega}-\widehat{J \Omega} \hat{\Omega}))-\delta U(R)\right) d t  \tag{3.3.22}\\
& =\int_{t_{0}}^{t_{f}}\left(-\frac{1}{2} \operatorname{tr}(\hat{\dot{\eta}} \widehat{J \Omega})+\frac{1}{2} \operatorname{tr}(\hat{\eta}(\Omega \times J \Omega))-\delta U(R)\right) d t .
\end{align*}
$$

The infinitesimal variation of the potential energy is given by

$$
\begin{align*}
\delta U(R) & =\left.\frac{d}{d \epsilon} U\left(R_{\epsilon}\right)\right|_{\epsilon=0}=\left.\sum_{i, j=1}^{3} \frac{\partial U}{\partial[R]_{i j}} \frac{\partial[R \exp \epsilon \hat{\eta}]_{i j}}{\partial \epsilon}\right|_{\epsilon=0}  \tag{3.3.23}\\
& =\sum_{i, j=1}^{3} \frac{\partial U}{\partial[R]_{i j}}[R \hat{\eta}]_{i j}=-\operatorname{tr}\left(\hat{\eta} R^{T} \frac{\partial U}{\partial R}\right)
\end{align*}
$$

where $[A]_{i j}$ denotes the $(i, j)$-th element of a matrix $A$, and $\frac{\partial U}{\partial R} \in \mathbb{R}^{3 \times 3}$ is defined such that $\left(\frac{\partial U}{\partial R}\right)_{i j}=\frac{\partial U(R)}{\partial[R]_{i j}}$. Substituting (3.3.22) and (3.3.23) and using integration by parts, we obtain

$$
\begin{equation*}
\delta \mathcal{A}=\int_{t_{0}}^{t_{f}} \frac{1}{2} \operatorname{tr}\left[\hat{\eta}\left((J \dot{\Omega}+\Omega \times J \Omega) \hat{}+2 R^{T} \frac{\partial U}{\partial R}\right)\right] d t . \tag{3.3.24}
\end{equation*}
$$

From Hamilton's principle, the above equation should be zero for all variations $\hat{\eta} \in \mathfrak{s o}(3)$. Given that $\hat{\eta}$ is skew-symmetric, the expression in the braces should be symmetric. Thus, we obtain the Euler-Lagrange equation

$$
(J \dot{\Omega}+\Omega \times J \Omega)-\frac{\partial U^{T}}{\partial R} R-R^{T} \frac{\partial U}{\partial R},
$$

or equivalently,

$$
J \dot{\Omega}+\Omega \times J \Omega=M
$$

where $M \in \mathbb{R}^{3}$ is determined by $S(M)=\frac{\partial U^{T}}{\partial R} R-R^{T} \frac{\partial U}{\partial R}$. More explicitly, it can be shown that the moment due to the attitude-dependent potential is given by

$$
\begin{equation*}
M=r_{1} \times u_{1}+r_{2} \times u_{2}+r_{3} \times u_{3}, \tag{3.3.25}
\end{equation*}
$$

where $r_{i}, u_{i} \in \mathbb{R}^{1 \times 3}$ are the $i$-th row vectors of $R$ and $\frac{\partial U}{\partial R}$, respectively.

$$
\begin{aligned}
\hat{M} & =\frac{\partial U^{T}}{\partial R} R-R^{T} \frac{\partial U}{\partial R} \\
& =\left(\begin{array}{lll}
u_{1}^{T} & u_{2}^{T} & u_{3}^{T}
\end{array}\right)\left(\begin{array}{c}
r_{1} \\
r_{2} \\
r_{3}
\end{array}\right)-\left(\begin{array}{lll}
r_{1}^{T} & r_{2}^{T} & r_{3}^{T}
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3}
\end{array}\right) \\
& =\left(u_{1}^{T} r_{1}-r_{1}^{T} u_{1}\right)+\left(u_{2}^{T} r_{2}-r_{2}^{T} u_{2}\right)+\left(u_{3}^{T} r_{3}-r_{3}^{T} u_{3}\right) .
\end{aligned}
$$

Since $\left(u^{T} r-r^{T} u\right)^{\widehat{x}}=\widehat{r \times u}$, we have

$$
\hat{M}=\left(r_{1} \times u_{1}+r_{2} \times u_{2}+r_{3} \times u_{3}\right) \hat{,}
$$

which is equivalent to (3.3.25).
The Legendre transformation $\mathbb{F} L:(S O(3) \times \mathfrak{s o}(3)) \rightarrow\left(S O(3) \times \mathfrak{s o}^{*}(3)\right)$ is defined as

$$
\mathbb{F} L(R, \hat{\Omega}) \cdot \hat{\eta}=\frac{1}{2} \operatorname{tr}\left[\widehat{J \Omega}^{T} \hat{\eta}\right]=\widehat{J \Omega} \cdot \hat{\eta}
$$

This gives the expression for the angular momenta expressed in the body fixed frame $\hat{\Pi}=\mathbb{F} L(R, \hat{\Omega})=\widehat{J \Omega}$ and from $\Pi=\frac{\partial L}{\partial \Omega}=J \Omega \in \mathbb{R}^{3}$ we obtain the Euler-Arnold's equations

$$
\dot{\Pi}+J^{-1} \Pi \times \Pi=M .
$$

Example 3.3.4 ( 50$]$ ). A 3D pendulum is a rigid body supported by a frictionless pivot acting under uniform gravitational potential. Let $\rho_{c} \in \mathbb{R}^{3}$ be the vector from the pivot to the mass center represented in the body fixed frame, and let $e_{3}=[0,0,1] \in \mathbb{R}^{3}$ be the gravity direction in the inertial frame. The gravitational potential energy is given by

$$
U(R)=-m g e_{3}^{T} R \rho_{c} .
$$

The derivative of the potential is

$$
\frac{\partial U}{\partial R}=-m g e_{3} \rho_{c}^{T},
$$

therefore the potential is $M=m g \rho_{c} \times R^{T} e_{3}$ becouse $u_{1}=u_{2}=0, u_{3}=-m g e_{3} \rho_{c}^{T}$.

### 3.4 Euler-Poincaré Equations

Let $G$ act on $T G$ by left-translation. A function $F: T G \rightarrow \mathbb{R}$ is called left invariant if and only if

$$
F(h(g, \dot{g}))=F(g, \dot{g}) \text { for all }(g, \dot{g}) \in T G
$$

where

$$
h(g, \dot{g}):=\left(g h, T_{e} L_{g}(\dot{h})\right) .
$$

When the lagrangian is left-invariant, we have the following identities

$$
L(g, \dot{g})=L\left(g^{-1} g, g^{-1} \dot{g}\right)=L\left(e, g^{-1} \dot{g}\right)=L(e, \xi) \text { for all }(g, \dot{g}) \in T G,
$$

where $\xi:=g^{-1} \dot{g}$. Note that in this case the Lagrangian satisfies

$$
L(g, \dot{g})=L(e, \xi),
$$

so it is independent of $g$. Then the Euler-Lagrange equations on $T G$ are rewritten as

$$
\frac{d}{d t}\left(\frac{\delta l}{\delta \xi}\right)=a d_{\xi}^{*} \frac{\delta l}{\delta \xi},
$$

where $l$ is defined to be the restriction of $L$ to $\mathfrak{g}$ :

$$
l: \mathfrak{g} \rightarrow \mathbb{R}, \quad l(\xi):=L(e, \xi) \text { for all } \xi \in \mathfrak{g}
$$

Now, we give the Euler-Poincaré reduction theorem [10, [40, [55].

## Theorem 3.4.1. (Euler-Poincaré Reduction Theorem)

Let $G$ be a Lie group and $L: T G \longrightarrow \mathbb{R}$ a left invariant Lagrangian. We define the reduced Lagrangian $l: \mathfrak{g} \longrightarrow \mathbb{R}$ as the restriction of $L$ to $\mathfrak{g}$. Then for a curve $g(t) \in G$, let $\xi(t)=T_{g(t)} L_{g(t)^{-1}} \dot{g}(t)$, the following statements are equivalent:
(i) $g(t)$ satisfies the Euler-Lagrange equations for $L$ defined on $T G$;
(ii) $g(t)$ extremize the functional

$$
g(\cdot) \longmapsto \int_{a}^{b} L(g(t), \dot{g}(t)) d t
$$

for variations among paths with fixed endpoints.
(iii) The (left-invariant) Euler-Poincaré equations hold:

$$
\frac{d}{d t} \frac{\delta l}{\delta \xi}=a d_{\xi}^{*} \frac{\delta l}{\delta \xi}
$$

(iv) $\xi(t)$ extremize

$$
\xi(\cdot) \longmapsto \int_{a}^{b} l(\xi(t)) d t
$$

for variations of the form

$$
\delta \xi=\dot{\eta}+[\xi, \eta]
$$

where $\eta(t)$ is an arbitrary path in $\mathfrak{g}$ that vanishes at the endpoints, i.e. $\eta(a)=\eta(b)=$ 0 .

Proof:
We assume the equivalence between (i) and (ii). The typical proof is given in an arbitrary configuration space $Q$ and therefore valid when $Q$ is a Lie group.

Now, we shall prove the equivalence between (ii) and (iv). For this, we need to compute the infinitesimal variations $\delta \xi$ where $\xi=g^{-1} \dot{g}$. Therefore, if we denote by $\delta g=d g_{\epsilon} / d \epsilon$ in $\epsilon=0$ for variations $g_{\epsilon}$ of $g$ and denote by $\eta=g^{-1} \delta g$ then the variations verify

$$
\delta \xi-\dot{\eta}=[\xi, \eta]
$$

and therefore, is easy to deduce the equivalence.
Finally, the equivalence between (iii) and (iv):

$$
\begin{aligned}
\delta \int l(\xi) d t & =\int \frac{\delta l}{\delta \xi} \delta \xi d t \\
& =\int \frac{\delta l}{\delta \xi}\left(\dot{\eta}+a d_{\xi} \eta\right) d t \\
& =\int\left[-\frac{d}{d t}\left(\frac{\delta l}{\delta \xi}\right)+a d_{\xi}^{*} \frac{\delta l}{\delta \xi}\right] \eta d t
\end{aligned}
$$

using integration by parts and the end points condition.
If we choose a basis $\left\{E_{1}, \ldots E_{n}\right\}$ of the Lie algebra $\mathfrak{g}$ such that

$$
\left[E_{i}, E_{j}\right]=C_{i j}^{k} E_{k}
$$

then any element is written as $\xi=\xi^{i} E_{i}$. In this coordinates, the Euler-Poincaré equations are written as

$$
\frac{d}{d t}\left(\frac{\delta l}{\delta \xi^{i}}\right)=C_{j i}^{k} \xi^{j} \frac{\delta l}{\delta \xi^{k}}
$$

Observe that if we take

$$
\alpha=\frac{\delta l}{\delta \xi} \in \mathfrak{g}^{*}
$$

we can define $\xi=\xi(\alpha)$ using this last expression. Then, we define

$$
h(\alpha)=\langle\alpha, \xi\rangle-l(\xi) .
$$

and therefore,

$$
\frac{\delta h}{\delta \alpha}=\xi+\left\langle\alpha, \frac{\delta \xi}{\delta \alpha}\right\rangle-\left\langle\frac{\delta l}{\delta \xi}, \frac{\delta \xi}{\delta \alpha}\right\rangle=\xi
$$

Then, we can write the equations of motion in the following form

$$
\dot{\alpha}=a d_{\delta h / \delta \alpha}^{*} \alpha
$$

that is, as the Lie-Poisson equations.
Remark 3.4.2. A similar statement holds, with obvious changes for right-invariant Lagrangian systems on TG. In this case the Euler-Poincaré equations are given by

$$
\frac{d}{d t} \frac{\delta l}{\delta \xi}=-a d_{\xi}^{*} \frac{\delta l}{\delta \xi}
$$

Reconstruction [39, [55]: The reconstruction of the solution $g(t)$ of the Euler-Lagrange equations, with initial conditions $g(0)=g_{0}$ and $\left.\dot{g}_{( } 0\right)=v_{0}$, is as follows: first, solve the initial value problem for the left invariant Euler-Poincaré equations:

$$
\frac{d}{d t} \frac{\delta l}{\delta \xi}=a d_{\xi}^{*} \frac{\delta l}{\delta \xi} \text { with } \xi(0)=\xi_{0}:=g_{0}^{-1} v_{0}
$$

Second, using the solution $\xi(t)$ of the above, find the curve $g(t) \in G$ by solving the reconstruction equation

$$
\dot{g}(t)=g(t) \xi(t) \text { with } g(0)=g_{0},
$$

which is a differential equation with time-dependent coefficients.

### 3.5 Euler-Poincaré Equations for the Motion of a Rigid Body

In this section we study in detail an example of mechanical system when defined on the Lie algebra $\mathfrak{s o}(3)$ of the Lie group $S O(3)$. This system describes the motion of a rigid body over a fixed point without external forces acting on it; understanding the rigid body as a system of particles where the distance between them is invariant. We start studying the kinematic, deducing the equations of motion applying a variational principle.

Consider a rigid body moving without external forces (that is, free) in the Euclidean space $\mathbb{R}^{3}$.

A configuration reference $\mathcal{B}$ consists on fixed the clausure of an open subset of $\mathbb{R}^{3}$ with smooth boundary.

Fix an orthonormal reference $\left(\mathbf{E}_{1}, \mathbf{E}_{2}, \mathbf{E}_{3}\right)$, then a point $X \in \mathcal{B}$ (material point) has coordinates $X=\left(X_{1}, X_{2}, X_{3}\right)$ known as material coordinates.

A configuration on $\mathcal{B}$ is an application $\varphi: \mathcal{B} \rightarrow \mathbb{R}^{3}$ which is $C^{1}$, preserving the orientation and with invertible image.

Fix an orthonormal reference $\left(\mathbf{e}_{1}, \mathbf{e}_{2}, \mathbf{e}_{3}\right)$ in $\mathbb{R}^{3}$, then a point $x=\varphi(X)$ (spacial point) has coordinates $x=\left(x_{1}, x_{2}, x_{3}\right)$ called spacial coordinates.

A motion of $\mathcal{B}$ is a family of time dependent configurations. We use indistinctly the following notations

$$
x=\varphi(t, X)=\varphi_{t}(X)=x(t, X)=x_{t}(X)
$$

identifying $x(0, X) \equiv X$.
A rigid body is a collection of particles in such a away the distance between all pair of particles is fixed, independently of the moving of the body or forces actuating in the body. We suppose that the rigid body has a fixed center of mass in the origin. All isometry of $\mathbb{R}^{3}$ that leaves fixed the origin is an element of the orthogonal group (this theorem has been proved by Mazur and Ulam in 1923). Then

$$
x_{t}(X)=g(t) X
$$

In coordinates, $x_{i}(t)=g^{i j}(t) X_{j}, i, j=1,2,3$; where $\left(g^{i j}\right)$ is the matrix of $g$ relative to the basis $\left(\mathbf{E}_{1}, \mathbf{E}_{2}, \mathbf{E}_{3}\right)$ and $x^{i}$ are the components of $x$ relative to the basis $\left(\mathbf{e}_{1}, \mathbf{e}_{2}, \mathbf{e}_{3}\right)$.

We suppose that the dynamics is continua and $g(0)$ is the identity, then $\operatorname{det} g(t)=1$ and, therefore, $g(t) \in S O(3)$.

Then we conclude that:
The configuration space of a rigid body is $S O(3)$, its phase space of velocities is TSO(3) and the phase space of momenta is $T^{*} S O(3)$.

There are a third type of coordinates which is of our interest, the coordinates of the body or convective coordinates. They are coordinates associated to a reference which is moving with the body. We consider the time-dependent basis

$$
\xi_{i}=g(t) \mathbf{E}_{i}, \quad i=1,2,3
$$

The coordinates in the body of an element of $\mathbb{R}^{3}$ are coordinates with respect to the basis $\left\{\xi_{1}, \xi_{2}, \xi_{3}\right\}$.

### 3.5.1 The Lagrangian of the Rigid Body

The trajectory of a material point $X \in \mathcal{B}$ is $x(t)=g(t) X$ with $g(t) \in S O(3)$.
Definition 3.5.1. We define the Lagrangian velocity or Material velocity as:

$$
V(X, t)=\frac{\partial x}{\partial t}(X, t)=\dot{g}(t) X
$$

The Eulerian velocity or spacial velocity is given by

$$
v(x, t)=V(X, t)=\dot{g}(t) g(t)^{-1} x
$$

and the velocity of the body, $\mathcal{V}(x, t)$ is defined by taking the velocity regarding $X$ as time-dependent and $x$ fixed; that is, we write $X(t, x)=g^{-1}(t) x$ and define

$$
\begin{aligned}
\mathcal{V}(X, t) & =-\frac{\partial X}{\partial t}(x, t)=g(t)^{-1} \dot{g}(t) g(t)^{-1} x \\
& =g(t)^{-1} \dot{g}(t) X(t) \\
& =g(t)^{-1} V(X, t) \\
& =g(t)^{-1} v(x, t)
\end{aligned}
$$

We suppose that the distribution of the mass in the body is described by a density $\rho_{0} d^{3} X$ in the reference configuration, which is zero at points outside the body.

The Lagrangian is taken to be the kinetic energy given by the following expression

$$
\begin{aligned}
L & =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|V(X, t)\|^{2} d^{3} X \\
& =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|\dot{g}(t) X\|^{2} d^{3} X
\end{aligned}
$$

Since $g(t) \in S O(3)$, then differentiating with respect to $t, g(t)^{T} g(t)=e$ and $g(t) g(t)^{T}=$ $e$, it follows that both $g(t)^{-1} \dot{g}(t)$ and $\dot{g}(t) g(t)^{-1}$ are skew-symmetric and therefore are in $\mathfrak{s o}(3)$. Define, the spacial angular velocity $\omega(t)$ :

$$
\widehat{\omega}(t)=\dot{g}(t) g(t)^{-1}
$$

and the connective angular velocity $\Omega(t)$ :

$$
\widehat{\Omega}(t)=g(t)^{-1} \dot{g}(t)
$$

Observe that, $v(x, t)=\widehat{\omega(t)} x=\omega(t) \times x$ y $\mathcal{V}(X, t)=\widehat{\Omega(t)} X=\Omega(t) \times X$.
Since

$$
\widehat{\omega}=g \widehat{\Omega} g^{-1}=A d_{g} \widehat{\Omega}
$$

then $\omega=g \Omega$.
We will see that the Lagrangian $L: T S O(3) \longrightarrow \mathbb{R}$ is left-invariant under $S O(3)$. Indeed, if $h \in S O(3)$, left translation by $h$ is $L_{h} g=h g$ and $T L_{h}(g, \dot{g})=(h g, h \dot{g})$, so

$$
\begin{aligned}
L(g h, h \dot{g}) & \left.=\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X) \| h \dot{g}(t) X\right) \|^{2} d^{3} X \\
& \left.=\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X) \| \dot{g}(t) X\right) \|^{2} d^{3} X \\
& =L(g, \dot{g})
\end{aligned}
$$

since $h$ is orthogonal. From this $L(g, \dot{g})=L\left(g^{-1} g, g^{-1} \dot{g}\right)=L(I, \hat{\Omega})=l(\Omega)$
We study different expressions of the Lagrangian which describe the dynamic of a rigid
body (free)

$$
\begin{aligned}
L & =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|V(X, t)\|^{2} d^{3} X \\
& =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|\dot{g} X\|^{2} d^{3} X \\
& =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\left\|g g^{-1} \dot{g} X\right\|^{2} d^{3} X \\
& =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|\hat{\Omega} X\|^{2} d^{3} X \\
& =\frac{1}{2} \int_{\mathcal{B}} \rho_{0}(X)\|\Omega \times X\|^{2} d^{3} X
\end{aligned}
$$

Introducing the inner product on $\mathbb{R}^{3}$ :

$$
\ll \mathbf{a}, \mathbf{b} \gg=\int_{\mathcal{B}} \rho_{0}(X)(\mathbf{a} \times X, \mathbf{b} \times X) d^{3} X
$$

then,

$$
l(\Omega)=\frac{1}{2} \ll \Omega, \Omega \gg
$$

and introducing the linear isomorphism $\mathbb{I}: \mathbb{R}^{3} \longrightarrow \mathbb{R}^{3}$ given by

$$
(\mathbb{I} \mathbf{a}, \mathbf{b})=\ll \mathbf{a}, \mathbf{b} \gg, \quad \text { for all } \mathbf{a}, \mathbf{b} \in \mathbb{R}^{3}
$$

By definition, $(\mathbb{I} \mathbf{a}, \mathbf{b})=(\mathbf{a}, \mathbb{I} \mathbf{b})$. Then $\mathbb{I}$ is symmetric and positive defined (under some initial hypotheses). Since $\mathbb{I}$ is symmetric, can be diagonalized. If we choose an orthonormal basis $\left(\mathbf{E}_{1}, \mathbf{E}_{2}, \mathbf{E}_{3}\right)$ in which II is diagonal then (see [55])

$$
l(\Omega)=\frac{1}{2} \ll \Omega, \Omega \gg=\frac{1}{2}\langle\mathbb{I} \Omega, \Omega\rangle=\frac{1}{2}\left(I_{1} \Omega_{1}^{2}+I_{2} \Omega_{2}^{2}+I_{3} \Omega_{3}^{2}\right) .
$$

If we denote by $\Pi=\mathbb{I} \Omega$ the we can define the function:

$$
H(\Pi)=\frac{1}{2}\left(\frac{\Pi_{1}^{2}}{I_{1}}+\frac{\Pi_{2}^{2}}{I_{2}}+\frac{\Pi_{3}^{2}}{I_{3}}\right)
$$

which represents the expression for the kinetic energy in $\mathfrak{s o}(3)^{*}$. Note that $\Pi=\mathbb{I} \Omega$ is the angular momentum in the body frame.

### 3.5.2 Euler-Poincaré Reduction Theorem for the Motion of a Rigid Body

Roughly speaking in this subsection we prove that a curve $g(t) \in S O(3)$ verify the EulerLagrange equations for $L(g, \dot{g})$ if and only if $\Omega(t)=g(t)^{-1} \dot{g}(t)$ verify Euler's equations, namely:

$$
\mathbb{I} \dot{\Omega}=\mathbb{I} \Omega \times \Omega
$$

From variational calculus we know that $L$ satisfies the Euler-Lagrange equations if and only if

$$
\delta \int L d t=0
$$

Consider $l: \mathfrak{s o}(3) \longrightarrow \mathbb{R}$ defined by $l(\Omega)=\frac{1}{2}(I \Omega, \Omega)$. Taking a variation $g_{\epsilon}(t)$ of a curve $g(t)$ we have a variation of the curve $\widehat{\Omega_{\epsilon}(t)}$ in the Lie algebra $\mathfrak{s o}(3)$. If denote by

$$
\begin{aligned}
\delta g(t) & =\frac{d}{d \epsilon} g_{\epsilon}(t) \\
\delta \Omega(t) & =\frac{d}{d \epsilon} \Omega_{\epsilon}(t)
\end{aligned}
$$

differentiating with respect to $\epsilon$ the relation $g^{-1} \dot{g}=\widehat{\Omega}$ we have

$$
-g^{-1}(\delta g) g^{-1} \dot{g}+g^{-1}(\delta \dot{g})=\widehat{\delta \Omega}
$$

Define the skew-symmetric matrix $\widehat{\Sigma}$ :

$$
\widehat{\Sigma}=g^{-1} \delta g
$$

and the vector $\Sigma=i(\widehat{\Sigma})$ in $\mathbb{R}^{3}$.
Observe that

$$
\dot{\widehat{\Sigma}}=-g^{-1} \dot{g} g^{-1} \delta g+g^{-1}(\delta \dot{g})
$$

Therefore,

$$
g^{-1} \delta \dot{g}=\dot{\widehat{\Sigma}}+g^{-1} \dot{g} \widehat{\Sigma}
$$

and then,

$$
-\widehat{\Sigma} \widehat{\Omega}+\dot{\hat{\Sigma}}+\widehat{\Omega} \widehat{\Sigma}=\widehat{\delta \Omega}
$$

that is,

$$
\widehat{\delta \Omega}=\dot{\hat{\Sigma}}+[\widehat{\Omega}, \widehat{\Sigma}]
$$

or,

$$
\delta \Omega=\dot{\Sigma}+\Omega \times \Sigma
$$

Then, the variational equations $\delta \int_{t_{0}}^{t_{1}} L d t=0$ in $T S O(3)$ are equivalent to the reduced variational principle

$$
\delta \int_{t_{0}}^{t_{1}} l d t=0
$$

in $\mathbb{R}^{3}$ where the variations $\delta \Omega$ have the form $\delta \Omega=\dot{\Sigma}+\Omega \times \Sigma$ with $\Sigma\left(t_{0}\right)=0$ and $\Sigma\left(t_{1}\right)=0$. In our case, since $l(\Omega)=\frac{1}{2}(\mathbb{I} \Omega, \Omega)$ then we have

$$
\begin{aligned}
\delta \int_{t_{0}}^{t_{1}} l d t & =\int_{t_{0}}^{t_{1}}(\mathbb{I} \Omega, \delta \Omega) d t \\
& =\int_{t_{0}}^{t_{1}}(\mathbb{I} \Omega, \dot{\Sigma}+\Omega \times \Sigma) d t \\
& =\int_{t_{0}}^{t_{1}}[-(\mathbb{I} \dot{\Omega}, \Sigma)+(\mathbb{I} \Omega, \Omega \times \Sigma)] d t \\
& =\int_{t_{0}}^{t_{1}}(-(\mathbb{I} \dot{\Omega}+I \Omega \times \Omega), \Sigma) d t
\end{aligned}
$$

where we integrate by parts and use the conditions $\Sigma\left(t_{0}\right)=0$ and $\Sigma\left(t_{1}\right)=0$. Since $\Sigma$ is an arbitrary element we have

$$
-\mathbb{I} \dot{\Omega}+\mathbb{I} \Omega \times \Omega=0
$$

which are the Euler equations for the rigid body:

$$
\begin{aligned}
I_{1} \dot{\Omega}_{1} & =\left(I_{2}-I_{3}\right) \Omega_{2} \Omega_{3} \\
I_{2} \dot{\Omega}_{2} & =\left(I_{1}-I_{3}\right) \Omega_{1} \Omega_{3} \\
I_{3} \dot{\Omega}_{3} & =\left(I_{1}-I_{2}\right) \Omega_{1} \Omega_{2}
\end{aligned}
$$

This result can be rewritten as
Theorem 3.5.2. (Euler's Rigid Body Equations [39]) Euler's rigid body equations are equivalent to Hamilton's principle

$$
\delta \mathcal{A}(\Omega)=\int_{t_{0}}^{t_{1}} l(\Omega) d t=0
$$

in which the Lagrangian $l(\Omega)$ appearing in the action integral $\mathcal{A}(\Omega)=\int_{t_{0}}^{t_{1}} l(\Omega) d t$ is given by the kinetic energy in principal axis coordinates,

$$
l(\Omega)=\frac{1}{2}(\mathbb{I} \Omega, \Omega):=\frac{1}{2} \mathbb{I} \Omega \cdot \Omega=\frac{1}{2}\left(I_{1} \Omega_{1}^{2}+I_{2} \Omega_{2}^{2}+I_{3} \Omega_{3}^{2}\right),
$$

and variations of $\Omega$ are restricted to be of the form

$$
\delta \Omega=\dot{\Sigma}+\Omega \times \Sigma
$$

where $\Sigma(t)$ is a curve in $\mathbb{R}^{3}$ that vanishes at the endpoints.
Remark 3.5.3. Reconstruction: The Euler solution is expressed in terms of the timedependent angular velocity vector in the body. The body angular velocity vector $\Omega(t)$ yields the tangent vector $\dot{g}(t) \in T_{g(t)} S O(3)$ along the integral curve in the rotation group $g(t) \in$ $S O(3)$ by the relation

$$
\begin{equation*}
\dot{g}(t)=g(t) \hat{\Omega}(t) \tag{3.5.1}
\end{equation*}
$$

where the left-invariant skew-symmetric $3 \times 3$ matrix $\hat{\Omega}$ is defined by the hat map. Equation (3.5.1) is the reconstruction equation for $g(t) \in S O(3)$. Once the time dependence of $\Omega(t)$, and hence $\hat{\Omega}(t)$, is determined from the Euler equations, solving (3.5.1) as a linear differential equation with time-dependent coefficients yields the integral curve $g(t) \in S O(3)$ for the orientation of the rigid body.

## Chapter 4

## Higher-Order Mechanical Systems

In the last decade many papers and books dealing with higher-order derivatives in Mechanics has appeared. An extensive analysis of these systems can be found in, for example, ([15], [51], [19], [20], [36], [47], [24], [25], [23], [58]. This kind of systems appear for example in electromagnetic theory, elasticity theory, the moving of a particle rotating around a point which is translated, the relativistic particle, optimization problems, control theory, etc.

The aim of this chapter is to build up the Lagrangian and Hamiltonian formalism for systems involving higher-order derivatives. The Lagrangian formalism is said to be of higher-order derivatives if it is described by a real (smooth) function $L$ which depends of higher-order derivatives. For sake of simplicity, we will say that $L$ is a Lagrangian of order $k$, where $k$ denotes the order of the derivative.

We give an introduction of the mechanics in higher-order tangent bundles and a corresponding variational principle. From this principle we derive the higher-order Euler Lagrange equations. Moreover, we also study systems subject to constraints. These constraints will be higher-order constraints. Using Lagrange multipliers theorem we can deduce the variational principle for systems with higher-order constraints. Finally, we study the Skinner and Rusk formalism for higher-order systems. This framework will be used to analyze constrained systems throughout this thesis.

Some examples are analyzed in this chapter as application to higher-order theories: An interpolation problem on Riemannian manifolds and the use of Hamilton's principle to construct numerical algorithms.

### 4.1 Higher-Order Tangent Bundles

In this section we recall some basic facts on the higher-order tangent bundles theory. For more details, see Refs. 51] and [24].

Let $Q$ be a manifold of dimension $n$. An equivalence relation is introduced in the set $C^{\infty}(\mathbb{R}, Q)$ of differentiable curves from $\mathbb{R}$ to $Q$. By definition, two given curves in $Q \gamma_{1}(t)$ and $\gamma_{2}(t)$ where $t \in(-a, a)$ with $a \in \mathbb{R}$ have contact of order $k$ at $q_{0}=\gamma_{1}(0)=\gamma_{2}(0)$ if there exists a local chart $(\varphi, U)$ of $Q$ such that $q_{0} \in U$ and

$$
\left.\frac{d^{s}}{d t^{s}}\left(\varphi \circ \gamma_{1}(t)\right)\right|_{t=0}=\left.\frac{d^{s}}{d t^{s}}\left(\varphi \circ \gamma_{2}(t)\right)\right|_{t=0}
$$

for $s=0, \ldots, k$. This is a well defined equivalence relation in $C^{\infty}(\mathbb{R}, Q)$ and the equivalence class of a curve $\gamma$ will be denoted by $[\gamma]_{0}^{(k)}$. The set of equivalence classes will be denoted by $T^{(k)} Q$ and it can be shown it is a differentiable manifold. Moreover, $\tau_{Q}^{k}: T^{(k)} Q \rightarrow Q$ where $\tau_{Q}^{k}\left([\gamma]_{0}^{(k)}\right)=\gamma(0)$ is a fiber bundle called the tangent bundle of order $k$ of $Q$.

We also may define the surjective mappings $\tau_{Q}^{(l, k)}: T^{(k)} Q \rightarrow T^{(l)} Q$, for $l \leq k$, given by $\tau_{Q}^{(l, k)}\left([\gamma]_{0}^{(k)}\right)=[\gamma]_{0}^{(l)}$. It is easy to see that $T^{(1)} Q \equiv T Q$, the tangent bundle of $Q$, $T^{(0)} Q \equiv Q$ and $\tau_{Q}^{(0, k)}=\tau_{Q}^{k}$.

Given a differentiable function $f: Q \longrightarrow \mathbb{R}$ and $l \in\{0, \ldots, k\}$, its $l$-lift $f^{(l, k)}$ to $T^{(k)} Q$, $0 \leq l \leq k$, is the differentiable function defined as

$$
f^{(l, k)}\left([\gamma]_{0}^{(k)}\right)=\left.\frac{d^{l}}{d t^{l}}(f \circ \gamma(t))\right|_{t=0}
$$

Of course, these definitions can be applied to functions defined on open sets of $Q$.
From a local chart $\left(q^{i}\right)$ on a neighborhood $U$ of $Q$, it is possible to induce local coordinates $\left(q^{(0) i}, q^{(1) i}, \ldots, q^{(k) i}\right)$ on $T^{(k)} U=\left(\tau_{Q}^{k}\right)^{-1}(U)$, where $q^{(s) i}=\left(q^{i}\right)^{(s, k)}$ if $0 \leq s \leq k$. Sometimes, we will use the standard conventions, $q^{(0) i} \equiv q^{i}, q^{(1) i} \equiv \dot{q}^{i}$ and $q^{(2) i} \equiv \ddot{q}^{i}$.

Given a vector field $X$ on $Q$, we define its $k$-lift $X^{(k)}$ to $T^{(k)} Q$ as the unique vector field on $T^{(k)} Q$ satisfying the following identities

$$
X^{(k)}\left(f^{(l, k)}\right)=(X(f))^{(l, k)}, \quad 0 \leq l \leq k
$$

for all differentiable function $f$ on $Q$. In coordinates, the $k$-lift of a vector field $X=X^{i} \frac{\partial}{\partial q^{i}}$ is

$$
X^{(k)}=\left(X^{i}\right)^{(s, k)} \frac{\partial}{\partial q^{(s) i}}
$$

Now, we consider the canonical immersion $j_{k}: T^{(k)} Q \rightarrow T\left(T^{(k-1)} Q\right)$ defined as $j_{k}\left([\gamma]_{0}^{(k)}\right)=$ $\left[\gamma^{(k-1)}\right]_{0}^{(1)}$, where $\gamma^{(k-1)}$ is the lift of the curve $\gamma$ to $T^{(k-1)} Q$; that is, the curve $\gamma^{(k-1)}: \mathbb{R} \rightarrow$ $T^{(k-1)} Q$ is given by $\gamma^{(k-1)}(t)=\left[\gamma_{t}\right]_{0}^{(k-1)}$ where $\gamma_{t}(s)=\gamma(t+s)$. In local coordinates

$$
j_{k}\left(q^{(0) i}, q^{(1) i}, q^{(2) i}, \ldots q^{(k) i}\right)=\left(q^{(0) i}, q^{(1) i}, \ldots, q^{(k-1) i} ; q^{(1) i}, q^{(2) i}, \ldots, q^{(k) i}\right) .
$$

We use the map $j_{k}$ to construct the differential operator $d_{T}$ which maps a function $f$ on $T^{(k)} Q$ into a function $d_{T} f$ on $T^{(k+1)} Q$

$$
d_{T} f\left([\gamma]_{0}^{k+1}\right)=j_{k+1}\left([\gamma]_{0}^{k+1}\right)(f)
$$

In the case when we have a group action $\phi: G \times Q \rightarrow Q$, it can be naturally lifted to a group action $\phi^{(k)}: G \times T^{(k)} Q \rightarrow T^{(k)} Q$ given by

$$
\phi_{g}^{(k)}\left([\gamma]_{0}^{(k)}\right):=\left[\phi_{g} \circ \gamma\right]_{\phi_{g}(0)}^{(k)} .
$$

This action endows $T^{(k)} Q$ with a principal G-bundle structure. The quotient $T^{(k)} Q / G$ is a fiber bundle over the base $Q / G$. The class of elements $[\gamma]_{0}^{(k)}$ in the quotient $\left(T^{(k)} Q / G\right)$ is denoted $\left[[\gamma]_{0}^{(k)}\right]_{G}$.

### 4.1.1 The Case of Lie Groups

Therefore, when the manifold $Q$ has a Lie group structure, we will denote $Q=G$ and we can also use the left trivialization to identify the higher-order tangent bundle $T^{(k)} G$ with $G \times k \mathfrak{g}$. That is, if $g: I \subset \mathbb{R} \rightarrow G$ is a curve in $C^{(k)}(\mathbb{R}, G)$ :

$$
\begin{aligned}
\Upsilon^{(k)}: T^{(k)} G & \longrightarrow G \times k \mathfrak{g} \\
{[g]_{0}^{(k)} } & \longmapsto\left(g(0), g^{-1}(0) \dot{g}(0),\left.\frac{d}{d t}\right|_{t=0}\left(g^{-1}(t) \dot{g}(t)\right), \ldots,\left.\frac{d^{k-1}}{d t^{k-1}}\right|_{t=0}\left(g^{-1}(t) \dot{g}(t)\right)\right) .
\end{aligned}
$$

It is clear that $\Upsilon^{(k)}$ is a diffeomorphism.
We will denote $\xi(t)=g^{-1}(t) \dot{g}(t)$. Therefore

$$
\Upsilon^{(k)}\left([g]_{0}^{(k)}\right)=\left(g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right),
$$

where

$$
\xi^{(l)}(t)=\frac{d^{l}}{d t^{l}}\left(g^{-1}(t) \dot{g}(t)\right), \quad 0 \leq l \leq k-1
$$

and $g(0)=g, \xi^{(l)}(0)=\xi^{(l)}, 0 \leq l \leq k-1$. We will indistinctly use the notation $\xi^{(0)}=\xi$, $\xi^{(1)}=\dot{\xi}$, where there is not danger of confusion.

We may also define the surjective mappings $\tau_{G}^{(l, k)}: T^{(k)} G \rightarrow T^{(l)} G$, for $l \leq k$, given by $\tau_{G}^{(l, k)}\left([g]_{0}^{(k)}\right)=[g]_{0}^{(l)}$. With the previous identifications we have that

$$
\tau_{G}^{(l, k)}\left(g(0), \xi(0), \dot{\xi}(0), \ldots, \xi^{(k-1)}(0)\right)=\left(g(0), \xi(0), \dot{\xi}(0), \ldots, \xi^{(l-1)}(0)\right)
$$

It is easy to see that $T^{(1)} G \equiv G \times \mathfrak{g}, T^{(0)} G \equiv G$ and $\tau_{G}^{(0, k)}=\tau_{G}^{k}$.
Now, we consider the canonical immersion $j_{k}: T^{(k)} G \rightarrow T\left(T^{(k-1)} G\right)$ defined as $j_{k}\left([g]_{0}^{(k)}\right)=$ $\left[g^{(k-1)}\right]_{0}^{(1)}$, where $g^{(k-1)}$ is the lift of the curve $g$ to $T^{(k-1)} G$; that is, the curve $g^{(k-1)}: \mathbb{R} \rightarrow$ $T^{(k-1)} G$ is given by $g^{(k-1)}(t)=\left[g_{t}\right]_{0}^{(k-1)}$ where $g_{t}(s)=g(t+s)$. Using the identification given by $\Upsilon^{(k)}$ we have that:

$$
\begin{aligned}
& j^{(k)}: \quad G \times k \mathfrak{g} \longrightarrow G \times(2 k-1) \mathfrak{g} \\
& \left(g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right) \longmapsto\left(g, \xi, \dot{\xi}, \ldots, \xi^{(k-2)} ; \xi, \dot{\xi}, \ldots, \xi^{(k)}\right)
\end{aligned}
$$

where we identify $T\left(T^{(k-1)} G\right) \equiv T(G \times(k-1) \mathfrak{g}) \equiv G \times(2 k-1) \mathfrak{g}$, in the natural way.

### 4.2 Hamilton's Principle and Euler-Lagrange Equations

Let us consider a mechanical system whose dynamic is described by a Lagrangian $L$ : $T^{(k)} Q \rightarrow \mathbb{R}$ that depends of higher-order derivatives up to order $k$. Given two points $x, y \in T^{(k-1)} Q$ we define the infinite-dimensional manifold $\mathcal{C}^{2 k}(x, y)$ of $2 k$-differentiable curves which connect $x$ and $y$ as

$$
\mathfrak{C}^{2 k}(x, y)=\left\{c:[0, T] \longrightarrow Q \mid c \text { is } C^{2 k}, c^{(k-1)}(0)=x \text { and } c^{(k-1)}(T)=y\right\} .
$$

Fixed a curve $c$ in $\mathcal{C}^{2 k}(x, y)$, the tangent space to $\mathcal{C}^{2 k}(x, y)$ at $c$ is given by

$$
\begin{aligned}
T_{c} \mathcal{C}^{2 k}(x, y)= & \left\{X:[0, T] \longrightarrow T Q \mid X \text { is } C^{2 k-1}, X(t) \in T_{c(t)} Q,\right. \\
& \left.X^{(k-1)}(0)=0 \text { and } X^{(k-1)}(T)=0\right\} .
\end{aligned}
$$

Let us consider the action functional $\mathcal{A}$ on $C^{2 k}$-curves in $Q$ given by

$$
\begin{aligned}
\mathcal{A}: C^{2 k}(x, y) & \longrightarrow \mathbb{R} \\
c & \longmapsto \int_{0}^{T} L\left(c^{(k)}(t)\right) d t .
\end{aligned}
$$

Definition 4.2.1. Hamilton's principle. A curve $c \in \mathcal{C}^{2 k}(x, y)$ is a solution of the Lagrangian system determined by $L: T^{(k)} Q \longrightarrow \mathbb{R}$ if and only if $c$ is a critical point of $\mathcal{A}$.

In order to find the critical points of $\mathcal{A}$, we need to characterize the curves $c$ such that $d \mathcal{A}(c)(X)=0$ for all $X \in T_{c} \mathcal{C}^{2 k}(x, y)$. Taking a family of curves $c_{s} \in \mathcal{C}^{2 k}(x, y)$ with $c_{0}=c$ and $s \in(-b, b) \subset \mathbb{R}$, the stationary condition can be written as

$$
\begin{equation*}
\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)=0 \tag{4.2.1}
\end{equation*}
$$

Since,

$$
d \mathcal{A}(c) \cdot(X)=\left.\frac{d}{d s}\right|_{s=0}\left(\mathcal{A} \circ c_{s}\right)=\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right) .
$$

Therefore,

$$
d \mathcal{A}(c) \cdot(X)=\left.0 \forall X \in T_{c} \mathcal{C}^{2 k}(x, y) \Longleftrightarrow \frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)=0
$$

Now, we analyze the derivative $\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)$

$$
\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)=\left.\frac{d}{d s}\left(\int_{0}^{1} L\left(c_{s}^{(k)}\right) d t\right)\right|_{s=0}=\left.\int_{0}^{1} \frac{d}{d s} L\left(c_{s}^{(k)}\right)\right|_{s=0} d t=\left.\int_{0}^{1} \sum_{l=0}^{k} \frac{\partial L}{\partial q^{(l) i}} \frac{\partial q^{(l) i}}{d s}\right|_{s=0} d t,(*)
$$

and using that the variations are given by $\delta c^{i}=\left.\frac{d}{d s} c_{s}^{(i)}\right|_{s=0} \mathrm{y} \delta^{(l)} c^{i}=\frac{d^{(l)}}{d t^{(l)}} \delta c^{i}$ it follows that,

$$
\begin{aligned}
\left.\frac{\partial q^{(0) i}}{d s}\right|_{s=0} & =c^{(i)}(t)=\delta c^{i}, \\
\left.\frac{\partial q^{(1) i}}{d s}\right|_{s=0} & =\frac{d}{d t} c^{(i)}(t)=\delta^{1} c^{i}, \\
& \cdots \\
\left.\frac{\partial q^{(l) i}}{d s}\right|_{s=0} & =\frac{d^{l}}{d t^{l}} c^{(i)}(t)=\delta^{l} c^{i},
\end{aligned}
$$

and, $\delta^{(l)} c^{i}=\frac{d}{d t} \delta^{(l-1)} c^{i}$.

Returning to (*), we have the following equalities

$$
\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)=\left.\int_{0}^{1} \sum_{l=0}^{k} \frac{\partial L}{\partial q^{(l) i}} \frac{\partial q^{(l) i}}{d s}\right|_{s=0} d t=\int_{0}^{1} \sum_{l=0}^{k} \frac{\partial L}{\partial q^{(l) i}} \delta^{(l)} c^{i} d t
$$

That is,

$$
\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}\left(c_{s}\right)=\int_{0}^{1} \frac{\partial L}{\partial q^{(0) i}} \delta c^{i} d t+\int_{0}^{1} \sum_{l=1}^{k} \frac{\partial L}{\partial q^{(l) i}} \delta^{(l)} c^{i} d t
$$

Using integration by parts,

$$
\begin{gathered}
\int_{0}^{1} \frac{\partial L}{\partial q^{(1) i}} \delta c^{(1) i} d t=\left.\frac{\partial L}{\partial q^{(1) i}} \delta c^{(0) i}\right|_{0} ^{1}-\int_{0}^{1} \frac{d}{d t} \frac{\partial L}{\partial q^{(1) i}} \delta c^{(0) i} d t, \\
\int_{0}^{1} \frac{\partial L}{\partial q^{(2) i}} \delta c^{(2) i} d t=\left.\left(\frac{\partial L}{\partial q^{(2) i}} \delta c^{(1) i} d t-\frac{d}{d t} \frac{\partial L}{\partial q^{(2) i}} \delta c^{(2) i}\right)\right|_{0} ^{1}+(-1)^{2} \int_{0}^{1} \frac{d^{2}}{d t^{2}} \frac{\partial L}{\partial q^{(2) i}} \delta c^{i} d t
\end{gathered}
$$

Making a constructive process over $l$ we obtain that

$$
\begin{aligned}
\left.\frac{d}{d s} \mathcal{A}\left(c_{s}(t)\right)\right|_{s=0}= & \int_{0}^{1} \sum_{l=0}^{k}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right) \cdot \delta c^{i} d t \\
& +\left[\sum_{l=0}^{k-1}\left[\sum_{s=0}^{k-l-1}(-1)^{l} \frac{d^{s}}{d t^{s}}\left(\frac{\partial L}{\partial q^{(l+s+1) i}}\right)\right] \cdot \delta^{(l)} c^{i}\right]_{0}^{1}
\end{aligned}
$$

Observe that the last term of the right side is equal to zero since $\delta^{(l)} c^{i}(0)=\delta^{(l)} c^{i}(1)=0$, $0 \leq l \leq k-1$, and $1 \leq i \leq n$.

From this we have the following theorem,
Theorem 4.2.2. Let $L: T^{(k)} Q \rightarrow \mathbb{R}$ be a higher-order Lagrangian and

$$
\mathcal{A}(c)=\int_{0}^{1} L\left(c^{(k)}(t)\right) d t
$$

the action of $L$ defined over $\mathcal{C}^{2 k}$.
Then, there exists an unique operator

$$
\mathcal{E} L: T^{(2 k)} Q \longrightarrow T^{*} Q
$$

and an unique $1-$ form $\Theta_{L}$ on $T^{(k)} Q$ such that for all variations of the form $\delta c_{s} \in T_{c} \mathcal{C}^{2 k}(x, y)$ with fixed endpoints we have that

$$
\left.\frac{d}{d s} \mathcal{A}\left(c_{s}(t)\right)\right|_{s=0}=\int_{0}^{1} \mathcal{E} L\left(c^{(2 k)}(t)\right) \cdot \delta c(t) d t+\left.\left(\Theta_{L}\left(c^{(2 k-1)}(t)\right) \cdot \delta^{(2 k-1)} c(t)\right)\right|_{0} ^{1}
$$

In local coordinates $\mathcal{E} L$ and $\Theta_{L}$ have the form

$$
\begin{gathered}
\mathcal{E} L=\sum_{l=0}^{k}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right) . \\
\Theta_{L}=\sum_{l=0}^{k-1} \hat{p}_{(l) i} d q^{(l) i}
\end{gathered}
$$

where the functions $\hat{p}_{(l) i}, 0 \leq l \leq k-1$, are the generalized Jacobi-Ostrogradski momenta defined by

$$
\hat{p}_{l(i)}=\sum_{s=0}^{k-l-1}(-1)^{l} \frac{d^{s}}{d t^{s}}\left(\frac{\partial L}{\partial q^{(l+s+1) i}}\right)
$$

The equations of motion are called Higher-Order Euler-Lagrange, and are written as

$$
\sum_{l=0}^{k}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right)=0
$$

Therefore, it is possible to define the 2 -form $\Omega_{L}=-d \Theta_{L}$. In local coordinates, we have that (Darboux's Theorem)

$$
\Omega_{L}=\sum_{l=0}^{k-1} d q^{(l) i} \wedge d \hat{p}_{(l) i}
$$

Is easy to see [51] that $\Omega_{L}$ is simplectic if and only if,

$$
\operatorname{det}\left(\frac{\partial^{2} L}{\partial q^{(l-1) j} \partial q^{(l-1) i}}\right) \neq 0 .
$$

We will say that the higher-order Lagrangian is regular if $\Omega_{L}$ is symplectic.
In the following, assume that the Lagrangian $L$ is regular. Take now the restriction $\mathcal{A}_{L}$ of the action functional $\mathcal{A}$ to the subspace $C_{L}$ of solutions of Euler-Lagrange equations. This space can be identified with the space of initial conditions $T^{(2 k-1)} Q$ of the Euler-Lagrange equations. Therefore is easy to show that

$$
d \mathcal{A}_{L}=F_{t}^{*} \Theta_{L}-\Theta_{L}
$$

where $F_{t}$ is the flow of the Euler-Lagrange vector field $X_{L}$, defined on $T^{(2 k-1)} Q$ by $\varepsilon L \circ X_{L}=$ 0 . Since $d^{2}=0$ we deduce that the flow is symplectic.

Moreover, if $G$ is a Lie group of symmetries preserving the action functional and $\mathfrak{g}$ its Lie algebra then

$$
0=i_{\xi_{Q}^{(2 k-1)}} d \mathcal{A}_{L}=i_{\xi_{Q}^{(2 k-1)}}\left(F_{t}^{*} \Theta_{L}-\Theta_{L}\right)=F_{t}^{*}\left(i_{\xi_{Q}^{(2 k-1)}} \Theta_{L}\right)-i_{\xi_{Q}^{(2 k-1)}} \Theta_{L}
$$

where $\xi_{Q}$ is the infinitesimal generator associated with $\xi \in \mathfrak{g}$. Therefore, $J_{\xi}=i_{\xi_{Q}^{(2 k-1)}} \Theta_{L}$ is a first integral of the flow.

Interpolation Problem on Riemannian Manifolds This example has been introduce by Noakes, Heinzigner and Paden in [60] and after studied by Crouch and Leite [34], Hussein and Bloch [12] for the application in interferometric imaging, and Gay-Balmaz, Holm, Meier, Ratiu and Vialard [35].

We consider a Riemannian manifold $(Q, \mathcal{G})$ where $\mathcal{G}$ is the Riemannian metric and $\frac{D}{D t}$ is the covariant derivative associated with the Levi-Civita connection $\nabla$ for the metric.

Consider a Lagrangian $L: T^{(2)} Q \rightarrow \mathbb{R}$ defined as,

$$
\begin{equation*}
L(q, \dot{q}, \ddot{q}):=\frac{1}{2} \mathcal{G}\left(\frac{D}{D t} \dot{q}, \frac{D}{D t} \dot{q}\right), \tag{4.2.2}
\end{equation*}
$$

where in local coordinates on $T^{(2)} Q$ the covariant derivative of the velocity is given by

$$
\frac{D}{D t} \dot{q}=\ddot{q}^{k}+\Gamma_{i j}^{k}(q) \dot{q}^{i} \dot{q}^{j}
$$

and $\Gamma_{i j}^{k}(q)$ are the Christoffel symbols of the metric $\mathcal{G}$ at point $q$ in the given basis.
The Riemannian cubic polynomials are defined as the minimizers of the action for $L$. These Riemannian cubic polynomials has been generalized the so-called elastic 2-splines thought the Lagrangian

$$
L(q, \dot{q}, \ddot{q}):=\frac{1}{2} \mathcal{G}\left(\frac{D}{D t} \dot{q}, \frac{D}{D t} \dot{q}\right)+\frac{\tau^{2}}{2} \mathcal{G}(\dot{q}, \dot{q})
$$

where $\tau$ is a real constant (see [12], [35] and references therein).
Given $N+1$ points in $Q, q_{i} \in Q$ with $i=0, \ldots, N$ and tangent vectors $v_{j} \in T_{q_{j}} Q$, $j=1, N$, the interpolation problem consists on finding a curve which minimizes the action

$$
\begin{equation*}
\mathcal{A}[q]:=\frac{1}{2} \int_{t_{0}}^{t_{N}} \mathcal{G}_{q(t)}\left(\frac{D}{D t} \dot{q}(t), \frac{D}{D t} \dot{q}(t)\right) d t \tag{4.2.3}
\end{equation*}
$$

among continuous curves on $\left[t_{0}, t_{N}\right]$, smooth on $\left[t_{i}, t_{i+1}\right]$, for $t_{0} \leq t_{1} \leq \ldots \leq t_{N}$ subject to some interpolating constraints

$$
q\left(t_{i}\right)=q_{i}
$$

for all $i \in\{1, \ldots, N-1\}$ and boundary conditions

$$
\begin{gathered}
q\left(t_{0}\right)=q_{0}, \quad q\left(t_{N}\right)=q_{N} \\
\frac{D q}{d t}\left(t_{0}\right)=v_{0}, \quad \frac{D q}{d t}\left(t_{N}\right)=v_{N}
\end{gathered}
$$

The extension of this problem to higher-order mechanics is giving by minimizing the mean-square of the $k-1$ covariant derivative of the velocity. For this, we consider the Lagrangian $L_{k}: T^{(k)} Q \rightarrow \mathbb{R}$ given by

$$
L_{k}\left(q, \dot{q}, \ldots, q^{(k)}\right):=\frac{1}{2} \mathcal{G}\left(\frac{D^{k-1}}{D t^{k-1}} \dot{q}, \frac{D^{k-1}}{D t^{k-1}} \dot{q}\right)
$$

for $k>2$ (see [35]).

The higher-order interpolation problem consists on, given $N+1$ points on $Q$ and tangent vectors $v_{j}^{(l)} \in T_{q_{j}}^{(l)} Q$, with $j=0, N$; minimizing

$$
\mathcal{A}[q]=\frac{1}{2} \int_{t_{0}}^{t_{N}} \mathcal{G}\left(\frac{D^{k-1}}{D t^{k-1}} \dot{q}, \frac{D^{k-1}}{D t^{k-1}} \dot{q}\right) d t
$$

among the curves $q(t) \in Q$, continuous in $\left[t_{0}, t_{N}\right]$ and $k-1$ piecewise smooths on $\left[t_{i}, t_{i+1}\right]$, for $t_{0} \leq t_{1} \leq \ldots \leq t_{N}$ subject to some interpolating constraints

$$
q\left(t_{i}\right)=q_{i}
$$

for all $i \in\{1, \ldots, N-1\}$ and the $2 k$ boundary conditions

$$
\begin{aligned}
q\left(t_{0}\right) & =q_{0}, \quad q\left(t_{N}\right)=q_{N} \\
\frac{D^{(l)} q}{d t^{l}}\left(t_{0}\right) & =v_{0}^{(l)}, \quad \frac{D^{(l)} q}{d t^{l}}\left(t_{N}\right)=v_{N}^{(l)} .
\end{aligned}
$$

for all $1 \leq l \leq k-1$. For the higher-order Lagrangians $L_{k}$, the Euler-Lagrange equations read

$$
\frac{D^{2 k-1}}{D t^{2 k-1}} \dot{q}(t)+\sum_{j=2}^{k}(-1)^{j} R\left(\frac{D^{2 k-j-1}}{D t^{2 k-j-1}} \dot{q}(t), \frac{D^{j-2}}{D t^{j-2}} \dot{q}(t)\right) \dot{q}(t)=0,
$$

where $R$ denotes the curvature tensor.
The higher-order Lagrangians are functions defined on $T^{(k)} Q$ and not on curves $q(t) \in Q$. Therefore, the notation $\frac{D}{D t} \dot{q}$ means the expression in terms of $\dot{q}$ and $\ddot{q}$ seen as independent elements in $T^{(2)} Q$.

### 4.3 Higher-order Mechanical Systems with Constraints

There are two different frameworks for dealing with systems with constraints; the nonholonomic mechanics and the variational calculus with constraints or vakonomic mechanics

The nonholonomic mechanics is based on Lagrange-D'Alembert principle and searches critical trajectories of the action which are compatible with the constraints. This approach has proven to be suitable for solving many interesting problems in different areas such as engineering and control theory.

On the other hand, the variational calculus with constraints is applied in optimal control problems, economy, physics, etc. Moreover, as we will see in this work, under regularity conditions, an optimal control problem is equivalent to a higher-order variational problem with higher-order constraints. Unlike the nonholonomic approach, this approach is purely variational and consist on finding a critical path which minimize the action restricted to the curves which satisfy the constraints.

It is well know that a variational problem with constraints is equivalent to a Lagrangian system defined by an extended Lagrangian with the constraints, but this system is found to be singular. Then for these kind of system one can use the geometric approach of the Dirac constraints theory given by Gotay and Nester in [37]

For the comparison of the equations of motion of these systems, we consider a mechanical system over a $n$-dimensional configuration manifold $Q$ subject to linear constraints in the velocities which define a distribution $\mathcal{D}$ of $T Q$ and whose dynamics is given by the Lagrangian $L$. We assume that $\mathcal{D}$ has constant rank and therefore, there exits, at least locally, $n-k:=m$ independent 1-forms $\left\{\omega^{a}\right\} 1 \leq a \leq m$ such that

$$
\mathcal{D}_{q}=\operatorname{Ker}\left\{\omega^{1}(q), \ldots, \omega^{n-k}(q)\right\}
$$

All solutions of the constrained system are requiered to satisfy

$$
\left\langle\omega^{a}(q(t)), \dot{q}(t)\right\rangle=0, \quad 1=1, \ldots, m
$$

### 4.3.1 Lagrange-D'Alembert Principle

In the following, we remember the Lagrange-D'Alembert principle in which is based the nonholonomic method.

A curve $q(t) \subset Q$ is a solution of the system if

$$
\delta \int_{0}^{T} L(q(t), \dot{q}(t)) d t=0, \quad \forall q(t) \subset Q \text { such that } \delta q(t) \in \mathcal{D}_{q(t)}
$$

for all $t, 0 \leq t \leq T$ and $\delta q(0)=\delta q(T)=0$.
This principle is equivalent to the so called Lagrange - $D^{\prime}$ 'Alembert equations. One can see that a curve $q(t) \in Q$ is a solution of the nonholonomic system if $(q(t), \dot{q}(t)) \subset T Q$ satisfying the following equations of motion ([14]),

$$
\begin{aligned}
\frac{d}{d t} \frac{\partial L}{\partial \dot{q}^{i}}-\frac{\partial L}{\partial q^{i}} & =\lambda_{a} \omega_{i}^{a}, \quad \forall i=1, \ldots, n ; \\
\omega_{i}^{a} \dot{q}^{i} & =0, \quad \forall a=1, \ldots, m .
\end{aligned}
$$

Consider a local system of coordinates such that $q^{i}=\left(r^{\alpha}, s^{a}\right) \in \mathbb{R}^{n-m} \times \mathbb{R}^{m}$ and

$$
\omega^{a}(q)=d s^{a}+A_{\alpha}^{a}(r, s) d r^{\alpha}
$$

with $a=1, \ldots, m$. Then, the variations $\delta q^{i}$ can be written as $\delta q^{i}=\left(\delta r^{\alpha}, \delta s^{a}\right)$ and if $\delta q \in \mathcal{D}_{q(t)}$ it satisfies that

$$
\delta s^{a}+A_{\alpha}^{a} \delta r^{\alpha}=0
$$

Then, the Lagrange-D'Alembert equations are written as

$$
\begin{aligned}
\left(\frac{d}{d t} \frac{\partial L}{\partial \dot{r}^{\alpha}}-\frac{\partial L}{\partial r^{\alpha}}\right) & =A_{\alpha}^{a}\left(\frac{d}{d t} \frac{\partial L}{\partial \dot{s}^{\alpha}}-\frac{\partial L}{\partial s^{\alpha}}\right) \\
\dot{s}^{a} & =-A_{\alpha}^{a} \dot{r}^{\alpha}, \quad a=1, \ldots, m
\end{aligned}
$$

On the other hand, the variational mechanics with constraints establishes that a curve $(q(t)) \subset Q$ is a solution of the system if $q(t)$ is a critical point of the action $\mathcal{A}(q)=$ $\int_{0}^{T} L(q(t), \dot{q}(t)) d t$ restricted to

$$
\mathcal{C}^{2}\left(q_{0}, q_{1},[a, b], \mathcal{D}\right)=\left\{q:[a, b] \rightarrow Q \mid q(a)=q_{0}, q(b)=q_{1} \text { and } \dot{q}(t) \in \mathcal{D}_{q(t)} \forall t \in[a, b]\right\} .
$$

For sake of simplicity we use the notation

$$
\mathcal{A}_{\mathcal{D}}=\left.\mathcal{A}\right|_{\mathcal{C}^{2}\left(q_{0}, q_{1},[a, b], \mathfrak{D}\right)}
$$

Consider $L_{\mathcal{D}}=\left.L\right|_{\mathcal{D}}: \mathcal{D} \rightarrow \mathbb{R}$. In local coordinates can be written as

$$
L_{\mathcal{D}}\left(r^{\alpha}, s^{a}, \dot{r}^{\alpha}\right)=L\left(r^{\alpha}, s^{a}, \dot{r}^{\alpha},-A_{\alpha}^{a}(r, s) \dot{r}^{\alpha}\right)
$$

If we applying the standard variational principle,

$$
\delta \int L_{\mathcal{D}}\left(r^{\alpha}, s^{a}, \dot{r}^{\alpha}\right) d t=0
$$

we obtain the equations

$$
\frac{d}{d t} \frac{\partial L_{\mathcal{D}}}{\partial \dot{r}^{\alpha}}-\frac{\partial L_{\mathcal{D}}}{\partial r^{\alpha}}+A_{\alpha}^{a} \frac{\partial L_{\mathcal{D}}}{\partial s^{a}}=-\frac{\partial L}{\partial \dot{s}^{a}} d \omega^{a}\left(\dot{q}, \frac{\partial}{\partial r^{\alpha}}\right) .
$$

In this approach, the well-known Lagrange multipliers theorem is fundamental (see [2] and [1])

Theorem 4.3.1. (Lagrangian multipliers theorem)
Let $N$ be a differentiable manifold and $F$ a Banach space with $g: N \rightarrow F$ a smooth submersion such as $g^{-1}(0)$ is a submanifold of $N$. Let $f: N \rightarrow \mathbb{R}$ be a differentiable function, then $c \in g^{-1}(0)$ is a critical point of $f_{\mid g^{-1}(0)}$ if and only if there exits $\lambda \in F$ such that $c$ is a critic point of $f-\lambda \circ g$.

From Th. 4.3.1, $q(t)$ is a solution of the variational problem with constraints if and only if $\exists \lambda(t)$ such that $(q(t), \lambda(t))$ verifying the Euler-Lagrange equations corresponding to the extended Lagrangian

$$
\mathcal{L}: T\left(Q \times \mathbb{R}^{n-k}\right) \rightarrow \mathbb{R}
$$

given by

$$
\mathcal{L}(q, \dot{q}, \lambda, \dot{\lambda})=L(q, \dot{q})-\lambda_{a} \omega_{i}^{a} \dot{q}^{i} .
$$

They are

$$
\begin{aligned}
\frac{d}{d t} \frac{\partial L}{\partial \dot{q}^{i}}-\frac{\partial L}{\partial q^{i}} & =-\dot{\lambda}_{a} \omega_{i}^{a}-\lambda_{a}\left(\frac{d}{d t} \omega_{i}^{a}-\frac{\partial \omega_{i}^{a}}{\partial q^{i}}\right) \\
\omega_{i}^{a} \dot{q}^{i} & =0 .
\end{aligned}
$$

Remark 4.3.2. As it is well known, if the constraints are holonomic, the equations of motion of the system derived by the nonholonomic method and the variational calculus with constraints, are equivalent [see [31], [14]].

Remark 4.3.3. One can see that if $\mathcal{D}$ is an integrable distribution, the connection $A$ on the tangent bundle vanishes. For this reason, in the equations

$$
\frac{d}{d t} \frac{\partial L_{\mathcal{D}}}{\partial \dot{r}^{\alpha}}-\frac{\partial L_{\mathcal{D}}}{\partial r^{\alpha}}+A_{\alpha}^{a} \frac{\partial L_{\mathcal{D}}}{\partial s^{a}}=-\frac{\partial L}{\partial \dot{s}^{a}} d \omega^{a}\left(\dot{q}, \frac{\partial}{\partial r^{\alpha}}\right)
$$

the term d $\omega^{a}$ doesn't appear. Thus they can be written as

$$
\left(\frac{d}{d t} \frac{\partial L}{\partial \dot{r}^{\alpha}}-\frac{\partial L}{\partial r^{\alpha}}\right)=A_{\alpha}^{a}\left(\frac{d}{d t} \frac{\partial L}{\partial \dot{s}^{\alpha}}-\frac{\partial L}{\partial s^{\alpha}}\right) \quad \alpha=1, \ldots, n-m ;
$$

which are the equations given by the nonholonomic method.
The constraints are generically defined by the vanishing of $m$ independent differentiable functions $\Phi^{\alpha}: T Q \rightarrow \mathbb{R}, \alpha=1, \ldots, m$. In this way, the equations of motion become

$$
\frac{d}{d t} \frac{\partial L}{\partial \dot{q}^{i}}-\frac{\partial L}{\partial q^{i}}=\lambda_{\alpha} \frac{\partial L}{\partial q^{i}}, \quad \alpha=1, \ldots, m ; \quad i=1, \ldots, n ;
$$

which are called Chetaev's equations.

### 4.3.2 Higher-Order Variational Calculus with Constraints

In this section we will consider higher-order Lagrangian mechanics for systems with higherorder constraints from the point of view of the variational calculus with constraints.

Consider the $m$ independent constraints

$$
\Phi=\left(\phi^{(j)}\right), \phi: T^{(k)} Q \rightarrow \mathbb{R}^{m}, \quad j=1, \ldots, m
$$

such that 0 is a regular value of $\Phi$. These constraints $\phi^{(j)}$ define a submanifold $\mathcal{M}=\Phi^{-1}(0)$ well-known as constraint submanifold.

We assume that the restriction of the projection $\left(\tau_{Q}^{(k-1, k)}\right)_{\mid \mathcal{M}}: \mathcal{M} \rightarrow T^{(k-1)} Q$ is a submersion. Locally, this condition means that the $m \times n$-matrix

$$
\left(\frac{\partial\left(\Phi^{1}, \ldots, \Phi^{m}\right)}{\partial\left(q_{1}^{(k)}, \ldots, q_{n}^{(k)}\right)}\right)
$$

has rank $m$ at all points of $\mathcal{M}$.
Consider now the subset $\mathcal{C}^{2 k}(x, y, \mathcal{M})$ of $\mathcal{C}^{2 k}(x, y)$ of curves that satisfies these constraint equations, that is

$$
\begin{aligned}
\mathcal{C}^{2 k}(x, y, \mathcal{M})= & \left\{c:[0, T] \longrightarrow Q \mid q \text { is } C^{2 k}, c^{(k-1)}(0)=x\right. \\
& \left.c^{(k-1)}(T)=y \text { and } c^{(k)}(t) \in \mathcal{M} \text { for all } t \in[0, T]\right\}
\end{aligned}
$$

Definition 4.3.4. A curve $c \in \mathcal{C}^{2 k}(x, y, \mathcal{M})$ is a solution of the higher-order variational problem with higher-order constraints if $c$ is a stationary point of $\left.\mathcal{A}\right|_{\mathrm{e}^{2 k}(x, y, \mathcal{M})}$.

As in the case of constraints in $T Q$, by the Lagrangian multipliers theorem, a higherorder variational problem with higher-order constraints is equivalent to solving a variational problem for the extended Lagrangian.

Definition 4.3.5. (Higher-order variational problem with higher-order constraints)
Let $c \in \mathfrak{C}^{2 k}(x, y, \mathcal{M})$, be a critical curve of the variational problem with higher-order constraints for the mechanical system given by $L: T^{(k)} Q \rightarrow \mathbb{R}$ if and only if $c$ is a critical point of the functional

$$
\mathcal{A}_{\mathcal{M}}(c)=\int_{0}^{1} L\left(q^{(k)}(t)\right) d t-\lambda_{\alpha} g^{\alpha}(c)
$$

where $\lambda_{\alpha} \in \mathcal{F}([0,1], \mathbb{R})^{*}$ and $g^{\alpha}: C^{2 k}(x, y) \rightarrow \mathcal{F}([0,1], \mathbb{R})$ given by $\left\{t \longrightarrow \Phi^{\alpha}\left(c^{(k)}(t)\right)\right\}$.
The stationary condition can be written as $\left.\frac{d}{d s}\right|_{s=0} \mathcal{J}_{M}\left(q_{s}(t)\right)=0$, for all variations $c_{s}$ of $c$, for $s \in(-b, b), b \in \mathbb{R}$.

Then we compute,

$$
\begin{aligned}
\left.\frac{d}{d s}\right|_{s=0} \mathcal{A}_{M}\left(q_{s}(t)\right) & =\left.\frac{d}{d s}\left(\int_{0}^{1} L\left(c_{s}^{k}\right) d t-\lambda_{\alpha} g^{\alpha}\left(c_{s}^{k}\right)\right)\right|_{s=0}=\left.\int_{0}^{1}\left(\frac{d}{d s} L\left(c_{s}^{k}\right)-\lambda_{\alpha} g^{\alpha}\left(c_{s}^{k}\right)\right)\right|_{s=0} d t \\
& =\int_{0}^{1} \sum_{l=0}^{k} \frac{\partial L}{\partial q^{(l) i}} \frac{\partial q^{(l) i}}{d s}-\left.\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}} \frac{\partial q^{(l) i}}{d s}\right|_{s=0} d t \\
& =\left.\int_{0}^{1} \sum_{l=0}^{k}\left(\frac{\partial L}{\partial q^{(l) i}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right) \frac{\partial q^{(l) i}}{d s}\right|_{s=0} d t=\int_{0}^{1} \sum_{l=0}^{k}\left(\frac{\partial L}{\partial q^{(l) i}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right) \delta^{(l)} c^{i} d t,
\end{aligned}
$$

and integrating $l$ times by parts,

$$
\begin{aligned}
& \int_{0}^{1} \sum_{l=0}^{k}\left(\frac{\partial L}{\partial q^{(l) i}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right) \delta^{(l)} c^{i} d t=\int_{0}^{1} \sum_{l=0}^{k}\left[(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right)-\frac{d^{l}}{d t^{l}}\left(\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right)\right] \delta c^{i} d t \\
& +\left[\sum_{l=0}^{k-1}\left[\sum_{s=0}^{k-l-1}(-1)^{l} \frac{d^{s}}{d t^{s}}\left(\frac{\partial L}{\partial q^{(l+s+1) i}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l+s+1) i}}\right)\right] \cdot \delta^{(l)} c^{i}\right]_{0}^{1}
\end{aligned}
$$

for all variation $c_{s}$ of $c$; and using the condition $\delta^{(l)} c^{i}(0)=\delta^{(l)} c^{i}(1)=0$ we can characterize the critical curves of the higher-order variational problem with constraints by the solutions of

$$
\sum_{l=0}^{k}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right)=\sum_{l=0}^{k} \frac{d^{l}}{d t^{l}}\left(\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right)
$$

over the curves which satisfies the constraints.

Remark 4.3.6. The equations

$$
\begin{aligned}
\sum_{l=0}^{k}(-1)^{l} \frac{d^{l}}{d t^{l}}\left(\frac{\partial L}{\partial q^{(l) i}}\right) & =\sum_{l=0}^{k} \frac{d^{l}}{d t^{l}}\left(\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(l) i}}\right) \quad i=1, \ldots, n \\
\phi^{(j)}\left(q^{(l) i}\right) & =0, \quad j=1, \ldots, m
\end{aligned}
$$

are called higher-order Euler-Lagrange equations with higher-order constraints.

### 4.3.3 Geometric Formulation for Higher-Order Constrained Mechanics.

Now, we develop a geometric characterization of higher-order constrained variational problems using, as an essential tool, the Skinner and Rusk formulation.

Let us consider the Whitney sum $T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q$ and the canonical projections

$$
\begin{gathered}
p r_{1}: T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q \longrightarrow T^{*}\left(T^{(k-1)} Q\right) \\
p r_{2}: T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q \longrightarrow T^{(k)} Q
\end{gathered}
$$

Let us take the submanifold $W_{0}=p r_{2}^{-1}(\mathcal{M})=T^{*}\left(T^{(k-1)} Q\right) \times \mathcal{M}$ and the restrictions to $W_{0}$ of the canonical projections $p r_{1}$ and $p r_{2}$

$$
\begin{gathered}
\pi_{1}=\left.p r_{1}\right|_{W_{0}}: W_{0} \subset T^{*}\left(T^{k-1} Q\right) \oplus T^{(k)} Q \rightarrow T^{*}\left(T^{(k-1)} Q\right) \\
\pi_{2}=\left.p r_{2}\right|_{W_{0}}: W_{0} \subset T^{*}\left(T^{k-1} Q\right) \oplus T^{(k)} Q \rightarrow \mathcal{M} .
\end{gathered}
$$

Now, we consider on $W_{0}$ the presymplectic 2-form

$$
\Omega_{W_{0}}=\pi_{1}^{*}\left(\omega_{T^{(k-1)} Q}\right)
$$

where $\omega_{T^{(k-1)} Q}$ is the canonical symplectic form on $T^{*}\left(T^{(k-1)} Q\right)$. Define also the function $H_{W_{0}}: W_{0} \rightarrow \mathbb{R}$ given by

$$
H_{W_{0}}(\alpha, p)=\left\langle\alpha, j_{k}(p)\right\rangle-\left.L\right|_{\mathcal{M}}(p)
$$

where $(\alpha, p) \in W_{0}=T^{*}\left(T^{(k-1)} Q\right) \times \mathcal{M}$. Here $\langle\cdot, \cdot\rangle$ denotes the natural paring between vectors and covectors on $T^{(k-1)} Q$ (observe that $j_{k}(p) \in T T^{(k-1)} Q$ ).

We will see that the dynamics of the higher-order constrained variational problem is intrinsically characterized as the solutions of the presymplectic hamiltonian equation

$$
\begin{equation*}
i_{X} \Omega_{W_{0}}=d H_{W_{0}} \tag{4.3.1}
\end{equation*}
$$

Let us consider $\Omega=p r_{1}{ }^{*}\left(\omega_{T^{(k-1)} Q}\right)$ and $H: T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q \rightarrow \mathbb{R}$ given by

$$
H=\left\langle p r_{1}, p r_{2}\right\rangle-p r_{2}^{*} L=\left\langle p r_{1}, p r_{2}\right\rangle-L \circ \pi_{2}
$$

Observe that locally

$$
\operatorname{ker} \Omega=\operatorname{span}\left\langle\mathcal{V}_{i}=\frac{\partial}{\partial q^{(k) i}}\right\rangle
$$

Then, it is easy to show that equations (7.2.1) are equivalent to (see 31])

$$
\left\{\begin{align*}
i_{X} \Omega-d H & \in\left(T W_{0}\right)^{0}  \tag{4.3.2}\\
X & \in T W_{0}
\end{align*}\right.
$$

where $\left(T W_{0}\right)^{0}$ is the annihilator of $T W_{0}$ locally spanned by $\left\{d \Phi^{\alpha}\right\}$, where $\Phi^{\alpha}: W_{0} \rightarrow \mathbb{R}$ denote the constraints $\Phi^{\alpha}=\Phi^{\alpha} \circ p r_{2}$ (for notational simplicity, we do not distinguish the notation between constraints on $\mathcal{M}$ and constraints on $W_{0}$ ).

Take coordinates $\left(q^{(0) i}, q^{(1) i}, \ldots, q^{(k-1) i} ; p_{i}^{(0)}, \ldots, p_{i}^{(k-1)}, q^{(k) i}\right)$ on $T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q$, then the local expressions of the presymplectic 2-form $\Omega$ and the hamiltonian $H$ are

$$
\begin{aligned}
\Omega & =\sum_{r=0}^{k-1} d q^{(r) i} \wedge d p_{i}^{(r)} \\
H & =\sum_{r=0}^{k-1} q^{(r+1) i} p_{i}^{(r)}-L\left(q^{(0) i}, q^{(1) i} \ldots, q^{(k) i}\right)
\end{aligned}
$$

Consider a vector field $X$ on $T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q$ with local expression

$$
X=\sum_{r=0}^{k} X^{(r) i} \frac{\partial}{\partial q^{(r) i}}+\sum_{r=0}^{k-1} Y_{i}^{(r)} \frac{\partial}{\partial p_{i}^{(r)}},
$$

and we analyze the equations $i_{X} \Omega=\Omega(X, \cdot)=d H(\cdot)+\lambda_{\alpha} d \Phi^{\alpha}(\cdot)$ : Given $v \in T^{*}\left(T^{(k-1)} Q\right) \oplus$ $T^{(k)} Q$

$$
v=\sum_{r=0}^{k} v^{(r) i} \frac{\partial}{\partial q^{(r) i}}+\sum_{r=0}^{k-1} \widetilde{v}_{i}^{(r)} \frac{\partial}{\partial p_{i}^{(r)}}
$$

we have that

$$
\begin{aligned}
\Omega(X, v) & =\sum_{r=0}^{k-1} d q^{(r) i} \wedge d p_{i}^{(r)}(X, v) \\
& =\sum_{r=0}^{k-1}\left[d q^{(r) i}(X) d p_{i}^{(r)}(v)-d q^{(r) i}(v) d p_{i}^{(r)}(X)\right]=\sum_{r=0}^{k-1}\left[X^{(r) i} \widetilde{v}_{i}^{(r)}-v^{(r) i} Y_{i}^{(r)}\right] .
\end{aligned}
$$

Therefore, since

$$
d H(v)=\sum_{r=0}^{k} \frac{\partial H}{\partial q^{(r) i} \partial v^{(r) i}}+\sum_{r=0}^{k-1} \frac{\partial H}{\partial p_{i}^{(r)}} \widetilde{v}_{i}^{(r)}
$$

and

$$
\lambda_{\alpha} d \Phi^{\alpha}(v)=\sum_{r=0}^{k} \lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}} v^{(r) i} .
$$

From the equation 7 7.2.1) we obtain

$$
\begin{equation*}
\sum_{r=0}^{k-1}\left[X^{(r)} \widetilde{v}_{i}^{(r)}-v^{(r) i} Y_{i}^{(r)}\right]=\sum_{r=0}^{k} \frac{\partial H}{\partial q^{(r) i}} v^{(r) i}+\sum_{r=0}^{k-1} \frac{\partial H}{\partial p_{i}^{(r)}} \widetilde{v}_{i}^{(r)}+\sum_{r=0}^{k} \lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}} v^{(r) i} \tag{4.3.3}
\end{equation*}
$$

On the other hand,

$$
\begin{aligned}
\frac{\partial H}{\partial p_{i}^{(0)}} & =q^{(r+1) 0} \\
\frac{\partial H}{\partial p_{i}^{(1)}} & =q^{(r+1) 1} \\
& \vdots \\
\frac{\partial H}{\partial p_{i}^{(r)}} & =q^{(r+1) i} ; \quad r=0, \ldots, k-1 .
\end{aligned}
$$

And also we have:

$$
\begin{aligned}
\frac{\partial H}{\partial q^{(0) i}} & =-\frac{\partial L}{\partial q^{(0) i}} \\
\frac{\partial H}{\partial q^{(1) i}} & =p_{i}^{(0)}-\frac{\partial L}{\partial q^{(1) i}} \\
\frac{\partial H}{\partial q^{(2) i}} & =p_{i}^{(1)}-\frac{\partial L}{\partial q^{(2) i}} \\
\frac{\partial H}{\partial q^{(r) i}} & =p_{i}^{(r-1)}-\frac{\partial L}{\partial q^{(r) i}} ; \quad r=1, \ldots, k-1 .
\end{aligned}
$$

As (4.3.3) holds for every vector $v$ in $T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q$, we obtain that

$$
-Y_{i}^{(r)}=\frac{\partial H}{\partial q^{(r) i}}-\frac{\partial L}{\partial q^{(r) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}} \quad r=0, \ldots, k-1
$$

Then,

$$
\begin{aligned}
-Y_{i}^{(0)} & =-\frac{\partial L}{\partial q^{(0) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(0) i}} \\
& \vdots \\
-Y_{i}^{(r)} & =p_{i}^{(r-1)}-\frac{\partial L}{\partial q^{(r) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}} ; \\
X_{i}^{(r)} & =\frac{\partial H}{\partial p_{i}^{(r)}}=q^{(r+1) i} ; \quad r=0, \ldots, k-1 \\
0 & =\frac{\partial H}{\partial q^{(k) i}}-\frac{\partial L}{\partial q^{(k) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(k) i}}=p_{i}^{(k-1)}-\frac{\partial L}{\partial q^{(k) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(k) i}} .
\end{aligned}
$$

The solutions of Equation $\sqrt{7.2 .1)}$ are defined on the first constraint submanifold given by the set of points $x \in W_{0}$ such that $\left(d H+\lambda_{\alpha} d \Phi^{\alpha}\right)(x)(Z)=0$, for all $Z \in \operatorname{ker} \Omega(x)$. Locally these restrictions are defined from the following relations

$$
\varphi_{i}^{1}=p_{i}^{(k-1)}-\frac{\partial L}{\partial q^{(k) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(k) i}}=0, \quad i=1, \ldots, n .
$$

The equations $\varphi_{i}^{1}=0$ (primary relations) determine the set of points $W_{1}$ of $W_{0}$ where (7.2.1) has a solution. $W_{1}$ is the primary constraint submanifold (assuming that it is a submanifold) for the presymplectic Hamiltonian system ( $W_{0}, \Omega_{W_{0}}, H_{W_{0}}$ ). (See, for instance, [37]).

Then, we have two different types of equations which restrict the dynamics on $T^{*}\left(T^{(k-1)} Q\right) \oplus$ $T^{(k)} Q$

$$
\begin{array}{rll}
\Phi^{\alpha} & =0 \quad \alpha=1, \ldots, m & \quad(\text { constraints determining } \mathcal{M}) \\
\varphi_{i}^{1} & =0 \quad i=1, \ldots, n . \quad \text { (primary relations) } \tag{4.3.5}
\end{array}
$$

Therefore, the equations of motion for an integral curve solution of $X$ are

$$
\begin{align*}
\frac{d}{d t} q^{(r) i} & =q^{(r+1) i}, \quad r=0, \ldots, k-1,  \tag{4.3.6}\\
-\frac{d}{d t} p_{i}^{(0)} & =-\frac{\partial L}{\partial q^{(0) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(0) i}}  \tag{4.3.7}\\
& \vdots  \tag{4.3.8}\\
-\frac{d}{d t} p_{i}^{(r)} & =p_{i}^{(r-1)}-\frac{\partial L}{\partial q^{(r) i}}+\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}}, \quad r=1, \ldots, k-1  \tag{4.3.9}\\
\left(\frac{d}{d t} q^{(r) i}\right. & \left.=X^{(r) i}, Y_{i}^{r}=\frac{d}{d t} p_{i}^{(r)}\right) \tag{4.3.10}
\end{align*}
$$

and the constraints equations (4.3.4) and (4.3.5).
Differentiating with respect to time the equations $\varphi_{i}^{1}$, substituting into (4.3.9) and proceeding further, we find the equations of motion for the higher-order variational problem analyzed in the last section, i.e.

$$
\begin{equation*}
\sum_{r=0}^{k}(-1)^{r} \frac{d^{r}}{d t^{r}}\left(\frac{\partial L}{\partial q^{(r) i}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(r) i}}\right)=0 . \tag{4.3.11}
\end{equation*}
$$

The solution of equation (7.2.1) on $W_{1}$ may not be tangent to $W_{1}$. In such a case, we have to restrict $W_{1}$ to the submanifold $W_{2}$ where there exists at least a solution tangent to $W_{1}$. Proceeding further, we obtain a sequence of submanifolds [37] (assuming that all the subsets generated by the algorithm are submanifolds)

$$
\cdots \hookrightarrow W_{k} \hookrightarrow \cdots \hookrightarrow W_{2} \hookrightarrow W_{1} \hookrightarrow W_{0} .
$$

Algebraically, these constraint submanifolds can be described as

$$
\begin{equation*}
W_{i}=\left\{x \in T^{*}\left(T^{(k-1)} Q\right) \times_{T^{(k-1)} Q} \mathcal{M} \mid d H_{W_{0}}(x)(v)=0 \quad \forall v \in\left(T_{x} W_{i-1}\right)^{\perp}\right\} \quad i \geq 1, \tag{4.3.12}
\end{equation*}
$$

where $\left(T_{x} W_{i-1}\right)^{\perp}=\left\{v \in T_{x} W_{0} \mid \Omega_{W_{0}}(x)(u, v)=0 \quad \forall u \in T_{x} W_{i-1}\right\}$.
If this constraint algorithm stabilizes, i.e., there exists a positive integer $k \in \mathbb{N}$ such that $W_{k+1}=W_{k}$ and $\operatorname{dim} W_{k} \geq 1$, then we will have at least a well defined solution $X$ on $W_{f}=W_{k}$ such that

$$
\left(i_{X} \Omega_{W_{0}}=d H_{W_{0}}\right)_{\mid W_{f}} .
$$

Now, denote by $\Omega_{W_{1}}$, the pullback of the presymplectic 2 -form $\Omega_{W_{0}}$ to $W_{1}$. In order to establish a necessary and sufficient condition for the symplecticity of the 2 -form $\Omega_{W_{1}}$, we define the extended Lagrangian

$$
\mathcal{L}=L-\lambda_{\alpha} \Phi^{\alpha} .
$$

Theorem 4.3.7. For any choice of coordinates $\left(q^{(0) i}, q^{(1) i}, \ldots, q^{(k-1) i} ; p_{i}^{(0)}, \ldots, p_{i}^{(k-1)}, q^{(k) i}\right)$ in $T^{*}\left(T^{(k-1)} Q\right) \oplus T^{(k)} Q$, we have that $\left(W_{1}, \Omega_{W_{1}}\right)$ is a symplectic manifold if and only if

$$
\operatorname{det}\left(\begin{array}{cc}
\frac{\partial^{2} \mathcal{L}}{\partial q^{(k)} \partial q^{(k) j}} & -\frac{\partial \Phi^{\alpha}}{\partial q^{(k) i}}  \tag{4.3.13}\\
\frac{\partial \Phi^{\beta}}{\partial q^{(k) j}} & \mathbf{0}
\end{array}\right)=\operatorname{det}\left(\begin{array}{cc}
\frac{\partial^{2} L}{\partial q^{(k)} \partial q^{(k) j}}-\lambda_{\alpha} \frac{\partial \Phi^{\alpha}}{\partial q^{(k) i} \partial q^{(k) j}} & -\frac{\partial \Phi^{\alpha}}{\partial q^{(k) i}} \\
\frac{\partial \Phi^{\beta}}{\partial q^{(k) j}} & \mathbf{0}
\end{array}\right) \neq 0
$$

## Proof:

Let us recall that $\Omega_{W_{1}}$ is symplectic if and only if $T_{x} W_{1} \cap\left(T_{x} W_{1}\right)^{\perp}=0 \quad \forall x \in W_{1}$, where

$$
\left(T_{x} W_{1}\right)^{\perp}=\left\{v \in T_{x}\left(T^{*} T Q\right) \times_{T Q} \mathcal{M} / \Omega_{W_{0}}(x)(v, w)=0, \text { for all } w \in T_{x} W_{1}\right\} .
$$

Suppose that ( $W_{1}, \Omega_{W_{1}}$ ) is symplectic and that

$$
\lambda^{a} \mathcal{R}_{a b}(x)=0 \text { for some } \lambda^{a} \in \mathbb{R} \text { and } x \in W_{1} .
$$

Hence

$$
\lambda^{b} \mathcal{R}_{a b}(x)=\lambda^{b} d \varphi_{a}(x)\left(\left.\frac{\partial}{\partial \ddot{q}^{b}}\right|_{x}\right)=0 .
$$

Therefore, $\left.\lambda^{b} \frac{\partial}{\partial \dot{q}^{b}}\right|_{x} \in T_{x} W_{1}$ but it is also in $T_{x} W_{1}^{\perp}$. This implies that $\lambda_{b}=0$ for all $b$ and that the matrix $\left(\mathcal{R}_{a b}\right)$ is regular.

Now, suppose that the matrix $\left(\mathcal{R}_{a b}\right)$ is regular. Since

$$
\mathcal{R}_{a b}(x)=d \varphi_{a}(x)\left(\left.\frac{\partial}{\partial \ddot{q}^{b}}\right|_{x}\right),
$$

then, $\left.\frac{\partial}{\partial \dot{q}^{6}}\right|_{x} \notin T_{x} W_{1}$ and, in consequence,

$$
T_{x} W_{1} \oplus \operatorname{span}\left\{\left.\frac{\partial}{\partial \ddot{q}^{b}}\right|_{x}\right\}=T_{x} W_{0} .
$$

Now, let $Z \in T_{x} W_{1} \cap\left(T_{x} W_{1}\right)^{\perp}$ with $x \in W_{1}$. It follows that

$$
0=i_{Z} \Omega_{W_{0}}(x)\left(\left.\frac{\partial}{\partial \ddot{q}^{a}}\right|_{x}\right), \text { for all } a \text { and } i_{Z} \Omega_{W_{0}}(x)(\bar{Z})=0, \text { for all } \bar{Z} \in T_{x} W_{1}
$$

Then, $Z \in \operatorname{ker} \Omega_{W_{0}}(x)$. This implies that

$$
Z=\left.\lambda_{b} \frac{\partial}{\partial \ddot{q}^{b}}\right|_{x}
$$

Since $Z \in T_{x} W_{1}$ then

$$
0=d \varphi_{a}(x)(Z)=d \varphi_{a}(x)\left(\left.\lambda_{b} \frac{\partial}{\partial \ddot{q}^{b}}\right|_{x}\right)=\lambda_{b} \mathcal{R}_{a b}
$$

and, consequently, $\lambda_{b}=0$, for all $b$, and $Z=0$.

Remark 4.3.8. Observe that if the determinant of the matrix in Theorem 5.3.1 is not zero, then we can apply the implicit function theorem to the constraint equation $\varphi_{i}^{1}=0$ and $\Phi^{\alpha}=0$, and we can express the Lagrange multipliers $\lambda_{\alpha}$ and higher-order velocities $q^{(k) i}$ in terms of coordinates $\left(q^{(0) i}, \ldots, q^{(k-1) i}, p_{i}^{(0)}, \ldots, p_{i}^{(k-1)}\right)$, i.e.,

$$
\begin{aligned}
\lambda_{\alpha} & =\lambda_{\alpha}\left(q^{(0)}, q^{(1)}, \ldots, q^{(k-1)}, p^{(0)}, \ldots, p^{(k-1)}\right) \\
q^{(k) i} & =q^{(k) i}\left(q^{(0)}, q^{(1)}, \ldots, q^{(k-1)}, p^{(0)}, \ldots, p^{(k-1)}\right) .
\end{aligned}
$$

Thus we can consider $\left(q^{(0) i}, q^{(1) i}, \ldots, q^{(k-1) i}, p_{i}^{(0)}, \ldots, p_{i}^{(k-1)}\right)$ as local coordinates in $W_{1}$. In this case,

$$
\Omega_{W_{1}}=\sum_{r=0}^{k-1} d q^{(r) i} \wedge d p_{i}^{(r)}
$$

which is obviously symplectic.

Application: Numerical algorithm based on Hamilton's principle [15] In Lewis and Kostelec 45 ] is given a discussion of the use of Hamilton's variational principle to derive numerical methods for systems of differential equations derived from a variational principle, in particular, Hamilton-s equations and Euler-Lagrange equations. The comparison between these methods and symplectic algorithms was treated in [45]. In order to apply numerical methods based in Hamilton's principle, first, it is necessary to determine a class of functions with undetermined parameters for approximating the solutions of the continuous equations over a fixed interval. Finally, Hamilton's principle is applying exactly.

Consider a Lagrangian function $L: T \mathbb{R}^{n} \rightarrow \mathbb{R}$. The extremals functionals

$$
\mathcal{A}(q)=\int_{t_{0}}^{t_{1}} L(q(t), \dot{q}(t)) d t
$$

with boundary conditions $q\left(t_{0}\right)=q_{0}$, and $q\left(t_{1}\right)=q_{1}$ are the solutions of the Euler-Lagrange equations

$$
\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{i}}\right)-\frac{\partial L}{\partial q^{i}}=0
$$

The numerical approximation is derived by applying at each step Hamilton's principle over a particular subset of $\mathcal{C}^{2}$-functions. Usually, they are polynomials taking fixed values at the initial point. For instance, taking a polynomial of degree $m$ satisfying the initial conditions; that is,

$$
P(t)=\frac{t_{1}-t}{t_{1}-t_{0}} q_{0}+\frac{t-t_{0}}{t_{1}-t_{0}} q_{1}+\left(t_{1}-t\right)\left(t-t_{0}\right) \sum_{j=0}^{m-2} \alpha_{j} t^{j}
$$

The constraints are given by

$$
\Phi^{i}=q^{(m+1) i}=0, \quad 1 \leq i \leq n
$$

The equations of an integral curve for the system are

$$
\begin{aligned}
\frac{d}{d t} q^{(r) i} & =q^{(r+1) i} \\
-\frac{d}{d t} p_{i}^{0} & =-\frac{\partial L}{\partial q^{i}} \\
-\frac{d}{d t} p_{i}^{(1)} & =p_{i}^{(0)}-\frac{\partial L}{\partial \dot{q}^{i}} \\
-\frac{d}{d t} p_{i}^{(r)} & =p_{i}^{(r-1)}, \quad r=2, \ldots, m,
\end{aligned}
$$

and the constraints

$$
\begin{aligned}
\Psi^{i} & =q^{(m+1) i}=0, \quad 1 \leq i \leq n \\
\varphi_{i}^{1} & =p_{i}^{(m)}+\lambda_{i}=0, \quad 1 \leq i \leq n
\end{aligned}
$$

The regularity condition given in the last theorem gives us a regular matrix. Then the constraint algorithm stops in the first constraint submanifold $W_{1}$ and $\Omega_{W_{1}}$ is symplectic.

## Chapter 5

## Higher-Order Mechanical Systems on Lie Groups

In the recent years, a strong effect has been put on the study of higher-order mechanical systems on Lie groups. This kind of mechanical systems (without symmetries) appear for example in computational anatomy and interpolation problems on Lie groups, where we need to minimize the mean-square covariant acceleration (these minimal curves are known as Riemannian cubic splines)(see [35], [34]). Our principal motivation concerns optimal control problems. If we need, for example, to solve an optimal control problem where the state manifold is a Lie group, under some regularity conditions, it will be solved as a variational problem with constraints depends on higher-order derivatives. The geometric point of view also is treated in this section, that is, a geometric formalism to solve underactuated mechanical systems can be developed using the Skinner-Rusk formalism. The idea is the following: to solve an optimal control problem is equivalent to solve a higher-order problem with higher-order constraints (under some regularity conditions). To solve a higher-order problem with higher-order constraints is equivalent to solve a presymplectic Hamiltonian problem. With the Skinner-Rusk formalism we solve a presymplectic Hamiltonian problem and therefore we solve the optimal control problem for underactuated mechanical systems.

In the case of forced systems, for example, we consider a mechanical system determined by a Lagrangian $L: T G \equiv G \times \mathfrak{g} \rightarrow \mathbb{R}$, where $\mathfrak{g}$ is the Lie algebra of a Lie group $G$, and external forces $f: G \times \mathfrak{g} \rightarrow G \times \mathfrak{g}^{*}$. The motion of the mechanical system is described applying the following principle

$$
\delta \int_{0}^{1} L(g(t), \xi(t)) d t+\int_{0}^{1} f(g(t), \xi(t)) \eta(t) d t=0
$$

for all variations $\delta \xi(t)$ of the form $\delta \xi(t)=\dot{\eta}(t)+[\xi(t), \eta(t)]$, where $\eta$ is an arbitrary curve on the Lie algebra with $\eta(0)=0$ and $\eta(1)=0$.

These equations give us the forced Euler-Poincaré equations:

$$
\frac{d}{d t}\left(\frac{\delta L}{\delta \xi}\right)=\operatorname{ad}_{\xi}^{*}\left(\frac{\delta L}{\delta \xi}\right)+l_{g}^{*} \frac{\partial L}{\partial g}+f
$$

where $\operatorname{ad}_{\xi} \eta=[\xi, \eta]$ and $l_{g}^{*}=\left(T_{e} L_{g}\right)^{*}$.

The force $f$ is chosen in such a way it minimizes the cost functional:

$$
\begin{equation*}
\int_{0}^{1} C(g(t), \xi(t), f(g(t), \xi(t))) d t \tag{5.0.1}
\end{equation*}
$$

where now $C: G \times \mathfrak{g} \times \mathfrak{g}^{*} \rightarrow \mathbb{R}$. In underactuated systems the forces $f$ are constrained, and then second-order constraints appear . This kind of forced systems can be seen as a higher-order variational problem with higher-order constraints, and for this reason is a motivation to study this class of higher-order systems.

First, we give a variational approach of higher-order systems on Lie groups and we obtain the higher-order Euler-Lagrange equations for Lagrangians defined on $G \times k \mathfrak{g}$. From Hamilton's principle, also, in the case when the higher-order Lagrangian is left invariant; we will obtain the higher-order Euler-Poincaré equations. Next, from the Hamiltonian point of view, we obtain the equations for the dynamics on $T^{*}\left(T^{(k-1)} G\right)$. These equations are the higher-order Euler-Arnold equations. Finally, we develop the unifying framework for mechanics using an adaptation of the Skinner-Rusk formalism. We deduce the $k$ order Euler-Lagrange equations and, as a particular example, the $k$-order Euler-Poincaré equations. Since the dynamics is presymplectic, it is necessary to analyze its consistency using a constraint algorithm [37].

### 5.1 Higher-Order Euler-Poincaré equations

In this section we derive the $k^{t h}$-order Euler-Poincaré equations by the variational principle associated with the Lagrangian $L: T^{(k)} G \rightarrow \mathbb{R}$. Let $L: T^{(k)} G \simeq G \times k \mathfrak{g} \rightarrow \mathbb{R}$ be a Lagrangian function, $L\left(g, \dot{g}, \ddot{g}, \ldots, g^{(k)}\right) \equiv L\left(g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right)$ where $\xi=g^{-1} \dot{g}$ (left-trivialization). The problem consists on finding the critical curves of the functional

$$
\mathcal{J}=\int_{0}^{T} L\left(g, \xi, \dot{\xi}, \ddot{\xi}, \ldots, \xi^{(k-1)}\right) d t
$$

among all curves satisfying the boundary conditions for arbitrary variations $\delta g=\left.\frac{d}{d \epsilon}\right|_{\epsilon=0} g_{\epsilon}$, $\delta^{(l)} g=\frac{d^{l}}{d t^{l}}\left(\left.\frac{d}{d \epsilon}\right|_{\epsilon=0} g_{\epsilon}\right), l=1, \ldots, k$; where $\epsilon \mapsto g_{\epsilon}$ is a smooth curve in $G$ such that $g_{0}=g$.

We define, for any $\epsilon, \xi_{\epsilon}:=g_{\epsilon}^{-1} \dot{g}_{\epsilon}$. The corresponding variations $\delta \xi$ induced by $\delta g$ are given by $\delta \xi=\dot{\eta}+[\xi, \eta]$ where $\eta:=g^{-1} \delta g \in \mathfrak{g}(\delta g=g \eta)$. Moreover, $\delta^{(l)} \xi=\frac{d}{d t}\left(\left.\frac{d}{d t}\right|_{\epsilon=0} \xi_{\epsilon}\right), l=$ $1, \ldots, k-2$. Therefore

$$
\begin{gathered}
\delta \int_{0}^{T} L\left(g(t), \xi(t), \dot{\xi}(t), \ldots, \xi^{(k-1)}\right) d t= \\
\left.\frac{d}{d \epsilon}\right|_{\epsilon=0} \int_{0}^{T} L\left(g_{\epsilon}(t), \xi_{\epsilon}(t), \dot{\xi}_{\epsilon}(t), \ldots, \xi_{\epsilon}^{(k-1)}\right) d t= \\
\int_{0}^{T}\left(\left\langle\frac{\partial L}{\partial g}, \delta g\right\rangle+\left\langle\frac{\delta L}{\delta \xi}, \delta \xi\right\rangle+\sum_{j=1}^{k-1}\left\langle\frac{\delta L}{\delta \xi^{(j)}}, \frac{d^{j}}{d t^{j}} \delta \xi\right\rangle\right) d t= \\
\int_{0}^{T}\left(\left\langle\frac{\partial L}{\partial g}, \delta g\right\rangle+\sum_{j=0}^{k-1}\left\langle\frac{\delta L}{\delta \xi^{(j)}}, \frac{d^{j}}{d t^{j}} \delta \xi\right\rangle\right) d t= \\
\int_{0}^{T}\left(\left\langle\frac{\partial L}{\partial g}, \delta g\right\rangle+,\left\langle\sum_{j=0}^{k-1}(-1)^{j} \frac{d^{j}}{d t^{j}} \frac{\delta L}{\delta \xi^{(j)}}, \delta \xi\right\rangle\right) d t= \\
\quad \int_{0}^{T}\left(\left\langle\frac{\partial L}{\partial g}, g \eta\right\rangle+\left\langle\sum_{j=0}^{k-1}(-1)^{j} \frac{d^{j}}{d t^{j}} \frac{\delta L}{\delta \xi^{(j)}}, \frac{d}{d t} \eta+[\xi, \eta]\right\rangle\right) d t= \\
\int_{0}^{T}\left\langle\left(-\frac{d}{d t}+a d_{\xi}^{*}\right) \sum_{j=0}^{k-1}(-1)^{j} \frac{d^{j}}{d t^{j}} \frac{\delta L}{\delta \xi^{(j)}}, \eta\right\rangle d t+\int_{0}^{T}\left\langle l_{g}^{*}\left(\frac{\partial L}{\partial g}\right), \eta\right\rangle d t=0,
\end{gathered}
$$

where we have used integration by parts and the endpoints condition. Thus, the stationary condition $\delta \mathcal{J}=0$ implies the higher-order Euler-Lagrange equations,

$$
l_{g}^{*} \frac{\partial L}{\partial g}+\left(-\frac{d}{d t}+a d_{\xi}^{*}\right) \sum_{j=0}^{k-1}(-1)^{j} \frac{d^{j}}{d t^{j}} \frac{\delta L}{\delta \xi^{(j)}}=0
$$

Observe that when $k=1$, we recover the Euler-Lagrange equations on Lie groups given in the previous chapter. If $k=2$ we have that,

$$
\begin{equation*}
l_{g}^{*} \frac{\partial L}{\partial g}-\frac{d}{d t} \frac{\delta L}{\delta \xi}+\frac{d^{2}}{d t^{2}} \frac{\delta L}{\delta \dot{\xi}}+a d_{\xi}^{*} \frac{\delta L}{\delta \xi}-a d_{\xi}^{*}\left(\frac{d}{d t} \frac{\delta L}{\delta \dot{\xi}}\right)=0 . \tag{5.1.1}
\end{equation*}
$$

They are the second-order Euler-Lagrange equations on Lie groups (see [[27] and [28]]).
If the Lagrangian is invariant under an action of the Lie group, the equations of motion are

$$
\left(-\frac{d}{d t}+a d_{\xi}^{*}\right) \sum_{j=0}^{k-1}(-1)^{j} \frac{d^{j}}{d t^{j}} \frac{\delta L}{\delta \xi^{(j)}}=0 .
$$

In the second-order case, we have that

$$
\begin{equation*}
\frac{d^{2}}{d t^{2}} \frac{\delta L}{\delta \dot{\xi}}-\frac{d}{d t} \frac{\delta L}{\delta \xi}+a d_{\xi}^{*} \frac{\delta L}{\delta \xi}-a d_{\xi}^{*}\left(\frac{d}{d t} \frac{\delta L}{\delta \dot{\xi}}\right)=0 . \tag{5.1.2}
\end{equation*}
$$

These equations are called second-order Euler-Poincaré equations.

In a recent paper [35], the authors studied invariant higher-order problems and obtain the equations (5.1.2) working in a reduced Lagrangian setting on $\mathfrak{g} \times \mathfrak{g}$.

The results obtained above are summarized in the following theorem.
Theorem 5.1.1. Let $L: T^{(k)} G \simeq G \times k \mathfrak{g} \rightarrow \mathbb{R}$ be a Lagrangian function and let $g(t)$ be a curve in $G$ and $\xi(t)=g(t)^{-1} \dot{g}(t)$ be a curve in the Lie algebra $\mathfrak{g}$. Then the following assertions are equivalent.
(i) The curve $g(t)$ is a solution of the $k^{\text {th }}$-order Euler-Lagrange equations for $L: T^{(k)} G \rightarrow$ $\mathbb{R}$.
(ii) Hamilton's variational principle

$$
\delta \int_{t_{1}}^{t_{2}} L\left(g, \dot{g}, \ldots, g^{(k)}\right) d t=0
$$

holds upon using variations $\delta g$ such that $\delta g^{(j)}$ vanish at the endpoints for $j=0, \ldots, k-$ 1.
(iii) If $L$ is left-invariant, the $k^{\text {th }}$-order Euler-Poincaré equations

$$
\begin{equation*}
\left(\partial_{t} \pm \operatorname{ad}_{\xi}^{*}\right) \sum_{j=0}^{k-1}(-1)^{j} \partial_{t}^{j} \frac{\delta L}{\delta \xi^{(j)}}=0 \tag{5.1.3}
\end{equation*}
$$

holds
(iv) The variational principle

$$
\delta \int_{t_{1}}^{t_{2}} L\left(g, \xi, \dot{\xi}, \ldots, \xi^{(k)}\right)=0
$$

holds for constrained variations of the form $\delta \xi=\partial_{t} \eta \mp[\xi, \eta], \delta^{(l)} \xi=\frac{d}{d t}\left(\left.\frac{d}{d t}\right|_{\epsilon=0} \xi_{\epsilon}\right), l=$ $1, \ldots, k-2$. where $\eta$ is an arbitrary curve in $\mathfrak{g}$ such that $\eta^{(j)}$ vanishes at the endpoints, for all $j=0, \ldots, k-1$.

### 5.2 Higher-order Euler-Arnold's equations on $T^{*}\left(T^{(k-1)} G\right)$

In the previous chapter we have given a geometric approach for a trivialization of a higherorder tangent bundle. Similarly, we can trivialize the higher-order cotangent bundle in the following way,

$$
T^{*}\left(T^{(k-1)} G\right) \equiv T^{*}(G \times(k-1) \mathfrak{g}) \equiv T^{*} G \times(k-1) T^{*} \mathfrak{g} \equiv G \times(k-1) \mathfrak{g} \times k \mathfrak{g}^{*}
$$

To develop our geometric formalism for higher-order problems on Lie groups we need to equip the previous space with a symplectic structure. Thus, we construct a Liouville 1-form $\theta_{G \times(k-1) \mathfrak{g}}$ and a canonical symplectic 2-form $\omega_{G \times(k-1) \mathfrak{g}}$ on $T^{*}(G \times(k-1) \mathfrak{g})$ after the left-trivialization that we are using. Denote by $\boldsymbol{\xi} \in(k-1) \mathfrak{g}$ and $\boldsymbol{\alpha} \in k \mathfrak{g}^{*}$ with components $\boldsymbol{\xi}=\left(\xi^{(0)}, \ldots, \xi^{(k-2)}\right)$ and $\boldsymbol{\alpha}=\left(\alpha_{0}, \ldots, \alpha_{k-1}\right)$.

As we work in a vector space, the Liouville $1-$ form $\theta_{G \times(k-1) \mathfrak{g}} \in \Lambda^{1}\left(G \times \mathfrak{g}^{*} \times(k-1)\left(\mathfrak{g} \times \mathfrak{g}^{*}\right)\right)$ is expressed as

$$
\theta_{G \times(k-1) \mathfrak{g}}=\theta_{G}+\theta_{(k-1) \mathfrak{g}} .
$$

We are interested to know $\theta_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}$; this 1-form is applied to elements $\left(\boldsymbol{\xi}_{1}, \boldsymbol{\nu}^{1}\right) \in T_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}(G \times$ $\mathfrak{g}^{*} \times(k-1)\left(\mathfrak{g} \times \mathfrak{g}^{*}\right)$, where $\boldsymbol{\xi}_{a} \in k \mathfrak{g}$ and $\boldsymbol{\nu}^{a} \in k \mathfrak{g}^{*}, a=1,2$ with components $\boldsymbol{\xi}_{a}=\left(\xi_{a}^{(i)}\right)_{0 \leq i \leq k-1}$ and $\boldsymbol{\nu}^{a}=\left(\nu_{(i)}^{a}\right)_{0 \leq i \leq k-1}$ where each component $\xi_{a}^{(i)} \in \mathfrak{g}$ and $\nu_{(i)}^{a} \in \mathfrak{g}^{*}$. Observe that $\alpha_{0}$ comes from the identification $T^{*} G=G \times \mathfrak{g}^{*}$.

To calculate $\theta_{G}$ we need to find the tangent application to $\tau \circ \operatorname{Pr}_{(1,2)} \mathcal{L}$ where $\operatorname{Pr}_{(1,2)}$ : $G \times \mathfrak{g}^{*} \times \mathfrak{g} \times \mathfrak{g}^{*} \rightarrow G \times \mathfrak{g}^{*}$ is a canonical projection of the first and second factors, and $\tau: T^{*} G \rightarrow G$ is the fibration which defines $T^{*} G$. We consider the application

$$
\varphi_{t}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}: G \times \mathfrak{g}^{*} \rightarrow G \times \mathfrak{g}^{*} .
$$

This is applied to an element $\left(g, \alpha_{0}\right) \in G \times \mathfrak{g}^{*}$ and return an element $\left(g \exp \left(t \xi_{1}^{0}\right), \alpha_{0}+t \nu_{0}^{1}\right) \in$ $G \times \mathfrak{g}^{*} . \varphi_{t}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}$ is a flow of the vector field $X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}\left(g, \alpha_{0}\right)=\left(g \xi_{1}^{0}, \nu_{0}^{1}\right)$.

Therefore the tangent application for $\tau \circ \operatorname{Pr}_{(1,2)} \mathcal{L}$ is

$$
\begin{aligned}
T_{\left(g, \alpha_{0}\right)}\left(\tau \circ \operatorname{Pr}_{(1,2)} \mathcal{L}\right)\left(g \xi_{1}^{0}, \nu_{0}^{1}\right) & =\left.\frac{d}{d t}\right|_{t=0} \tau \circ \operatorname{Pr}_{(1,2)} \mathcal{L}\left(\varphi_{t}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}\left(g, \alpha_{0}\right)\right) \\
& =\left.\frac{d}{d t}\right|_{t=0} g \exp \left(t \xi_{1}^{0}\right)=g \xi_{1}^{0}
\end{aligned}
$$

where $\mathcal{L}: T^{*} G \rightarrow G \times \mathfrak{g}^{*}$ is the left-trivialization.
Now, we can calculate $\theta_{G}$

$$
\begin{aligned}
\left\langle\theta_{\left(g, \alpha_{0}\right)},\left(g \xi_{1}^{0}, \nu_{0}^{1}\right)\right\rangle & =\left\langle\theta_{\left(P r_{(1,2)} \mathcal{L}\right)}\left(g, \alpha_{0}\right), T_{\left(g, \alpha_{0}\right)}\left(\operatorname{Pr}_{(1,2)} \mathcal{L}\right)\left(g \xi_{1}^{0}, \nu_{0}^{1}\right)\right\rangle \\
& =\left\langle\alpha_{0}, \xi_{1}^{0}\right\rangle=\alpha_{0}\left(\xi_{1}^{0}\right) .
\end{aligned}
$$

In the same way as before, we calculate $\theta_{(k-1) \mathfrak{g}}$. This is given by $\sum_{i=1}^{k-1} \alpha_{i}\left(\xi_{1}^{(i)}\right)$. Then,

$$
\left(\theta_{G \times(k-1) \mathfrak{g}}\right)_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}\left(\boldsymbol{\xi}_{1}, \boldsymbol{\nu}^{1}\right)=\left\langle\boldsymbol{\alpha}, \boldsymbol{\xi}_{1}\right\rangle
$$

In the next, we will find the expression of the 2 -form $\omega_{G \times(k-1) \mathfrak{g}}$. For this, we will use the followings formulae

$$
-d \theta_{G \times(k-1) \mathfrak{g}}=-d\left(\theta_{G}+\theta_{(k-1) \mathfrak{g}}\right)=-d\left(\theta_{G}\right)-d\left(\theta_{(k-1) \mathfrak{g}}\right) .
$$

And to calculate $-d \theta_{G}$ we use the formula

$$
\begin{aligned}
-d \theta_{G}\left(X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}, X^{\left(\xi_{0}^{1}, \nu_{0}^{1}\right)}\right) & =-i_{X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}} d\left(i_{X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}} \theta_{G}\right) \\
& +i_{X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}} d\left(i_{\left.X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right.}\right)} \theta_{G}\right) \\
& +i_{\left[X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}, X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}\right]} \theta_{G}
\end{aligned}
$$

We calculate each term of the equality,

$$
\begin{aligned}
& i_{X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}} d\left(i_{\left.X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)} \theta_{G}\right)}\left(g, \alpha_{0}\right)=\right. \\
& L_{X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}}\left(i_{\left.X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right.}\right)} \theta_{G}\right)\left(g, \alpha_{0}\right)= \\
& \left.\frac{d}{d t}\right|_{t=0}\left(\varphi_{t}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}\right)^{*}\left(i_{X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)}} \theta_{G}\right)\left(g, \alpha_{0}\right)= \\
& \left.\frac{d}{d t}\right|_{t=0}\left\langle\theta_{G}\left(g \exp \left(t \xi_{1}^{0}\right), \alpha_{0}+t \nu_{0}^{1}\right), X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)}\left(g \exp \left(t \xi_{1}^{0}\right), \alpha_{0}+t \nu_{0}^{1}\right)\right\rangle= \\
& \left.\frac{d}{d t}\right|_{t=0}\left\langle\theta_{G}\left(g \exp \left(t \xi_{1}^{0}\right), \alpha_{0}+t \nu_{0}^{1}\right),\left(g \exp \nu_{0}^{1} \xi_{2}^{0}, \nu_{0}^{2}\right)\right\rangle= \\
& \left.\frac{d}{d t}\right|_{t=0}\left(\alpha_{0}+t \nu_{0}^{1}\right)\left(\nu_{0}^{2}\right)=\nu_{0}^{1}\left(\xi_{2}^{0}\right) .
\end{aligned}
$$

The second term is computed in a similar form, and is given by $\nu_{0}^{2}\left(\xi_{1}^{0}\right)$. To calculate the third term, we observe that

$$
\begin{aligned}
& {\left[X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}, X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)}\right]\left(g, \alpha_{0}\right)=\left.\frac{d}{d t}\right|_{t=0}\left(\varphi_{-\sqrt{t}}^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)} \circ \varphi_{-\sqrt{t}}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)} \circ \varphi_{\sqrt{t}}^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)} \circ \varphi_{\sqrt{t}}^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}\right)\left(g, \alpha_{0}\right)} \\
& =\left.\frac{d}{d t}\right|_{t=0}\left(g \exp \left(\sqrt{t} \xi_{1}^{0}\right) \exp \left(\sqrt{t} \xi_{2}^{0}\right) \exp -\sqrt{t} \xi_{1}^{0} \exp -\sqrt{t} \xi_{2}^{0}, \alpha_{0}\right)=\left(T_{e} L_{g}\left[\xi_{1}^{0}, \xi_{2}^{0}\right], 0\right)= \\
& \left(g\left[\xi_{1}^{0}, \xi_{2}^{0}\right], 0\right) .
\end{aligned}
$$

Then,

$$
\begin{aligned}
\theta_{G}\left(\left[X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}, X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)}\right]\right)\left(g, \alpha_{0}\right)= & =\theta_{G}\left(g, \alpha_{0}\right)\left(g\left[\xi_{1}^{0}, \xi_{2}^{0}\right], 0\right) \\
& =\alpha_{0}\left(\left[\xi_{1}^{0}, \xi_{2}^{0}\right]\right) .
\end{aligned}
$$

Therefore,

$$
-d \theta_{G}\left(X^{\left(\xi_{1}^{0}, \nu_{0}^{1}\right)}, X^{\left(\xi_{2}^{0}, \nu_{0}^{2}\right)}\right)=\omega_{\left(g, \alpha_{0}\right)}\left(\left(g \xi_{1}^{0}, \nu_{0}^{1}\right),\left(g \xi_{2}^{0}, \nu_{0}^{2}\right)\right)=-\nu_{0}^{1}\left(\xi_{2}^{0}\right)+\nu_{0}^{2}\left(\xi_{1}^{0}\right)+\alpha_{0}\left(\left[\xi_{1}^{0}, \xi_{2}^{0}\right]\right) .
$$

Applying, as before, the same formulae we also have

$$
\omega_{(k-1) \mathfrak{g}}=\sum_{i=1}^{k-1}\left\langle\nu_{(i)}^{1}, \xi_{2}^{(i)}\right\rangle+\left\langle\nu_{(i)}^{2}, \xi_{1}^{(i)}\right\rangle
$$

Then, since $\omega_{G \times(k-1) \mathfrak{g}}=\omega_{G}+\omega_{(k-1) \mathfrak{g}}$, we have the identities,

$$
\begin{aligned}
\left(\theta_{G \times(k-1) \mathfrak{g}}\right)_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}\left(\boldsymbol{\xi}_{1}, \boldsymbol{\nu}^{1}\right) & =\left\langle\boldsymbol{\alpha}, \boldsymbol{\xi}_{1}\right\rangle, \\
\left(\omega_{G \times(k-1) \mathfrak{g}}\right)_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}\left(\left(\boldsymbol{\xi}_{1}, \boldsymbol{\nu}^{1}\right),\left(\boldsymbol{\xi}_{2}, \boldsymbol{\nu}^{2}\right)\right) & =-\left\langle\boldsymbol{\nu}^{1}, \boldsymbol{\xi}_{2}\right\rangle+\left\langle\boldsymbol{\nu}^{2}, \boldsymbol{\xi}_{1}\right\rangle+\left\langle\alpha_{0},\left[\xi_{1}^{(0)}, \xi_{2}^{(0)}\right]\right\rangle \\
& =-\sum_{i=0}^{k-1}\left[\left\langle\nu_{(i)}^{1}, \xi_{2}^{(i)}\right\rangle+\left\langle\nu_{(i)}^{2}, \xi_{1}^{(i)}\right\rangle\right]+\left\langle\alpha_{0},\left[\xi_{1}^{(0)}, \xi_{2}^{(0)}\right]\right\rangle .
\end{aligned}
$$

Now, given the Hamiltonian $H: T^{*} T^{(k-1)} G \equiv G \times(k-1) \mathfrak{g} \times k \mathfrak{g}^{*} \longrightarrow \mathbb{R}$, we compute

$$
\begin{aligned}
d H_{(g, \boldsymbol{\xi}, \boldsymbol{\alpha})}\left(\boldsymbol{\xi}_{2}, \boldsymbol{\nu}^{2}\right)= & \left\langle £_{g}^{*}\left(\frac{\delta H}{\delta g}(g, \boldsymbol{\xi}, \boldsymbol{\alpha})\right), \xi_{2}^{(0)}\right\rangle+\sum_{i=0}^{k-2}\left\langle\frac{\delta H}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \xi_{2}^{(i+1)}\right\rangle \\
& +\left\langle\boldsymbol{\nu}^{2}, \frac{\delta H}{\delta \boldsymbol{\alpha}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha})\right\rangle
\end{aligned}
$$

As in the previous section, we can derive the Hamilton's equations which are satisfied by the integral curves of the Hamiltonian vector field $X_{H}$ defined by $X_{H}(g, \boldsymbol{\xi}, \boldsymbol{\alpha})=\left(\boldsymbol{\xi}_{1}, \boldsymbol{\nu}^{1}\right)$. Therefore, we deduce that

$$
\begin{aligned}
\boldsymbol{\xi}_{1} & =\frac{\delta H}{\delta \boldsymbol{\alpha}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \\
\nu_{(0)}^{1} & =-£_{g}^{*}\left(\frac{\delta H}{\delta g}(g, \boldsymbol{\xi}, \boldsymbol{\alpha})\right)+a d_{\xi_{1}^{(0)}}^{*} \alpha_{0}, \\
\nu_{(i+1)}^{1} & =-\frac{\delta H}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \quad 0 \leq i \leq k-2 .
\end{aligned}
$$

In other words, taking $\dot{g}=g \xi^{(0)}$ we obtain the higher-order Euler-Arnold's equations:

$$
\begin{aligned}
\dot{g} & =g \frac{\delta H}{\delta \alpha_{0}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \\
\frac{d \xi^{(i)}}{d t} & =\frac{\delta H}{\delta \alpha_{i}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \quad 1 \leq i \leq k-1 \\
\frac{d \alpha_{0}}{d t} & =-£_{g}^{*}\left(\frac{\delta H}{\delta g}(g, \boldsymbol{\xi}, \boldsymbol{\alpha})\right)+a d_{\delta H / \delta \alpha_{0}}^{*} \alpha_{0} \\
\frac{d \alpha_{i+1}}{d t} & =-\frac{\delta H}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \boldsymbol{\alpha}), \quad 0 \leq i \leq k-2
\end{aligned}
$$

### 5.3 Higher-Order Unified Mechanics on Lie Groups

In this section, we describe the main results of this chapter. First, we intrinsically derive the equations of motion for Lagrangian systems defined on higher-order tangent bundles of a Lie group and finally, we will extend the results to the cases of variationally constrained problems.

### 5.3.1 Unconstrained problem

The equations of motion: Now, we will give an adaptation of the Skinner-Rusk algorithm to the case of higher-order theories on Lie groups. We use the identifications

$$
\begin{aligned}
T^{(k)} G & \equiv G \times k \mathfrak{g} \\
T^{*} T^{(k-1)} G & \equiv G \times(k-1) \mathfrak{g} \times k \mathfrak{g}^{*}
\end{aligned}
$$

We define as in [28] the higher-order Pontryaguin bundle

$$
W_{0}=T^{(k)} G \times_{T^{(k-1)} G} T^{*} T^{(k-1)} G \equiv G \times k \mathfrak{g} \times k \mathfrak{g}^{*},
$$

with induced projections

$$
\begin{aligned}
p r_{1}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right) & =\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right) \\
p r_{2}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right) & =(g, \boldsymbol{\xi}, \boldsymbol{\alpha})
\end{aligned}
$$

where, $\boldsymbol{\xi}=\left(\xi^{(0)}, \ldots, \xi^{(k-2)}\right) \in(k-1) \mathfrak{g}$ and $\boldsymbol{\alpha}=\left(\alpha_{0}, \ldots, \alpha_{k-1}\right) \in k \mathfrak{g}^{*}$.


To developing the Skinner and Rusk formalism, it is only necessary to construct the presymplectic 2-form $\Omega_{W_{0}}$ by $\Omega_{W_{0}}=p r_{2}^{*} \omega_{G \times(k-1) \mathfrak{g}}$ and the Hamiltonian function $H: W_{0} \rightarrow$ $\mathbb{R}$ by

$$
H\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right)=\sum_{i=0}^{k-1}\left\langle\alpha_{i}, \xi^{(i)}\right\rangle-L\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right) .
$$

Therefore

$$
\begin{aligned}
& \left(\Omega_{W_{0}}\right)_{\left(g, \boldsymbol{\xi}, \xi\left(\xi^{(k-1)}, \boldsymbol{\alpha}\right)\right.}\left(\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right),\left(\boldsymbol{\xi}_{2}, \xi_{2}^{(k)}, \boldsymbol{\nu}^{2}\right)\right)=-\left\langle\boldsymbol{\nu}^{1}, \boldsymbol{\xi}_{2}\right\rangle+\left\langle\boldsymbol{\nu}^{2}, \boldsymbol{\xi}_{1}\right\rangle \\
& \quad+\left\langle\alpha_{0},\left[\xi_{1}^{(0)}, \xi_{2}^{(0)}\right]\right\rangle=-\sum_{i=0}^{k-1}\left[\left\langle\nu_{(i)}^{1}, \xi_{2}^{(i)}\right\rangle-\left\langle\nu_{(i)}^{2}, \xi_{1}^{(i)}\right\rangle\right]+\left\langle\alpha_{0},\left[\xi_{1}^{(0)}, \xi_{2}^{(0)}\right]\right\rangle
\end{aligned}
$$

where $\boldsymbol{\xi}_{a} \in k \mathfrak{g}, \boldsymbol{\nu}^{a} \in k \mathfrak{g}^{*}$, and $\xi_{a}^{(k)} \in \mathfrak{g}, a=1,2$. Observe that $\xi_{a}^{(k)}$ does not appear on the right-hand side of the previous expression, as a consequence of the presymplectic character of $\Omega_{W_{0}}$. Moreover,

$$
\begin{aligned}
d H_{\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right)}\left(\boldsymbol{\xi}_{2}, \xi_{2}^{(k)}, \boldsymbol{\nu}^{2}\right)= & \left\langle-£_{g}^{*}\left(\frac{\delta L}{\delta g}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)\right), \xi_{2}^{(0)}\right\rangle \\
& +\sum_{i=0}^{k-2}\left\langle\alpha_{i}-\frac{\delta L}{\delta \xi^{(i)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right), \xi_{2}^{(i+1)}\right\rangle \\
& +\left\langle\boldsymbol{\nu}^{2}, \boldsymbol{\xi}\right\rangle
\end{aligned}
$$

Therefore, the intrinsic equations of motion of a higher-order problem on Lie groups are now

$$
\begin{equation*}
i_{X} \Omega_{W_{0}}=d H \tag{5.3.1}
\end{equation*}
$$

If we look for a solution $X\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right)=\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k-1)}, \boldsymbol{\nu}^{1}\right)$ of Equation 5.3.1 we deduce:

$$
\begin{aligned}
\xi_{1}^{(i)} & =\xi^{(i)}, \quad 0 \leq i \leq k-1, \\
\nu_{(0)}^{1} & =£_{g}^{*}\left(\frac{\delta L}{\delta g}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)\right)+a d_{\xi_{1}^{(0)}} \alpha_{0}, \\
\nu_{(i+1)}^{1} & =\frac{\delta L}{\delta \xi^{(i)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)-\alpha_{i}, \quad 0 \leq i \leq k-2,
\end{aligned}
$$

and the constraint functions

$$
\alpha_{k-1}-\frac{\delta L}{\delta \xi^{(k-1)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)=0 .
$$

Observe that the coefficients $\xi_{1}^{k}$ are still undetermined.
An integral curve of $X$, that is a curve of the type

$$
t \longrightarrow\left(g(t), \xi(t), \ldots, \xi^{(k-1)}(t), \alpha_{0}(t), \ldots, \alpha_{k-1}(t)\right),
$$

must satisfy the following system of differential-algebraic equations (DAEs):

$$
\begin{align*}
\dot{g} & =g \xi  \tag{5.3.2}\\
\frac{d \xi^{(i-1)}}{d t} & =\xi^{(i)}, \quad 1 \leq i \leq k-1  \tag{5.3.3}\\
\frac{d \alpha_{0}}{d t} & =£_{g}^{*}\left(\frac{\delta L}{\delta g}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)\right)+a d_{\xi}^{*} \alpha_{0},  \tag{5.3.4}\\
\frac{d \alpha_{i+1}}{d t} & =\frac{\delta L}{\delta \xi^{(i)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)-\alpha_{i}, \quad 0 \leq i \leq k-2  \tag{5.3.5}\\
\alpha_{k-1} & =\frac{\delta L}{\delta \xi^{(k-1)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right) \tag{5.3.6}
\end{align*}
$$

If $k \geq 2$, combining Equation 5.3.6 with the 5.3.5 for $i=k-2$, we obtain

$$
\frac{d}{d t} \frac{\delta L}{\delta \xi^{(k-1)}}=\frac{\delta L}{\delta \xi^{(k-2)}}-\alpha_{k-2} .
$$

Proceeding successively, now with $i=k-3$ and ending with $i=0$ we obtain the following relation:

$$
\alpha_{0}=\sum_{i=0}^{k-1}(-1)^{i} \frac{d^{i}}{d t^{i}} \frac{\delta L}{\delta \xi^{(i)}} .
$$

This last expression is also valid for $k \geq 1$. Substituting in the Equation (5.3.4) we finally deduce the $k$-order trivialized Euler-Lagrange equations:

$$
\begin{equation*}
\left(\frac{d}{d t}-a d_{\xi}^{*}\right) \sum_{i=0}^{k-1}(-1)^{i} \frac{d^{i}}{d t^{i}} \frac{\delta L}{\delta \xi^{(i)}}=£_{g}^{*}\left(\frac{\delta L}{\delta g}\right) . \tag{5.3.7}
\end{equation*}
$$

Of course if the Lagrangian $L: T^{(k)} G \equiv G \times k \mathfrak{g} \longrightarrow \mathbb{R}$ is left-invariant, that is

$$
L\left(g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right)=L\left(h, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right)
$$

for all $g, h \in G$, then defining the reduced Lagrangian $l: k \mathfrak{g} \longrightarrow \mathbb{R}$ by

$$
l\left(\xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right)=L\left(e, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}\right)
$$

we write Equations (5.3.7) as

$$
\begin{equation*}
\left(\frac{d}{d t}-a d_{\xi}^{*}\right) \sum_{i=0}^{k-1}(-1)^{i} \frac{d^{i}}{d t^{i}} \frac{\delta l}{\delta \xi^{(i)}}=0 \tag{5.3.8}
\end{equation*}
$$

which are the $k$-order Euler-Poincaré equations (see, for instance, [35]).

## The constraint algorithm

Since $\Omega_{W_{0}}$ is presymplectic, then (5.3.1) has not solution along $W_{0}$ then it is necessary to identify the unique maximal submanifold $W_{f}$ which (5.3.1) possesses tangent solutions on $W_{f}$. This final constraint submanifold $W_{f}$ is detected using the Gotay-Nester-Hinds algorithm [?]. This algorithm prescribes that $W_{f}$ is the limit of a string of sequentially constructed constraint submanifolds

$$
\cdots \hookrightarrow W_{k} \hookrightarrow \cdots \hookrightarrow W_{2} \hookrightarrow W_{1} \hookrightarrow W_{0} .
$$

where

$$
\begin{aligned}
W_{i}= & \left\{x \in G \times k \mathfrak{g} \times k \mathfrak{g}^{*} \mid d H(x)\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right)=0\right. \\
& \left.\forall\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right) \in\left(T_{x} W_{i-1}\right)^{\perp}\right\}
\end{aligned}
$$

with $i \geq 1$ and where

$$
\begin{aligned}
\left(T_{x} W_{i-1}\right)^{\perp} & =\left\{\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right) \in(k+1) \mathfrak{g} \times k \mathfrak{g}^{*} \mid \Omega_{W_{0}}(x)\left(\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right),\left(\boldsymbol{\xi}_{2}, \xi_{2}^{(k)}, \boldsymbol{\nu}^{2}\right)\right)=0\right. \\
& \left.\forall\left(\boldsymbol{\xi}_{2}, \xi_{2}^{(k)}, \boldsymbol{\nu}^{2}\right) \in T_{x} W_{i-1}\right\} .
\end{aligned}
$$

where we are using the previously defined identifications. If this constraint algorithm stabilizes, i.e., there exists a positive integer $k \in \mathbb{N}$ such that $W_{k+1}=W_{k}$ and $\operatorname{dim} W_{k} \geq 1$, then we will have at least a well defined solution $X$ on $W_{f}=W_{k}$ such that

$$
\left(i_{X} \Omega_{W_{0}}=d H\right)_{\mid W_{f}}
$$

From these definitions, we deduce that the first constraint submanifold $W_{1}$ is defined by the vanishing of the constraint functions

$$
\alpha_{k-1}-\frac{\delta L}{\delta \xi^{(k-1)}}=0 .
$$

Applying the constraint algorithm we deduce that the following condition, if $k>2$ :

$$
\frac{\delta L}{\delta \xi^{(k-2)}}-\alpha_{k-2}=\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} \xi_{1}^{(k)}+\sum_{i=0}^{k-2} \frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(i)}} \xi^{i+1}+£_{g}^{*}\left(\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta g}\right) \xi
$$

In the particular case $k=1$, we deduce the equation

$$
£_{g}^{*}\left(\frac{\delta L}{\delta g}\right)+a d_{\xi}^{*} \alpha_{0}=\frac{\delta^{2} L}{\delta \xi^{2}} \xi_{1}^{(1)}+£_{g}^{*}\left(\frac{\delta^{2} L}{\delta \xi \delta g}\right) \xi .
$$

In both cases, these equations impose restrictions over the remainder coefficients $\xi_{1}^{(k)}$ of the vector field $X$.

If the bilinear form $\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}}: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$ defined by

$$
\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)(\xi, \tilde{\xi})=\left.\left.\frac{d}{d t}\right|_{t=0} \frac{d}{d s}\right|_{s=0} L\left(g, \boldsymbol{\xi}, \xi^{(k-1)}+t \xi+s \tilde{\xi}\right)
$$

is nondegenerate, we have a special case when the constraint algorithm finishes at the first step $W_{1}$. More precisely, if we denote by $\Omega_{W_{1}}$ the restriction of the presymplectic 2-form $\Omega$ to $W_{1}$, then we have the following theorem,

Theorem 5.3.1. $\left(W_{1}, \Omega_{W_{1}}\right)$ is a symplectic manifold if and only if

$$
\begin{equation*}
\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} \tag{5.3.9}
\end{equation*}
$$

is nondegenerate.

### 5.3.2 Constrained problem

## The equations of motion

The geometrical interpretation of constrained problems determined by a submanifold $\mathcal{M}$ of $G \times k \mathfrak{g}$, with inclusion $i_{\mathcal{M}}: \mathcal{M} \hookrightarrow G \times k \mathfrak{g}$ and a Lagrangian function defined on it, $L_{\mathcal{M}}: \mathcal{M} \rightarrow \mathbb{R}$, is an extension of the previous framework. First, it is necessary to note that for constrained systems, we understand a variational problem subject to constraints, being this analysis completely different in the case of nonholonomic constraints (see [10, 32, 21]).

Given the pair ( $\mathcal{M}, L_{\mathcal{M}}$ ) we can define the space

$$
\bar{W}_{0}=\mathcal{M} \times k \mathfrak{g}^{*} .
$$

Take the inclusion $i_{\bar{W}_{0}}: \bar{W}_{0} \hookrightarrow G \times k \mathfrak{g} \times k \mathfrak{g}^{*}$, then we can construct the following presymplectic form

$$
\Omega_{\bar{W}_{0}}=\left(p r_{2} \circ i_{\bar{W}_{0}}\right)^{*} \Omega_{G \times(k-1) \mathfrak{g} \times k \mathfrak{g}^{*}},
$$

and the function $\bar{H}: \bar{W}_{0} \rightarrow \mathbb{R}$ defined by

$$
\bar{H}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right)=\sum_{i=0}^{k-1}\left\langle\alpha_{i}, \xi^{(i)}\right\rangle-L_{\mathcal{M}}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right),
$$

where $\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right) \in \mathcal{M}$.
With these two elements it is possible to write the following presymplectic system:

$$
\begin{equation*}
i_{X} \Omega_{\bar{W}_{0}}=d \bar{H} \tag{5.3.10}
\end{equation*}
$$

This then justifies the use of the following terminology.
Definition 5.3.2. The presymplectic Hamiltonian system $\left(\bar{W}_{0}, \Omega_{\bar{W}_{0}}, \bar{H}\right)$ will be called the variationally constrained Hamiltonian system.

To characterize the equations we will adopt an "extrinsic point of view", that is, we will work on the full space $W_{0}$ instead of in the restricted space $\overline{W_{0}}$. Consider an arbitrary extension $L: G \times k \mathfrak{g} \rightarrow \mathbb{R}$ of $L_{\mathcal{M}}: \mathcal{M} \rightarrow \mathbb{R}$. The main idea is to take into account that Equation (5.3.10) is equivalent to

$$
\left\{\begin{aligned}
i_{X} \Omega_{W_{0}}-d H & \in \operatorname{ann} T \bar{W}_{0} \\
X & \in T \bar{W}_{0}
\end{aligned}\right.
$$

where ann denotes the annihilator of a distribution and $H$ is the function defined in Section 5.3.1.

Assuming that $\mathcal{M}$ is determined by the vanishing of $m$-independent constraints

$$
\Phi^{A}\left(g, \boldsymbol{\xi}, \xi^{(k-1)}\right)=0, \quad 1 \leq A \leq m
$$

then, locally, ann $T \bar{W}_{0}=\operatorname{span}\left\{d \Phi^{A}\right\}$, and therefore the previous equations is rewritten as

$$
\left\{\begin{aligned}
i_{X} \Omega_{W_{0}}-d H & =\lambda_{A} d \Phi^{A}, \\
X\left(\Phi^{A}\right) & =0,
\end{aligned}\right.
$$

where $\lambda_{A}$ are Lagrange multipliers to be determined.
If $X\left(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}\right)=\left(\boldsymbol{\xi}_{1}, \xi_{1}^{(k)}, \boldsymbol{\nu}^{1}\right)$ then, as in the previous subsection, we obtain the following prescription about these coefficients:

$$
\begin{aligned}
\xi_{1}^{(i)} & =\xi^{(i)}, \quad 0 \leq i \leq k-1 \\
\nu_{(0)}^{1} & =£_{g}^{*}\left(\frac{\delta L}{\delta g}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta g}\right)+a d_{\xi_{1}^{(0)}} \alpha_{0}, \\
\nu_{(i+1)}^{1} & =\frac{\delta L}{\delta \xi^{(i)}}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(i)}}-\alpha_{i}, \quad 0 \leq i \leq k-2, \\
0 & =£_{g}^{*}\left(\frac{\delta \Phi^{A}}{\delta g}\right) \xi+\sum_{i=1}^{k-2} \frac{\delta \Phi^{A}}{\delta \xi^{(i)}} \xi^{(i+1)}+\frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}} \xi_{1}^{(k)}, \quad 1 \leq A \leq m
\end{aligned}
$$

and the algebraic equations:

$$
\begin{aligned}
\alpha_{k-1}-\frac{\delta L}{\delta \xi^{(k-1)}}+\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}} & =0 \\
\Phi^{A} & =0
\end{aligned}
$$

The integral curves of $X$ satisfy the system of differential-algebraic equations with additional unknowns $\left(\lambda_{A}\right)$ :

$$
\begin{aligned}
\dot{g} & =g \xi, \\
\frac{d \xi^{(i-1)}}{d t} & =\xi^{(i)}, \quad 1 \leq i \leq k-1, \\
\frac{d \alpha_{0}}{d t} & =£_{g}^{*}\left(\frac{\delta L}{\delta g}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta g}\right)+a d_{\xi}^{*} \alpha_{0}, \\
\frac{d \alpha_{i+1}}{d t} & =\frac{\delta L}{\delta \xi^{(i)}}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(i)}}-\alpha_{i}, \\
0 & =£_{g}^{*}\left(\frac{\delta \Phi^{A}}{\delta g}\right) \xi+\sum_{i=1}^{k-2} \frac{\delta \Phi^{A}}{\delta \xi^{(i)}} \xi^{(i+1)}+\frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}} \xi_{1}^{(k-1)} \\
\alpha_{k-1} & =\frac{\delta L}{\delta \xi^{(k-1)}}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}}, \\
\Phi^{A} & =0 .
\end{aligned}
$$

As a consequence we finally obtain the $k$-order trivialized constrained Euler-Lagrange equations,

$$
\begin{equation*}
\left(\frac{d}{d t}-a d_{\xi}^{*}\right) \sum_{i=0}^{k-1}(-1)^{i} \frac{d^{i}}{d t^{i}}\left[\frac{\delta L}{\delta \xi^{(i)}}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(i)}}\right]=£_{g}^{*}\left(\frac{\delta L}{\delta g}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta g}\right) . \tag{5.3.11}
\end{equation*}
$$

If the Lagrangian $L: T^{(k)} G \equiv G \times k \mathfrak{g} \longrightarrow \mathbb{R}$ and the constraints $\Phi^{A}: G \times k \mathfrak{g} \longrightarrow \mathbb{R}$, $1 \leq A \leq m$ are left-invariant then defining the reduced lagrangian $l: k \mathfrak{g} \longrightarrow \mathbb{R}$ and the reduced constraints $\phi^{A}: k \mathfrak{g} \rightarrow \mathbb{R}$ we write Equations (5.3.11) as

$$
\begin{equation*}
\left(\frac{d}{d t}-a d_{\xi}^{*}\right) \sum_{i=0}^{k-1}(-1)^{i} \frac{d^{i}}{d t^{i}}\left[\frac{\delta l}{\delta \xi^{(i)}}-\lambda_{A} \frac{\delta \phi^{A}}{\delta \xi^{(i)}}\right]=0 . \tag{5.3.12}
\end{equation*}
$$

## The constraint algorithm

As in the previous subsection it is possible to apply the Gotay-Nester algorithm to obtain a final constraint submanifold where we have at least a solution which is dynamically compatible. The algorithm is exactly the same but applied to the equation (5.3.10).

The first constraint submanifold $\bar{W}_{1}$ is determined by the conditions

$$
\begin{aligned}
\alpha_{k-1} & =\frac{\delta L}{\delta \xi^{(k-1)}}-\lambda_{A} \frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}}, \\
\Phi^{A} & =0
\end{aligned}
$$

If we denote by $\Omega_{\bar{W}_{1}}$ the pullback of the presymplectic 2 -form $\Omega_{\bar{W}_{0}}$ to $\bar{W}_{1}$, we have the following theorem,

Theorem 5.3.3. ( $\left.\bar{W}_{1}, \Omega_{\bar{W}_{1}}\right)$ is a symplectic manifold if and only if

$$
\left(\begin{array}{ll}
\frac{\delta^{2} L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} & \frac{\delta \Phi^{A}}{\delta \xi^{(k-1)}}  \tag{5.3.13}\\
\frac{\delta \Phi^{A} A}{\delta \xi^{(k-1)}} & \mathbf{0}
\end{array}\right)
$$

is nondegenerate, considered as a bilinear form on the vector space $\mathfrak{g} \times \mathbb{R}^{m}$.

## Chapter 6

## Optimal Control of Mechanical Systems

As we said in the introduction, the goal of control theory is determine the behavior of a dynamical system by external forces acting. In this chapter we gives a selection of techniques and results in optimal control theory that are optimization problems for mechanical systems based on [10], and [56].

### 6.1 Optimal Control

Given a set of constraints, there are two type of associated problems. One of them is not variational (Lagrange - D'Alambert principle) but very appropriated to study the dynamics of certain class of mechanical systems [14], [13]. The other one is the variational approach. This framework is the suitable framework to study the class of optimal control systems.

Recall that the variational problems are equivalent to the classical problems of minimization, which consist on minimizing the Lagrangian action over a set of curves which satisfies the condition of fixed endpoints.

That is, let $Q$ be a configuration manifold and $T Q$ its tangent bundle with local coordinates $\left(q^{i}, \dot{q}^{i}\right)$. Let $L: T Q \rightarrow \mathbb{R}$ be a Lagrangian function and $\Phi: T Q \rightarrow \mathbb{R}^{n-m}$ given smooth constraints.

Definition 6.1.1. [10] The Lagrange problem is given by

$$
\min _{q(\cdot)} \int_{0}^{T} L\left(q^{i}, \dot{q}^{i}\right) d t
$$

subject to the condition of fixed endpoints, $q(0)=0, q(T)=q_{T}$, and subject to the constraint functions,

$$
\Phi\left(q^{i}, \dot{q}^{i}\right)=0 .
$$

### 6.1.1 Optimal Control and Maximum Principle

In this subsection we will discuss the maximum principle, which gives rise necessary conditions for the existence of a solution curve of the optimal control problem. Optimal control
problems include, for example, the minimum time problem given by Bernoulli, which is not given in a classical variational procedure.

The difference between both, the optimal control problem and the variational problems with constraints, is the extremal conditions, which are expressed in Hamiltonian form and Pontryagin's maximum principle ; nevertheless, the variational approach gives us a Lagrangian setting.

## General Setting of Optimal Control Problems

Suppose that we have a classical optimal control problem,

$$
\begin{equation*}
\min _{u(\cdot)} \int_{0}^{T} g(q, u) d t \tag{6.1.1}
\end{equation*}
$$

subject to the conditions:

- a differential equation $\dot{q}=f(q, u)$, and the state space contains $q \in Q$ and the controls in $\Omega \in \mathbb{R}^{k}$;
- $q(0)=q_{0}, q(T)=q_{T}$
where $f$ and $g \geq 0$ are smooth functions, $\Omega$ is a closed subset of $\mathbb{R}^{k}$, and $Q$ is a $n$-dimensional differentiable manifold, called state space of the system. The function $g$ is the cost function or objective.

Pontryagin's Maximum Principle Consider a Hamiltonian parameterized on $T^{*} Q$ and given by

$$
\widehat{H}(q, p, u)=\langle p, f(q, u)\rangle-p_{0} g(q, u),
$$

where $p_{0} \geq 0$ is a fixed positive constant and $p \in T^{*} Q$. We observe that $p_{0}$ is a multiplier of the cost functional and that $\widehat{H}$ is linear in $p$. We denote by $t \mapsto u^{*}(t)$ a curve that satisfies the following relationship along the trajectory $t \mapsto(q(t), p(t)) \in T^{*} Q$ :

$$
\begin{equation*}
H\left(q(t), p(t), u^{*}(t)\right)=\max _{u \in \Omega} \widehat{H}(q(t), p(t), u) . \tag{6.1.2}
\end{equation*}
$$

Then, if $u^{*}$ defines implicitly a function depending of $q$ and $p$ by the equation the equation (6.1.2), we can define $H^{*}$ by

$$
H^{*}(q(t), p(t), t)=H\left(q(t), p(t), u^{*}(t)\right)
$$

The time-varying function $H^{*}$ defines a time-varying Hamiltonian vector field $X_{H^{*}}$ on $T^{*} Q$ with respect to the canonical symplectic structure on $T^{*} Q$.

Pontryagin's maximum principle gives necessary conditions for extremals of the optimal control problem as follows: An extremal trajectory $t \mapsto q(t)$ for the optimal control problem is a projection onto $Q$ of a trajectory of the flow of the vector field $X_{H^{*}}$ that satisfies the boundary conditions $q(0)=q_{0}, q(T)=q_{T}$ and for which $t \mapsto\left(p(t), p_{0}\right) \neq 0$ for all $t \in[0, T]$.

The extremal is called normal when $p_{0} \neq 0$. When $p_{0}=0$ we said that the extremal is abnormal. Moreover, $u^{*}$ is determined in unique form under the condition

$$
0=\frac{\partial \widehat{H}}{\partial u}\left(q(t), p(t), u^{*}(t)\right), t \in[0, T] .
$$

That is, $u^{*}$ minimize the function $\widehat{H}$.
By the implicit function theorem, there exists a function $k$ such that $u^{*}(t)=k(q(t), p(t))$. We establish that

$$
H(q, p)=\widehat{H}(q, p, k(q, p))
$$

then, along the extremal curves,

$$
H(q(t), p(t))=H^{*}(q(t), p(t), t)
$$

### 6.2 Variational Problems and Optimal Control

Variational problems with constraints are equivalent to an optimal control problem under some regularity conditions.

Consider the modified Lagrangian,

$$
\begin{equation*}
\Lambda(q, \dot{q}, \lambda)=L(q, \dot{q})+\lambda \Phi(q, \dot{q}) \tag{6.2.1}
\end{equation*}
$$

The Euler-Lagrange equations are given by

$$
\begin{align*}
\frac{d}{d t} \frac{\partial}{\partial \dot{q}} \Lambda(q, \dot{q}, \lambda)-\frac{\partial}{\partial q} \Lambda(q, \dot{q}, \lambda) & =0  \tag{6.2.2}\\
\Phi(q, \dot{q}) & =0 \tag{6.2.3}
\end{align*}
$$

We rewrite these equations in a Hamiltonian form and we prove that these equations are equivalent to the equations of motion given by the maximum principle for a suitable optimal control problem.

Let

$$
\begin{equation*}
p=\frac{\partial}{\partial \dot{q}} \Lambda(q, \dot{q}, \lambda) \tag{6.2.4}
\end{equation*}
$$

and consider this equation with the constraints

$$
\begin{equation*}
\Phi(q, \dot{q})=0 \tag{6.2.5}
\end{equation*}
$$

Then we wish to solve (6.2.4) and (6.2.5) for $(\dot{q}, \lambda)$. We assume that on an open set $U \subset Q$, the matrix

$$
\left(\begin{array}{cc}
\frac{\partial^{2}}{\partial \dot{q}^{2}} \Lambda(q, \dot{q}, \lambda) & \frac{\partial}{\partial \dot{q}} \Phi(q, \dot{q})^{T} \\
\frac{\partial}{\partial \dot{q}} \Phi(q, \dot{q}) . & 0
\end{array}\right)
$$

has full rank. Then, by the implicit function theorem we can clear $\dot{q}$ and $\lambda$ as a function of $q$ and $p$

$$
\begin{align*}
\dot{q} & =\phi(q, p)  \tag{6.2.6}\\
\lambda & =\psi(q, p) \tag{6.2.7}
\end{align*}
$$

Theorem 6.2.1 (Caratheódory (1967) ,Rund (1966), Arnold, Kozlov and Neishtadt (1988), Bloch and Crouch (1994)). Under the transformations (6.2.6) and (6.2.7), the EulerLagrange system (6.2.2) is transformed in the Hamiltonian system

$$
\begin{aligned}
\dot{q} & =\frac{\partial}{\partial p} H(q, p) \\
\dot{p} & =-\frac{\partial}{\partial q} H(q, p)
\end{aligned}
$$

where

$$
\begin{equation*}
H(q, p)=p \cdot \phi(q, p)-L(q, \phi(q, p)) \tag{6.2.8}
\end{equation*}
$$

Proof: The fact that $\Phi(q, \phi(q, p))=0$ implies that

$$
\begin{gathered}
\frac{\partial \Phi}{\partial q}+\frac{\partial \Phi}{\partial \dot{q}} \frac{\partial \phi}{\partial q}=0 \\
\frac{\partial \Phi}{\partial \dot{q}} \frac{\partial \phi}{\partial p}=0
\end{gathered}
$$

Then, using (6.2.4), we obtain that

$$
\frac{\partial H}{\partial p}=\phi+\left(p-\frac{\partial L}{\partial \dot{q}}\right) \frac{\partial \phi}{\partial p}=\dot{q}+\lambda\left(\frac{\partial \Phi}{\partial \dot{q}} \frac{\partial \phi}{\partial p}\right)=\dot{q}
$$

In a similar form,

$$
\frac{\partial H}{\partial q}=-\frac{\partial L}{\partial q}+\left(p-\frac{\partial L}{\partial \dot{q}}\right) \frac{\partial \phi}{\partial p}=-\frac{\partial L}{\partial \dot{q}}+\lambda\left(\frac{\partial \Phi}{\partial \dot{q}} \frac{\partial \phi}{\partial p}\right)=-\left(\frac{\partial L}{\partial \dot{q}}+\lambda \frac{\partial \Phi}{\partial q}\right)=-\frac{\partial \Lambda}{\partial q}=-\dot{p} .
$$

Definition 6.2.2. Let $q \in \mathbb{R}^{n}$, $u \in \mathbb{R}^{m}$ the Optimal Control Problem is given by

$$
\begin{equation*}
\min _{u(\cdot)} \int_{0}^{T} g(q, u) d t \tag{6.2.9}
\end{equation*}
$$

subject to

$$
\dot{q}=f(q, u),
$$

with boundary conditions $q(0)=0, q(T)=q_{T}$.
With this definition we have the following theorem,
Theorem 6.2.3. [10] The Lagrange problem and optimal control problem generate the same extremal trajectories if and only if,

1. $\Phi(q, \dot{q})=0$ if and only if there exits $u$ such that $\dot{q}=f(q, u)$.
2. $L(q, f(q, u))=g(q, u)$.
3. The optimal control $u^{*}$ is uniquely determined by the condition

$$
\frac{\partial \widehat{H}}{\partial u}\left(q, p, u^{*}\right)=0
$$

where

$$
\frac{\partial^{2} \widehat{H}}{\partial u^{2}}\left(q, p, u^{*}\right)
$$

has full rank and

$$
\begin{equation*}
\widehat{H}(q, p, u)=\langle p, f(q, u)\rangle-g(q, u) \tag{6.2.10}
\end{equation*}
$$

is the Hamiltonian given by the maximum principle.

Proof: By (3), we can use the equation

$$
\frac{\partial \widehat{H}}{\partial u}\left(q, p, u^{*}\right)=p \cdot \frac{\partial f}{\partial u}\left(q, u^{*}\right)-\frac{\partial g}{\partial u}\left(q, u^{*}\right)=0
$$

to deduce that, there exists a function $r$ such that $u^{*}=r(q, p)$. The extremal trajectories are now generated by the Hamiltonian

$$
\begin{equation*}
\bar{H}(q, p)=\widehat{H}(q, p, r(x, p))=p \cdot f(q, r(q, p))-g(q, r(q, p)) . \tag{6.2.11}
\end{equation*}
$$

Then, the result follows, and we have

$$
\begin{align*}
\bar{H}(q, p) & =H(q, p)  \tag{6.2.12}\\
f(q, r(q, p)) & =\phi(q, p)  \tag{6.2.13}\\
g(q, r(q, p)) & =L(q, \phi(q, p)) \tag{6.2.14}
\end{align*}
$$

### 6.3 Lagrangian and Hamiltonian Control Systems

The extension of the notion of Hamiltonian and Lagrangian systems to control theory was formally proposed by Brockett, Willems and van der Schaft, among others.

The simplest form of Lagrangian control system is a Lagrangian system with external forces: in local coordinates we have

$$
\begin{align*}
& \frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{i}}\right)-\frac{\partial L}{\partial q^{i}}=u_{i}, \quad i=1, \ldots, m,  \tag{6.3.1}\\
& \frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{i}}\right)-\frac{\partial L}{\partial q^{i}}=0, \quad i=m+1, \ldots, n . \tag{6.3.2}
\end{align*}
$$

More generally, we have the system

$$
\frac{d}{d t}\left(\frac{\partial L(q, \dot{q}, u)}{\partial \dot{q}^{i}}\right)-\frac{\partial L(q, \dot{q}, u)}{\partial q^{i}}
$$

for $q \in \mathbb{R}^{n}$ and $u \in \mathbb{R}^{m}$ (see [10] and reference therein). When $m<n$ we say that the system is underactuated.

Similarly, one can define a Hamiltonian control system. In local coordinates, these have the form

$$
\begin{aligned}
& \dot{q}^{i}=\frac{\partial H(q, p, u)}{\partial p_{i}} \\
& \dot{p}^{i}=-\frac{\partial H(q, p, u)}{\partial q_{i}},
\end{aligned}
$$

for $q, p \in \mathbb{R}^{n}$ and $u \in \mathbb{R}^{m}$.

### 6.4 Optimal Control of Mechanical Systems on Lie Groups

### 6.4.1 Left-invariant control systems

Let $G$ be a Lie group and $\mathfrak{g}$ its Lie algebra. A left-invariant control system $\Gamma$ in a Lie group $G$ is a subset of $\mathfrak{g}$;

$$
\Gamma \subset \mathfrak{g}
$$

The most typical examples are the affine control systems

$$
\Gamma=\left\{\xi+\sum_{i=1}^{m} u_{i} \xi_{i} \mid u=\left(u_{1}, \ldots, u_{m}\right) \in U \subset \mathbb{R}^{m}\right\}
$$

where $\xi, \xi_{1}, \ldots, \xi_{m}$ are elements of $\mathfrak{g}$. The classical form in which are written this kind of control systems is

$$
\dot{g}=g \xi+\sum_{i=1}^{m} u_{i} g \xi_{i}, \text { with } g \in G .
$$

### 6.4.2 Accessibility and Controllability

A trajectory of a left-invariant control system $\Gamma$ on $G$ is a continuous curve $g(t)$ in $G$ defined in an interval $\left[t_{0}, T\right] \subset \mathbb{R}$ such that there exists a partition

$$
t_{0}<t_{1}<\ldots<t_{N}=T
$$

and elements,

$$
\xi_{1}, \ldots, \xi_{N} \in \Gamma
$$

such that the restriction of $g(t)$ in each open interval $\left(t_{i-1}, t_{i}\right)$ is smooth and

$$
\dot{g}(t)=g(t) \xi_{i}, \quad i=1, \ldots, N .
$$

Definition 6.4.1. For each $T \geq 0$ and $g \in G$, the reachable set at time $T$ from $g$ of the left-invariant control system $\Gamma \in \mathfrak{g}$ is the set

$$
\mathcal{A}_{\Gamma}(g, T)=\{g(T) \mid g(t) \text { trajectory of } \Gamma, g(0)=g\}
$$

Definition 6.4.2. The reachable set at time less than or equal to $T$ is defined by

$$
\mathcal{A}_{\Gamma}(g, \leq T)=\bigcup_{0 \leq t \leq T} \mathcal{A}_{\Gamma}(g, t) .
$$

Definition 6.4.3. The reachable set is defined as

$$
\mathcal{A}_{\Gamma}(g)=\bigcup \bigcup_{T \geq 0} \mathcal{A}_{\Gamma}(g, T)
$$

Definition 6.4.4. The system $\Gamma \subset \mathfrak{g}$ is controllable if for each pair of points $g_{0}$ and $g_{1}$ in G:

$$
g_{1} \in \mathcal{A}_{\Gamma}\left(g_{0}\right)
$$

### 6.4.3 Properties of Reachable Sets

First, we observe that for all $\xi \in \mathfrak{g}$ and $g_{0} \in G$ the Cauchy problem

$$
\dot{g}=g \xi, \quad g\left(t_{0}\right)=g_{0}
$$

has a solution $g(t)=g_{0} \exp \left(\left(t-t_{0}\right) \xi\right)$.
Lemma 6.4.5. Let $g(t), t \in\left[t_{0}, T\right]$ be a trajectory of a left-invariant control system $\Gamma \subset \mathfrak{g}$ with $g(0)=g_{0}$. Then there exists $N \in \mathbb{N}$ and

$$
\tau_{1}, \ldots, \tau_{N}>0, \quad \xi_{1}, \ldots, \xi_{N} \in \mathfrak{g}
$$

such that

$$
g(T)=g_{0} \exp \left(\tau_{1} \xi_{1}\right) \cdots \exp \left(\tau_{N} \xi_{N}\right)
$$

and $T-t_{0}=\tau_{1}+\ldots+\tau_{N}$.
Proof:
By definition of trajectory, there exists a partition

$$
t_{0}<t_{1}<\ldots<t_{N}=T
$$

and elements,

$$
\xi_{1}, \ldots, \xi_{N} \in \Gamma
$$

such that the restriction of $g(t)$ in each open subset $\left(t_{i-1}, t_{i}\right)$ is differentiable and

$$
\dot{g}(t)=g(t) \xi_{i}, \quad i=1, \ldots, N .
$$

In the first interval, we have that

$$
t \in\left(t_{0}, t_{1}\right), \quad \dot{g}=g \xi_{1}, \quad g\left(t_{0}\right)=g_{0}
$$

Therefore,

$$
g(t)=g_{0} \exp \left(\left(t-t_{0}\right) \xi_{1}\right), \quad g\left(t_{1}\right)=g_{0} \exp \left(\left(t_{1}-t_{0}\right) \xi_{1}\right),
$$

The next interval, $\left(t_{1}, t_{2}\right)$ :

$$
\dot{g}=g \xi_{2}, \quad g\left(t_{1}\right)=g_{0} \exp \left(\left(t_{1}-t_{0}\right) \xi_{1}\right)
$$

then

$$
g(t)=g_{0} \exp \left(\left(t_{1}-t_{0}\right) \xi_{1}\right) \exp \left(\left(t-t_{1}\right) \xi_{2}\right),
$$

and

$$
g\left(t_{2}\right)=g_{0} \exp \left(\left(t_{1}-t_{0}\right) \xi_{1}\right) \exp \left(\left(t_{2}-t_{1}\right) \xi_{2}\right),
$$

If we denote by $\tau_{1}=t_{1}-t_{0}$ and $\tau_{2}=t_{2}-t_{0}$ we have that

$$
g\left(t_{2}\right)=g_{0} \exp \left(\tau_{1} \xi_{1}\right) \exp \left(\tau_{2} \xi_{2}\right)
$$

Proceeding in the same way,

$$
g\left(t_{N}\right)=g(T)=g_{0} \exp \left(\tau_{1} \xi_{1}\right) \cdots \exp \left(\tau_{N} \xi_{N}\right)
$$

with $\tau_{1}=t_{i}-t_{i-1}, i=1, \ldots, N$ y $\tau_{N}+\ldots+\tau_{1}=T$.
From this lemma, we deduce that

- $\mathcal{A}_{\Gamma}(g)=\left\{g \exp \left(t_{1} \xi_{1}\right) \cdots \exp \left(t_{N} \xi_{N}\right) \mid \xi_{i} \in \Gamma, t_{i}>0, N \geq 0\right\}$
- $\mathcal{A}_{\Gamma}(g)=g \mathcal{A}_{\Gamma}(e)$
- $\mathcal{A}_{\Gamma}(g)$ is simply connected .

Then, the control system $\Gamma$ is controllable if and only if $\mathcal{A}_{\Gamma}(e)=G$.
Definition 6.4.6. The orbit of $\Gamma$ in a point $g \in G$ is the set

$$
\mathcal{O}_{\Gamma}(g)=\left\{g \exp \left(t_{1} \xi_{1}\right) \cdots \exp \left(t_{N} \xi_{N}\right) \mid \xi_{i} \in \Gamma, t_{i} \in \mathbb{R}, N \geq 0\right\}
$$

Obviously,

$$
\mathcal{O}_{\Gamma}(g)=g \mathcal{O}_{\Gamma}(e)
$$

Denote by Lie $\Gamma$ the Lie algebra generated by $\Gamma$.
Theorem 6.4.7 (Hermann-Nagano theorem). It is verified by $\mathcal{O}_{\Gamma}(\boldsymbol{I}) \subset G$ is a differentiable submanifold of $G$ with tangent space $T_{e} \mathcal{O}_{\Gamma}(e)=L i e \Gamma$.

Therefore, is easy to deduce that $\mathcal{O}_{\Gamma}(e)$ is a Lie subgroup of $G$ with Lie algebra Lie $\Gamma$.
Theorem 6.4.8. A control system $\Gamma \subset \mathfrak{g}$ is controllable if and only if

1. $G$ is connected;
2. $\operatorname{Lie} \Gamma=\mathfrak{g}$;
3. $\mathcal{A}_{\Gamma}(e)$ is a Lie subgroup of $G$

Proof:
The necessary condition is trivial. Then we proof the sufficient condition.
If $\mathcal{A}_{\Gamma}(e)$ is a subgroup of $G$, then for all $g \in \mathcal{A}_{\Gamma}(e)$ we have that $g^{-1} \in \mathcal{A}_{\Gamma}(e)$
Since $\left(\exp \left(t_{i} \xi_{i}\right)\right)^{-1}=\exp \left(-t_{i} \xi_{i}\right)$ we deduce that $\mathcal{A}_{\Gamma}(e)=\mathcal{O}_{\Gamma}(e)$. Therefore, $\mathcal{O}_{\Gamma}(e)$ is a connected Lie subgroup with Lie algebra $\mathfrak{g}$. Therefore $\mathcal{A}_{\Gamma}(e)=G$.

### 6.4.4 Optimal Control of the Position of a Rigid Body

Motivated by the control of a satellite control, we will study the maneuver problem of a rigid body which moves from a initial position to a desired final position at a fixed time and minimizing a cost function [66], [56].

Working directly in $S O(3)$ avoids problems that appear when we work with Euler angles (singularities and ambiguities in the description).

Control Equations of the Rigid Body: Recall that the equations of motion of a rigid body are

$$
\begin{aligned}
\dot{g}(t) & =g(t) \widehat{\Omega(t)^{\prime}} \\
& =g(t)\left(\begin{array}{lll}
0 & -\Omega_{3}(t) & \Omega_{2}(t) \\
\Omega_{3}(t) & 0 & -\Omega_{1}(t) \\
-\Omega_{2}(t) & \Omega_{1}(t) & 0
\end{array}\right) \\
& =g(t)\left(\Omega_{1}(t) E_{1}+\omega_{2}(t) E_{2}+\Omega_{3}(t) E_{3}\right)
\end{aligned}
$$

where the angular velocities ( $\left.\Omega_{1}(t), \Omega_{2}(t), \Omega_{3}(t)\right)$ verify the Euler's equations

$$
\begin{align*}
I_{1} \dot{\Omega}_{1}(t) & =\left(I_{2}-I_{3}\right) \Omega_{2}(t) \Omega_{3}(t)+T_{1}(t) \\
I_{2} \dot{\Omega}_{2}(t) & =\left(I_{3}-I_{1}\right) \Omega_{3}(t) \Omega_{1}(t)+T_{2}(t)  \tag{6.4.1}\\
I_{3} \dot{\Omega}_{3}(t) & =\left(I_{1}-I_{2}\right) \Omega_{1}(t) \Omega_{2}(t)+T_{3}(t)
\end{align*}
$$

where $I_{1}, I_{2}, I_{3}$ are the inertia moments and $T_{1}, T_{2}$ and $T_{3}$ the torques applied to the body which allow us to move and control.

Suppose that we want to pass of the matrix $g\left(t_{0}\right)$ in a time $t_{0}$ to a matrix $g\left(t_{1}\right)$ in a time $t_{1}$. The cost functional is:

$$
\int_{t_{0}}^{t_{1}}\left(c_{1} \Omega_{1}(t)^{p_{1}}+c_{2} \Omega_{2}(t)^{p_{2}}+c_{3} \Omega_{3}(t)^{p_{3}}\right) d t
$$

with $c_{1}, c_{2}, c_{3}>0$ and $p_{1}, p_{2}, p_{3} \in \mathbb{N}$.
In our problem, we choose angular velocities $\left(\Omega_{1}, \Omega_{2}, \Omega_{3}\right)$ as controls, instead of torques $\left(T_{1}, T_{2}, T_{3}\right)$. Using the equation (6.4.1) we will obtain the torques from the angular velocities.

Then, the optimal control problem is

$$
\min \int_{t_{0}}^{t_{1}}\left(c_{1} \Omega_{1}(t)^{p_{1}}+c_{2} \Omega_{2}(t)^{p_{2}}+c_{3} \Omega_{3}(t)^{p_{3}}\right) d t
$$

with control equation

$$
\dot{g}(t)=g(t)\left(\Omega_{1}(t) E_{1}+\omega_{2}(t) E_{2}+\Omega_{3}(t) E_{3}\right)
$$

and $g\left(t_{0}\right)=g_{0}$ and $g\left(t_{1}\right)=g_{1}$ fixed.
The Pontryagin Hamiltonian will be a function

$$
H: T^{*} S O(3) \times \mathbb{R}^{3} \longrightarrow \mathbb{R}
$$

defined by

$$
H=p_{0}\left(\left(c_{1} \Omega_{1}^{p_{1}}+c_{2} \Omega_{2}^{p_{2}}+c_{3} \Omega_{3}^{p_{3}}\right)+\Omega_{1} H_{1}+\Omega_{2} H_{2}+\Omega_{3} H_{3}\right.
$$

where $H_{i}: T^{*} S O(3) \longrightarrow \mathbb{R}$ is defined by $H_{i}(g, \alpha)=\alpha\left(E_{i}\right), i=1,2,3$.
Suppose that the optimal control is given by the functions $t \longmapsto\left(\Omega_{1}^{*}(t), \Omega_{2}^{*}(t), \Omega_{3}^{*}(t)\right)$ then the trajectory $t \longmapsto \xi^{*}(t) \in T^{*} S O(3)$ verify the maximum principle that

$$
\begin{equation*}
\dot{\xi}^{*}(t)=\Omega_{1}^{*}(t) X_{H_{1}}\left(\xi^{*}(t)\right)+\Omega_{2}^{*}(t) X_{H_{2}}\left(\xi^{*}(t)\right)+\Omega_{3}^{*}(t) X_{H_{3}}\left(\xi^{*}(t)\right) \tag{6.4.2}
\end{equation*}
$$

## Studying the abnormal solutions [56]

Suppose that $p_{0}=0$ and there exits an optimal solution. Then since,

$$
H\left(\xi^{*}(t), \Omega^{*}(t)\right) \geq H\left(\xi^{*}(t), \Omega\right), \quad \text { para todo } \quad \Omega \in \mathbb{R}^{3}
$$

that is,

$$
\begin{aligned}
& \Omega_{1}^{*}(t) H_{1}\left(\xi^{*}(t)\right)+\Omega_{2}^{*}(t) H_{2}\left(\xi^{*}(t)\right)+\Omega_{3}^{*}(t) H_{3}\left(\xi^{*}(t)\right) \\
& \geq \Omega_{1} H_{1}\left(\xi^{*}(t)\right)+\Omega_{2} H_{2}\left(\xi^{*}(t)\right)+\Omega_{3} H_{3}\left(\xi^{*}(t)\right)
\end{aligned}
$$

for all $\Omega=\left(\Omega_{1}, \Omega_{2}, \Omega_{3}\right) \in \mathbb{R}^{3}$.
From this, we deduce that $H_{i}\left(\xi^{*}(t)\right)=0$. If we denote by $\left(g^{*}(t), \alpha^{*}(t)\right) \in S O(3) \times \mathfrak{s o}(3)^{*}$ the trajectory such that $\mathcal{L}\left(g^{*}(t), \alpha^{*}(t)\right)=\xi^{*}(t)$ then

$$
0=H_{i}\left(\xi^{*}(t)\right)=\alpha^{*}(t)\left(E_{i}\right), \quad i=1,2,3
$$

then $\alpha^{*}(t)=0$ and we have a contradiction.
Therefore, we don't have abnormal solutions.

## Studying the regular solutions [56]

Take $p_{0}=-1$.
We start from the equations $\frac{\partial H}{\partial \Omega_{i}}\left(\xi^{*}(t)\right)=0, i=1,2,3$, that is

$$
H_{i}\left(\xi^{*}(t)\right)=c_{i} p_{i} \Omega_{i}^{*}(t)^{q_{i}}
$$

where $q_{i}=p_{i}-1$. Then clear the controls in the equation (6.4.2) we obtain that
$\dot{\xi}^{*}(t)=\left[\frac{H_{1}\left(\xi^{*}(t)\right)}{c_{1} p_{1}}\right]^{1 / q_{1}} X_{H_{1}}\left(\xi^{*}(t)\right)+\left[\frac{H_{2}\left(\xi^{*}(t)\right)}{c_{2} p_{2}}\right]^{1 / q_{2}} X_{H_{2}}\left(\xi^{*}(t)\right)+\left[\frac{H_{3}\left(\xi^{*}(t)\right)}{c_{3} p_{3}}\right]^{1 / q_{3}} X_{H_{3}}\left(\xi^{*}(t)\right)$
which are the Hamilton equations for the Hamiltonian function

$$
H_{0}=\sum_{j=1}^{3} \frac{q_{j}}{p_{j}\left(c_{j} p_{j}\right)^{1 / q_{j}}} H_{j}^{p_{j} / q_{j}}
$$

Trivializing we obtain that the Hamiltonian in $S O(3) \times \mathbb{R}^{3}$ is written as:

$$
H_{0}(g, \Pi)=\sum_{j=1}^{3} \frac{q_{j}}{p_{j}\left(c_{j} p_{j}\right)^{1 / q_{j}}} \Pi_{j}^{p_{j} / q_{j}}
$$

Since

$$
\frac{\partial H}{\partial \Pi_{j}}=\frac{1}{\left(c_{j} p_{j}\right)^{1 / q_{j}}} \Pi_{j}^{1 / q_{j}}
$$

the Euler-Arnold's equations are:

$$
\left(\dot{\Pi}_{1}, \dot{\Pi}_{2}, \dot{\Pi}_{3}\right)=\left(\Pi_{1}, \Pi_{2}, \Pi_{3}\right) \times\left(\frac{1}{\left(c_{1} p_{1}\right)^{1 / q_{1}}} \Pi_{1}^{1 / q_{1}}, \frac{1}{\left(c_{2} p_{2}\right)^{1 / q_{2}}} \Pi_{2}^{1 / q_{2}}, \frac{1}{\left(c_{3} p_{3}\right)^{1 / q_{3}}} \Pi_{3}^{1 / q_{3}}\right)
$$

Computing and rewritting the last equation we obtain that

$$
\begin{aligned}
& c_{1} p_{1} q_{1} \Omega_{1}^{q_{1}-1} \dot{\Omega}_{1}=c_{2} p_{2} \Omega_{3} \Omega_{2}^{q_{2}}-c_{3} p_{3} \Omega_{2} \Omega_{3}^{q_{3}} \\
& c_{2} p_{2} q_{2} \Omega_{2}^{q_{2}-1} \dot{\Omega}_{2}=c_{3} p_{3} \Omega_{1} \Omega_{3}^{q_{3}}-c_{1} p_{1} \Omega_{3} \Omega_{1}^{q_{1}} \\
& c_{3} p_{3} q_{3} \Omega_{3}^{q_{3}-1} \dot{\Omega}_{3}=c_{1} p_{1} \Omega_{2} \Omega_{1}^{q_{1}}-c_{2} p_{2} \Omega_{1} \Omega_{2}^{q_{2}}
\end{aligned}
$$

Observe that when $p_{1}=p_{2}=p_{3}=2$ we have to minimize

$$
\min \int_{t_{0}}^{t_{1}}\left(c_{1} \Omega_{1}^{2}(t)+c_{2} \Omega_{2}^{2}(t)+c_{3} \Omega_{3}^{2}(t)\right) d t
$$

and obtain the Euler's equations for the rigid body

$$
\begin{align*}
c_{1} \dot{\Omega}_{1}(t) & =\left(c_{2}-c_{3}\right) \Omega_{2}(t) \Omega_{3}(t) \\
c_{2} \dot{\Omega}_{2}(t) & =\left(c_{3}-c_{1}\right) \Omega_{3}(t) \Omega_{1}(t)  \tag{6.4.3}\\
c_{3} \dot{\Omega}_{3}(t) & =\left(c_{1}-c_{2}\right) \Omega_{1}(t) \Omega_{2}(t)
\end{align*}
$$

Finally, when $c_{1}=c_{2}=c_{3}$, we obtain that the unique solutions are $\Omega(t)=\Omega_{1}=$ constant. Then we wish to find a curve $g(t)$ verifying

$$
\dot{g}(t)=g(t)\left(\Omega_{1} E_{1}+\Omega_{2} E_{2}+\Omega_{3} E_{3} \quad g\left(t_{0}\right), g\left(t_{1}\right)\right. \text { are fixed. }
$$

That is,

$$
g(t)=g\left(t_{0}\right) e^{\left(t-t_{0}\right) \hat{\Omega}}
$$

with

$$
g\left(t_{1}\right)=g\left(t_{0}\right) e^{\left(t_{1}-t_{0}\right) \hat{\Omega}}
$$

Therefore

$$
g^{-1}\left(t_{0}\right) g\left(t_{1}\right)=e^{\left(t_{1}-t_{0}\right) \hat{\Omega}}
$$

Denoting by $r=\sqrt{\Omega_{1}^{2}+\Omega_{2}^{2}+\Omega_{3}^{2}}$ we have that

$$
\begin{aligned}
g\left(t_{1}\right) g\left(t_{0}\right)^{-1}= & I+\frac{\sin \left(t_{1}-t_{0}\right) r}{\left(t_{1}-t_{0}\right) r} \hat{\Omega}+\frac{1-\cos \left(t_{1}-t_{0}\right) r}{\left(t_{1}-t_{0}\right)^{2} r^{2}} \\
= & \left(\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{array}\right)+\frac{\sin \left(t_{1}-t_{0}\right) r}{\left(t_{1}-t_{0}\right) r}\left(\begin{array}{ccc}
0 & -\Omega_{3} & \Omega_{2} \\
\Omega_{3} & 0 & -\Omega_{1} \\
-\Omega_{2} & \Omega_{1} & 0
\end{array}\right) \\
& -\frac{1-\cos \left(t_{1}-t_{0}\right) r}{\left(t_{1}-t_{0}\right)^{2} r^{2}}\left(\begin{array}{ccc}
\Omega_{2}^{2}+\Omega_{3}^{2} & -\Omega_{1} \Omega_{2} & -\Omega_{1} \Omega_{3} \\
-\Omega_{1} \Omega_{2} & \Omega_{1}^{2}+\Omega_{3}^{2} & -\Omega_{2} \Omega_{3} \\
-\Omega_{1} \Omega_{3} & -\Omega_{2} \Omega_{3} & \Omega_{1}^{2}+\Omega_{2}^{2}
\end{array}\right)
\end{aligned}
$$

which can be solved numerically.

## Chapter 7

## Optimal Control of Underactuated Mechanical Systems

The class of underactuated mechanical systems are abundant in real life for different reasons, for instance, as a result of design choices motivated by the search of less cost engineering devices or as a result of a failure regime in fully actuated mechanical systems. The underactuated systems include spacecraft, underwater vehicles, mobile robots, helicopters, wheeled vehicles, mobile robots, underactuated manipulators...

On the other hand, there are many papers in which optimal control problems are addressed using geometric techniques (see, for instance, [13, 42, 43, 68] and references therein). Now, we introduce an optimization strategy in an underactuated mechanical system, that is, we are interested in studying the implementation of devices in which a controlled quantity is used to influence the behavior of the undeactuated system in order to achieve a desired goal (control) using the most economical strategy (optimization). Thus, in this section we develop a new geometric setting for optimal control of underatuated Lagrangian systems strongly inspired on the Skinner and Rusk formulation for singular Lagrangians systems [65]. Since in this setting the controlled Euler-Lagrange equation are second-order differential equations we will need to implement an higher-order version of this classical Skinner and Rusk formalism [15]. This geometric procedure gives us an intrinsic version of the differential equations for optimal trajectories and permits us to detect the preservation of geometric properties (symplecticity, preservation of the hamiltonian, etc.). For expository simplicity, we restrict ourselves in Section 7.1 to the so-called optimal control of superarticulated mechanical systems, in which only some of the degrees of freedom are controlled directly, with the remaining variables freely evolving subject only to dynamic interactions with the actuated degrees of freedom (see [3, 67]). Obviously, our theory can be easily extended to more general class of underactuated Lagrangian systems.

### 7.1 Optimal Control of underactuated mechanical systems

After introducing the geometry of higher-order Lagrangian system with constraints in the previous chapters, we may turn to the geometric framework for optimal control of underac-
tuated mechanical systems. We recall that a Lagrangian control system is underactuated if the number of the control inputs is less than the dimension of the configuration space. We assume, in the sequel, that the considered systems are controllable.

Consider the class of underactuated Lagrangian control system (superarticulated mechanical system following the nomenclature by [3]) where the configuration space $Q$ is the cartesian product of two differentiable manifolds, $Q=Q_{1} \times Q_{2}$. Denote by $\left(q^{A}\right)=\left(q^{a}, q^{\alpha}\right)$, $1 \leq A \leq n$, local coordinates on $Q$ where $\left(q^{a}\right), 1 \leq a \leq r$ and $\left(q^{\alpha}\right), r+1 \leq \alpha \leq n$, are local coordinates on $Q_{1}$ and $Q_{2}$, respectively.

Given a Lagrangian $L: T Q \equiv T Q_{1} \times T Q_{2} \rightarrow \mathbb{R}$, we assume that the controlled external forces can be applied only to the coordinates $\left(q^{a}\right)$. Thus, the equations of motion are given by

$$
\begin{align*}
& \frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{a}}\right)-\frac{\partial L}{\partial q^{a}}=u^{a} \\
& \frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{\alpha}}\right)-\frac{\partial L}{\partial q^{\alpha}}=0 \tag{7.1.1}
\end{align*}
$$

where $a=1, \ldots, r$, and $\alpha=r+1, \ldots, n$.
We study the optimal control problem that consists on finding a trajectory $\left(q^{a}(t), q^{\alpha}(t), u^{a}(t)\right)$ of state variables and control inputs satisfying equations (7.1.1) from given initial and final conditions, $\left(q^{a}\left(t_{0}\right), q^{\alpha}\left(t_{0}\right), \dot{q}^{a}\left(t_{0}\right), \dot{q}^{\alpha}\left(t_{0}\right)\right),\left(q^{a}\left(t_{f}\right), q^{\alpha}\left(t_{f}\right), \dot{q}^{a}\left(t_{f}\right), \dot{q}^{\alpha}\left(t_{f}\right)\right)$ respectively, minimizing the cost functional

$$
\mathcal{A}=\int_{t_{0}}^{t_{f}} C\left(q^{a}, q^{\alpha}, \dot{q}^{a}, \dot{q}^{\alpha}, u^{a}\right) d t
$$

This optimal control problem is equivalent to the following constrained variational problem:

Extremize

$$
\widetilde{\mathcal{A}}=\int_{t_{0}}^{t_{f}} \widetilde{L}\left(q^{a}(t), q^{\alpha}(t), \dot{q}^{a}(t), \dot{q}^{\alpha}(t), \ddot{q}^{a}(t), \ddot{q}^{\alpha}(t)\right) d t
$$

subject to the second order constraints given by

$$
\Phi^{\alpha}\left(q^{a}, q^{\alpha}, \dot{q}^{a}, \dot{q}^{\alpha}, \ddot{q}^{a}, \ddot{q}^{\alpha}\right)=\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{\alpha}}\right)-\frac{\partial L}{\partial q^{\alpha}}=0
$$

and the boundary conditions, where $\widetilde{L}: T^{(2)} Q \rightarrow \mathbb{R}$ is defined as

$$
\widetilde{L}\left(q^{a}, q^{\alpha}, \dot{q}^{a}, \dot{q}^{\alpha}, \ddot{q}^{a}, \ddot{q}^{\alpha}\right)=C\left(q^{a}, q^{\alpha}, \dot{q}^{a}, \dot{q}^{\alpha}, \frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{a}}\right)-\frac{\partial L}{\partial q^{a}}\right) .
$$

Now, according to the formulation given in Section 4. the dynamics of this second order constrained variational problem is determined by the solution of a presymplectic Hamiltonian system. In the following we repeat some of the constructions given in 4 but specialized to this particular setting, obtaining new insights for the optimal control problem under study.

If $\mathcal{M} \subset T^{(2)} Q$ is the submanifold given by annihilation of the functions $\Phi^{\alpha}$, we will see how to define local coordinates on $\mathcal{M}$.

From the constraint equations we have

$$
\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{\alpha}}\right)-\frac{\partial L}{\partial q^{\alpha}}=0 \Longleftrightarrow \frac{\partial^{2} L}{\partial \dot{q}^{\beta} \partial \dot{q}^{\alpha}} \ddot{q}^{\beta}=F_{\alpha}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)
$$

Let us assume that the matrix $\left(W_{\alpha \beta}\right)=\left(\frac{\partial^{2} L}{\partial \dot{q}^{\alpha} \partial \dot{q}^{\beta}}\right)$ is non-singular and denote by $\left(W^{\alpha \beta}\right)$ its inverse. Thus,

$$
\ddot{q}^{\alpha}=W^{\alpha \beta} F_{\alpha}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)=G^{\alpha}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right) .
$$

Therefore, we can consider $\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)$ as a system of local coordinates on $\mathcal{M}$. The canonical inclusion $i_{\mathcal{M}}: \mathcal{M} \hookrightarrow T T Q$ can be written as

$$
\begin{aligned}
\mathcal{M} & \stackrel{i \mathbb{M}}{\rightarrow} T T Q \\
\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right) & \mapsto\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}, G^{\alpha}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)\right) .
\end{aligned}
$$

Define the restricted lagrangian $\left.\widetilde{L}\right|_{\mathcal{M}}: \mathcal{M} \rightarrow \mathbb{R}$.


Figure 7.1: Second order Skinner and Rusk formalism
We will consider $W_{0}=T^{*}(T Q) \times_{T Q} \mathcal{M}$ whose coordinates are $\left(q^{A}, \dot{q}^{A} ; p_{A}^{0}, p_{A}^{1}, \ddot{q}^{a}\right)$.
Let us define the 2 -form $\Omega_{W_{0}}=\pi_{1}^{*}\left(\omega_{T Q}\right)$ on $W_{0}$ and $H_{W_{0}}\left(\alpha_{x}, v_{x}\right)=\left\langle\alpha_{x}, i_{\mathcal{M}}\left(v_{x}\right)\right\rangle-\widetilde{L}_{\mathcal{M}}\left(v_{x}\right)$ where $x \in T Q, v_{x} \in \mathcal{M}_{x}=\left(\left.\left(\tau_{Q}^{(1,2)}\right)\right|_{\mathcal{M}}\right)^{-1}(x)$ and $\alpha_{x} \in T_{x}^{*} T Q$. In local coordinates,

$$
\begin{aligned}
& \Omega_{W_{0}}=d q^{A} \wedge d p_{A}^{0}+d \dot{q}^{A} \wedge d p_{A}^{1} \\
& H_{W_{0}}=p_{A}^{0} \dot{q}^{A}+p_{a}^{1} \ddot{q}^{a}+p_{\alpha}^{1} G^{\alpha}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)-\widetilde{L}_{\mathcal{M}}\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}\right)
\end{aligned}
$$

The dynamics of this variational constrained problem is determined by the solution of the equation

$$
\begin{equation*}
i_{X} \Omega_{W_{0}}=d H_{W_{0}} \tag{7.1.2}
\end{equation*}
$$

It is clear that $\Omega_{W_{0}}$ is a presymplectic form on $W_{0}$ and locally

$$
\operatorname{ker} \Omega_{W_{0}}=\operatorname{span}\left\langle\frac{\partial}{\partial \ddot{q}^{a}}\right\rangle
$$

Following the Gotay-Nester-Hinds algorithm we obtain the primary constraints

$$
d H_{W_{0}}\left(\frac{\partial}{\partial \ddot{q}^{a}}\right)=0 .
$$

That is,

$$
\varphi_{a}^{1}=\frac{\partial H_{W_{0}}}{\partial \ddot{q}^{a}}=p_{a}^{1}+p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}}=0
$$

These new constraints $\varphi_{a}^{1}=0$ give rise to a submanifold $W_{1}$ of dimension $4 n$ with local coordinates $\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}, p_{A}^{0}, p_{\alpha}^{1}\right)$.

Consider a solution curve $\left(q^{A}(t), \dot{q}^{A}(t), \ddot{q}^{a}(t), p_{A}^{0}(t), p_{A}^{1}(t)\right)$ of Equation (7.1.2). Then, this curve satisfies the following system of differential equations

$$
\begin{align*}
\frac{d q^{A}}{d t} & =\dot{q}^{A}, \quad \frac{d^{2} q^{a}}{d t^{2}}=\ddot{q}^{a}  \tag{7.1.3}\\
\frac{d^{2} q^{\alpha}}{d t^{2}} & =G^{\alpha}\left(q^{A}, \frac{d q^{A}}{d t}, \frac{d^{2} q^{a}}{d t^{2}}\right)  \tag{7.1.4}\\
\frac{d p_{A}^{0}}{d t} & =-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial q^{A}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{A}}  \tag{7.1.5}\\
\frac{d p_{A}^{1}}{d t} & =-p_{A}^{0}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \dot{q}^{A}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{A}}  \tag{7.1.6}\\
p_{a}^{1} & =-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}} . \tag{7.1.7}
\end{align*}
$$

From Equations (7.2.5) and 7.2.6 we deduce

$$
\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}\right)=-p_{a}^{0}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \dot{q}^{a}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{a}} .
$$

Differentiating with respect to time, replacing in the previous equality and using (7.2.4) we obtain the following system of 4 -order differential equations

$$
\begin{equation*}
\frac{d^{2}}{d t^{2}}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}\right)-\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \dot{q}^{a}}\right)+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial q^{a}}=0 \tag{7.1.8}
\end{equation*}
$$

Also, using (7.2.4 and 7.2.5 we deduce

$$
\begin{equation*}
\frac{d^{2} p_{\alpha}^{1}}{d t^{2}}=\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{\alpha}}-p_{\beta}^{1} \frac{\partial G^{\beta}}{\partial \dot{q}^{\alpha}}\right)-\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{\alpha}}-p_{\beta}^{1} \frac{\partial G^{\beta}}{\partial q^{\alpha}}\right) \tag{7.1.9}
\end{equation*}
$$

If we solve the implicit system of differential equations given by (7.1.8) and (7.1.9) then from Equations 7.2.5 and 7.2.6 we deduce that the values of $p_{a}^{0}$ and $p_{\alpha}^{0}$ are

$$
\begin{align*}
p_{a}^{0} & =\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \dot{q}^{a}}-\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}}-p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}\right)  \tag{7.1.10}\\
p_{\alpha}^{0} & =\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{\alpha}}-p_{\beta}^{1} \frac{\partial G^{\beta}}{\partial \dot{q}^{\alpha}}-\frac{d p_{\alpha}^{1}}{d t} \tag{7.1.11}
\end{align*}
$$

Since, from our initial problem, we are only interested in the values $q^{A}(t)$, it is uniquely necessary to solve the coupled system of implicit differential equations given by (7.1.8), (7.1.9) and 7.2.3) without explicitly calculate the values $p_{a}^{0}(t)$.

Now, we are interested in the geometric properties of the dynamics. First, consider the submanifold $W_{1}$ of $W_{0}$ determined by

$$
W_{1}=\left\{x \in T^{*} T Q \times_{T Q} \mathcal{M} \mid d H_{W_{0}}(x)(V)=0 \forall V \in \operatorname{ker} \Omega(x)\right\}
$$

and the 2 -form $\Omega_{W_{1}}=i_{W_{1}}^{*} \Omega_{W_{0}}$, where $i_{W_{1}}: W_{1} \hookrightarrow W_{0}$ denotes the canonical inclusion. Locally, $W_{1}$ is determined by the vanishing of the constraint equations

$$
\varphi_{a}^{1}=p_{a}^{1}+p_{\alpha}^{1} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}}=0
$$

Therefore, we can consider local coordinates $\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}, p_{A}^{0}, p_{\alpha}^{1}\right)$ on $W_{1}$.
Proposition 7.1.1. $\left(W_{1}, \Omega_{W_{1}}\right)$ is symplectic if and only if for any choice of local coordinates $\left(q^{A}, \dot{q}^{A}, \ddot{q}^{a}, p_{A}^{0}, p_{A}^{1}\right)$ on $W_{0}$,

$$
\begin{equation*}
\operatorname{det}\left(\mathcal{R}_{a b}\right)=\operatorname{det}\left(\frac{\partial^{2} \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}}-p_{\alpha}^{1} \frac{\partial^{2} G^{\alpha}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}}\right)_{(n-r) \times(n-r)} \neq 0 \quad \text { along } W_{1} \tag{7.1.12}
\end{equation*}
$$

In the case where the matrix $(7.1 .12)$ is regular then the equations (7.1.8), (7.1.9) and (7.2.3) can be written as an explicit system of differential equations of the form

$$
\begin{align*}
\frac{d^{4} q^{a}}{d t^{4}} & =\Gamma^{a}\left(q^{A}, \frac{d q^{A}}{d t}, \frac{d^{2} q^{a}}{d t^{2}}, \frac{d^{3} q^{a}}{d t^{2}}, p_{\alpha}^{1}, \frac{d p_{\alpha}^{1}}{d t}\right)  \tag{7.1.13}\\
\frac{d^{2} q^{\alpha}}{d t^{2}} & =G^{\alpha}\left(q^{A}, \frac{d q^{A}}{d t}, \frac{d^{2} q^{a}}{d t^{2}}\right)  \tag{7.1.14}\\
\frac{d^{2} p_{\alpha}^{1}}{d t^{2}} & =\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^{\alpha}}-p_{\beta}^{1} \frac{\partial G^{\beta}}{\partial \dot{q}^{\alpha}}\right)-\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{\alpha}}-p_{\beta}^{1} \frac{\partial G^{\beta}}{\partial q^{\alpha}}\right) \tag{7.1.15}
\end{align*}
$$

Remark 7.1.2. The Pontryaguin's maximun principle gives us necessary conditions for optimality for an optimal control problem. In our case, we are analyzing a particular case of optimal control problem (an underactuated mechanical system) and under some regularity conditions, the necessary conditions of maximum principle are written in terms of expressions (7.2.9), (7.1.14) and (7.1.15), jointly with constraints. The dynamic evolution of the problem is determined as the integral curves of a unique vector field determined by the symplectic Hamiltonian equations:

$$
i_{X} \Omega_{W_{1}}=d H_{W_{1}}
$$

This is the case of a regular optimal control problem (4]. From (7.2.9), (7.1.14) and (7.1.15) we deduce a unique curve $\left(q^{A}(t)\right)$ (fixed appropriate initial conditions) which determine the controls from

$$
\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{a}}\right)-\frac{\partial L}{\partial q^{a}}=u^{a}
$$

Obviously, if the boundary conditions are given by a initial and final states then it is not guaranteed the existence and uniqueness of an optimal trajectory satisfying the transversallity conditions.

Remark 7.1.3. Now, we will analyze an alternative characterization of the condition (7.1.12) and its relationship with the matrix condition that appears in Theorem 5.3.1. Using the chain rule

$$
\begin{aligned}
\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a}} & =\frac{\partial \widetilde{L}}{\partial \ddot{q}^{a}}+\frac{\partial \widetilde{L}}{\partial \ddot{q}^{\alpha}} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}} \\
\frac{\partial^{2} \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}} & =\frac{\partial^{2} \widetilde{L}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}}+\frac{\partial^{2} \widetilde{L}}{\partial \ddot{q}^{a} \partial \ddot{q}^{\beta}} \frac{\partial G^{\beta}}{\partial \ddot{q}^{b}}+\frac{\partial^{2} \widetilde{L}}{\partial \ddot{q}^{\alpha} \partial \ddot{q}^{b}} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}}+\frac{\partial^{2} \widetilde{L}}{\partial \ddot{q}^{\alpha} \partial \ddot{q}^{\beta}} \frac{\partial G^{\alpha}}{\partial \ddot{q}^{a}} \frac{\partial G^{\beta}}{\partial \ddot{q}^{b}}+\frac{\partial^{2} \widetilde{L}}{\partial \ddot{q}^{\alpha}} \frac{\partial^{2} G^{\alpha}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}} .
\end{aligned}
$$

Define $W_{i j}=\left(\frac{\partial^{2} \tilde{L}}{\partial \dot{q}^{2} \partial \ddot{q}^{j}}\right)$, where $\Phi^{\alpha}=\ddot{q}^{\alpha}-G^{\alpha}$. Then we can write (7.1.12) as

$$
\mathcal{R}_{a b}=W_{a b}-W_{a \beta} \frac{\partial \Phi^{\beta}}{\partial \ddot{q}^{b}}-W_{\alpha b} \frac{\partial \Phi^{\alpha}}{\partial \ddot{q}^{a}}+W_{\alpha \beta} \frac{\partial \Phi^{\alpha}}{\partial \ddot{q}^{a}} \frac{\partial \Phi^{\beta}}{\partial \ddot{q}^{b}}+\left(p_{\alpha}^{1}-\frac{\partial \widetilde{L}}{\partial \dot{q}^{\alpha}}\right) \frac{\partial^{2} \Phi^{\alpha}}{\partial \ddot{q}^{a} \partial \ddot{q}^{b}} .
$$

Consider now the extended lagrangian $\mathcal{L}=\widetilde{L}-\lambda_{\alpha} \Phi^{\alpha}$ where $\lambda_{\alpha}=\frac{\partial \widetilde{L}}{\partial \dot{q}^{\alpha}}-p_{\alpha}^{1}$.
Then, the matrix $\left(\bar{W}_{i j}\right)=\left(\frac{\partial^{2} \mathcal{L}}{\partial \dot{q}^{\imath} \partial \dot{q}^{j}}\right)$ is equal, along $W_{1}$, to

$$
\bar{W}_{i j}=\left(\begin{array}{cc}
\bar{W}_{a b} & W_{a \beta}  \tag{7.1.16}\\
W_{\alpha b} & W_{\alpha \beta}
\end{array}\right)
$$

where $\bar{W}_{a b}=\frac{\partial^{2} \tilde{L}}{\partial \dot{q}^{a} \partial \tilde{q}^{b}}-\lambda_{\alpha} \frac{\partial^{2} \Phi^{\alpha}}{\partial \dot{q}^{a} \partial \dot{q}^{q}}$.
The elements of the matrix (7.1.12) are given by

$$
\begin{equation*}
\mathcal{R}_{a b}=\bar{W}_{a b}-\bar{W}_{a \beta} \frac{\partial \Phi^{\beta}}{\partial \ddot{q}^{b}}-\bar{W}_{\alpha b} \frac{\partial \Phi^{\alpha}}{\partial \ddot{q}^{a}}+\bar{W}_{\alpha \beta} \frac{\partial \Phi^{\alpha}}{\partial \ddot{q}^{a}} \frac{\partial \Phi^{\beta}}{\partial \ddot{q}^{b}} . \tag{7.1.17}
\end{equation*}
$$

Now, using elemental linear algebra the matrix (7.1.17) is regular if and only if the matrix of elements (7.1.16) is regular.

Remark 7.1.4. Condition (7.1.12) implies that the final constraint submanifold is $W_{1}$ and, moreover, there exists a unique vector field on $W_{1}$ determining the dynamics of our initial optimal control problem. Of course, this symplectic case is the more useful for many concrete applications. But it is possible to think in situations where the constraint algorithm does not stop in $W_{1}$ and it is necessary to find a proper subset of $W_{1}$ where there exists a well-defined solution of the problem. For instance, and as a trivial mathematical example, consider the following the system determined by the control equations $\ddot{x}=u_{1}, \ddot{y}=u_{2}$ and cost function $C\left(x, y, \dot{x}, \dot{y}, u_{1}, u_{2}\right)=\frac{1}{2}\left(u_{1}^{2}+2 u_{1} u_{2}+u_{2}^{2}\right)$. If we apply our techniques we deduce that $W_{1}$ is determined by the constraints

$$
p_{x}^{1}-(\ddot{x}+\ddot{y})=0, \quad p_{y}^{1}-(\ddot{x}+\ddot{y})=0 .
$$

But the solution of the dynamics is only consistently defined on the submanifold $W_{2}$ of $W_{1}$ determined by the additional (secondary) constraint

$$
p_{x}^{0}-p_{y}^{0}=0 .
$$

Example 7.1.5. Cart with Pendulum or Cart-Pole System (see [10] and references therein). A Cart-Pole System consists of a cart and an inverted pendulum on it. The coordinate $x$ denotes the position of the cart on the $x$-axis and $\theta$ denotes the angle of the pendulum with the upright vertical. The configuration space is $Q=\mathbb{R} \times \mathbb{S}^{1}$.

First, we describe the Lagrangian function describing this system. The inertia matrix of the cart-pole system is given by

$$
\begin{aligned}
m_{11} & =M+m \\
m_{12}\left(q_{2}\right) & =m l \cos (\theta) \\
m_{22} & =m l^{2}
\end{aligned}
$$

where $M$ is the mass of the cart and $m, l$ are the mass, and length of the center of mass of pendulum, respectively. The potential energy of the cart-pole system is $V(\theta)=m g l \cos (\theta)$.

The Lagrangian of the system (kinetic energy minus potential energy) is given by

$$
L(q, \dot{q})=L(x, \theta, \dot{x}, \dot{\theta})=\frac{1}{2} M \dot{x}^{2}+\frac{1}{2} m\left(\dot{x}^{2}+2 \dot{x} l \dot{\theta} \cos \theta+l^{2} \dot{\theta}^{2}\right)-m g l \cos \theta-m g \widetilde{h}
$$

where $\widetilde{h}$ is the car height.
The controller can apply a force $F$, the control input, parallel to the track remaining the joint angle $\theta$ unactuated. Therefore, the equations of motion of the controlled system are

$$
\begin{aligned}
(M+m) \ddot{x}-m l \dot{\theta}^{2} \sin \theta+m l \ddot{\theta} \cos \theta & =u \\
\ddot{x} \cos \theta+l \ddot{\theta}-g \sin \theta & =0
\end{aligned}
$$

Now we look for trajectories $(x(t), \theta(t), u(t))$ on the state variables and the controls inputs with initial and final conditions, $(x(0), \theta(0), \dot{x}(0), \dot{\theta}(0)),(x(T), \theta(T), \dot{x}(T), \dot{\theta}(T))$ respectively, and minimizing the cost functional

$$
\mathcal{A}=\frac{1}{2} \int_{0}^{T} u^{2} d t
$$

Following our formalism this optimal control problem is equivalent to the constrained second-order variational problem determined by

$$
\widetilde{\mathcal{A}}=\int_{0}^{T} \widetilde{L}(x, \theta, \dot{x}, \dot{\theta}, \ddot{x}, \ddot{\theta})
$$

and the second-order constraint

$$
\Phi(x, \theta, \dot{x}, \dot{\theta}, \ddot{x}, \ddot{\theta})=\ddot{x} \cos \theta+l \ddot{\theta}-g \sin \theta=0
$$

where

$$
\widetilde{L}(x, \theta, \dot{x}, \dot{\theta}, \ddot{x}, \ddot{\theta})=\frac{1}{2}\left(\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{x}}\right)-\frac{\partial L}{\partial x}\right)^{2}=\frac{1}{2}\left[(M+m) \ddot{x}-m l \dot{\theta}^{2} \sin \theta+m l \ddot{\theta} \cos \theta\right]^{2} .
$$

We rewrite the second-order constraint as

$$
\ddot{\theta}=\frac{g \sin \theta-\ddot{x} \cos \theta}{l} .
$$

Thus, the submanifold $\mathcal{M}$ of $T^{(2)}\left(\mathbb{R} \times \mathbb{S}^{1}\right)$ is given by

$$
\mathcal{M}=\{(x, \theta, \dot{x}, \dot{\theta}, \ddot{x}, \ddot{\theta}) \mid \ddot{x} \cos \theta+l \ddot{\theta}-g \sin \theta=0\} .
$$

Let us consider the submanifold $W_{0}=T^{*}\left(T\left(\mathbb{R} \times \mathbb{S}^{1}\right)\right) \times_{T\left(\mathbb{R} \times \mathbb{S}^{1}\right)} \mathcal{M}$ with induced coordinates $\left(x, \theta, \dot{x}, \dot{\theta} ; p_{x}^{0}, p_{\theta}^{0}, p_{x}^{1}, p_{\theta}^{1}, \ddot{x}\right)$.

Now, we consider the restriction of $\widetilde{L}$ to $\mathcal{N}$ given by

$$
\begin{gathered}
\left.\widetilde{L}\right|_{\mathcal{M}}=\frac{1}{2}\left[(M+m) \ddot{x}-m l \sin \theta \dot{\theta}^{2}+m l \cos \theta\left(\frac{g \sin \theta-\ddot{x} \cos \theta}{l}\right)\right]^{2} \\
=\frac{1}{2}\left[(M+m) \ddot{x}-m i \dot{\theta}^{2} \sin \theta+m g \cos \theta \sin \theta-m \ddot{x} \cos ^{2} \theta\right]^{2}
\end{gathered}
$$

For simplicity, denote by

$$
G^{\theta}=\frac{g \sin \theta-\ddot{x} \cos \theta}{l} .
$$

Now, the presymplectic 2-form $\Omega_{W_{0}}$, the Hamiltonian $H_{W_{0}}$ and the primary constraint $\varphi_{x}^{1}$ are, respectively

$$
\begin{aligned}
\Omega_{W_{0}}= & d x \wedge d p_{x}^{0}+d \theta \wedge d p_{\theta}^{0}+d \dot{x} \wedge d p_{x}^{1}+d \dot{\theta} \wedge d p_{\theta}^{0}, \\
H_{W_{0}}= & p_{x}^{0} \dot{x}+p_{\theta}^{0} \dot{\theta}+p_{x}^{1} \ddot{x}+p_{\theta}^{1}\left[\frac{g \sin \theta-\ddot{x} \cos \theta}{l}\right] \\
& -\frac{1}{2}\left[(M+m) \ddot{x}-m l \dot{\theta}^{2} \sin \theta+m g \cos \theta \sin \theta-m \ddot{x} \cos ^{2} \theta\right]^{2}, \\
\varphi_{x}^{1}= & \frac{\partial \widetilde{H}}{\partial \ddot{x}}=p_{x}^{1}+p_{\theta}^{1} \frac{\partial G^{\theta}}{\partial \ddot{x}}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{x}}=0,
\end{aligned}
$$

i.e.,

$$
p_{x}^{1}=-p_{\theta}^{1} \frac{\partial G^{\theta}}{\partial \ddot{x}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{x}} .
$$

This constraint determines the submanifold $W_{1}$. Applying Proposition 7.1.1 we deduce that the 2-form $\Omega_{W_{1}}$, restriction of $\Omega_{W_{0}}$ to $W_{1}$, is symplectic since

$$
\frac{\partial^{2} \widetilde{L}_{\mathcal{M}}}{\partial \ddot{x}^{2}}-p_{\theta}^{1} \frac{\partial^{2} G^{\theta}}{\partial \ddot{x}^{2}}=\left[(M+m)-m \cos ^{2} \theta\right]^{2} \neq 0
$$

Therefore, the algorithm stabilizes at the first constraint submanifold $W_{1}$. Moreover, there exists a unique solution of the dynamics, the vector field $X \in \mathfrak{X}\left(W_{1}\right)$ which satisfies $i_{X} \Omega_{W_{1}}=d H_{W_{1}}$. In consequence, we have a unique control input which extremizes (minimizes) the objective function $\mathcal{A}$ and then the force exerted to the car is the minimum possible. If we take the flow $F_{t}: W_{1} \rightarrow W_{1}$ of the vector field $X$ then we have that $F_{t}^{*} \Omega_{W_{1}}=\Omega_{W_{1}}$. Obviously, the Hamiltonian function

$$
\begin{aligned}
\left.\widetilde{H}\right|_{W_{1}}= & p_{x}^{0} \dot{x}+p_{\theta}^{0} \dot{\theta}+\left[-p_{\theta}^{1} \frac{\partial G^{\theta}}{\partial \ddot{x}}+\frac{\partial \widetilde{L}_{N}}{\partial \ddot{x}}\right] \ddot{x}+p_{\theta}^{1}\left[\frac{g \sin \theta-\ddot{x} \cos \theta}{l}\right]- \\
& \frac{1}{2}\left[(M+m) \ddot{x}-m l \sin \theta \dot{\theta}^{2}+m g \cos \theta \sin \theta-m \ddot{x} \cos ^{2} \theta\right]^{2}
\end{aligned}
$$

### 7.2. QUASIVELOCITIES AND OPTIMAL CONTROL OF UNDERACTUATED SYSTEMS101

is preserved by the solution of the optimal control problem, that is $\left.\widetilde{H}\right|_{W_{1}} \circ F_{t}=\left.\widetilde{H}\right|_{W_{1}}$. Both properties, symplecticity and preservation of energy, are important geometric invariants. In next section, we will construct, using discrete variational calculus, numerical integrators which inherit some of the geometric properties of the optimal control problem (symplecticity, momentum preservation and, in consequence, a very good energy behavior).

The resulting equations of the optimal dynamics of the cart-pole system are

$$
\begin{aligned}
\frac{d^{4} x}{d t^{4}}= & -\frac{1}{\left[(M+m)-m \cos ^{2} \theta\right]^{2}}\{[4 m \dot{\theta} \cos \theta \sin \theta] \times \\
& {\left[(M+m) \dddot{x}-m \dot{\theta}^{3} \cos \theta-2 m \sin \theta \dot{\theta}(g \sin \theta-\ddot{x} \cos \theta)+m g \dot{\theta} \cos (2 \theta)\right] } \\
+ & 2 m\left[(M+m) \ddot{x}-m l \dot{\theta}^{2} \sin \theta+m g \cos \theta \sin \theta-m \ddot{x} \cos ^{2} \theta\right] \times \\
& \times\left[\dot{\theta}^{2} \cos (f r m-e \theta)+\cos \theta \sin \theta \ddot{\theta}\right] \\
+ & \left.\frac{1}{l}\left(\frac{d^{2}}{d t^{2}} p_{\theta} 1 \cos \theta-2 \frac{d}{d t} p_{\theta}^{1} \dot{\theta} \sin \theta-p_{\theta} 1\left(\dot{\theta}^{2} \cos \theta+\ddot{\theta} \sin \theta\right)\right)\right\} \\
+ & \frac{1}{\left[(M+m)-m \cos ^{2} \theta\right]}\left\{\frac{2 m \sin \theta}{l}(g \dot{\theta} \cos \theta-\dddot{x} \cos \theta+\ddot{x} \dot{\theta} \sin \theta)^{2}\right. \\
+ & 2 m \dot{\theta} \sin \theta(g \dot{\theta} \cos \theta-\dddot{x} \cos \theta+\ddot{x} \dot{\theta} \sin \theta)+4 m \dot{\theta}^{2} \cos \theta(q \sin \theta-\ddot{x} \cos \theta) \\
- & m l \dot{\theta}^{4} \sin \theta-4 m g \sin \theta \cos \theta \dot{\theta}^{2}+\frac{m g}{l} \cos (2 \theta)(g \sin \theta-\ddot{x} \cos \theta) \\
- & \left.4 m \dddot{x} \dot{\theta} \cos \theta+2 m \ddot{x} \dot{\theta}^{2} \cos (2 \theta)+\frac{2 m}{l} \ddot{x} \cos \theta \sin \theta(g \sin \theta-\ddot{x} \cos \theta)\right\} \\
\frac{d^{2} \theta}{d t^{2}}= & \frac{1}{l}(g \sin \theta-\ddot{x} \cos \theta) \\
\frac{d^{2} p_{\theta}^{1}}{d t^{2}=}= & \left\{(M+m) \dddot{x}-m l\left(2 \dot{\theta} \sin \theta\left(\frac{g \sin \theta-\ddot{x} \cos \theta}{l}\right)+\dot{\theta}^{3} \cos \theta\right)+m g \dot{\theta} \cos (2 \theta)\right. \\
& \left.-m \dddot{x} \cos { }^{2} \theta+2 m \ddot{x} \dot{\theta} \cos \theta \sin \theta\right\} \\
& \times\left(-2 m l \dot{\theta} \sin \theta-m g \cos (2 \theta)+m l \dot{\theta}^{2} \cos \theta-2 m \ddot{x} \cos \theta \sin \theta\right) \\
+ & \left\{(M+m) \ddot{x}-m l \dot{\theta}^{2} \sin \theta+m g \cos \theta \sin \theta-m \ddot{x} \cos \theta\right\} \times \\
& \times\left(-2 m l\left(\frac{(g \sin \theta-\ddot{x} \cos \theta)}{l} \sin \theta+\dot{\theta}^{2} \cos \theta\right)+2 m \dot{\theta} \cos \theta(g \sin \theta-\ddot{x} \cos \theta)\right. \\
& \left.-m l \dot{\theta}^{3} \sin \theta-2 m \dddot{x} \cos \theta \sin \theta-2 m \ddot{x} \dot{\theta}+2 m g \dot{\theta} \sin \theta \cos \theta(1+m \ddot{x})\right) \\
+ & \frac{1}{l}\left[-p_{\theta}^{0}+(M+m) \ddot{x}-m l \sin \theta \dot{\theta}^{2}+m \cos \theta(g \sin \theta-\ddot{x} \cos \theta)\right](-2 m l \sin \theta) \\
& \left.+p_{\theta} 1(\dddot{x} \sin \theta+\ddot{x} \dot{\theta} \cos \theta-g \dot{\theta} \sin \theta)\right]
\end{aligned}
$$

### 7.2 Quasivelocities and Optimal Control of Underactuated Systems

Geometrically, quasivelocities are the components of velocities, describing a mechanical system, relative to a set of vector fields (in principle, local) that span on each point the
fibers of the tangent bundle of the configuration space. The main point is that these vector fields don't need to be associated with (local) configuration coordinates on the configuration space. In this section we will use quasivelocities as a tool to describe optimal control problem for underactuated mechanical systems.

### 7.2.1 Quasivelocities

Let $Q$ be a $n$ dimensional differentiable manifold, and $L: T Q \rightarrow \mathbb{R}$ a Lagrangian function determining the dynamics. Let $\left(q^{A}\right), 1 \leq A \leq n$, be local coordinates on $Q$ and choose a local basis of vector fields $X_{B}$ with $1 \leq B \leq n$, defined in the same coordinate neighborhood. The components of $X_{B}$ relative to the standard basis $\frac{\partial}{\partial q^{j}}$ will be denoted $X_{B}^{A}$, that is $X_{B}=X_{B}^{A}(q) \frac{\partial}{\partial q^{A}}$.

Let $\left(y^{1}, \ldots, y^{n}\right)$ (the quasivelocities) be the components of a velocity vector $v$ on $T Q$ relative to the basis $X_{B}$, then

$$
v=y^{B} X_{B}(q)=y^{B} X_{B}^{A}(q) \frac{\partial}{\partial q^{A}} .
$$

Therefore, $\dot{q}^{A}=y^{B} X_{B}^{A}(q)$, then

$$
L(q, \dot{q})=L\left(q, y^{B} X_{B}^{A}(q)\right):=l(q, y) .
$$

On $T Q$ we have induced coordinates $\left\{\left(q^{A}, y^{A}\right) \mid A=1, \ldots, n\right\}$.
The Lie bracket of the vector fields $X_{A}$ is $\left[X_{A}, X_{B}\right]=\mathcal{C}_{A B}^{D} X_{D}$, where $\mathcal{C}_{A B}^{D}$ are called Hamel's transpositional symbols or structure coefficients.

Given a Lagrangian function $L: T Q \rightarrow \mathbb{R}$, the Euler-Lagrange equations in quasivelocities or Hamel equations are

$$
\begin{aligned}
\dot{q}^{A} & =y^{B} X_{B}^{A}(q) \\
\frac{d}{d t}\left(\frac{\partial l}{\partial y^{A}}\right) & =\frac{\partial l}{\partial q^{B}} X_{B}^{A}-\mathfrak{C}_{A B}^{D} y^{B} \frac{\partial l}{\partial y^{D}}
\end{aligned}
$$

These equations were introduced by [38] (see also [59]). It is interesting to note that these equations admit a nice, useful and intrinsic interpretation in terms of mechanics on Lie algebroids (see 53].

### 7.2.2 Optimal Control for Underactuated Mechanical Systems

We recall that a Lagrangian Control System is underactuated if the number of the control inputs is less than the dimension of the configuration space. We assume, in the sequel, that the system is controllable [10].

Consider a Lagrangian function $L: T Q \rightarrow \mathbb{R}$. Adding external forces and controlled forces we have that the equations of motion are:

$$
\frac{d}{d t}\left(\frac{\partial L}{\partial \dot{q}^{A}}\right)-\frac{\partial L}{\partial q^{A}}=F_{A}+u_{a} \bar{X}_{A}^{a}
$$

where $F=F^{A}(q, \dot{q}) d q^{A}$ represents given external forces and $\bar{X}^{a}=\bar{X}_{A}^{a}(q) d q^{a}, 1 \leq a \leq m<$ $n$, the control forces.

Complete with independent 1-forms $\bar{X}^{\alpha}$ to obtain a local basis $\left\{\bar{X}^{a}, \bar{X}^{\alpha}\right\}$ of $\Lambda^{1} Q$ and take its dual basis that we denote by $\left\{X_{a}, X_{\alpha}\right\}$. Now, considering the quasivelocities induced by the local basis $\left\{X_{a}, X_{\alpha}\right\}$, the control equations are written as

$$
\begin{aligned}
\dot{q}^{A} & =y^{B} X_{B}^{A}(q) \\
\frac{d}{d t}\left(\frac{\partial l}{\partial y^{a}}\right)-\frac{\partial l}{\partial q^{B}} X_{a}^{B}+\mathfrak{C}_{a B}^{D} y^{B} \frac{\partial l}{\partial y^{D}} & =F_{A} X_{a}^{A}+u_{a} \\
\frac{d}{d t}\left(\frac{\partial l}{\partial y^{\alpha}}\right)-\frac{\partial l}{\partial q^{B}} X_{\alpha}^{B}+\mathcal{C}_{\alpha B}^{D} y^{B} \frac{\partial l}{\partial y^{D}} & =F_{A} X_{\alpha}^{A}
\end{aligned}
$$

where $1 \leq a \leq m, m+1 \leq \alpha \leq n$, and $u(t)=\left(u_{1}(t), \ldots, u_{m}(t)\right) \in U$ where $U$ is and open subset of $\mathbb{R}^{m}$ containing $\mathbf{0}$.

To solve an optimal control problem we need to find a trajectory $\left(q^{A}(t), u^{a}(t)\right)$ (called an optimal curve) of the configuration variables and control inputs satisfying the control equations from given initial and final conditions: $\left(q^{A}\left(t_{0}\right), y^{A}\left(t_{0}\right)\right),\left(q^{A}\left(t_{f}\right), y^{A}\left(t_{f}\right)\right)$ and minimizing the cost functional

$$
\mathcal{A}=\int_{t_{0}}^{t_{f}} C\left(q^{A}(t), y^{A}(t), u^{a}(t)\right) d t
$$

On the other hand, a second order variational Lagrangian problem with constraints is given by

$$
\min _{q(\cdot)} \int_{0}^{T} L\left(q^{A}, \dot{q}^{A}, \ddot{q}^{A}\right) d t
$$

subject to the constraints

$$
\Phi\left(q^{A}, \dot{q}^{A}, \ddot{q}^{A}\right)=0
$$

In the sequel we will show the equivalence of both theories (optimal control for underactuated systems and second order variational problems with constraints) under some regularity conditions (see 10 and references therein). Indeed, our initial optimal control problem is equivalent to the following constrained variational problem

$$
\text { Minimize } \overline{\mathcal{A}}=\int_{t_{0}}^{t_{f}} \widetilde{L}\left(q^{A}(t), y^{A}(t), \dot{y}^{A}(t)\right) d t
$$

subject to constraints

$$
\Phi^{\alpha}\left(q^{A}, y^{A}, \dot{y}^{A}\right)=\frac{d}{d t}\left(\frac{\partial l}{\partial y^{\alpha}}\right)-\frac{\partial l}{\partial q^{B}} X_{\alpha}^{B}+\mathfrak{C}_{\alpha B}^{D} y^{B} \frac{\partial l}{\partial y^{D}}-F_{A} X_{\alpha}^{A}=0
$$

and where $\widetilde{L}$ is defined as

$$
\widetilde{L}\left(q^{A}, y^{A}, \dot{y}^{A}\right)=C\left(q^{A}, y^{A}, \frac{d}{d t}\left(\frac{\partial l}{\partial y^{a}}\right)-\frac{\partial l}{\partial q^{B}} X_{a}^{B}+\mathcal{C}_{a B}^{D} y^{B} \frac{\partial l}{\partial y^{D}}-F_{A} X_{a}^{A}\right)
$$

More geometrically, we have that $\left(q^{A}, y^{A}, \dot{y}^{A}\right)$ are coordinates on $T^{(2)} Q$ (the second order tangent bundle) and the constraints $\Phi^{\alpha}$ determine a submanifold $\mathcal{M}$ of $T^{(2)} Q$ and $\widetilde{L}: T^{(2)} Q \rightarrow \mathbb{R}$.

The canonical immersion $j_{2}: T^{(2)} Q \rightarrow T(T Q)$ in the induced coordinates $\left(q^{A}, y^{A}, \dot{y}^{A}\right)$ is

$$
\begin{aligned}
T^{(2)} Q & \xrightarrow{j_{2}} T T Q \\
\left(q^{A}, y^{A}, \dot{y}^{A}\right) & \mapsto\left(q^{A}, y^{A}, X_{B}^{A} y^{B}, \dot{y}^{A}\right)
\end{aligned}
$$

Assume that the matrix $\left(\frac{\partial^{2} l}{\partial y^{\alpha} \partial y^{\beta}}\right)_{m+1 \leq \alpha, \beta \leq n}$ is regular, then we can rewrite the constraints in the form $\dot{y}^{\beta}=G^{\alpha}\left(q^{A}, y^{A}, \dot{y}^{a}\right)$ and select coordinates $\left(q^{A}, y^{A}, \dot{y}^{a}\right)$ on $\mathcal{M}$.

Hence, $\left(j_{2}\right)_{\mid \mathcal{M}}: \mathcal{M} \rightarrow T(T Q)$ is

$$
\begin{aligned}
\mathcal{M} & \xrightarrow{\left(j_{2}\right)_{\text {M }}} \\
\left(q^{A}, y^{A} ; \dot{y}^{a}\right) & \xrightarrow{\mapsto} \\
\mapsto & \left(q^{A}, y^{A} ; X_{B}^{A} y^{B}, \dot{y}^{a}, G^{\alpha}\left(q^{A}, y^{A}, \dot{y}^{a}\right)\right)
\end{aligned}
$$

Let us define $\widetilde{L}_{\mathcal{M}}$ by $\widetilde{L}_{\mathcal{M}}=\left.\widetilde{L}\right|_{\mathcal{M}}: \mathcal{M} \rightarrow \mathbb{R}$ and consider $W_{0}=\mathcal{M} \times{ }_{T Q} T^{*} T Q$ with induced coordinates $\left(q^{A}, y^{A}, \dot{y}^{a}, p_{A}, \widetilde{p}_{A}\right)$.

Now, we will describe geometrically the problem based on the Skinner and Rusk formalism (see [65]).


Figure 7.2: Skinner and Rusk Formalism
Let us define the 2-form $\Omega=p r_{2}^{*}\left(\omega_{T Q}\right)$ on $W_{0}$, where $\omega_{T Q}$ is the canonical symplectic form on $T^{*} T Q$, and $\widetilde{H}\left(v_{x}, \alpha_{x}\right)=\left\langle\alpha_{x},\left.\left(j_{2}\right)\right|_{\mathcal{M}}\left(v_{x}\right)\right\rangle-\widetilde{L}_{\mathcal{M}}\left(v_{x}\right)$ where $x \in T Q, v_{x} \in \mathcal{M}_{x} \cap\left(\left.\tau_{T Q}\right|_{\mathcal{M}}\right.$ $)^{-1}(x)$ and $\alpha_{x} \in T_{x}^{*} T Q$.

In coordinates

$$
\begin{gathered}
\Omega=d q^{A} \wedge d p_{A}+d y^{A} \wedge d \widetilde{p}_{A}, \\
\widetilde{H}=p_{A} X_{B}^{A}(q) y^{B}+\widetilde{p}_{a} \dot{y}^{a}+\widetilde{p}_{\alpha} G^{\alpha}\left(q^{A}, y^{A}, \dot{y}^{a}\right)-\widetilde{L}_{\mathcal{M}}\left(q^{A}, y^{A}, \dot{y}^{a}\right) .
\end{gathered}
$$

The intrinsic expression of this constrained problem is given by the following presymplectic equation

$$
\begin{equation*}
i_{X} \Omega=d \widetilde{H} \tag{7.2.1}
\end{equation*}
$$

Observe that $\operatorname{ker} \Omega=\operatorname{span}\left\langle\frac{\partial}{\partial \dot{y}^{a}}\right\rangle$.
Following the Gotay-Nester-Hinds algorithm [?] for presymplectic Hamiltonian systems we obtain the primary constraints $d \widetilde{H}\left(\frac{\partial}{\partial \dot{j}^{a}}\right)=0$, that is

$$
\varphi_{a}=\frac{\partial \widetilde{H}}{\partial \dot{y}^{a}}=\widetilde{p}_{a}+\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial \dot{y}^{a}}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}=0 .
$$

### 7.2. QUASIVELOCITIES AND OPTIMAL CONTROL OF UNDERACTUATED SYSTEMS105

Therefore the dynamics is restricted to the manifold $W_{1}$ determined by the vanishing of the constraints $\varphi_{a}=0$. Observe that $\operatorname{dim} W_{1}=4 n$ with induced coordinates $\left(q^{A}, y^{A}, \dot{y}^{a}, p_{A}, \widetilde{p}_{\alpha}\right)$.

A curve $t \rightarrow\left(q^{A}(t), y^{A}(t), \dot{y}^{a}(t), p_{A}(t), \widetilde{p}_{A}(t)\right)$ solution of the equations 7.2.1 must verify the following system of differential-algebraic equations.

$$
\begin{align*}
\frac{d q^{A}}{d t}= & X_{B}^{A}(q(t)) y^{B}(t)  \tag{7.2.2}\\
\frac{d y^{\alpha}}{d t}= & G^{\alpha}\left(q^{A}(t), y^{A}(t), \frac{d y^{a}}{d t}(t)\right), \quad \frac{d y^{a}}{d t}=\dot{y}^{a}(t)  \tag{7.2.3}\\
\frac{d p_{A}}{d t}= & -p_{C}(t) \frac{\partial X_{B}^{C}}{\partial q^{A}}(q(t)) y^{B}(t)-\widetilde{p}_{\alpha}(t) \frac{\partial G^{\alpha}}{\partial q^{A}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right) \\
& +\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{A}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right)  \tag{7.2.4}\\
\frac{d \widetilde{p}_{A}}{d t}= & -p_{C}(t) X_{A}^{C}(q(t))-\widetilde{p}_{\alpha}(t) \frac{\partial G^{\alpha}}{\partial y^{A}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right) \\
& +\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{A}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right)  \tag{7.2.5}\\
\widetilde{p}_{a}(t)= & -\widetilde{p}_{\alpha}(t) \frac{\partial G^{\alpha}}{\partial \dot{y}^{a}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right)+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}\left(q^{B}(t), y^{B}(t), \dot{y}^{b}\right) \tag{7.2.6}
\end{align*}
$$

From Equations (7.2.5 and (7.2.6 we deduce

$$
\begin{equation*}
\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}-\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial \dot{y}^{a}}\right)=-p_{C} X_{a}^{C}-\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial y^{a}}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{a}} \tag{7.2.7}
\end{equation*}
$$

Differentiating with respect to time, replacing in the previous equality and using (7.2.4), we obtain the following system of equations

$$
\begin{align*}
& \frac{d^{2}}{d t^{2}}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}-\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial \dot{y}^{a}}\right)-\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{a}}-\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial y^{a}}\right) \\
& +X_{a}^{A}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^{A}}-\widetilde{p}_{\alpha} \frac{\partial G^{\alpha}}{\partial q^{A}}\right)-p_{C} y^{B}\left[X_{a}^{D} \frac{\partial X_{B}^{C}}{\partial q^{D}}-X_{B}^{D} \frac{\partial X_{a}^{C}}{\partial q^{D}}\right]=0 \tag{7.2.8}
\end{align*}
$$

Let us consider the 2-form $\Omega_{W_{1}}=i_{W_{1}}^{*} \Omega$ where $i_{W_{1}}: W_{1} \hookrightarrow W_{0}$ is the canonical inclusion.
Theorem 7.2.1. The submanifold $\left(W_{1}, \Omega_{W_{1}}\right)$ is symplectic if and only if for any system of local coordinates $\left(q^{A}, y^{A}, \dot{y}^{a}, p_{A}, \widetilde{p}_{A}\right)$ on $W_{0}$

$$
\operatorname{det}\left(\mathcal{R}_{a b}\right)=\operatorname{det}\left(\frac{\partial^{2} \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a} \partial \dot{y}^{b}}-\widetilde{p}_{\alpha} \frac{\partial^{2} G^{\alpha}}{\partial \dot{y}^{a} \partial \dot{y}^{b}}\right)_{m \times m} \neq 0 \text { along } W_{1} .
$$

Under the hypothesis of Theorem 7.2.1, we can rewrite the necessary conditions for optimality as an explicit system of differential equations where Equation 7.2.8 is replaced
by

$$
\begin{equation*}
\frac{d^{3} y^{a}}{d t^{3}}=\Gamma^{a}\left(q^{A}, y^{A}, \frac{d y^{a}}{d t}, \frac{d^{2} y^{a}}{d t^{2}}, p_{A}, \tilde{p}_{\alpha}\right) \tag{7.2.9}
\end{equation*}
$$

Example 7.2.2. The Planar Rigid Body
The configuration space for this system on $Q=\mathbb{R}^{2} \times S^{1}$ and it can be considered as the simplest example in the category of rigid body dynamics. The three degrees of freedom describe the translations in $\mathbb{R}^{2}$ and the rotation about its center of mass. The configuration is given by the following variables: $\theta$ describes the relative orientation of the body reference frame with respect to the inertial reference frame. The vector $(x, y)$ denotes the position of the center of mass measured with respect to the inertial reference frame. The Lagrangian is of kinetical type

$$
L=\frac{1}{2} \dot{q}^{T} \mathcal{G}(q) \dot{q} \text {, where } \mathcal{G}(q)=\left(\begin{array}{ccc}
m & 0 & 0 \\
0 & m & 0 \\
0 & 0 & J
\end{array}\right)
$$

and where $m$ is the mass of the body and $J$ is its moment of inertia about the center of mass. If we assume that the body moves in a plane perpendicular to the direction of the gravitational forces being the potential energy zero. For the planar body, the control forces that we consider are applied to a point on the body with distance $h>0$ from the center of mass, along the body $x$-axis (see [17] for more details about this example).

The equations of motion are

$$
\begin{aligned}
m \ddot{x} & =u_{1} \cos \theta-u_{2} \sin \theta \\
m \ddot{y} & =u_{1} \sin \theta+u_{2} \cos \theta \\
J \ddot{\theta} & =-h u_{2}
\end{aligned}
$$

The control fields are

$$
\begin{aligned}
& X_{1}=\frac{\cos \theta}{m} \frac{\partial}{\partial x}+\frac{\sin \theta}{m} \frac{\partial}{\partial y} \\
& X_{2}=-\frac{\sin \theta}{m} \frac{\partial}{\partial x}+\frac{\cos \theta}{m} \frac{\partial}{\partial y}-\frac{h}{J} \frac{\partial}{\partial \theta},
\end{aligned}
$$

and we complete the basis of vector fields with $X_{3}=h \sin \theta \frac{\partial}{\partial x}-h \cos \theta \frac{\partial}{\partial y}-\frac{\partial}{\partial \theta}$
The nonzero structure functions are

$$
\begin{array}{ll}
\mathfrak{C}_{12}^{2}=\frac{h}{m h^{2}+J}=-\mathfrak{C}_{21}^{2}, & \mathfrak{C}_{12}^{3}=-\frac{h^{2}}{\left(m h^{2}+J\right) J}=-\mathfrak{C}_{21}^{3} \\
\mathcal{C}_{23}^{1}=-\frac{m h^{2}+J}{J}=-\mathfrak{C}_{32}^{1}, & \mathfrak{C}_{13}^{2}=\frac{J}{m h^{2}+J}=-\mathfrak{C}_{31}^{2} \\
\mathfrak{C}_{13}^{3}=-\frac{h}{m h^{2}+J}=-\mathfrak{C}_{31}^{3} &
\end{array}
$$

### 7.2. QUASIVELOCITIES AND OPTIMAL CONTROL OF UNDERACTUATED SYSTEMS107

Taking the corresponding quasivelocities $\left\{y^{1}, y^{2}, y^{3}\right\}$, we have that

$$
\begin{aligned}
\dot{x} & =y^{1} \frac{\cos \theta}{m}-y^{2} \frac{\sin \theta}{m}+y^{3} h \sin \theta \\
\dot{y} & =y^{1} \frac{\sin \theta}{m}+y^{2} \frac{\cos \theta}{m}-y^{3} h \cos \theta \\
\dot{\theta} & =-y^{2} \frac{h}{J}-y^{3} .
\end{aligned}
$$

The Lagrangian of this system is

$$
l\left(x, y, \theta, y^{1}, y^{2}, y^{3}\right)=\frac{1}{2}\left[\frac{1}{m}\left(y^{1}\right)^{2}+\frac{m h^{2}+J}{m J}\left(y^{2}\right)^{2}+\left(m h^{2}+J\right)\left(y^{3}\right)^{2}\right]
$$

then the Hamel equations with controls are:

$$
\begin{aligned}
u_{1} & =\dot{y}^{1}+\frac{h}{J}\left(y^{2}\right)^{2}-h m\left(y^{3}\right)^{2}+\frac{J-m h^{2}}{J} y^{2} y^{3} \\
u_{2} & =\frac{J+m h^{2}}{J} \dot{y}^{2}--\frac{h}{J} y^{1} y^{2}-y^{1} y^{3} \\
0 & =\left(J+m h^{2}\right) \dot{y}^{3}+\frac{h^{2}}{J} y^{1} y^{2}+h y^{1} y^{3} .
\end{aligned}
$$

Consider the following cost functional

$$
\mathcal{A}=\frac{1}{2} \int_{0}^{T}\left(u_{1}^{2}+u_{2}^{2}\right) d t
$$

Following our formalism this optimal control problem is equivalent to the constrained second-order variational problem determined by:

$$
\widetilde{\mathcal{A}}=\int_{0}^{T} \widetilde{L}\left(x, y, \theta, y^{1}, y^{2}, y^{3}, \dot{y}^{1}, \dot{y}^{2}, \dot{y}^{3}\right) d t
$$

and the second-order constraint

$$
\Phi\left(x, y, \theta, y^{1}, y^{2}, y^{3}, \dot{y}^{1}, \dot{y}^{2}, \dot{y}^{3}\right)=\left(J+m h^{2}\right) \dot{y}^{3}+\frac{h^{2}}{J} y^{1} y^{2}+h y^{1} y^{3}=0
$$

where

$$
\begin{aligned}
\widetilde{L}\left(x, y, \theta, y^{1}, y^{2}, y^{3}, \dot{y}^{1}, \dot{y}^{2}, \dot{y}^{3}\right) & =\frac{1}{2}\left[\dot{y}^{1}+\frac{h}{J}\left(y^{2}\right)^{2}-h m\left(y^{3}\right)^{2}+\frac{J-m h^{2}}{J} y^{2} y^{3}\right]^{2} \\
& +\frac{1}{2}\left[\frac{J+m h^{2}}{J} \dot{y}^{2}--\frac{h}{J} y^{1} y^{2}-y^{1} y^{3}\right]^{2}
\end{aligned}
$$

Now, we rewrite the second-order constraint in the form

$$
\dot{y}^{3}=-\frac{h^{2}}{\left(J+m h^{2}\right) J} y^{1} y^{2}-\frac{h}{\left(J+m h^{2}\right)} y^{1} y^{3} .
$$

Take now $W_{0}=\mathcal{M} \times T^{*}\left(T\left(\mathbb{R}^{2} \times \mathbb{S}^{1}\right)\right)$ with coordinates $\left(x, y, \theta, y^{1}, y^{2}, y^{3}, \dot{y}^{1}, \dot{y}^{2}, p_{1}, p_{2}, p_{3}, \tilde{p}_{1}, \tilde{p}_{2}, \tilde{p}_{3}\right)$.
Now, the presymplectic 2-form $\Omega$, the Hamiltonian $\widetilde{H}$ and the primary constraints, $\varphi_{1}, \varphi_{2}$, are:

$$
\begin{aligned}
\Omega= & d x \wedge d p_{1}+d y \wedge d p_{2}+d \theta \wedge d p_{3}+d y^{1} \wedge d \tilde{p}_{1}+d y^{2} \wedge d \tilde{p}_{2}+d y^{3} \wedge d \tilde{p}_{3}, \\
\widetilde{H}= & p_{1}\left[y^{1} \frac{\cos \theta}{m}-y^{2} \frac{\sin \theta}{m}+y^{3} h \sin \theta\right]+p_{2}\left[y^{1} \frac{\sin \theta}{m}+y^{2} \frac{\cos \theta}{m}-y^{3} h \cos \theta\right] \\
& -p_{3}\left[y^{2} \frac{h}{J}+y^{3}\right]+\tilde{p}_{1} \dot{y}^{1}+\tilde{p}_{2} \dot{y}^{2}-\tilde{p}_{3}\left(\frac{h^{2}}{\left(J+m h^{2}\right) J} y^{1} y^{2}+\frac{h}{\left(J+m h^{2}\right)} y^{1} y^{3}\right) \\
& -\frac{1}{2}\left[\dot{y}^{1}+\frac{h}{J}\left(y^{2}\right)^{2}-h m\left(y^{3}\right)^{2}+\frac{J-m h^{2}}{J} y^{2} y^{3}\right]^{2}-\frac{1}{2}\left[\frac{J+m h^{2}}{J} \dot{y}^{2}--\frac{h}{J} y^{1} y^{2}-y^{1} y^{3}\right]^{2} \\
\varphi_{1}= & \frac{\partial \widetilde{H}}{\partial \dot{y}^{1}}=\tilde{p}_{1}-\left[\dot{y}^{1}+\frac{h}{J}\left(y^{2}\right)^{2}-h m\left(y^{3}\right)^{2}+\frac{J-m h^{2}}{J} y^{2} y^{3}\right]=0, \\
\varphi_{2}= & \tilde{p}_{2}-\frac{J+m h^{2}}{J}\left[\frac{J+m h^{2}}{J} \dot{y}^{2}--\frac{h}{J} y^{1} y^{2}-y^{1} y^{3}\right], \\
\text { i.e., } &
\end{aligned}
$$

$$
\begin{aligned}
& \tilde{p}_{1}=\dot{y}^{1}+\frac{h}{J}\left(y^{2}\right)^{2}-h m\left(y^{3}\right)^{2}+\frac{J-m h^{2}}{J} y^{2} y^{3} \\
& \tilde{p}_{2}=\frac{J+m h^{2}}{J}\left[\frac{J+m h^{2}}{J} \dot{y}^{2}--\frac{h}{J} y^{1} y^{2}-y^{1} y^{3}\right] .
\end{aligned}
$$

These constraints determine the submanifold $W_{1}$. Applying the Theorem 7.2.1, we deduce that the 2-form $\Omega_{W_{1}}$, restriction of $\Omega$ to $W_{1}$, is symplectic since

$$
\left(\begin{array}{ll}
\mathcal{R}_{11} & \mathcal{R}_{12} \\
\mathcal{R}_{21} & \mathcal{R}_{22}
\end{array}\right)=\left(\begin{array}{rr}
1 & 0 \\
0 & \left(J+m h^{2}\right) / J
\end{array}\right)
$$

is regular.
Therefore, the algorithm stabilizes in the first constraint submanifold $W_{1}$. Moreover, there exists a unique solution of the dynamics, a vector field $X$ which satisfies $i_{X} \Omega_{W_{1}}=$ $d \widetilde{H}_{W_{1}}$. In consequence, we have a unique control input which extremizes the objective function $\mathcal{A}$. If we take the flow $F_{t}: W_{1} \rightarrow W_{1}$ of the vector field $X$ then we have that $F_{t}^{*} \Omega_{W_{1}}=\Omega_{W_{1}}$, then the evolution is symplectic preserving. Obviously, the Hamiltonian function $\left.\widetilde{H}\right|_{W_{1}}$ is preserved by the solution of the optimal control problem, that is $\left.\widetilde{H}\right|_{W_{1}} \circ F_{t}=\left.\widetilde{H}\right|_{W_{1}}$. Both properties, symplecticity and preservation of energy, are important geometric invariants. In [30], we construct, using discrete variational calculus, numerical integrators which inherit some of the geometric properties of the optimal control problem (symplecticity, momentum preservation and a very good energy behavior).

### 7.3 Underactuated Mechanical Control Systems on Lie Groups

In this section we analyze the case of underactuated control of mechanical systems on Lie groups and, as a particular example, a family of underactuated problems for the rigid body
on $S O(3)$. Finally we study the control equation of the Cosserat Rod, an static rod where the configuration space is the Lie group $S E(3)$, the group of rotations and translations in the space.

In the following we assume that the controlled equations are trivialized where $L: G \times \mathfrak{g} \rightarrow$ $\mathbb{R}$

$$
\frac{d}{d t}\left(\frac{\delta L}{\delta \xi}\right)-a d_{\xi}^{*}\left(\frac{\delta L}{\delta \xi}\right)-£_{g}^{*} \frac{\partial L}{\partial g}=u_{a} e^{a}
$$

where we are assuming that $\left\{e^{a}\right\}$ are independent elements on $\mathfrak{g}^{*}$ and $\left(u_{a}\right)$ are the admissible controls. Complete it to a basis $\left\{e^{a}, e^{A}\right\}$ of the vector space $\mathfrak{g}^{*}$. Take its dual basis $\left\{e_{i}\right\}=\left\{e_{a}, e_{A}\right\}$ on $\mathfrak{g}$ with bracket relations:

$$
\left[e_{i}, e_{j}\right]=\mathcal{C}_{i j}^{k} e_{k}
$$

The basis $\left\{e_{i}\right\}=\left\{e_{a}, e_{A}\right\}$ induced coordinates $\left(y^{a}, y^{A}\right)=\left(y^{i}\right)$ on $\mathfrak{g}$, that is, if $e \in \mathfrak{g}$ then $e=y^{i} e_{i}=y^{a} e_{a}+y^{A} e_{A}$. In $\mathfrak{g}^{*}$, we have induces coordinates ( $p_{a}, p_{\alpha}$ ) for the previous fixed basis $\left\{e^{i}\right\}$

In these coordinates, the equations of motion are rewritten as

$$
\begin{aligned}
\frac{d}{d t}\left(\frac{\partial L}{\partial y^{a}}\right)-\mathfrak{C}_{i a}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{a}\right\rangle & =u_{a} \\
\frac{d}{d t}\left(\frac{\partial L}{\partial y^{A}}\right)-\mathfrak{C}_{i A}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{A}\right\rangle & =0
\end{aligned}
$$

With these equations we can study the optimal control problem that consists on finding trajectories $\left(g(t), u^{a}(t)\right)$ of state variables and control inputs satisfying the previous equations from given initial and final conditions $\left(g\left(t_{0}\right), y^{i}\left(t_{0}\right)\right)$ and $\left(g\left(t_{f}\right), y^{i}\left(t_{f}\right)\right)$, respectively, and extremizing the functional

$$
\mathcal{J}=\int_{t_{0}}^{t_{f}} C\left(g(t), y^{i}(t), u^{a}(t)\right) d t
$$

Obviously the proposed optimal control problem is equivalent to a variational problem with second order constraints, determined by the Lagrangian $\widetilde{L}: G \times 2 \mathfrak{g} \rightarrow \mathbb{R}$ given, in the selected coordinates, by

$$
\widetilde{L}\left(g, y^{i}, \dot{y}^{i}\right)=C\left(g, y^{i}, \frac{d}{d t}\left(\frac{\partial L}{\partial y^{a}}\right)-\mathcal{C}_{i a}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{a}\right\rangle\right) .
$$

subjected to the second-order constraints

$$
\Phi^{A}\left(g, y^{i}, \dot{y}^{i}\right)=\frac{d}{d t}\left(\frac{\partial L}{\partial y^{A}}\right)-\mathfrak{C}_{i A}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{A}\right\rangle=0 .
$$

which determine the submanifold $\mathcal{M}$ of $G \times 2 \mathfrak{g}$.
Observe that from the constraint equations we have that

$$
\frac{\partial^{2} L}{\partial y^{A} \partial y^{B}} \dot{y}^{B}+\frac{\partial^{2} L}{\partial y^{A} \partial y^{b}} \dot{y}^{b}-\mathcal{C}_{i A}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{A}\right\rangle=0 .
$$

Therefore, assuming that the matrix $\left(W_{A B}\right)=\left(\frac{\partial^{2} L}{\partial y^{A} \partial y^{B}}\right)$ is regular we can write the constraint equations as

$$
\begin{aligned}
\dot{y}^{B} & =-W^{B A}\left(\frac{\partial^{2} L}{\partial y^{A} \partial y^{b}} \dot{y}^{b}-\mathfrak{C}_{i A}^{j} y^{i} \frac{\partial L}{\partial y^{j}}-\left\langle £_{g}^{*} \frac{\delta L}{\delta g}, e_{A}\right\rangle\right) \\
& =G^{B}\left(g, y^{i}, \dot{y}^{a}\right)
\end{aligned}
$$

where $W^{B A}=\left(W_{B A}\right)^{-1}$.


Figure 7.3: Skinner and Rusk formalism on Lie groups

This means that we can identify $T \mathcal{M} \equiv G \times \operatorname{span}\left\{\left(e_{i}, \mathbf{0}, \mathbf{0}\right),\left(\mathbf{0}, e_{i}, \mathbf{0}\right),\left(\mathbf{0}, \mathbf{0}, e_{a}\right)\right\}$ where $\left(e_{i}, \mathbf{0}, \mathbf{0}\right),\left(\mathbf{0}, e_{i}, \mathbf{0}\right),\left(\mathbf{0}, \mathbf{0}, e_{a}\right) \in 3 \mathfrak{g}$.

Therefore, we can choose coordinates $\left(g, y^{i}, \dot{y}^{a}\right)$ on $\mathcal{M}$. This choice allows us to consider an "intrinsic point view", that is, to work directly on $\bar{W}_{0}=\mathcal{M} \times 2 \mathfrak{g}^{*}$ avoiding the use of Lagrange multipliers.

Define the restricted lagrangian $\widetilde{L}_{\mathcal{M}}$ by $\widetilde{L}_{\mathcal{M}}=\left.\widetilde{L}\right|_{\mathcal{M}}: \mathcal{N} \rightarrow \mathbb{R}$ and take induced coordinates on $\bar{W}_{0}$ are $\gamma=\left(g, y^{i}, \dot{y}^{a}, p_{i}, \tilde{p}_{i}\right)$. Consider the presymplectic 2-form on $\bar{W}_{0}, \Omega_{\bar{W}_{0}}=\left(p r_{2} \circ\right.$ $\left.i_{\bar{W}_{0}}\right)^{*}\left(\omega_{G \times \mathfrak{g}}\right)$.

Using the notation $\left(e_{i}\right)_{0}=\left(e_{i}, \mathbf{0}, \mathbf{0} ; \mathbf{0}, \mathbf{0}\right) \in 3 \mathfrak{g} \times 2 \mathfrak{g}^{*}$ and, in the same way $\left(e_{i}\right)_{1}=$ $\left(\mathbf{0}, e_{i}, \mathbf{0} ; \mathbf{0}, \mathbf{0}\right),\left(e_{a}\right)_{2}=\left(\mathbf{0}, \mathbf{0}, e_{a} ; \mathbf{0}, \mathbf{0}\right),\left(e^{i}\right)_{3}=\left(\mathbf{0}, \mathbf{0}, \mathbf{0} ; e^{i}, \mathbf{0}\right)$ and $\left(e^{i}\right)_{4}=\left(\mathbf{0}, \mathbf{0}, \mathbf{0} ; \mathbf{0}, e^{i}\right)$ then the unique nonvanishing elements on the expression of $\Omega_{\bar{W}_{0}}$ are:

$$
\begin{aligned}
\left(\Omega_{\bar{W}_{0}}\right)_{\gamma}\left(\left(e_{i}\right)_{0},\left(e_{j}\right)_{0}\right) & =\mathfrak{C}_{i j}^{k} p_{k}, \\
\left(\Omega_{\bar{W}_{0}}\right)_{\gamma}\left(\left(e_{i}\right)_{0},\left(e^{j}\right)_{3}\right) & =-\left(\Omega_{\bar{W}_{0}}\right)_{\gamma}\left(\left(e^{i}\right)_{3},\left(e_{j}\right)_{0}\right)=\delta_{i}^{j}, \\
\left(\Omega_{\bar{W}_{0}}\right)_{\gamma}\left(\left(e_{i}\right)_{1},\left(e^{j}\right)_{4}\right) & =-\left(\Omega_{\bar{W}_{0}}\right)_{\gamma}\left(\left(e^{i}\right)^{4},\left(e_{j}\right)_{1}\right)=\delta_{i}^{j} .
\end{aligned}
$$

Taking the dual basis $\left(e^{i}\right)_{0}=\left(e^{i}, \mathbf{0}, \mathbf{0} ; \mathbf{0}, \mathbf{0}\right) \in 3 \mathfrak{g}^{*} \times 2 \mathfrak{g}$ and, in the same way $\left(e^{i}\right)_{1}=$ $\left(\mathbf{0}, e^{i}, \mathbf{0} ; \mathbf{0}, \mathbf{0}\right),\left(e^{a}\right)_{2}=\left(\mathbf{0}, e^{a}, \mathbf{0} ; \mathbf{0}, \mathbf{0}\right),\left(e_{i}\right)_{3}=\left(\mathbf{0}, \mathbf{0}, \mathbf{0} ; e_{i}, \mathbf{0}\right)$ and $\left(e_{i}\right)_{4}=\left(\mathbf{0}, \mathbf{0}, \mathbf{0} ; \mathbf{0}, e_{i}\right)$ we deduce that

$$
\left(\Omega_{\bar{W}_{0}}\right)=\left(e^{i}\right)_{0} \wedge\left(e_{i}\right)_{3}+\left(e^{i}\right)_{1} \wedge\left(e_{i}\right)_{4}+\frac{1}{2} \mathrm{C}_{i j}^{k} p_{k}\left(e^{i}\right)_{0} \wedge\left(e^{j}\right)_{0}
$$

Moreover

$$
\bar{H}=y^{i} p_{i}+\dot{y}^{a} \tilde{p}_{a}+G^{A}\left(g, y^{i}, \dot{y}^{a}\right) \tilde{p}_{A}-\widetilde{L}_{\mathcal{M}}\left(g, y^{i}, \dot{y}^{a}\right)
$$

and, in consequence,

$$
\begin{aligned}
d \bar{H}= & -\left\langle £_{g}^{*}\left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g}+\tilde{p}_{B} \frac{\delta G^{B}}{\delta g}\right), e_{i}\right\rangle\left(e^{i}\right)_{0}+\left(p_{i}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{i}}+\tilde{p}_{B} \frac{\partial G^{B}}{\partial y^{i}}\right)\left(e^{i}\right)_{1} \\
& +\left(\tilde{p}_{a}-\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}+\tilde{p}_{B} \frac{\partial G^{B}}{\partial \dot{y}^{a}}\right)\left(e^{a}\right)_{2}+y^{i}\left(e_{i}\right)_{3}+\dot{y}^{a}\left(e_{a}\right)_{4}+G^{A}\left(e_{A}\right)_{4}
\end{aligned}
$$

The conditions for the integral curves $t \rightarrow\left(g(t), y^{i}(t), \dot{y}^{a}(t), p_{A}(t), \tilde{p}_{A}(t)\right)$ of a vector field $X$ satisfying equations $i_{X} \Omega_{\bar{W}_{0}}=d \bar{H}$ are

$$
\begin{align*}
\frac{d g}{d t} & =g\left(y^{i}(t) e_{i}\right)  \tag{7.3.1}\\
\frac{d y^{a}}{d t} & =\dot{y}^{a}  \tag{7.3.2}\\
\frac{d y^{A}}{d t} & =G^{A}\left(g, y^{i}, \dot{y}^{a}\right)  \tag{7.3.3}\\
\frac{d p_{i}}{d t} & =\left\langle £_{g}^{*}\left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g}-\widetilde{p}_{B} \frac{\delta G^{B}}{\delta g}\right), e_{i}\right\rangle+\mathcal{C}_{i j}^{k} p_{k} y^{j}  \tag{7.3.4}\\
\frac{d \tilde{p}_{i}}{d t} & =-p_{i}+\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{i}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{i}}  \tag{7.3.5}\\
\widetilde{p}_{a} & =\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial \dot{y}^{a}}=\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}+W^{B A} \widetilde{p}_{B} \frac{\partial^{2} L}{\partial y^{A} \partial y^{a}} \tag{7.3.6}
\end{align*}
$$

As a consequence we obtain the following set of differential equations:

$$
\begin{aligned}
\frac{d g}{d t}= & g\left(y^{i}(t) e_{i}\right) \\
\frac{d y^{A}}{d t}= & G^{A}\left(g, y^{i}, \dot{y}^{a}\right) \\
0= & \frac{d^{2}}{d t^{2}}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial \dot{y}^{a}}\right]-\mathcal{C}_{i a}^{b} y^{i}\left(\frac{d}{d t}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{b}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial \dot{y}^{b}}\right]\right) \\
& -\frac{d}{d t}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{a}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{a}}\right)+\mathfrak{C}_{i a}^{k} y^{i}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{k}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{k}}\right) \\
& +\left\langle £_{g}^{*}\left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g}-\widetilde{p}_{B} \frac{\delta G^{B}}{\delta g}\right), e_{a}\right\rangle-\mathfrak{C}_{i a}^{C} y^{i} \frac{d \tilde{p}_{C}}{d t} \\
0= & \frac{d^{2} \widetilde{p}_{A}}{d t^{2}}+\mathfrak{C}_{i A}^{B} y^{i} \frac{\widetilde{p}_{B}}{d t}-\mathfrak{C}_{i A}^{k} y^{i}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{k}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{k}}\right] \\
& -\frac{d}{d t}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{A}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{A}}\right]+\left\langle £_{g}^{*}\left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g}-\widetilde{p}_{B} \frac{\delta G^{B}}{\delta g}\right), e_{A}\right\rangle \\
& +\mathfrak{C}_{i A}^{b} y^{i}\left(\frac{d}{d t}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{b}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial \dot{y}^{b}}\right]\right)-\mathcal{C}_{i A}^{b} y^{i}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^{b}}-\widetilde{p}_{B} \frac{\partial G^{B}}{\partial y^{b}}\right]
\end{aligned}
$$

which determine completely the dynamics.
If the matrix

$$
\left(\frac{\partial^{2} \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^{a} \partial \dot{y}^{b}}\right)
$$

is regular then we can write the previous equations as a explicit system of third-order differential equations. It is easy to show that this regularity assumption is equivalent to the condition that the constrain algorithm stops at the first constraint submanifold $\bar{W}_{1}$ (see [5], [30], [29] and reference therein for more details).

### 7.3.1 Optimal Control of an Underactuated Rigid Body

We consider the motion of a rigid body where the configuration space is the Lie group $G=S O(3)$ (see [11, 49]). Therefore, $T S O(3) \simeq S O(3) \times \mathfrak{s o}(3)$, where $\mathfrak{s o}(3) \equiv \mathbb{R}^{3}$ is the Lie algebra of the Lie group $S O(3)$. The Lagrangian function for this system is given by $L: S O(3) \times \mathfrak{s o}(3) \rightarrow \mathbb{R}$,

$$
L\left(R, \Omega_{1}, \Omega_{2}, \Omega_{3}\right)=\frac{1}{2}\left(I_{1} \Omega_{1}^{2}+I_{2} \Omega_{2}^{2}+I_{3} \Omega_{3}^{2}\right)
$$

Now, denote by $t \rightarrow R(t) \in S O(3)$ a curve. The columns of the matrix $R(t)$ represent the directions of the principal axis of the body at time $t$ with respect to some reference system. Now, we consider the following control problem. First, we have the reconstruction equations:

$$
\dot{R}(t)=R(t)\left(\begin{array}{ccc}
0 & -\Omega_{3}(t) & \Omega_{2}(t) \\
\Omega_{3}(t) & 0 & -\Omega_{1}(t) \\
-\Omega_{2}(t) & \Omega_{1}(t) & 0
\end{array}\right)=R(t)\left(\Omega_{1}(t) E_{1}+\Omega_{2}(t) E_{2}+\Omega_{3}(t) E_{3}\right)
$$

where

$$
E_{1}:=\left(\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & 1 & 0
\end{array}\right), \quad E_{2}:=\left(\begin{array}{ccc}
0 & 0 & 1 \\
0 & 0 & 0 \\
-1 & 0 & 0
\end{array}\right), \quad E_{3}:=\left(\begin{array}{ccc}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right)
$$

and the equations for the angular velocities $\Omega_{i}$ with $i=1,2,3$ :

$$
\begin{aligned}
I_{1} \dot{\Omega}_{1}(t) & =\left(I_{2}-I_{3}\right) \Omega_{2}(t) \Omega_{3}(t)+u_{1}(t) \\
I_{2} \dot{\Omega}_{2}(t) & =\left(I_{3}-I_{1}\right) \Omega_{3}(t) \Omega_{1}(t)+u_{2}(t) \\
I_{3} \dot{\Omega}_{3}(t) & =\left(I_{1}-I_{2}\right) \Omega_{1}(t) \Omega_{2}(t)
\end{aligned}
$$

where $I_{1}, I_{2}, I_{3}$ are the moments of inertia and $u_{1}, u_{2}$ denote the applied torques playing the role of controls of the system.

The optimal control problem for the rigid body consists on finding the trajectories ( $R(t), \Omega(t), u(t))$ with fixed initial and final conditions $\left(R\left(t_{0}\right), \Omega\left(t_{0}\right)\right),\left(R\left(t_{f}\right), \Omega\left(t_{f}\right)\right)$ respectively and minimizing the cost functional

$$
\mathcal{A}=\int_{0}^{T} \mathcal{E}\left(\Omega, u_{1}, u_{2}\right) d t=\int_{0}^{T}\left[c_{1}\left(u_{1}^{2}+u_{2}^{2}\right)+c_{2}\left(\Omega_{1}^{2}+\Omega_{2}^{2}+\Omega_{3}^{2}\right)\right] d t
$$

with $c_{1}, c_{2} \geq 0$. The constants $c_{1}$ and $c_{2}$ represent weights on the cost functional. For instance, $c_{1}$ is the weight in the cost functional measuring the fuel expended by an attitude manoeuver of a spacecraft modeled by the rigid body and $c_{2}$ is the weight given to penalize high angular velocities.

This optimal control problem is equivalent to solve the following variational problem with constraints ([10], [33]),

$$
\min \widetilde{\mathcal{J}}=\int_{0}^{T} \widetilde{L}(\Omega, \dot{\Omega}) d t
$$

subject to constraints $I_{3} \dot{\Omega}_{3}-\left(I_{1}-I_{2}\right) \Omega_{1} \Omega_{2}=0$, where

$$
\widetilde{L}(\Omega, \dot{\Omega})=\mathfrak{C}\left(\Omega, I_{1} \dot{\Omega}_{1}-\left(I_{2}-I_{3}\right) \Omega_{2} \Omega_{3}, I_{2} \dot{\Omega}_{2}-\left(I_{3}-I_{1}\right) \Omega_{3} \Omega_{1}\right)
$$

Thus, the submanifold $\mathcal{M}$ of $G \times 2 \mathfrak{s o}$ (3), is given by

$$
\mathcal{M}=\left\{(R, \Omega, \dot{\Omega}) \left\lvert\, \dot{\Omega}_{3}=\left(\frac{I_{1}-I_{2}}{I_{3}}\right) \Omega_{1} \Omega_{2}\right.\right\} .
$$

We consider the submanifold $\bar{W}_{0}=\mathcal{M} \times 2 \mathfrak{s o}^{*}(3)$ with induced coordinates

$$
\left(g, \Omega_{1}, \Omega_{2}, \Omega_{3}, \dot{\Omega}_{1}, \dot{\Omega}_{2}, p_{1}, p_{2}, p_{3}, \tilde{p}_{1}, \tilde{p}_{2}, \tilde{p}_{3}\right)
$$

Now, we consider the restriction $L_{\mathcal{M}}$ given by

$$
\tilde{L}_{\mathcal{M}}=c_{1}\left[\left(I_{1} \dot{\Omega}_{1}-\left(I_{2}-I_{3}\right) \Omega_{2} \Omega_{3}\right)^{2}+\left(I_{2} \dot{\Omega}_{2}-\left(I_{3}-I_{1}\right) \Omega_{3} \Omega_{1}\right)^{2}\right]+c_{2}\left(\Omega_{1}^{2}+\Omega_{2}^{2}+\Omega_{3}^{2}\right)
$$

For simplicity we denote by $G^{3}=\frac{I_{1}-I_{2}}{I_{3}} \Omega_{1} \Omega_{2}$.
Then, we can write the equations of motion of the optimal control for this underactuated system. For simplicity, we consider the particular case $I_{1}=I_{2}=I_{3}=1$ then the equations of motion of the optimal control system are:

$$
\begin{aligned}
\Omega_{2}(t) \frac{d \tilde{p}_{3}}{d t}-2\left(c_{2} \frac{d \Omega_{1}}{d t}+c_{1} \Omega_{3}(t) \frac{d^{2} \Omega_{2}}{d t^{2}}-c_{1} \frac{d^{3} \Omega_{1}}{d t^{3}}\right) & =0 \\
-\Omega_{1}(t) \frac{d \tilde{p}_{3}}{d t}-2\left(c_{2} \frac{d \Omega_{2}}{d t}-c_{1} \Omega_{3}(t) \frac{d^{2} \Omega_{1}}{d t^{2}}-c_{1} \frac{d^{3} \Omega_{2}}{d t^{3}}\right) & =0 \\
\frac{d^{2} \tilde{p}_{3}}{d t^{2}}-2 c_{2} \frac{d \Omega_{3}}{d t}-2 c_{1} \Omega_{2}(t) \frac{d^{2} \Omega_{1}}{d t^{2}}-2 c_{1} \Omega_{1}(t) \frac{d^{2} \Omega_{2}}{d t^{2}} & =0 \\
\frac{d \Omega_{3}}{d t} & =0
\end{aligned}
$$

If we consider the rigid body as a model of a spacecraft then we observe that this particular cost function is taking into account both, the fuel expenditure $\left(c_{1}\right)$ and also is trying to minimize the overall angular velocity ( $c_{2}$ ). In Figures (7.4) and (7.5) we compare their behavior in two particular cases: $c_{1}=1 / 2$ and $c_{2}=1 / 2$ and $c_{1}=1 / 2$ and $c_{2}=0$.

In all cases we additionally have the reconstruction equation

$$
\dot{R}(t)=R(t)\left(\Omega_{1}(t) E_{1}+\Omega_{2}(t) E_{2}+\Omega_{3}(t) E_{3}\right)
$$

with boundary conditions $R\left(t_{0}\right)$ and $R\left(t_{f}\right)$.


Figure 7.4: Angular velocity values for initial conditions satisfying $\Omega_{i}(0)=\Omega_{i}(4)=0$, $i=1,2$ and fixed values of $R(0)$ and $R(4)$.



Figure 7.5: Comparison of the functions $1 / 2\left(\Omega_{1}^{2}(t)+\Omega_{2}^{2}(t)+\Omega_{3}^{2}(t)\right)$ (left) and $1 / 2\left(u_{1}^{2}(t)+u_{2}^{2}(t)\right)$ (right) in both cases

The case $c_{1}=0$ and $c_{2}=1$, that is, we only try to minimize the overall angular velocity (see [66] for the underactuated case) is singular. We obtain the following system of equations:

$$
\begin{aligned}
\Omega_{2}(t) \frac{d \tilde{p}_{3}}{d t}-2 \frac{d \Omega_{1}}{d t} & =0 \\
-\Omega_{1}(t) \frac{d \tilde{p}_{3}}{d t}-2 \frac{d \Omega_{2}}{d t} & =0, \\
\frac{d^{2} \tilde{p}_{3}}{d t^{2}}-2 \frac{d \Omega_{3}}{d t} & =0 \\
\frac{d \Omega_{3}}{d t} & =0 .
\end{aligned}
$$

Observe that in this case it is not possible to impose arbitrary boundary conditions $\left(R\left(t_{0}\right), \Omega\left(t_{0}\right)\right)$ and $\left(R\left(t_{f}\right), \Omega\left(t_{f}\right)\right)$ although it is always possible to find a trajectory verifying initial and final attitude conditions $R\left(t_{0}\right)$ and $R\left(t_{f}\right)$.

### 7.3.2 Optimal control of a Cosserat rod

A static road corresponds to a Lagrangian system where the energy density takes the role of the Lagrangian.

Let $r:[0, T] \rightarrow \mathbb{R}^{3}$ a vector function and we consider the vector functions $d^{(1)}, d^{(2)}$ : $[0, T] \rightarrow \mathbb{R}^{3}$ such satisfies the orthonormality condition $\left\langle d^{(1)}(t), d^{(2)(t)}\right\rangle=0,\left\|d^{(l)}(t)\right\|_{2}=1$ for $l=1,2, t \in[0, T]$.
$d^{(l)}$ describe the orientation of the cross-section ${ }^{1}$ along $[0, T]$ and $r$ addresses points on the centraline. Moreover, let $d^{(3)}(t)=d^{(1)}(t) \times d^{(2)}(t)$ the normal cross-section.

The deformed configuration can be described taking coordinates on $S E(3)=\mathbb{R}^{3} \times S O(3)$. Let us $(r, R):[0, T] \rightarrow \mathbb{R}^{3} \times S O(3)$, where $R(t) e_{k}=\left(d^{(1)}(t), d^{(2)}(t), d^{(3)}(t)\right) \in S O(3)$, is the matrix representation of $R(t)$ with respect to basis ( $e_{1}, e_{2}, e_{3}$ ) and ( $d^{(1)}, d^{(2)}, d^{(3)}$ ) is the orthonormal basis for the Euclidean space $\mathbb{E}^{3}$.

We denote by $W=W^{\text {int }}+W^{e x t}: T Q \rightarrow \mathbb{R}$ the potential energy of the mechanical system and we assume that $W^{\text {int }}$ is frame independent, that is,

$$
W^{i n t}(R, r, \dot{R}, \dot{r})=\bar{W}^{i n t}\left(R^{-1} \dot{r}, R^{-1} \dot{R}\right)=\bar{W}^{i n t}(u, v)
$$

where $u=R^{-1} \dot{R} \in \mathfrak{s o}(3), v=R^{-1} \dot{r}, \in \mathbb{R}^{3}, r, \dot{r} \in \mathbb{R}^{3}, R \in S O(3)$ and $\dot{R} \in T_{R} S O(3)$.
We assume that $W^{e x t}$ depends only of the position, that is, $W^{e x t}=W^{e x t}(R, r)$. Therefore, our new problem is defined in the left-trivialized tangent space $S E(3) \times \mathfrak{s e}(3)$ as

$$
W=\bar{W}^{\text {int }}(u, v)+W^{e x t}(R, r)=W^{i n t}(R, r, \dot{R}, \dot{r})+W^{e x t}(R, r) .
$$

With some abuse of notation, let define the elements of $S E(3)$ and $\mathfrak{s e}(3)=\mathfrak{s o}(3) \times \mathbb{R}^{3}$ as

$$
\Phi=(R, r)=\left(\begin{array}{cc}
R & r  \tag{7.3.8}\\
0_{3} & 1
\end{array}\right) \in S E(3), \quad \phi=(u, v)=\left(\begin{array}{cc}
\hat{u} & v \\
0_{3} & 0
\end{array}\right) \in \mathfrak{s e}(2),
$$

where $0_{3}$ is the null $1 \times 3$ matrix (both $\Phi$ and $\phi$ are $4 \times 4$ matrices).
The potential total energy is

$$
V=\int_{t_{0}}^{t_{f}}\left[\bar{W}^{\text {int }}(u, v)+W^{e x t}(R, r)\right] d t
$$

The equilibrium configurations of any static system coincide with the critical points of the potential energy. Observe that $\delta u=\delta(R \dot{R})=-R^{-1} \delta R R^{-1} \dot{R}+R^{-1} \delta \dot{R}$. We denote by $\Sigma_{u}=R^{-1} \delta R$, then

$$
\dot{\Sigma}_{u}=-R^{-1} \dot{R} R^{-1} \delta R+R^{-1}(\delta \dot{R})=-u \Sigma_{u}+R^{-1}(\delta \dot{R})
$$

[^1]Therefore $\dot{\Sigma}_{u}+u \Sigma_{u}=R^{-1}(\dot{R})$ and

$$
\delta u=-\Sigma_{u} u+u \Sigma_{u}+\dot{\Sigma}_{u}=\left[u, \Sigma_{u}\right]+\dot{\Sigma}_{u}=\dot{\Sigma}_{u}+u \times \Sigma_{u} .
$$

Where we identify the Lie algebra $\mathfrak{s o}(3) \mathbb{R}^{3}$ and the bracket with the cross product.
Also, $\delta v=\delta(R \dot{r})=-R^{-1} \delta R R^{-1} \dot{r}+R^{-1} \delta \dot{r}$. Denote by $\Sigma_{v}=R^{-1} \delta r$, then

$$
\dot{\Sigma}_{v}=-R^{-1} \dot{R} R^{-1} \delta r+R^{-1}(\delta \dot{r})=-u \Sigma_{v}+R^{-1}(\delta \dot{r})
$$

Finally,

$$
\delta v=-\Sigma_{u} v+u \Sigma_{v}+\dot{\Sigma}_{v} .
$$

Observe that $\Sigma_{u} \in \mathfrak{s o}(3)$ and $\Sigma_{v} \in \mathbb{R}^{3}$. From the definition of $\Sigma_{u}$ and $\Sigma_{v}$ we deduce that $R \Sigma_{u}=\delta R$ and $R \Sigma_{v}=\delta r$.

The equilibrium configurations are characterized by

$$
\begin{aligned}
0 & =\int_{t_{0}}^{t_{f}} \frac{\partial \bar{W}^{\text {int }}}{\partial v} \delta v+\frac{\partial \bar{W}^{\text {int }}}{\partial u} \delta u+\frac{\partial W^{e x t}}{\partial r} \delta r+\frac{\partial W^{e x t}}{\partial R} \delta R d t \\
& =\int_{t_{0}}^{t_{f}} \frac{\partial \bar{W}^{\text {int }}}{\partial v}\left(\dot{\Sigma}_{v}-\Sigma_{u} v-u \Sigma_{v}\right)+\frac{\partial \bar{W}^{\text {int }}}{\partial u}\left(u \times \Sigma_{u}+\dot{\Sigma}_{u}\right)+\frac{\partial W^{e x t}}{\partial r}\left(R \Sigma_{v}\right)+\frac{\partial W^{e x t}}{\partial R}\left(R \Sigma_{u}\right) d t .
\end{aligned}
$$

Taking the redefinition

$$
\begin{equation*}
n=\frac{\partial W^{i n t}(u, v)}{\partial v}, \quad m=\frac{\partial W^{i n t}(u, v)}{\partial u} \tag{7.3.9}
\end{equation*}
$$

and

$$
\begin{equation*}
f=\frac{\partial W^{e x t}(R, r)}{\partial r} \quad l=\frac{\partial W^{e x t}(R, r)}{\partial R} \tag{7.3.10}
\end{equation*}
$$

which we consider the control forces, we finally arrive to the equations of motion

$$
\begin{align*}
& \dot{n}+n \times u+f=0, \\
& \dot{m}+n \times v+m \times u+l=0 . \tag{7.3.11}
\end{align*}
$$

For more details see [41]
The optimal control problem consists on finding a trajectory of the state variables and control inputs that minimize the cost functional

$$
\mathcal{C}=\int_{0}^{T}\left(f^{2}+\rho_{1}^{2} l^{2}\right) d t
$$

where $\rho_{1}$ is a weight constant. The control problem is subject to the following boundary conditions $\Phi(0)=(R(0), r(0)), \phi(0)=(u(0), v(0))$ and $\Phi(T)=(R(T), r(T)), \phi(T)=$ $(u(T), v(T))$ belonging to $S E(3) \times \mathfrak{s e}(2)$.

As in the rigid body example, from eqs. (7.3.11) we can obtain an expression of $f$ and $l$ in terms of the other variables. Furthermore, differentiating equations (7.3.9) with respect to time, we can find out $\dot{n}$ and $\dot{m}$ in terms of $((u, v),(\dot{u}, \dot{v}))$ if we assume $W^{\text {int }}(u, v)$ twice differentiable, i.e., $\binom{\dot{n}}{\dot{m}}=\mathcal{H}(u, v)\binom{\dot{u}}{\dot{v}}$, where $\mathcal{H}$ is the Hessian matrix of $W^{\text {int }}(u, v)$.

Now, setting the function $L: \mathfrak{s e}(2) \times \mathfrak{s e}(2) \rightarrow \mathbb{R}$ as $L((u, v),(\dot{u}, \dot{v}))=[f((u, v),(\dot{u}, \dot{v}))]^{2}+$ $\rho_{1}^{2}[l((u, v),(\dot{u}, \dot{v}))]^{2}$, our problem reduces to extremize the control functional

$$
\begin{equation*}
\mathcal{C}=\int_{0}^{T} L((u, v),(\dot{u}, \dot{v})) d t=\int_{0}^{T} L(\phi, \dot{\phi}) d t \tag{7.3.12}
\end{equation*}
$$

subject to the boundary conditions above. For sake of completeness we can write down the explicit form of $L$, namely

$$
\begin{aligned}
& L((u, v),(\dot{u}, \dot{v}))=f((u, v),(\dot{u}, \dot{v}))^{2}+\rho_{1}^{2} l((u, v),(\dot{u}, \dot{v}))^{2}= \\
& \left(\mathcal{H}_{11}(u, v) \dot{u}+\mathcal{H}_{12}(u, v) \dot{v}+\partial_{v} W^{\text {int }}(u, v) \times u\right)^{2}+ \\
& +\rho_{1}^{2}\left(\mathcal{H}_{21}(u, v) \dot{u}+\mathcal{H}_{22}(u, v) \dot{v}+\right. \\
& \left.+\partial_{v} W^{\text {int }}(u, v) \times v+\partial_{u} W^{\text {int }}(u, v) \times u\right)^{2} .
\end{aligned}
$$

## REFERENCES

[1] Abraham R and Marsden JE. Foundations of Mechanics. Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., second edition, (1978)
[2] Abraham R, Marsden TE and Ratiu T. Manifolds, Tensor Analysis and Applications, vol 75 if Applied Mathematical Sciencis. Springer-Verlag, New York, second edition, (1988)
[3] J. Baillieul: The geometry of controlled mechanical systems. Mathematical control theory, Springer, New York, 1999, 322-354.
[4] M. Barbero-Liñán, D. Martín de Diego, M. C. Muñoz-Lecanda: Lie algebroids and optimal control: abnormality. Geometry and Physics: XVII International Fall Workshop on Geometry and Physics. AIP Conference Proceedings, 1130 (2009), 113-119.
[5] M. Barbero-Liñán, A.Echeverría Enríquez, D. Martín de Diego, M.C Muñoz-Lecanda and N. Román-Roy. Skinner-Rusk unified formalism for optimal control systems and applications. J. Phys. A: Math Theor. 40 (2007), 12071-12093.
[6] Barbero-Liñán M and Muõz-Lecanda MC.Constraint algorithm for extremals in optimal control problems. Int. J. Geom. Methods Mod. Phys. 6(7), pp. 1221-1233, (2009)
[7] Barbero-Liñán M and Muõz-Lecanda MC. Strict abnormal extremals in nonholonomic and kinematic control systems. Discrete Contin. Dyn. Syst. Ser. S3(1), pp. 1-17, (2010)
[8] Barth E and Leimkuhler B. Symplectic methods for conservative multibody systems. in Integration Algorithms and Classical Mechanics, American Mathematical Society, Providence, RI, pp. 25-43, (1996)
[9] Berndt R. An introduction to Symplectic Geometry. Graduate studies in Mathematics, vol. 26, American Mathematical Society.
[10] A.M. Bloch: Nonholonomic Mechanics and Control, Interdisciplinary Applied Mathematics Series, 24, Springer-Verlag, New York (2003).
[11] A.M. Bloch, I.I. Hussein, M. Leok, A.K. Sanyal. Geometric Structure-Preserving Optimal Control of the Rigid Body, Journal of Dynamical and Control Systems, 15(3), 307-330, 2009.
[12] Bloch, Anthony M, Hussein, Islam I. Dynamic coverage optimal control for multiple spacecraft interferometric imaging. J. Dyn. Control Syst. 13 (2007), no. 1, 69-93.
[13] A.M. Bloch, P.E. Crouch, Nonholonomic and vakonomic control systems on Riemannian manifolds, in Dynamics and Control of Mechanical Systems, Michael J. Enos, ed., Fields Inst. Commun. 1, AMS, Providence, RI, 1993, pp. 2552.
[14] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, R.M. Murray: Nonholonomic mechanical systems with symmetry, Arch. Rational Mech. Anal., 136 (1996), 21-99.
[15] R. Benito, D. Martín de Diego: Hidden symplecticity in Hamilton's principle algorithms, Proc. Conf. Prague, August 30 - September 3, 2004. Charles University, Prague (Czech Republic) (2005), 411-419.
[16] Boothby WA. An introduction to differentiable Manifolds and Riemannian geometry. Pure and Applied Mathematics Series of Monographs and Textbooks. Academic Press, New York, San Francisco, London, (1975)
[17] Bullo F and Lewis AD. Geometric control of mechanical systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems. Texts in Applied Mathematics, Springer Verlang, New York, (2005)
[18] J.F. Cariñena, C. López: Geometric study of Hamilton's variational principle. Rev. Math. Phys. 3 (4) (1991), 379-401.
[19] Cariñena, José F.; López, Carlos; Martínez, Eduardo Sections along a map applied to higher-order Lagrangian mechanics. Noether's theorem. Acta Appl. Math. 25 (1991), no. 2, 127-151
[20] Cariñena, José F.; López, Carlos The time-evolution operator for higher-order singular Lagrangians. Internat. J. Modern Phys. A 7 (1992), no. 11, 2447-2468.
[21] M. Crampin, T. Mestdag, Anholonomic frames in constrained dynamics. Dynamical Systems. An International Journal 25 159-187 (2010).
[22] do Carmo MP. Riemannian geometry. Birkhäuser, Boston-Basel- Berlin, (1992)
[23] Cantrijn, Frans, Sarlet, Willy. Higher-order Noether symmetries and constants of the motion. J. Phys. A 14 (1981), no. 2, 479-492.
[24] M. Crampin, W. Sarlet, F. Cantrijn. Higher order differential equations and higher order Lagrangian Mechanics. Math. Proc. Camb. Phil. Soc. 99, 565-587, (1986).
[25] Crampin, M.; Saunders, D. J. On the geometry of higher-order ordinary differential equations and the Wuenschmann invariant. Groups, geometry and physics, 79-92, Monogr. Real Acad. Ci. Exact. Fís.-Quím. Nat. Zaragoza, 29, Acad. Cienc. Exact. Fís. Quím. Nat. Zaragoza, Zaragoza, 2006.
[26] L. Colombo, F. Jimenez. Continuous and discrete mechanics for the attitude dynamics of the rigid body on SO(3). GMC Notes, Number 1, 2012. (With Fernando Jimenez). http://gmcnetwork.org/drupal/sites/default/files/GMCNotes1_OK.pdf
[27] L. Colombo, F. Jimenez, D. Martin de Diego Discrete Second-Order Euler-Poincaré Equations. An application to optimal control . International Journal of Geometric Methods in Modern Physics. Vol 9, No ${ }^{\circ}$ (2012).
[28] L. Colombo, D. Martín de Diego. On the Goemetry of Higher-Order Problems on Lie Groups. http://arxiv.org/abs/1104.3221.
[29] L. Colombo, D. Martín de Diego. Quasivelocities and Optimal Control of Underactuated Mechanical Systems. Geometry and Physics: XVIII FallWorkshop on Geometry and Physics. AIP Conference Proceedings, no. 1260, 133-140 (2010).
[30] L. Colombo, D. Martin de Diego and M. Zuccalli. Optimal Control for Underactuated Mechanical Systems: A Geometrical Approach. Journal Mathematical Physics 51 (2010) 083519.
[31] J. Cortés. Geometric Control and Numerical Aspects of Nonholonomic Systems, Lec. Notes in Math., 1793, Springer-Verlag, Berlin (2002).
[32] J. Cortés, M. de León, D. Martín de Diego, S. Martínez. Geometric description of vakonomic and nonholonomic dynamics, SIAM J. Control Optim. 41, no. 5, 1389-1412, (2002).
[33] P. Crouch, F. Silva-Leite. Geometry and the dynamic interpolation problem. American Control Conference, 1131-1136 (1991).
[34] Crouch, P.; Silva Leite, F. The dynamic interpolation problem: on Riemannian manifolds, Lie groups, and symmetric spaces. J. Dynam. Control Systems 1 (1995), no. 2, 177-202.
[35] F. Gay-Balmaz, D. D. Holm, D. M. Meier, T. S. Ratiu, F.-X. Vialard.Invariant higherorder variational problems, arXiv:1012.5060v1.
[36] Gràcia, X.; Pons, J. M.; Román-Roy, N. Higher-order Lagrangian systems: geometric structures, dynamics, and constraints. J. Math. Phys. 32 (1991), no. 10, 2744-2763.
[37] Gotay MJ, Nester JM and Hinds G. Presymplectic manifolds and the Dirac-Bergmann theory of constraints. J. Math. Phys., 32, pp. 2744-2763, (1991)
[38] G. Hamel. Die Lagrange-Eulersche Gleichungen der Mechanik, Z. Math. Phys. 50, 1-57 (1904).
[39] D. D. Holm: Geometric mechanics. Part I and II, Imperial College Press, London; distributed by World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.
[40] Holm D, Marsden JE and Ratiu T The Euler-Poincaré equations and semidirect products with applications to continuum theories. Adv. Math. 137, 1, pp. 1-25, (1998)
[41] P. Jung, S. Leyendecker, J. Linn, M. Ortiz, A discrete mechanics approach to Cosserat rod theory. Part 1: static equilibria, International Journal for numerical methods in engineering 1 (2010) 101-130.
[42] V. Jurdjevic: Geometric Control Theory. Cambridge Studies in Advanced Mathematics, 52, Cambridge University Press, 1997.
[43] V. Jurdjevic: Optimal Control, Geometry and Mechanics. In Mathematical Control Theory, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, 227-267.
[44] Kobayashi S and Nomizu K. Foundations of Differential Geometry. Vol I and II. Wiley Classics Library. John Wiley and Sons Inc., New York, (1996)
[45] Kostelec, Peter J.; Lewis, H. Ralph. The use of Hamilton's principle to derive timeadvance algorithms for ordinary differential equations. Comput. Phys. Comm. 96 (1996), no. 2-3, 129-151.
[46] O. Krupková, The Geometry of Ordinary Variational Equations, Lecture Notes in Mathematics 1678, Springer, Berlin, 1997.
[47] Francaviglia, Mauro; Krupka, Demeter The Hamiltonian formalism in higher order variational problems. Ann. Inst. H. Poincaré Sect. A (N.S.) 37 (1982), no. 3, 295-315 (1983).
[48] Lagrange J.L., Mechanique Analytique. Chez la Venue Desaint (1788)
[49] T. Lee, M. Leok, N.H. McClamroch. Optimal Attitude Control of a Rigid Body using Geometrically Exact Computations on SO(3), Journal of Dynamical and Control Systems, 14 (4), 465-487, (2008).
[50] T. Lee Thesis, Computational Geometric Mechanics and Control of the Rigid Bodies" Universidad of Michigan, [2008].
[51] M. de León, P. R. Rodrigues: Generalized Classical Mechanics and Field Theory, North-Holland Mathematical Studies 112, North-Holland, Amsterdam, 1985.
[52] de León M and Martín de Diego D. On the geometry of nonholonomic Lagrangian systems. J. Math. Phys., 37(7), pp. 3389-3414, (1996)
[53] M. de León, J.C. Marrero, E. Martínez. Lagrangian submanifolds and dynamics on Lie algebroids, J. Phys. A: Math. Gen. 38 (2005), R241-R308.
[54] Libermann P and Marle CM. Symplectic Geometry and Analytical Mechanics. D. Reidel Publishing Company, Holland, (1987)
[55] Marsden JE and Ratiu TS Introduction to mechanics and symmetry. $\mathbf{1 7}$ of Texts in Appied Mathematics, Springer-Verlag, New York, (1999).
[56] David Martín de Diego Control Optimo en Grupos de Lie. Courses notes, UPC, (2007).
[57] C. Martinez Campos: Geometric methods in classical field theory and continuous media. Universidad Autonoma de Madrid, 2010.
[58] Morimoto, Akihiko Prolongations of G-structures to tangent bundles of higher order. Nagoya Math. J. 381970 153-179.
[59] J. I. Neimark, N. A. Fufaev. Dynamics of Nonholonomic Systems. Translations of Mathematical Monographs, AMS, 33 (1972).
[60] Noakes, Lyle; Heinzinger, Greg; Paden, Brad Cubic splines on curved spaces. IMA J. Math. Control Inform. 6 (1989), no. 4, 465-473.
[61] Pere Daniel Prieto Martínez: Estudi de les estructures geomètriques dels sistemes dinàmics d'ordre superior. Universidad Politecnica de Cataluna, 2011.
[62] H. Poincaré. Sur une forme nouvelle des équations de la méchanique, C. R. Acad. Sci.,132, 369-371, (1901).
[63] H. Poincaré Les formes d' equilibre d' une masse fluide en rotation, Revue Generale des Sciences 3, 809-815 (1892).
[64] H. Poincaré Sur une forme nouvelle des équations de la mécanique. C.R. Acad. Sci. Paris, 132, pp. 369-371, (1901)
[65] R. Skinner, R. Rusk: Generalized Hamiltonian dynamics I. Formulation on $T^{*} Q \oplus T Q$, Journal of Mathematical Pyhsics, 24 (11), 2589-2594 and 2595-2601, (1983).
[66] K. Spindler: Optimal attitude control of a rigid body, Applied Mathematics\& Optimization 34 (1), 79-90 (1996).
[67] D. Seto, J. Baillieul: Control problems in super-articulated mechanical systems. IEEE Trans. Automat. Control 39, no. 12 (1994) 2442-2453.
[68] H.J. Sussmann: Geometry and Optimal Control. In Mathematical Control Theory, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, pp. 140-198.


[^0]:    ${ }^{1}$ A Lagrangian $L$ is hyperregular if $\mathbb{F} L$ is a global diffeomorphism.

[^1]:    ${ }^{1}$ In geometry, a cross-section is the intersection of a figure in 2-dimensional space with a line, or of a body in 3-dimensional space with a plane, etc. More plainly, when cutting an object into slices one gets many parallel cross-sections

