Universidad Autónoma de Madrid

Facultad de Ciencias

Departamento de Matemáticas

Instituto de Ciencias Matemáticas

ICMAT-CSIC

UAM-UC3M-UCM

Geometric and Numerical Methods for Optimal Control of Mechanical Systems

Leonardo Jesús Colombo

Madrid, April 2014

Madrid, Abril de 2014

Dr. D. David Martín de Diego, Investigador Científico del Consejo Superior de Investigaciones Científicas y miembro del Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM en Madrid

HACE CONSTAR:

que la presente memoria de Tesis Doctoral presentada por Leonardo Jesús Colombo y titulada **Geometric and Numerical Methods for Optimal Control of Mechanical Systems** ha sido realizada bajo su dirección en el Instituto de Ciencias Matemáticas y tutelada en el Departamento de Matemática de la Universidad Autónoma de Madrid por el Dr. Marco Zambón. El trabajo recogido en dicha memoria se corresponde con el planteado en el proyecto de tesis doctoral aprobado en su día por el órgano responsable del programa de doctorado.

Y para que conste, en cumplimiento de la legislación vigente, informo favorablemente sobre la referida Tesis Doctoral y autorizo su presentación para su admisión a trámite.

El Director de la Tesis

Fdo.: David Martín de Diego

Universidad Autónoma de Madrid

Facultad de Ciencias
Departamento de Matemáticas

Instituto de Ciencias Matemáticas

ICMAT-CSIC
UAM-UC3M-UCM

Geometric and Numerical Methods for Optimal Control of Mechanical Systems

Memoria realizada por
Leonardo Jesús Colombo
presentada ante el Departamento de Matemática
de la Universidad Autónoma de Madrid
para la obtención del grado de Doctor en Matemáticas.

Proyecto dirigido por el
Dr. David Martín de Diego
del Instituto de Ciencias Matemáticas (ICMAT-CISC-UAM-UC3M-UCM)
del Consejo Superior de Investigaciones Científicas.

Madrid, Abril 2014.

To my parents
Salvador Colombo and Nilda Ortiz
and my wife
Estefania Hochleitner

“El científico no estudia la naturaleza por la utilidad que le pueda reportar; la estudia por el gozo que le proporciona, y este gozo se debe a la belleza que hay en ella... me refiero a aquella profunda belleza que surge de la armonía del orden en sus partes y que una pura inteligencia puede captar. La belleza intelectual se basta a sí misma, y es por ella, más que quizá por el bien futuro de la humanidad, por lo que el científico consagra su vida a un trabajo largo y difícil.”

Henri Poincaré.

Resumen

Las aplicaciones de técnicas provenientes de la Geometría Diferencial moderna y la Topología han ayudado a una mayor comprensión de los problemas provenientes de la teoría de Sistemas Dinámicos. Estas aplicaciones han reformulado la mecánica analítica y clásica en un lenguaje geométrico que junto a nuevos métodos analíticos, topológicos y numéricos conforman una nueva área de investigación en matemáticas y física teórica llamada *Mecánica Geométrica*.

La Mecánica Geométrica se configura como un punto de encuentro de disciplinas diversas como la Mecánica, la Geometría, el Análisis, el Álgebra, el Análisis Numérico, las Ecuaciones en Derivadas Parciales... Actualmente, la Mecánica Geométrica es un área de investigación pujante con fructíferas conexiones con otras disciplinas como la Teoría de Control no-lineal y el Análisis Numérico.

El objetivo de la Teoría de Control es determinar el comportamiento de un sistema dinámico por medio de acciones externas de forma que se cumplan ciertas condiciones prefijadas como, por ejemplo, que haya un extremo fijo, los dos, que ciertas variables no alcancen algunos valores u otro tipo de situaciones más o menos complicadas. Las aplicaciones de la Mecánica Geométrica en Teoría de Control han causado grandes progresos en esta área de investigación. Por ejemplo, la formulación geométrica de los sistemas mecánicos de control sujetos a ligaduras no holónomas han ayudado a la comprension de problemas en locomoción, contrabilidad y planificación de trayectorias, problemas de control con obstáculos e interpolación.

Uno de los mayores objetivos del Análisis Numérico y de la Matemática Computacional ha sido traducir los fenómenos físicos en algoritmos que producen aproximaciones numéricas suficientemente precisas, asequibles y robustas. En los últimos años, el campo de la Integración Geométrica surgió con el objetivo de diseñar y analizar métodos numéricos para ecuaciones diferenciales ordinarias y, más recientemente, para ecuaciones diferenciales en derivadas parciales, que preservan, tanto como sea posible, la estructura geométrica subyacente.

La Mecánica Discreta, entendida como el punto de encuentro de la Mecánica Geométrica y la Integración Geométrica, es un área de investigación bien fundamentada y una herramienta poderosa a la hora de entender los sistemas dinámicos y físicos, más concretamente, aquellos relacionados con la Mecánica y la Teoría de Control. Una herramienta clave en Mecánica Discreta, y muy utilizada en este trabajo, son los integradores variacionales, i.e., integradores geométricos basados en la discretización de los principios variacionales.

El presente trabajo de investigación incluye nuevos resultados en el área de la Mecánica Geométrica que permiten el estudio de sistemas mecánicos, su aplicación a la teoría de control óptimo y la construcción de integradores geométricos que preservan ciertas estructuras subyacentes de gran interés para el análisis numérico de los sistemas de control. Más precisamente, presentamos una nueva formulación geométrica para la dinámica de sistemas mecánicos de orden superior sujeto a ligaduras, también de orden superior, debido a que un problema de control óptimo de sistemas mecánicos puede ser resuelto como un problema variacional

de orden superior con ligaduras de orden superior. Hemos estudiado la relación entre los sistemas Lagrangianos de orden superior con ligaduras (noholónomas y vakónomas) y los sistemas Hamiltonianos asociados, la reducción por simetrías de esta clase de sistemas y la integración geométrica de problemas de control. El trabajo desarrollado en esta tesis también contribuye con nuevos desarrollos en Mecánica Discreta y su interrelación con la teoría de control, algebroides de Lie y grupoides de Lie.

Abstract

The applications of techniques from the modern Differential Geometry and Topology have helped a new way of understanding the problems which come from the theory of Dynamical Systems. These applications have reformulated the analytic mechanics and classical mechanics in a geometric language which attracted new analytic, topologic and numerical methods given rise to a new research line in mathematics and theoretical physics, called *Geometric Mechanics*.

Geometric Mechanics is a meeting point for different areas such as, Analysis, Algebra, Numerical Analysis, Partial Differential Equations... Currently, Geometric Mechanics is a research area with a strong relationship with Nonlinear Control Theory and Numerical Analysis.

The applications of Geometric Mechanics in control theory have given great progress in this area. For example, the geometric formulation of mechanical systems subject to nonholonomic constraints has helped the understanding of problems in locomotion, controllability and trajectory planning, control problems with obstacles and interpolation problems.

One of the main goals of the numerical analysis and computational mathematics has been rendering physical phenomena into algorithms that produces sufficiently accurate, affordable, and robust numerical approximations. In the last years, the field of Geometric Integration arose to design and to analyze numerical methods for ordinary differential equations and, more recently, for partial differential equations, that preserves exactly, as much as possible, the underlying geometrical structures.

The Discrete Mechanics, understood as the confluence of Geometric Mechanics and Geometric Integration, is both a well-founded research area and a powerful tool in the understanding of dynamical and physical systems, more concretely of those related to mechanics. A key tool of Discrete Mechanics, which has been strongly used in this work, is the variational integrators, i.e., geometric integrators for mechanical problems based on the discretization of variational principles.

The work developed in this thesis includes new valuable developments in Geometric Mechanics which permits the understanding about mechanical systems, its applications in control theory and the construction of geometric integrators which preserves underlying geometrical structures of great interest to the numerical analysis of control systems. More precisely, we give a new geometric formulation for the dynamics of higher-order mechanical systems subject also to higher-order constraints since an optimal control problem for mechanical systems can be seen as higher-order variational problem with higher-order constraints. We have studied the relation between higher-order Lagrangian systems with constraints (nonholonomics and vakonomics) and higher-order Hamiltonian systems, the reduction by symmetries of this kind of mechanical systems and the geometric integration of control problems. The work developed in this thesis also is in line with new developments in Discrete Mechanics and its relation with control theory, Lie groupoids and Lie algebroids.

Contents

Introduction

In a nutshell, this thesis deals with *new developments on geometric mechanics and geometric integration for optimal control of mechanical systems.* Optimal control, geometric mechanics and geometric integration have at least one important property in common: all three are natural described using the language of *differential geometry.*

The emphasis in the geometry is an attempt to understand qualitatively the dynamics of mechanical systems giving advantages in the analysis and the design of numerical methods.

The geometrical description of Lagrangian and Hamiltonian mechanics have the advantage to give global and intrinsic equations which are invariant with respect to any change of coordinates in the configuration space. For a system with n degrees of freedom, Lagrangian mechanics gives rise a system of n ordinary second-order equations called Euler-Lagrange equations which determine completely the evolution of the system given initial boundary conditions and assuming the regularity of the Lagrangian function.

An alternative description is the so-called Hamiltonian mechanics (locally equivalent to the Lagrangian ones when the Lagrangian function is regular) which describes the evolution of the system through a system of $2n$ first order ordinary differential equations.

During the last past half century, mathematicians, physicists and engineering have been extended and generalized this framework to mechanical systems with friction, time dependent systems, with constraints (holonomic and nonholonomic), classical field theory, etc. But, one of the areas that has been most successfully by its applications in engineering was control theory. The main purpose of control theory is to study systems in which we can influence the dynamics externally using control variables. In control systems appearing in engineering studies, 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

$$\dot{x} = f(x, u), \tag{1}$$

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 can not be expressed globally as an Euclidean space. In this work, dynamics and control problems for mechanical systems are studied, incorporating careful consideration of their geometric features. The goal is to find control inputs moving the initial state of the system to a prefixed target state. In this work we will focus in the case when initial and final states are fixed.

The study of control systems is a research area with a lot of activity in the last sixty years studying topics like controllability, accessibility, design of trajectories, design of numerical methods, among others.

In optimal control theory, we want also that the system verifies an extra condition which consists on minimizing a cost functional, that is, an optimal control problem which consists on finding a trajectory $\gamma(t) = (x(t), u(t))$ of the state variables and control inputs, satisfying the control equation (1) given initial and final boundary conditions $x(0) = x_0$, $x(T) = x_T$ and minimizing the cost functional

$$\mathcal{J}(\gamma(t)) = \int_0^T C(x(y), u(t))dt. \tag{2}$$

The trajectory $\gamma(t)$ verifying all these conditions will be called *optimal*. Optimal control is a technique in mathematics useful to solve optimization problems evolving in time and susceptible to have an external influence. Optimal control theory is a young research area that appears in a wide variety of fields such as medicine, economics, traffic flow, engineering and astronomy. However, applications and understanding do not always come together. In order to gain insight, differential geometry has been used in control theory, giving rise to geometric control theory in the 70's (see works of Sussman, Jurdjevic Nijmeijer and van der Schaft [145], [165], [166]).

Looking back in time to the birth of optimal control theory and calculus of variations we should go to the year 1696 when the solution of the brachistochrone curve problem's solved by Johann Bernoulli was published in Acta Eruditorum. Here we will focus in the great contribution of the Russian mathematician L.S. Pontryagin. In 1950's, Pontryagin organized a seminar at Steklov Institute of Mathematics about some problems in applied mathematics inviting to some engineers as speakers. The seminar ends with the discover of the so called Pontryagin maximum principle [157]. The problem they tried to solve was a system of five ordinary differential equations with three control parameters modeling the maneuvers of a fighter jet. After it researchers found that the applications of this principle could be applied to others research fields as control of spacecrafts and satellites, biomechanics, economy, robotic, etc; currently becoming in an interesting area of research in mathematics [17],[105],[99] for example.

This work deals with mechanical control systems, giving emphasis to a particular class of mechanical control systems: underactuated mechanical systems. Underactuated mechanical systems are characterized by the fact that they have more degrees of freedom than actuators. 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. Underactuated systems include spacecrafts, underwater vehicles, mobile robots, helicopters, wheeled vehicles and underactuated manipulators.

To analyze geometrically these control systems we will need an unifying concept: the notion of Lie algebroid. Lie algebroids have deserved a lot of interest in recent years. Since a Lie algebroid is a concept which unifies tangent bundles and Lie algebras, one can suspect their relation with mechanics. More precisely, a Lie algebroid over a manifold Q is a vector bundle $\tau_E : E \to Q$ over Q with a Lie algebra structure over the space $\Gamma(\tau_E)$ of sections of E and an application $\rho : E \to TQ$ called anchor map satisfying some compatibility conditions (see [118]). Examples of Lie algebroids are the tangent bundle over a manifold Q where the Lie bracket is the usual Lie bracket of vector fields and where the anchor map is the identity function; the real finite dimensional Lie algebras as vector bundles over a point, where the anchor map is the null application; the action Lie algebroids of the type $pr_1 : M \times \mathfrak{g} \to M$ where $\mathfrak{g}$ is a Lie algebra acting infinitesimally over the manifold M with a Lie bracket over the space of sections induced by the Lie algebra structure and whose anchor map is the action of $\mathfrak{g}$ over M; and finally, the Lie-Atiyah algebroid $\tau_{TQ/G} : TQ/G \to \widehat{M} = Q/G$ associated with the G-principal bundle $p : Q \to \widehat{M}$ where the anchor map is induced by the tangent application of p, $Tp : TQ \to T\widehat{M}$ [111],[118],[134],[176].

In [176] Alan Weinstein developed a generalized theory of Lagrangian mechanics on Lie algebroids and obtained the equations of motion using the linear Poisson structure on the dual of the Lie algebroid and the Legendre transformation associated with a regular Lagrangian $L : E \to \mathbb{R}$. In [176] also he ask about whether it is possible to develop a formalism similar on Lie algebroids to Klein's formalism [96] in Lagrangian mechanics. This task was obtained by Eduardo Martínez in [134] ([133] and [158]). The main notion is that of prolongation of a Lie algebroid over a mapping introduced by Higgins ans Mackenzie in [118]. A more general situation, the prolongation of an anchored bundle $\tau_E : E \to Q$ was also considered by Popescu in [158].

The importance of Lie algebroids in mathematics is beyond doubt and in the last years Lie algebroids has been a lot of applications in theoretical physics and other related sciences. More concretely in Classical Mechanics and Classical Field Theory. One of the main things that Lie algebroids are interesting in Classical Mechanics lie in the fact that there are many different situations that can be understand in a general framework using the theory of Lie algebroids as systems with symmetries, systems over semidirect products, Hamiltonian and Lagrangian systems, systems with constraints (nonholonomic and vakonomic) and Classical Fields theory.

In [111] M. de León, J.C Marrero and E. Martínez have developed a Hamiltonian description for the mechanics on Lie algebroids and they have shown that the dynamics is obtained solving an equation in the same way than in Classical Mechanics (see also [133] and [176]). Moreover, they shown that the Legendre transformation $leg_L : E \to E^*$ associated to the Lagrangian $L : E \to \mathbb{R}$ induces a Lie algebroid morphism and when the Lagrangian is regular both formalisms are equivalent. Also they have extended the Tulczyjew's contruction [169], [170] to the framework of Lie algebroids and they have been introduced the notion of Lagrangian Lie subalgebroid of a symplectic Lie algebroid. Then they have shown that Euler-Lagrange equations and Hamilton equations over a Lie algebroids are just the local equations defined by certain Lagrangian submanifolds of a symplectic Lie algebroid associated to E. As a consequence they have deduced the Lagrange-Poincaré and Hamilton-Poincaré equations associated to a G-invariant Lagrangian and Hamiltonian, respectively.

Marrero and collaborators also have analyzed the case of non-holonomic mechanics on Lie algebroids [55]. In another direction, in [87] D. Iglesias, J.C. Marrero, D. Martín de Diego and D. Sosa have studied singular Lagrangian systems and vakonomic mechanics from the point of view of Lie algebroids obtained through the application of a constrained variational principle. They have developed a constraint algorithm for presymplectic Lie algebroids generalizing the well know constraint algorithm of Gotay, Nester and Hinds [73] and they also have established the Skinner and Rusk formalism on Lie algebroids.

Recently, higher-order variational problems have been studied for their important applications in aeronautics, robotics, computer-aided design, air traffic control and trajectory planning. There are variational principles which involves higher-order derivatives [68],[69],[70], [112], [155] since from it one can obtain the equations of motion for Lagrangians where the configuration space is a higher-order tangent bundle.

In this thesis we will consider higher-order mechanics from the point of view of the Skinner and Rusk formalism to obtain higher-order Euler-Lagrange equations, higher-order Euler-Poincaré equations and higher-order Lagrange-Poincaré equations. Also, we will focus in the case of systems with higher-order constraints and their extension to the natural and unifying setting of Lie algebroids. One of the main objectives is to characterize geometrically the equations of motion of an optimal control problem for an underactuated mechanical system. In this last system, the trajectories are parameterized by the admissible controls and the necessary conditions for extremals in the optimal control problem are expressed using a pseudo-Hamiltonian formulation based on the Pontryagin maximun principle. Many of the concrete examples under study have additional geometric properties as, for instance, the configuration space is not only a differentiable manifold but it also has a compatible structure of group, that is, the configuration space is a Lie group. We will take advantage of this property to give an intrinsic expression of the equations of motion for higher-order mechanical systems and for optimal control problems with symmetries (see also [68],[69],[70] and [88],[89],[155],[154]).

Other characterization of the higher-order mechanics in this work is due by the well known Tulczyjew's triple extending the program started in [109]. Regarding Geometric Mechanics, the theory of Lagrangian submanifolds gives a geometric and intrinsic description of Lagrangian and Hamiltonian dynamics (see the work by W.M. Tulczyjew [169, 170] and, Grabowska and Grabowski [74]). Given a mechanical system defined by a Lagrangian function $L : TQ \to \mathbb{R}$, then the Lagrangian dynamics will be "generated" by the Lagrangian submanifold $\mathrm{d}L(TQ) \subset T^*TQ$. On the other hand, if the mechanical system is defined by the Hamiltonian function $H : T^*Q \to \mathbb{R}$, the Hamiltonian dynamics will be "generated" by the Lagrangian submanifold $\mathrm{d}H(T^*Q) \subset T^*T^*Q$. A way to perform the relationship between these two formalisms is by the so-called Tulczyjew's triple:

$$T^*TQ \xleftarrow{\alpha_Q} TT^*Q \xrightarrow{\beta_Q} T^*T^*Q,$$

where α_Q and β_Q are both isomorphisms and T^*TQ, TT^*Q, T^*T^*Q, are double vector bundles equipped with symplectic structures.

In the higher-order setting, roughly speaking, a second-order Lagrangian system is defined by a second-order Lagrangian function $L : T^{(2)}Q \to \mathbb{R}$, where $T^{(2)}Q$ is the second-order tangent bundle of Q with inclusion onto TTQ denoted by $j_2 : T^{(2)}Q \hookrightarrow TTQ$. A Lagrangian submanifold $\Sigma_L \subset T^*TQ$ can be built as

$$\Sigma_L = \{\mu \in T^*TTQ \mid j_2^*\,\mu = dL\}.$$

Thus, one can obtain via α_{TQ} (where α_{TQ} is the generalization to second-order tangent bundles of the isomorphism α_Q) a new Lagrangian submanifold of the tangent bundle TT^*TQ

which completely determines the equations of motion for the Lagrangian dynamics which are, in a regular case, of Hamiltonian type. Taking this into account, it is clear that Lagrangian systems and Hamiltonian systems are closely related (relationship which will be further studied in this work). Also, in this work, we will comment how to extend this construction in the framework of Lie algebroids.

Many important problems in robotics, the dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics, which typically means mechanical systems with rolling constraints.

A nonholonomic system is a mechanical system subjected to constraint functions which are, roughly speaking, functions on the velocities that are not derivable from position constraints. They arise, for instance, in mechanical systems that have rolling or certain kinds of sliding contact. Traditionally, the equations of motion for nonholonomic mechanics are derived from the Lagrange-d'Alembert principle which restricts the set of infinitesimal variations (or constrained forces) in terms of the constraint functions. In such systems, some differences between unconstrained classical Hamiltonian and Lagrangian systems and nonholonomic dynamics appear. For instance, nonholonomic systems are non-variational in the classical sense, since they arise from the Lagrange-d'Alembert principle and not from Hamilton's principle. Moreover, when the nonholonomic constrains are linear in velocities, then energy is preserved but momentum is not always preserved when a symmetry arises. Nonholonomic systems are described by an almost-Poisson structure but not Poisson (i.e., there is a bracket that together with the energy on the phase space defines the motion, but the bracket generally does not satisfy the Jacobi identity); and finally, unlike the Hamiltonian setting, volume may not be preserved in the phase space, leading to interesting asymptotic stability in some cases, despite energy conservation(see [123],[65]).

As we have commented before, the application of tools from modern differential geometry in the fields of mechanics and control theory has caused an important progress in these research areas. For example, the study of the geometrical formulation of the nonholonomic equations of motion has led to a better comprehension of locomotion generation, controllability, motion planning, and trajectory tracking, raising new interesting questions in these subjects (see [17], [19], [20], [21], [25], [31], [34], [94], [100], [113], [141], [149] and references therein). On the other hand, there are by now many papers in which optimal control problems are addressed using geometric techniques (references [21], [90], [91], and [165] are good examples). In this context, we present a geometrical formulation of the dynamics of higher-order mechanical systems with nonholonomic constraints as higher-order constrained systems.

Thus, in this thesis we will also study optimal control problems of mechanical systems subject to nonholonomic constraints. Of much interest in the present work are the recent developments that utilize a geometric approach and in particular the theory of Lagrangian submanifolds and Lie algebroids. The class of nonholonomic systems we study in this work includes, in particular, any wheeled-type vehicle, such as robots on wheels and or tracks. The fact that most of these robotic systems apply torques and forces internal to the system, which makes these system move in an undulatory fashion (see [149] and references therein for more on undulatory locomotion), without the application of any external forces, makes the system underactuated. Hence, including underactuated systems in our study is crucial in covering a wide range of robotic applications. Moreover, we can easily extend our framework to an arbitrary Lie algebroid.

Discrete mechanics has become a field of intensive research activity in the last decades [35],[59],[131],[171],[172], [178]. This area allows the construction of integration schemes, the so-called geometric integrators. Many of the geometric properties of mechanical systems in the continuous case admit an appropriate counterpart in the discrete setting, which makes it a rich area to be explored. Mechanical integrators preserve some of the invariants of the mechanical system, such as energy, momentum or the symplectic form (see [92], [93], [131], [178], [179]). In the last years, the variational approach in the construction of geometric integration for mechanical systems has been of great interest within the framework of Geometric Integration (see [131, 171]). This point of view is a clear consequence of a deeper insight into the geometric structure of numerical methods and the geometry of the mechanical systems that they approximate. In particular, this effort has been concentrated on the case of discrete Lagrangian functions L_d on the cartesian product $Q \times Q$ of a differentiable manifold. This cartesian product plays the role of a discretized version of the standard velocity space TQ. Applying a natural discrete variational principle and assuming a regularity condition, one obtains a second order recursion operator $F_{L_d} : Q \times Q \to Q \times Q$ assigning to each input pair (q_k, q_{k+1}) the output pair (q_{k+1}, q_{k+2}). When the discrete Lagrangian is an approximation of the integral action we obtain a numerical integrator which inherits some of the geometric properties of the continuous Lagrangian (symplecticity, momentum preservation).

For instance, let us consider the following discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$

$$L_d(q_0, q_1) = \frac{h}{2}\left(\frac{q_1 - q_0}{h}\right)^T M \left(\frac{q_1 - q_0}{h}\right) - hV(q_0),$$

where $Q = \mathbb{R}^n$, which is the very simple approximation to the action integral $\mathcal{A}_L$ for $L : TQ \to \mathbb{R}$ given by

$$\mathcal{A}_L = \int_0^T L(q, \dot{q})\, \mathrm{d}t,$$

using the rectangle rule. Here, $q_0 \approx q(0)$ and $q_1 \approx q(h)$ shall be thought of as being two points on a curve in Q at time h apart. Consider a discrete curve of points $\{q_k\}_{k=0}^N$, also belonging to Q, and calculate the discrete action along this sequence by summing the discrete Lagrangian on each adjacent pair, that is

$$\mathcal{A}_{L_d} = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}),$$

which are the discrete counterpart of $\mathcal{A}_L$. Following the continuous derivation above, we compute variations of this action sum with the boundary points q_0 and q_N held fixed. This gives the discrete Euler-Lagrange equations:

$$D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) = 0,$$

which is the discrete counterpart of Euler-Lagrange equations

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

and must hold for each k. For the particular L_d chosen above, the discrete Euler-Lagrange equations give

$$M\left(\frac{q_2 - 2q_1 + q_0}{h^2}\right) = -\nabla V(q_0).$$

This is clearly a discretization of Newton's equations using a simple finite difference rule for the derivative. This kind of integrators are called variational integrators because of its procedure of derivation. Furthermore, as mentioned above, and also due to its variational nature, these integrators are symplectic and have the property of conserving momentum maps arising from symmetry group actions.

Although this type of geometric integrators have been mainly considered for conservative systems, the extension to geometric integrators for more involved situations is relatively easy, since, in some sense, many of the constructions mimic the corresponding ones for the continuous counterpart. In this sense, it has been recently shown how discrete variational mechanics can include forced or dissipative systems, holonomic constraints, explicitly time-dependent systems, frictional contact, nonholonomic constraints, etc. All these geometric integrators have demonstrated, in worked examples, an exceptionally good longtime behavior and obviously this research is of great interest for numerical and geometric considerations ([79, 160]). In addition, there are several extensions of variational integrators for systems defined in spaces different from $Q \times Q$, such as Lie algebras, reduced spaces, etc, which are of great interest in realistic systems coming from physics, engineering and other applied sciences. The generalization of variational integrators to more involved geometric scenarios can be enshrined in the program initiated by Alan Weinstein [176], which will be detailed below.

The variational view of discrete mechanics and its numerical implementation is further developed in Wendlandt and Marsden ([178, 179]) and then extended in Kane, Marsden and Ortiz ([92]), Marsden, Pekarsky and Shkoller ([128, 129]), Bobenko and Suris ([28, 29]) and Kane, Marsden, Ortiz and West ([93]). Central references of this thesis are based on the works by Marsden and West ([131]) and Marrero, Martinez and Martín de Diego ([124, 125]).

A step further, Alan Weinstein began the study of discrete mechanics on Lie groupoids. His attention was called by the work by Moser and Veselov [140], where the authors study the complete integrability of certain discrete dynamical systems. Moreover the authors describe the Lagrangian and Hamiltonian formalisms for discrete mechanics in two different settings: $Q \times Q$ and a Lie group. Therefore, in [176] Weinstein described versions of the Lagrangian formalism for discrete and continuous time which are general enough to include both constructions used by Moser and Veselov, as well as a Lagrangian formalism on Lie algebras due essentially to Poincaré [156]. In the discrete version, the Lagrangian function is defined on a Lie groupoid.

A Lie groupoid G is a natural generalization of the concept of a Lie group, where now not all elements are composable. The product $g_1 g_2$ of two elements is only defined on the set of composable pairs $G_2 = \{(g, h) \in G \times G \,|\, \beta(g) = \alpha(h)\}$, where $\alpha : G \to Q$ and $\beta : G \to Q$ are the source and target maps over a base manifold Q. This concept was introduced in differential geometry by Ereshmann in the 1950's. The infinitesimal version of a Lie groupoid G is the Lie algebroid $AG \to Q$, which is the restriction of the vertical bundle of α to the submanifold of the identities.

A complete description of the discrete Lagrangian and Hamiltonian mechanics on Lie groupoids was given in the work by Marrero, Martín de Diego and Martínez [124]. In this work, we generalize the theory of discrete second-order Lagrangian mechanics and variational integrators for a second-order discrete Lagrangian $L : G_2 \to \mathbb{R}$ in two main directions. First, we develop variational principles for higher-order variational problems on Lie groupoids and we show how to apply this theory to the construction of variational integrators for optimal control problems of mechanical systems. Secondly, we show that Lagrangian submanifolds of a symplectic groupoid (cotangent groupoid) give rise to discrete dynamical second-order systems, and we study the properties of these systems, including their regularity and reversibility, from the perspective of symplectic and Poisson geometry. We also develop a reduction by Noether symmetries, and study the relationship between the dynamics and variational principles for these second-order variational problems. Next, we use this framework along with a generalized notion of generating function due to Tulczyjew to develop a theory of discrete constrained Lagrangian mechanics. This allows for systems with arbitrary constraints, including those which are nonholonomic (in an appropriate discrete, variational sense). Our results are strongly based on the paper of J.C. Marrero, D. Martín de Diego and A. Stern [126] but in higher-order theory. We will show the following result in Theorem 5.3.5

Theorem: *Let G be a Lie groupoid over a manifold Q. Let $L : G_2 \to \mathbb{R}$ be a discrete second-order Lagrangian. The discrete second order Euler-Lagrange equation are*

$$\ell^*_{g_k}\left(D_1L(g_k, g_{k+1}) + D_2L(g_{k-1}, g_k)\right) + (r_{g_{k+1}} \circ i)^*\left(D_1L(g_{k+1}, g_{k+2}) + D_2L(g_k, g_{k+1})\right) = 0,$$

for $k = 2, \ldots, N-2$ where ℓ_g and r_g denotes the left and right translation of an element g of the Lie groupoid G.

Finally, we want to point out that in previous approaches (see for example [14] and [52]), the theory of discrete variational mechanics for higher-order systems was derived using a discrete lagrangian $L_d : Q^{k+1} \to \mathbb{R}$ where Q^{k+1} is the cartesian product of $k+1$-copies of the configuration manifold Q. In some sense, this is a very natural discretization since we are using $k+1$-points to approximate the positions and the higher-order velocities (such as the standard velocities, accelerations, jerks...) which represents the higher-order tangent bundle $T^{(k)}Q$.

We will see at the end of this this thesis other possibility to work taking a Lagrangian function $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ since the discrete variational calculus is not based on the discretization of the Lagrangian itself, but on the discretization of the associated action. We will see that the appropriate approximation of the action

$$\int_0^T L(q, \dot{q}, \ldots, q^{(k)})\, dt \tag{3}$$

is given by a Lagrangian of the form $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$. Moreover, we will derive a particular choice of discrete Lagrangian which gives an exact correspondence between discrete and continuous systems, the exact discrete Lagrangian.

In this sense the theory of variational integrators for higher-order system is even simpler, since it fits directly into the standard discrete mechanics theory of Marsden and West [131] for a discrete lagrangian of the form $L_d : M \times M \to \mathbb{R}$ where $M = T^{(k-1)}Q$. We will see in some numerical simulations the numerical efficiency of these methods.

Outline of the thesis

Here let us point out the organization of the present thesis and give a brief description of every chapter:

- Chapter 1 gives a brief review of several differential geometric tools used throughout this work.

- Chapter 2 explore some geometric techniques to describe the formulation of first-order and higher-order mechanics. We give a brief review of the description of classical mechanics in terms of Lie algebroids following [111] and we will introduce the constraint algorithm for presymplectic Lie algebroids constructed in [87] which generalizes the well-known Gotay-Nester-Hinds algorithm [73]. We will derive first-order and higher-order dynamics respectively in terms of Lagrangian submanifolds using the Tulczyjew triple constructed in [109] and we will construct a double vector bundle antisymplectomorphism to obtain in an alternative way the dynamics. Also we will give an alternative way to describe the higher-order dynamics in terms of the solution of the Euler-Lagrange equations in Theorem 2.5.6.

- In Chapter 3 we will consider higher-order mechanics from the point of view of the Skinner and Rusk formalism to obtain higher-order Euler-Lagrange equations, higher-order Euler-Poincaré equations and higher-order Lagrange-Poincaré equations. Also, we will focus in the case of systems with higher-order constraints. The extension of this theory to the natural setting of Lie algebroids will be also developed. Moreover, we will characterize geometrically the equations of motion of an optimal control problem for an underactuated mechanical system. Many of the concrete examples under study have additional geometric properties, as for instance, the configuration space is not only a differentiable manifold but it also has a compatible structure of group, that is, the configuration space is a Lie group. We will take advantage of this property to give an intrinsic expression of the equations of motion for higher-order mechanical systems and for optimal control problems with symmetries. The main results in this chapter are given in Equations 3.7, 3.17, 3.27, 3.28, 3.34, 3.35, 3.77, 3.79, 3.90 and 3.91, 3.93, 3.94 and 3.95; Theorems 3.2.1, 3.2.3, 3.3.1, 3.3.3, 3.4.1, 3.6.2; Proposition 3.5.5; and Examples 3.6.6, 3.6.3 and 3.6.10.

- In Chapter 4 we will study optimal control problems of mechanical systems subject to nonholonomic constraints. Of much interest in this chapter are the recent developments that utilize a geometric approach and in particular the theory of Lagrangian submanifolds and Lie algebroids. Hence, including under-actuated systems in our study is crucial in covering a wide range of robotic applications. The main results in this Chapter are shown in Equations 4.3, 4.5 and 4.7; Definition 4.3.4; and Propositions 4.3.1 and 4.3.6.

- In Chapter 5 we will generalize the theory of discrete higher-order Lagrangian mechanics and variational integrators. We will develop variational principles for second-order variational problems on Lie groupoids and we will show how to apply this theory to the construction of variational integrators for optimal control problems of mechanical systems. Also, we will show that Lagrangian submanifolds of a symplectic groupoid give rise to discrete dynamical second-order systems, and we will study the properties of these systems, including their regularity and reversibility, from the perspective of symplectic and Poisson geometry. Finally we will develop a theory of reduction by Noether symmetries, and we will study the relationship between the dynamics and variational principles for these second-order variational problems. We will extend this framework to the case of higher-order constrained systems. These results are given in Equations 5.20, 5.26, 5.36 and 5.37; Lemma 5.3.4; Theorems 5.3.2, 5.3.5, 5.3.16, 5.3.19; Propositions 5.3.8 and 5.3.9; and Corollary 5.3.17.

- Chapter 6 develop the design of geometric integrators for higher-order variational systems. We show that a regular higher-order Lagrangian system has a unique solution for given nearby endpoint conditions using a direct variational proof of existence and uniqueness of the local boundary values problem using a regularization procedure which it results in the replacement of the variational problem with an equivalent one which is regular at the initial singular point of the problem. The argument follows closely the proof by Patrick [150] for first-order Lagrangians; the formulas, of course, reduce to those in [150] for order 1, but we introduce an additional modification using orthonormal polynomials. We will give the notion of exact discrete Lagrangian for higher-order Lagrangian systems and we will show that if the original Lagrangian is regular then it is also the exact discrete Lagrangian, in the sense of [131]. We will apply this theory to the construction of geometric integrators for optimal control problems of mechanical systems. The new results given in this chapter are refereed in Definition 6.2.1, Theorems 6.3.1, 6.3.2, 6.3.4 and 6.3.6, and Corollary 6.3.5.

- Chapter Conclusions and future work exposes a summary of the main results presented in this thesis, together with some conclusions and the future work which could come from it.

This memory is based on original results published in international journal, papers in review process and some papers in preparation as we comment in Chapter, Conclusions and future work.

Chapter 1

Mathematical background

This chapter gives a brief review of several differential geometric tools used throughout this work. We refer to [1, 2, 17, 30, 59, 98, 112, 117, 130] for more specifications about the topics studied in this chapter.

1.1 Manifolds and tensor calculus

A minimum knowledge in linear algebra, topology and differential geometry is assumed in the following. For further understanding in this topic, references [1, 2, 98, 174] are very useful.

The basic idea of a manifold is to introduce spaces which are locally like Euclidean spaces and with structure enough so that differential calculus can be carried over. The manifolds we deal with will be assumed to belong to the C^∞-category. We shall further suppose that all manifolds are finite-dimensional, paracompact and Hausdorff.

Two interesting examples of manifolds which will be extensively used throughout this dissertation are the *tangent* and *cotangent bundles* and their generalizations, Lie algebroids and the corresponding duals.

The tangent bundle of a manifold Q is the collection of all the tangent vectors to Q at each point. We will denote it by TQ. The tangent bundle projection, which assigns to each tangent vector its base point is denoted by $\tau_{TQ} : TQ \to Q$. Given a tangent space T_qQ, we denote the dual space, i.e. the space of linear functions from T_qQ to $\mathbb{R}$ by T_q^*Q. The cotangent bundle T^*Q of a manifold Q is the space formed by the collection of all the dual spaces T_q^*Q. Elements $\alpha \in T_q^*Q$ are called dual vectors or covectors. The cotangent bundle projection, which assigns to each covector its base point, is denoted by $\pi_{T^*Q} : T^*Q \to Q$.

Let $f : Q \to N$ be a smooth mapping between manifolds Q and N. We write $Tf : TQ \to TN$ to denote the tangent map or differential of f. The set of all smooth mappings from Q to N will be denoted by $C^\infty(Q, N)$. When $N = \mathbb{R}$ we shall denote the set of smooth real-valued functions on Q by $C^\infty(Q)$.

A *vector field* X on Q is a smooth mapping $X : Q \to TQ$ which assigns to each point $q \in Q$ a tangent vector $X(q) \in T_qQ$ or, $\tau_{TQ} \circ X = Id_Q$. The set of all vector fields over Q is denoted by $\mathfrak{X}(Q)$. An *integral curve* of a vector field X is a curve satisfying $\dot{c}(t) = X(c(t))$.

Given $q \in Q$, let $\phi_t(q)$ denote the maximal integral curve of X, $c(t) = \phi_t(q)$ starting at q, i.e. $c(0) = q$. Here "maximal" means that the interval of definition of $c(t)$ is maximal. It is easy to verify that $\phi_0 = Id_Q$ and $\phi_t \circ \phi_s = \phi_{t+s}$, whenever the composition is defined. The flow of a vector field X is then determined by the collection of mappings $\phi_t : Q \to Q$. From the definition, they satisfy

$$\frac{d}{dt}(\phi_t(q)) = X(\phi_t(q)), \quad t \in (-\epsilon_1(q), \epsilon_2(q)) \quad \forall q \in Q.$$

Locally, a curve $t \mapsto (c_1(t), \ldots, c_n(t))$ is an integral curve of X when the following system of ordinary differential equations hold

$$\begin{aligned} \frac{dc^1}{dt}(t) &= X^1(c^1(t), \ldots, c^n(t)) \\ &\vdots \\ \frac{dc^n}{dt}(t) &= X^n(c^1(t), \ldots, c^n(t)). \end{aligned}$$

The previous system is called *autonomous* since there are not explicit dependence of time on the right hand side. If $X : \mathbb{R} \times Q \to TQ$ verifies $\tau_{TQ} \circ X = pr_Q$ where $pr_Q : \mathbb{R} \times Q \to Q$ is the natural projection we will say that X is a time-dependent vector field. The integral curves are the solutions of an explicit time-dependent system

$$\frac{dc^i}{dt}(t) = X^i(t, c(t))$$

where $X = X^i(t, q)\dfrac{\partial}{\partial q^i}$.

In a similar way to the definition of vector fields, a *one-form* α on Q is a mapping $\alpha : Q \to T^*Q$ such that $\pi_{T^*Q} \circ \alpha = \mathrm{Id}_Q$. In other words, it assigns to each point $q \in Q$ a covector $\alpha(q) \in T_q^*Q$. The set of all the one-forms over Q is denoted by $\Omega^1(Q)$. As is well established in linear algebra, there always exists a bilinear natural pairing between a vector space V and its dual vector space V^*. Here $\langle \cdot\, , \cdot \rangle$ denotes such a pairing: $\langle \cdot\, , \cdot \rangle : V \times V^* \to \mathbb{R}$. In consequence, one always can define a natural pairing between elements of the tangent and contangent bundles; $\langle \cdot\, , \cdot \rangle_q : T_qQ \times T_q^*Q \to \mathbb{R}$.

Vector fields and one-forms are particular cases of a more general object, called *tensor fields.* Given $r, s \in \mathbb{N} \cup \{0\}$, an *$r$-contravariant and s-covariant tensor field* t on Q is a C^∞-section of $T_s^r(Q)$; that is, it associates to each point $q \in Q$ an $\mathbb{R}$-multilinear mapping

$$t(q) : \underbrace{(T_q^*Q \times \ldots \times T_q^*Q)}_{r-times} \times \underbrace{(T_qQ \times \ldots \times T_qQ)}_{s-times} \to \mathbb{R}.$$

It is common to say that t is a (r, s)-tensor field on Q. Thus, a vector field is a $(1, 0)$-tensor field on Q and a 1-form is a $(0, 1)$-tensor field on Q.

The *tensor product* of a $(r, s)-$tensor field, t, and a $(r', s')-$tensor field, t', is the $(r+r', s+s')-$tensor field $t \otimes t'$ defined by

$$(t \otimes t')(q)(\omega_1, \dots, \omega_r, \mu_1, \dots, \mu_{r'}, v_1, \dots, v_s, w_1, \dots, w_{s'}) =$$
$$t(q)(\omega_1, \dots, \omega_r, v_1, \dots, v_s) \cdot t'(q)(\mu_1, \dots, \mu_{r'}, w_1, \dots, w_{s'}),$$

where $q \in Q, v_i, w_i \in T_qQ$ and $\omega_j, \mu_j \in T_q^*Q$.

A special subset of tensor fields is $\Omega^k(Q) \subset T_k^0 Q$, the set of all $(0, k)$ skew-symmetric tensor fields. The elements of $\Omega^k(Q)$ are called $k-$forms.

The *alternation map* $A : T_0^k Q \to \Omega^k(Q)$ is defined by

$$A(t)(v_1, ..., v_k) = \frac{1}{k!} \sum_{\sigma \in \Sigma_k} \operatorname{sign}(\sigma) t(v_{\sigma(1)}, ..., v_{\sigma(k)}),$$

where Σ_k is the set of $k-$permutations. Is easy to see that A is linear, $A\big|_{\Omega^k(Q)} = \mathrm{Id}$ and $A \circ A = A$.

The *wedge product* or *exterior product* between $\alpha \in \Omega^k(Q)$ and $\beta \in \Omega^l(Q)$ is the form $\alpha \wedge \beta \in \Omega^{k+l}(Q)$ defined by

$$\alpha \wedge \beta = \frac{(k+l)!}{k!l!} A(\alpha \otimes \beta).$$

Some important properties of the wedge product are the following:

1. $\wedge$ is bilinear and associative.
2. $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$, where $\alpha \in \Omega^k(Q)$ and $\beta \in \Omega^l(Q)$.

The *algebra of exterior differential forms*, represented by $\Omega(Q)$, is the direct sum of $\Omega^k(Q)$, $k = 0, 1, \dots$, together with its structure as an infinite-dimensional real vector space and with the multiplication $\wedge$.

The *exterior derivative,* represented by d, is defined as the unique family of mappings $\mathrm{d}^k(U) : \Omega^k(U) \to \Omega^{k+1}(U)$ ($k = 0, 1, ...$ and $U \subset Q$ open) such that (see [1, 174]):

1. d is a $\wedge$-antiderivation, that is, d is $\mathbb{R}-$linear and $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$, where $\alpha \in \Omega^k(Q)$ and $\beta \in \Omega^l(Q)$.
2. $\mathrm{d}f = p_2 \circ Df$, for $f \in C^\infty(U)$, with p_2 the canonical projection of $T\mathbb{R} \simeq \mathbb{R} \times \mathbb{R}$ onto the second factor.
3. $\mathrm{d} \circ \mathrm{d} = 0$.
4. d is natural with respect to inclusions, that is, if $U \subset V \subset Q$ are open sets of Q, then $\mathrm{d}(\alpha\big|_U) = \mathrm{d}(\alpha)\big|_U$, where $\alpha \in \Omega^k(V)$.

Let $f : Q \to N$ be a smooth mapping and $\omega \in \Omega^k(N)$. Define the *pull-back* $f^*\omega$ of ω by f as

$$(f^*\omega(q))(v_1, ..., v_k) = \omega(f(q))(T_qf(v_1), ..., T_qf(v_k)),$$

where $v_i \in T_qQ$ with $i = 1, \dots, k$. Note that the pull-back defines the mapping $f^* : \Omega^k(N) \to \Omega^k(Q)$. The main properties related with the pull-back are the following:

1. $(g \circ f)^* = f^* \circ g^*$, where $f \in C^\infty(Q, N)$ and $g \in C^\infty(N, W)$.
2. $\mathrm{Id}_Q^*\big|_{\Omega^k(Q)} = \mathrm{Id}_{\Omega^k(Q)}$.
3. If $f \in C^\infty(Q, N)$ is a diffeomorphism, then f^* is a vector bundle isomorphism and $(f^*)^{-1} = (f^{-1})^*$.
4. $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$, where $f \in C^\infty(Q, N)$, $\alpha \in \Omega^k(N)$ and $\beta \in \Omega^l(N)$.
5. d is natural with respect to mappings, i.e., for $f \in C^\infty(Q, N)$, $f^*\mathrm{d}\omega = \mathrm{d}f^*\omega$.

Given a vector field $X \in \mathfrak{X}(Q)$ and a function $f \in C^\infty(Q)$ the *Lie derivative of f with respect to X*, $\mathcal{L}_X f \in C^\infty(Q)$, is defined as

$$\mathcal{L}_X f(q) = \mathrm{d}f(q)[X(q)].$$

The operation $\mathcal{L}_X : C^\infty(Q) \to C^\infty(Q)$ is a derivation, i.e. it is $\mathbb{R}-$linear and $\mathcal{L}_X(fg) = \mathcal{L}_X(f)g + f\mathcal{L}_X(g)$, for any $f, g \in C^\infty(Q)$.

Given two vector fields $X, Y \in \mathfrak{X}(Q)$ we may define the $\mathbb{R}-$linear derivation

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X.$$

This enables us to define the *Lie derivative of Y with respect to X*, $\mathcal{L}_X Y = [X, Y]$ as the unique vector field such that $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$. Some important properties are:

1. $\mathcal{L}_X$ is natural with respect to restrictions, i.e., for $U \subset Q$ open, $[X\big|_U, Y\big|_U] = [X, Y]\big|_U$ and $(\mathcal{L}_X f)\big|_U = \mathcal{L}_X\big|_U(f\big|_U)$, for $f \in C^\infty(Q)$.
2. $\mathcal{L}_X(f\,Y) = (\mathcal{L}_X f)Y + f(\mathcal{L}_X Y)$, for $f \in C^\infty(Q)$.

There is also another natural operator associated with a vector field X. Let $\omega \in \Omega^k(Q)$. The *inner product* or *contraction of X* and ω, $i_X\omega \in \Omega^{k-1}(Q)$, is defined by

$$i_X\omega(q)(v_1, ..., v_{k-1}) = \omega(q)(X(q), v_1, ..., v_{k-1}),$$

where $v_i \in T_qQ$ with $i = 1, \ldots, k$. The operator i_X is an $\wedge-$antiderivation, that is, it is $\mathbb{R}-$linear and $i_X(\alpha \wedge \beta) = (i_X\alpha) \wedge \beta + (-1)^k\alpha \wedge (i_X\beta)$, where $\alpha \in \Omega^k(Q)$ and $\beta \in \Omega^l(Q)$. Also, for $f \in C^\infty(Q)$, we have that $i_{(f\,X)}\alpha = f(i_X\alpha)$.

Finally, we conclude this section by stating some relevant properties involving d, i_X and $\mathcal{L}_X$. For arbitrary $X, Y \in \mathfrak{X}(Q)$, $f \in C^\infty(Q)$ and $\alpha \in \Omega^k(Q)$, we have

1. $\mathrm{d}\mathcal{L}_X\alpha = \mathcal{L}_X\mathrm{d}\alpha$.
2. $i_X\mathrm{d}f = \mathcal{L}_X f$.
3. $\mathcal{L}_X\alpha = i_X\mathrm{d}\alpha + \mathrm{d}i_X\alpha$.
4. $\mathcal{L}_{(f\,X)}\alpha = f\mathcal{L}_X\alpha + \mathrm{d}f \wedge i_X\alpha$.
5. $i_{[X,Y]}\alpha = \mathcal{L}_X i_Y\alpha - i_Y\mathcal{L}_X\alpha$.

1.2 Distributions and codistributions

Definition 1.2.1. *Let Q be an $n-$dimensional differentiable manifold. A k-dimensional distribution $\mathcal{D}$ on a manifold Q, is a k-dimensional subspace $\mathcal{D}_q$ of T_qQ for each $q \in Q$. $\mathcal{D}$ is smooth if for each $q \in Q$ there is a neighborhood U of q and there are k C^∞ vector fields $X_1, \dots, X_k$ on U which span $\mathcal{D}$ at each point of U; that is,*

$$\mathcal{D}_q = span\{X_1, \dots, X_k\}$$

In other words, for every $q \in Q$, $\mathcal{D}_q$ is a vector subspace of T_qQ. The *rank of* $\mathcal{D}$ *at* $q \in Q$ is the dimension of the subspace $\mathcal{D}_q$, i.e. $\varrho : Q \to \mathbb{R}$, $\varrho(q) = \dim \mathcal{D}_q$. For any $q_0 \in Q$ it is clear that $\varrho(q) \geq \varrho(q_0)$ in a neighborhood of q_0. If ϱ is a constant function, then $\mathcal{D}$ is called a *regular distribution.*

The following diagram illustrates the situation

$$\begin{tikzcd} \mathcal{D} \arrow[rr, hook, "i"] \arrow[dr, "\tau_{\mathcal{D}}"'] & & TQ \arrow[dl, "\tau_{TQ}"] \\ & Q, & \end{tikzcd}$$

where $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ is the restriction of τ_{TQ} to $\mathcal{D}$, that is, $\tau_{\mathcal{D}} = \tau_{TQ} \mid_{\mathcal{D}}$ and i represents the inclusion map.

Definition 1.2.2. 1. *A submanifold $S \hookrightarrow Q$ is said to be an* **integral submanifold** *of a smooth distribution $\mathcal{D} \hookrightarrow TQ$ if $TS = \mathcal{D}$ along the points of S.*

2. *Let $\mathcal{D}$ be a smooth distribution on Q such that through each point of Q there passes an integral manifold of $\mathcal{D}$. Then $\mathcal{D}$ is* **completely integrable**.

3. *A smooth distribution $\mathcal{D}$ is* **involutive** *if $[X, Y] \in \Gamma(\tau_{\mathcal{D}})$ for every $X, Y \in \Gamma(\tau_{\mathcal{D}})$, that is, it is closed under the Lie bracket operation. Here $\Gamma(\tau_D)$ denotes the set of sections of $\mathcal{D}$.*

Theorem 1.2.3 (Frobenius' Theorem). *A smooth distribution $\mathcal{D}$ is completely integrable if and only if it is involutive.*

In a equivalent fashion as for distributions, it is possible to define *codistributions.* Let Q be a manifold. A *smooth regular codistribution* $\widetilde{\mathcal{D}}$ on T^*Q is a subbundle of T^*Q with k-dimensional fiber. The following diagram show the situation:

$$\begin{tikzcd} \widetilde{\mathcal{D}} \arrow[rr, hook, "i"] \arrow[dr, "\pi_{\widetilde{\mathcal{D}}}"'] & & T^*Q \arrow[dl, "\pi_{T^*Q}"] \\ & Q, & \end{tikzcd}$$

where $\pi_{\widetilde{\mathcal{D}}} : \widetilde{\mathcal{D}} \to Q$ is the restriction of π_{T^*Q} to $\widetilde{\mathcal{D}}$, that is, $\pi_{\widetilde{\mathcal{D}}} = \pi_{T^*Q} \mid_{\widetilde{\mathcal{D}}}$.

Given the concept of codistribution, it is possible to define the *annihilator* of a distribution. Let $\mathcal{D} \hookrightarrow TQ$ be a distribution, the annihilator of $\mathcal{D}$ is a codistribution $\mathcal{D}^\circ : \mathrm{Dom}\,\mathcal{D} \subset Q \to T^*Q$, given by

$$\mathcal{D}^\circ_q = (\mathcal{D}_q)^\circ = \left\{\alpha \in T^*_q Q \mid \alpha(v) = \langle \alpha, v\rangle = 0,\ \forall v \in \mathcal{D}_q\right\}$$

for every $q \in Q$.

A submanifold S of Q will be an integral submanifold of a distribution $\mathcal{D}$ if

$$T_s S^\circ = \mathcal{D}^\circ_q,\ \text{for all } s \in S.$$

In particular, this implies that the rank of $\mathcal{D}$ is constant along S.

1.3 Lie groups and Lie algebras

Lie groups arise in discussing conservation laws for mechanical and control systems and in the analysis of systems with some underlying symmetry [17, 95]. In this section, we will recall the key notations and facts from the theory of Lie group and Lie algebra.

1.3.1 Lie groups

Roughly speaking, a Lie group is a manifold on which the group operations, *product* and *inverse*, are defined.

Definition 1.3.1. *A nonempty collection G of transformations of some set is called a (transformation) group if along every two transformations $g, h \in G$ belonging to the collection, the composition $g \circ h$ and the inverse transformation g^{-1} belong to the same collection G.*

It follows from this definition that every group contains the *identity* transformation e. Also the composition of transformations is an associative operation. These properties, associativity and the existence of the unit and inverse of each element, are often taken as the definition of an abstract group. Here we employ the point of view of V.I. Arnold, that every group should be viewed as the group of transformations of some set, and the usual "axiomatic" definition of a group only obscures its true meaning (cf.[5] p.58).

The groups we are concerned in this thesis are so-called Lie groups. In addition to being a group, they carry the structure of a smooth manifold such that the multiplication and inversion respect this structure.

Definition 1.3.2. *A Lie group is a differentiable manifold G with a group structure such that the multiplication $G \times G \to G$ and the inversion $G \to G$ are smooth maps.*

The dimension of a Lie group G is defined to be the dimension of G as a manifold. The product symbol may be omitted and $g_1 \cdot g_2$ is usually written as $g_1 g_2$. The inverse element will be denoted by g^{-1} and the identity element by e.

Let G be a Lie group and $H \subset G$ a Lie subgroup of G. Define the equivalence relation $\sim$ by $g \sim g'$ if there exists an element $h \in H$ such that $g' = gh$. An equivalence class $[g]$

is a set $\{gh \,|\, h \in H\}$. The *coset space* G/H is a manifold (not necessarily a Lie group) with dim $G/H =$ dim $G-$ dim H. G/H is a Lie group if H is a *normal subgroup* of G, that is, if $ghg^{-1} \in H$ for any $g \in G$ and $h \in H$. In fact, take equivalence classes $[g], [g'] \in G/H$ and construct the product $[g][g'] = [gg']$. If the group structure is well defined in G/H, the product must be independent of the choice of the representatives. Let gh and $g'h'$ be the representatives of $[g]$ and $[g']$ respectively. Then $ghg'h' = gg'h''h' \in [gg']$ where the equality follows since there exists $h'' \in H$ such that $hg' = g'h''$.

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 1.3.3. 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 $GL(n, R)$ with the matrix multiplication, and several of its Lie subgroups: the group of rigid motions in $3-$dimensional Euclidean space, $SE(3)$; the group of rigid motions in the plane, $SE(2)$; and the group of rotations in $\mathbb{R}^3, SO(3)$.

1.3.2 Lie algebras

Definition 1.3.4. *A Lie algebra over* $\mathbb{R}$ *is a real vector space* $\mathfrak{g}$ *together with a bilinear operation* $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$*, called Lie bracket, such that, for all* $\xi_1, \xi_2, \xi_3 \in \mathfrak{g}$*,*

1. $[\xi_1, \xi_2] = -[\xi_2, \xi_1]$ *skew-symmetry.*
2. $[\xi_1, [\xi_2, \xi_3]] + [\xi_3, [\xi_1, \xi_2]] + [\xi_2, [\xi_3, \xi_1]] = 0$ *Jacobi identity.*

Definition 1.3.5. *A map* $\varphi : \mathfrak{g} \to \mathfrak{h}$ *between two Lie algebras is a Lie algebra homomorphism if it satisfies*

$$\varphi([X, Y]) = [\varphi(X), \varphi(Y)],$$

for all $X, Y \in \mathfrak{g}$.

Locally, if we denote by $\{e_1, \ldots, e_n\}$ a basis of the Lie algebra $\mathfrak{g}$ then we have the relation

$$[e_i, e_j] = \mathcal{C}_{ij}^k e_k,$$

where $\mathcal{C}_{ij}^k$ are called the *structure constants* of the Lie algebra $\mathfrak{g}$.

We will also need another important class of maps between Lie algebras called *derivations:*

Definition 1.3.6. *A linear map* $\delta : \mathfrak{g} \to \mathfrak{g}$ *of a Lie algebra to itself is called a derivation if it satisfies*

$$\delta([X, Y]) = [\delta(X), Y] + [X, \delta(Y)],$$

for all $X, Y \in \mathfrak{g}$.

The map $ad_X : \mathfrak{g} \to \mathfrak{g}$ associated to a fixed vector $X \in \mathfrak{g}$ via

$$ad_X(Y) = [X, Y]$$

is a derivation for any choice of X as a consequence of the Jacobi identity (see [95]). If a derivation of a Lie algebra $\mathfrak{g}$ can be expressed in the form ad_X for some $X \in \mathfrak{g}$, it is called an *inner derivation.*

Definition 1.3.7. *A subalgebra of a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h} \subset \mathfrak{g}$ invariant under the Lie bracket in $\mathfrak{g}$. An ideal of a Lie algebra $\mathfrak{g}$ is a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ such that $[X, \mathfrak{h}] \subset \mathfrak{h}$ for all $X \in \mathfrak{g}$.*

The importance of ideals comes from the fact that if $\mathfrak{h} \subset \mathfrak{g}$ is an ideal then the quotient space $\mathfrak{g}/\mathfrak{h}$ is again a Lie algebra.

Now, we will associate to any Lie group a Lie algebra. For it, we will need the following definition.

Definition 1.3.8. *Let h and g be elements of a Lie group G. The right-translation $R_h : G \to G$ and the left-translation $L_h : G \to G$ are defined by*

$$R_h(g) = gh \text{ and } L_h(g) = hg. \tag{1.1}$$

By definition, R_h and L_h are diffeomorphisms from G to G. Hence, the maps $L_h : G \to G$ and $R_h : G \to G$ induce the applications $T_g L_h : T_g G \to T_{hg} G$ and $T_g R_h : T_g G \to T_{gh} G$. Since these translations give equivalent theories, we are concerned mainly with the left-translation in the following. The analysis based on the right-translation can be carried out in a similar way.

Given a Lie group G, there exists a special class of vector fields characterized by the *invariance* under group action.

Definition 1.3.9. *Let X be a vector field on a Lie group G. X is said to be a left-invariant vector field if*

$$(T_g L_h) X(g) = X(hg)$$

and X is said to be right-invariant if

$$(T_g R_h) X(g) = X(gh)$$

A vector $\xi \in T_e G$ defines unique left-invariant and right-invariant vector field $\overleftarrow{\xi}$ and $\overrightarrow{\xi}$ respectively, throughout G by

$$\overleftarrow{\xi}(g) = T_e L_g \xi, \quad g \in G,$$
$$\overrightarrow{\xi}(g) = T_e R_g \xi, \quad g \in G.$$

Observe that,

$$\overleftarrow{\xi}(hg) = T_e L_{hg} \xi = T_e(L_h \circ L_g)\xi = (T_e L_h) \circ (T_e L_g)\xi = T_g L_h \overleftarrow{\xi}(g).$$

Conversely, a left-invariant vector field $\overleftarrow{\xi}$ defines an unique vector $\xi = \overleftarrow{\xi}(e) \in T_e G$. Let us denote the set of left-invariant vector fields on G by $\mathfrak{g}$. The map $T_e G \to \mathfrak{g}$ defined by $\xi \mapsto \overleftarrow{\xi}$ is a isomorphism, and it follows that the set of left-invariant vector fields is a vector space isomorphic to $T_e G$. In particular, dim $\mathfrak{g}$ = dim G. Moreover, the following property holds

$$[\overleftarrow{\xi}, \overleftarrow{\eta}] = \overleftarrow{[\xi, \eta]},$$

that is, the Lie bracket of two left-invariant vector fields is itself a left-invariant vector field.

Since $\mathfrak{g}$ is identified with a set of vector fields, it is a subset of $\mathfrak{X}(G)$ and the Lie bracket is also defined in $\mathfrak{g}$. We show now that $\mathfrak{g}$ is closed under the Lie bracket. Take two points g and $hg = L_h(g)$ in G. If we apply $T_g L_h$ to the Lie bracket $[\xi, \eta]$ of $\xi, \eta \in \mathfrak{g}$, we have that

$$T_g L_h \left(\overleftarrow{[\xi, \eta]}(g) \right) = [T_g L_h \overleftarrow{\xi}(g), T_g L_h \overleftarrow{\eta}(g)] = \overleftarrow{[\xi, \eta]}(hg),$$

where the left-invariance of $\overleftarrow{\xi}, \overleftarrow{\eta}$ has been used. Thus $[\xi, \eta] \in \mathfrak{g}$, i.e. $\mathfrak{g}$ is closed under the Lie bracket. Finally, the *Lie algebra* of G is defined as the set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket.

Definition 1.3.10. *The set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is called the Lie algebra of a Lie group G.*

We denote the Lie algebra of a Lie group by the corresponding lower-case German gothic letter. For instance, $\mathfrak{so}(3)$ is the Lie algebra of the Lie group $SO(3)$.

To complete this subsection is necessary to introduce the definition of *exponential map.*

Definition 1.3.11. *Let G be a Lie group and $\mathfrak{g}$ its associated Lie algebra. For all $\xi \in \mathfrak{g}$, let $\gamma_\xi : \mathbb{R} \to G$ denote the integral curve of the left-invariant vector field $\overleftarrow{\xi}$ induced by ξ, which is defined uniquely by claiming*

$$\overleftarrow{\xi}(e) = \xi, \quad \gamma_\xi(0) = e, \quad \gamma'_\xi(t) = \overleftarrow{\xi}(\gamma_\xi(t)) \text{ for all } t \in \mathbb{R}.$$

The map

$$\exp : \mathfrak{g} \to G, \quad \exp(\xi) = \gamma_\xi(1)$$

is called the exponential map of the Lie algebra $\mathfrak{g}$ in the Lie group G.

1.3.3 Action of a Lie group on a manifold

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

Let Q be a manifold and let G be a Lie group. A *(left) action* of a Lie group G is a smooth mapping $\Phi : G \times Q \to Q$ such that

i) $\Phi(e, q) = q$ for all $q \in Q$, and

ii) $\Phi(g, \Phi(h, q)) = \Phi(gh, q)$ for all $g, h \in G$ and $q \in Q$.

A *right action* is a smooth mapping $\Psi : Q \times G \to Q$ that satisfies $\Psi(q, e) = q$ and $\Psi(\Psi(q, g), h) = \Psi(q, gh)$ for all $g, h \in G$ and $q \in Q$.

Definition 1.3.12. *A function F is invariant with respect to an action Φ of a Lie group G if, for every $g \in G$, the map Φ_g is a symmetry of F, that is, $F \circ \Phi_g = F$.*

Normally, we will only be interested in the action as a mapping from Q to Q, and so we will write the action as $\Phi_g : Q \to Q$, 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 gq instead of $\Phi_g(q)$.

The *orbit* of the $G-$action through a point q is $Orb_G(q) = \{gq | g \in G\}$. 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 $\tilde{\Phi} : G \times Q \to Q \times Q$ 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 Φ is a free and proper action, then Φ is regular, and therefore Q/G is a smooth manifold and $\pi : Q \to Q/G$ is a submersion.

Let ξ 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 Φ^{ξ} as a flow of a vector field on the manifold Q. Consequently, it determines a vector field on Q, given by

$$\xi_Q(q) = \left.\frac{d}{dt}\right|_{t=0} (\Phi(exp(t\xi), q)), \tag{1.2}$$

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

$$\xi_G(h) = \left.\frac{d}{dt}\right|_{t=0} (exp(t\xi) \cdot h) = T_e R_h \xi,$$

that is, the right-invariant vector field defined by ξ.

Remark 1.3.13. For the standard action of a matrix Lie group on $\mathbb{R}^n$, the expression $exp(t\xi)q$ is just the matrix product of ξ and q, that is,

$$\xi_Q(q) = \left.\frac{d}{dt}\right|_{t=0} (exp(t\xi)q) = \left(\left.\frac{d}{dt}\right|_{t=0} exp(t\xi) \right) q = \xi q \text{ (matrix product)}$$

$\diamond$

Remark 1.3.14. Given a G action on Q, the set of infinitesimal vector fields forms a subalgebra of $\mathfrak{X}(Q)$ since the map $\mathfrak{g} \to \mathfrak{X}(Q)$, $\xi \mapsto \xi_Q$, is linear and satisfies

$$[\xi, \eta] = -[\xi_Q, \eta_Q]$$

for all $\xi, \eta \in \mathfrak{g}$. Such map is called a *Lie algebra anti-homomorphism.* $\diamond$

Example 1.3.15. An action Φ (left or right) of G on a manifold Q induces an action of the Lie group on the tangent bundle of Q, $\hat{\Phi} : G \times TQ \to TQ$ defined by $\hat{\Phi}(g, v_q) = T\Phi_g(v_q) = (\Phi_g(q), T_q\Phi_g(v_q))$ for any $g \in G$ and $v_q \in T_qQ$. $\hat{\Phi}$ is called the *tangent lift* of the action Φ.

Also from Φ one can induce an action of the Lie group on the cotangent bundle of Q, $\widetilde{\Phi} : G \times T^*Q \to T^*Q$ defined by $\widetilde{\Phi}(g, \alpha_q) = T\Phi_{g^{-1}}(\alpha_q) = (\Phi_g(q), T^*_{\Phi_g(q)}\Phi_{g^{-1}}(\alpha_q))$ for any $g \in G$ and $\alpha_q \in T^*_qQ$. $\widetilde{\Phi}$ is called the *cotangent lift* of the action Φ.

If Φ is a left action (resp. right), then the tangent lift and cotangent lift actions are left (resp. right) actions (see [82]).

1.3.4 The adjoint and coadjoint representations

A *representation* of a Lie group G on a real vector space V is a linear action Φ of the group G on V that is smooth in the sense that the map $G \times V \to V$, $(g, v) \mapsto gv$, is smooth. Every Lie group has two distinguished representations: the adjoint and the coadjoint representations.

Any element $g \in G$ defines an automorphism c_g of the group G by conjugation:

$$c_g : h \in G \mapsto ghg^{-1}.$$

The differential of c_g at the identity $e \in G$ maps the Lie algebra of G to itself and thus defines an element $Ad_g \in \mathrm{Aut}(\mathfrak{g})$, the group of all automorphisms of the Lie algebra $\mathfrak{g}$.

Definition 1.3.16. *The map $Ad : G \to Aut(\mathfrak{g})$, $g \mapsto Ad_g$ defines a representation of the group G on the space $\mathfrak{g}$ and is called the group of the adjoint representation.*

Roughly speaking, the adjoint representation measures the non-commutativity of the multiplication of the Lie group: if G is Abelian, then the adjoint action Ad_h 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 differential of $Ad : G \to \mathrm{Aut}(\mathfrak{g})$ at the group identity e defines a map $ad : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, the *adjoint representation of the Lie algebra* $\mathfrak{g}$.

The dual object to the adjoint representation of a Lie group G on its Lie algebra $\mathfrak{g}$ is called the coadjoint representation of G on $\mathfrak{g}^*$.

Definition 1.3.17. *The coadjoint representation Ad^* of the group G on the space $\mathfrak{g}^*$ is the dual of the adjoint representation. Let $\langle \cdot, \cdot \rangle$ denote the pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$. Then the coadjoint action of the group G on the dual space $\mathfrak{g}^*$ is given by the operators $Ad^*_g : \mathfrak{g}^* \to \mathfrak{g}^*$ for any $g \in G$ that are defined by the relation*

$$\langle Ad^*_g(\alpha), \xi \rangle := \langle \alpha, Ad_{g^{-1}}(\xi) \rangle$$

for all $\alpha \in \mathfrak{g}^*$ and $\xi \in \mathfrak{g}$.

The differential $ad^* : \mathfrak{g} \to End(\mathfrak{g}^*)$ of the adjoint $Ad^* : G \to Aut(\mathfrak{g}^*)$ at the group identity $e \in G$ is called the coadjoint representation of the Lie algebra $\mathfrak{g}$. Explicitly is defined by the relation

$$\langle ad^*_\eta \alpha, \xi \rangle = \langle \alpha, ad_\eta \xi \rangle$$

for $\alpha \in \mathfrak{g}^*$ and $\xi, \eta \in \mathfrak{g}$.

1.4 Connections on principal bundles

Roughly speaking, a *connection* tells us how a quantity associated with a manifold changes as we move from one point to another - it "connects" neighboring spaces. In terms of fiber bundles, a connection tells us how movement in the total space induces change along the fiber. Remember that a bundle is a triple (E, p, B), where $p : E \to B$ is a surjective map. The space B is called the base space, the space E is called the total space, and the map p is called projection of the bundle. For each $b \in B$, the space $p^{-1}(b) \in E$ is called the fibre of the bundle over $b \in B$.

With the previous setup, we can define an *Ehresmann connection* on any fiber bundle. Specifically, consider tangent vectors on the total space that lie "along" fibers, i.e., all the vectors in the kernel of Tp, the *vertical subbundle* V of TE. An *Ehresmann connection* A is a vertical $1-$form $A_q : T_qE \to V_q$ which leaves vertical vectors fixed, i.e., $A(v) = v$ for all $v \in V$ [17]. The only other requirement is that this map is linear, i.e., if $A_q : T_qE \to V_q$ is a connection $1-$form then for any scalar values a, b and tangent vectors $v, u \in T_qE$ we must have $A_q(au + bv) = aA_q(u) + bA_q(v)$ at each point $q \in E$.

In mechanics and control problems an important class of connections are *principal connections.* Let $\Phi : G \times Q \to Q$, $(g, q) \mapsto \Phi_g(q)$ be a free and proper left action of a Lie group G on a manifold Q. Thus we get the principal bundle $\pi : Q \to \widehat{Q} := Q/G$, where $\widehat{Q}$ is endowed with the unique manifold structure for which π is a submersion (see [98]).

To any element $\xi \in \mathfrak{g}$ there corresponds a vector field ξ_Q on Q the infinitesimal generator defined in (1.2). Then, for any $q \in Q$ these vector fields generate the *vertical subspace* $V_qQ := \{\xi_Q(q) | \xi \in \mathfrak{g}\} = Ker(T_q\pi)$.

The *adjoint bundle* is defined by $Ad(Q) := Q \times_G \mathfrak{g} \to \widehat{Q}$, where the quotient is taken relative to the action $(g, (q, \xi)) \mapsto (\Phi_g(q), Ad_{g^{-1}}(\xi))$. For $(q, \xi) \in Q \times \mathfrak{g}$ the corresponding element in $Ad(Q)$ is denoted by $[q, \xi]_G$. Moreover, in each fiber $(Ad(Q))_{\pi(q)}$ (depending smoothly for each $x = \pi(q) \in \widehat{Q}$) there is a Lie bracket operation $[\cdot, \cdot]_{\pi(q)}$ given by

$$[[q, \xi]_G, [q, \eta]_G]_{\pi(q)} := [q, [\xi, \eta]]_G$$

for $[q, \xi]_G, [q, \eta]_G \in Ad(Q)$.

Remark 1.4.1. The map $j : Ad(Q) \to TQ/G, \quad [(q, \xi)] \mapsto [\xi_Q(q)]$ induces a monomorphism between the vector bundles $Ad(Q)$ and TQ/G. Thus $Ad(Q)$ may be considered as a vector subbundle of TQ/G. In addition, the space of sections $\Gamma(Ad(Q))$ may be identified with the set of vector fields which are vertical and G-invariant (see [118]).

The tangent bundle to π, $T\pi : TQ \to T\widehat{Q}$ induces an epimorphism $[T\pi] : TQ/G \to T\widehat{Q}$ and $\mathrm{im} j = \ker[T\pi]$. Therefore, we have an exact sequence of vector fields

$$Ad(Q) \xrightarrow{j} TQ/G \xrightarrow{[T\pi]} T\widehat{Q}$$

which is just the *Atiyah sequence associated with the principal bundle* $\pi : Q \to \widehat{Q}$ (see [118] for more details).

$\diamond$

Definition 1.4.2. *Denoting by $\Omega^1(Q,\mathfrak{g})$ the space of $\mathfrak{g}-$valued 1-forms on Q, a principal connection $\mathcal{A}$ on the principal bundle $\pi : Q \to Q/G$ is a 1-form $\mathcal{A} \in \Omega^1(Q,\mathfrak{g})$ such that*

$$\mathcal{A}(\xi_Q(q)) = \xi, \text{ and } \Phi_g^*\mathcal{A} = Ad_g \circ \mathcal{A},$$

where ξ_Q is the infinitesimal generator associated to $\xi \in \mathfrak{g}$ for $q \in Q$.

A connection induces a splitting $T_qQ = V_qQ \oplus H_qQ$ on the tangent space into the vertical and *horizontal subspace* defined by

$$H_qQ := Ker(\mathcal{A}(q)).$$

Finally, the *Cartan structure equations* state that for all vector fields $u, v \in \mathfrak{X}(Q)$ the following identity holds

$$\mathcal{B}(u,v) = d^{\mathcal{A}}\mathcal{A}(u,v) - [\mathcal{A}(u), \mathcal{A}(v)]_{\mathfrak{g}}.$$

This equation introduce the notion of curvature associated with the principal connection $\mathcal{A}$ denoted by $\mathcal{B}$. Moreover, the definition of curvature and exterior differential implies the *Bianchi identity* (see [2] and [98] for example)

$$d^{\mathcal{A}}\mathcal{B} = 0.$$

1.5 Riemannian geometry

In this subsection we recall some facts about Riemannian geometry that will be used later on [30, 44, 102].

Definition 1.5.1. *Let Q be a n-dimensional differentiable manifold. A Riemannian metric $\mathcal{G}$ on Q is a $(0,2)-$tensor on Q which satisfies the following at each point $q \in Q$:*

1. *$\mathcal{G}(v_q, w_q) = \mathcal{G}(w_q, v_q)$, where $v_q, w_q \in T_qQ$, (symmetry).*
2. *$\mathcal{G}(v_q, v_q) \geq 0$, where the equality holds only when $v_q = 0$, (positive-definite).*

In short, a *Riemannian metric* $\mathcal{G}$ is a symmetric positive-definite bilinear form at each $q \in Q$.

The pair $(Q,\mathcal{G})$, where $\mathcal{G}$ is a Riemannian metric, is called *Riemannian manifold*. Just as in Euclidean geometry, and as its extension, we define the *norm* of any tangent vector $v_q \in T_qQ$ to be $||v_q|| = \mathcal{G}(v_q, v_q)^{\frac{1}{2}}$. In addition, the metric defines the natural *musical isomorphisms*

$$\sharp\mathcal{G} : \Omega^1(Q) \to \mathfrak{X}(Q), \qquad \flat\mathcal{G} : \mathfrak{X}(Q) \to \Omega^1(Q),$$

where the mapping $\flat\mathcal{G}$ is defined by $\flat\mathcal{G}(X) = \mathcal{G}(X,\cdot) : \mathfrak{X}(Q) \to \mathbb{R}$, such that $\flat\mathcal{G}(X)(Y) = \mathcal{G}(X,Y)$ and $\sharp\mathcal{G}$ is its inverse, i.e., $\sharp\mathcal{G} := \left(\flat\mathcal{G}\right)^{-1}$. If $f \in C^\infty(Q)$, we define its *gradient* as $\text{grad } f :=^\sharp \mathcal{G}(\mathrm{d}f) \in \mathfrak{X}(Q)$.

Given a local chart (U,φ) of Q and local coordinates (q^i) for $U \subset Q$, the metric has the form

$$\mathcal{G} = \mathcal{G}_{ij} \ \mathrm{d}q^i \otimes \mathrm{d}q^j,$$

where $\mathcal{G}_{ij} = \mathcal{G}\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right)$.

A Riemannian manifold $(Q, \mathcal{G})$ has associated an *affine connection*, that is, a mapping

$$\begin{aligned} \nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) &\rightarrow \mathfrak{X}(Q) \\ (X, Y) &\mapsto \nabla_X Y, \end{aligned}$$

where $X, Y \in \mathfrak{X}(Q)$, which satisfies the following properties:

1. it is $\mathbb{R}$-bilinear,
2. $\nabla_{f\,X} Y = f\,\nabla_X Y$, where $f \in C^\infty(Q)$,
3. $\nabla_X f\,Y = f\,\nabla_X Y + (\mathcal{L}_X f)\,Y$, where $f \in C^\infty(Q)$.

The mapping $\nabla_X Y$ is called the *covariant derivative of* Y *with respect to* X. Given local coordinates (q^i) on Q, the *Christoffel symbols for the affine connection* are defined by:

$$\nabla_{\frac{\partial}{\partial q^k}} \frac{\partial}{\partial q^j} = \Gamma^i_{k\,j} \frac{\partial}{\partial q^i}.$$

From the above properties of the affine connection and for two vector fields defined by $X = X^i \frac{\partial}{\partial q^i}$ and $Y = Y^i \frac{\partial}{\partial q^i}$, we have the *coordinate expression for the covariant derivative:*

$$\nabla_X Y = \left(X^j \frac{\partial Y^i}{\partial q^j} + \Gamma^i_{j\,k} X^j\, Y^k \right) \frac{\partial}{\partial q^i}.$$

Given a curve in a manifold Q, we may define the *parallel transport* of a vector along the curve. Let $c : I \rightarrow Q$ be that curve locally given by $c(t) = (q^1(t), \ldots, q^n(t))$ where (q^i) $i = 1, \ldots, n$ are local coordinates on Q given by a local chart (U, φ) of Q. Let X be a vector field defined (at least) along $c(t)$

$$X|_{c(t)} = X^i\,(c(t)) \left.\frac{\partial}{\partial q^i}\right|_{c(t)}.$$

If X satisfies the condition

$$\nabla_V X = 0, \quad \text{for any } \ t \in I,$$

X is said to be *parallel transported* along $c(t)$, where $V(t) = \frac{dc}{dt}(t)$. Locally, the previous condition is written as

$$\frac{\mathrm{d}X^i}{\mathrm{d}t} + \Gamma^i_{j\,k} \frac{\mathrm{d}q^j}{\mathrm{d}t} X^k = 0.$$

If the vector V itself is parallel transported along $c(t)$, namely if $\nabla_V V = 0$, then the curve $c(t)$ is called a *geodesic*. Geodesics are, in a sense, the *straightest possible curves* in a Riemannian manifold. Locally, the geodesic equations becomes

$$\frac{d^2 q^i}{dt^2} + \Gamma^i_{j\,k} \frac{\mathrm{d}q^j}{\mathrm{d}t} \frac{\mathrm{d}q^k}{\mathrm{d}t} = 0, \tag{1.3}$$

where, as before, q^i are the coordinates of the curve $c(t)$ (see [44]).

We have considered the affine connection ∇_X as a mapping between two vector fields on Q. On the other hand, it can be considered as a derivation and consequently one can naturally wonder about the definition of such a derivative on function and tensors. The covariant derivative of $f \in \mathbb{C}^\infty(Q)$ is the ordinary directional derivative, namely $\nabla_X f = \mathcal{L}_X f$. Then the condition

$$\nabla_X(f\,Y) = f\,\nabla_X Y + (\mathcal{L}_X f)\,Y,$$

can be exactly rewritten as the Leibniz rule

$$\nabla_X(f\,Y) = f\,\nabla_X Y + (\nabla_X f)Y.$$

There exists a natural connection on each Riemannian manifold that is particularly suited to computations in mechanics and optimal control applications. In order to define it is necessary to introduce some extra concepts.

Definition 1.5.2. *Let $(Q, \mathcal{G})$ be a Riemannian manifold, an affine connection ∇ is said to be metric with respect to $\mathcal{G}$ if $\nabla\mathcal{G} = 0$, that is, it satisfies the rule*

$$Z(\mathcal{G}(X,Y)) = \mathcal{G}(\nabla_Z X, Y) + \mathcal{G}(X, \nabla_Z Y),$$

for all vector fields $X, Y, Z \in \mathfrak{X}(Q)$.

It turns out that requiring a connection to be compatible with the metric is not enough to determine a unique connection, so we turn to another key property. It involves the *torsion tensor* of the connection, which is the $(2,1)$ tensor field $T : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ defined by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y].$$

Locally, if we write

$$T\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right) = T_{ij}^k \frac{\partial}{\partial q^k} \text{ then } T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k.$$

Definition 1.5.3. *A connection is said to be symmetric if its torsion vanishes identically, that is, if*

$$\nabla_X Y - \nabla_Y X = [X,Y].$$

Theorem 1.5.4 (*Fundamental Lemma of Riemannian Geometry*)**.** *Let $(Q, \mathcal{G})$ be a Riemannian manifold. There exists a unique connection $\nabla^{\mathcal{G}}$ on Q which is metric and symmetric.*

See [30, 44, 102] for the proof. This connection is called *Riemannian connection* or *Levi-Civita connection* of $\mathcal{G}$. In this case, the Christoffel symbols are given in terms of the components of the metric as:

$$\Gamma^i_{j\,k} = \frac{1}{2}\mathcal{G}^{il}\left(\frac{\partial \mathcal{G}_{l\,j}}{\partial q^k} + \frac{\partial \mathcal{G}_{l\,k}}{\partial q^j} - \frac{\partial \mathcal{G}_{j\,k}}{\partial q^l}\right), \tag{1.4}$$

where $\mathcal{G}^{il}$ denotes the inverse matrix of $\mathcal{G}_{li}$.

Remark 1.5.5. The Christoffel symbols for the Levi-Civita connection $\nabla^{\mathcal{G}}$ satisfies the relation $\Gamma^C_{AB} = \Gamma^C_{BA}$. ◇

Remark 1.5.6. Alternatively given a Riemannian manifold $(Q, \mathcal{G})$ one can define the Levi-Civita $\nabla^{\mathcal{G}}$ by the formula

$$2\mathcal{G}(Z, \nabla_X Y) = X(\mathcal{G}(Z, Y)) + Y(\mathcal{G}(Z, X)) - Z(\mathcal{G}(Y, X)) + \mathcal{G}(X, [Z, Y]) + \mathcal{G}(Y, [Z, X]) - \mathcal{G}(Z, [Y, X]),$$

where $X, Y, Z \in \mathfrak{X}(Q)$. ◇

Finally it is important to note that a metric connection preserve the kinetic energy associated with the metric $\mathcal{G}$. An affine connection is *energy-preserving* if for every geodesic c of ∇ we have that $\dfrac{d}{dt}K_c(t) := \dfrac{d}{dt}\left(\dfrac{1}{2}\mathcal{G}(\dot{c}(t), \dot{c}(t))\right) = 0$.

To prove the energy-preserving property, we consider local coordinates $(q^1, \ldots, q^n)$ on Q. Therefore $K_c = \frac{1}{2}\mathcal{G}_{ij}\dot{q}^i\dot{q}^j$ for a curve $c : t \mapsto (q^1(t), \ldots, q^n(t))$. If c is a geodesic for the Levi-Civita connection $\nabla^{\mathcal{G}}$ then by (1.3), $\ddot{q}^i = -\Gamma^i_{jk}\dot{q}^j\dot{q}^k$. Now for a geodesic c of $\nabla^{\mathcal{G}}$ we compute

$$\frac{d}{dt}K_c = \mathcal{G}_{ij}\dot{q}^i\ddot{q}^j + \frac{1}{2}\frac{\partial \mathcal{G}_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j\dot{q}^k = \frac{1}{2}\left(\frac{\partial \mathcal{G}_{ij}}{\partial q^k} - 2\mathcal{G}_{il}\Gamma^l_{jk}\right)\dot{q}^i\dot{q}^j\dot{q}^k.$$

Using (1.4) we have that

$$\frac{\partial \mathcal{G}_{ij}}{\partial q^k} - 2\mathcal{G}_{il}\Gamma^l_{jk} = \frac{\partial \mathcal{G}_{jk}}{\partial q^i} - \frac{\partial \mathcal{G}_{ik}}{\partial q^j}.$$

Skew-symmetry of this expression in the indices i and j implies that $\frac{d}{dt}K_c = 0$.

The computation we perform comes from [142] to derive stability results for certain control laws in robotic. Also, in [115] this result was generalized to the case of any affine connection on a Riemannian manifold related to mechanical systems with constraints.

1.6 Symplectic geometry

While Riemannian geometry is based on the study of smooth manifolds that are endowed with a non-degenerate symmetric tensor, i.e. the metric; symplectic geometry covers the study of smooth manifolds equipped with a non-degenerate skew-symmetric tensor. For deeper understanding see [1, 7, 16, 117].

1.6.1 Symplectic vector spaces

Definition 1.6.1. *A symplectic structure on a vector space V is a $2-$form $\omega : V \times V \to \mathbb{R}$ on V which is non-degenerate. The pair (V, ω) is called symplectic vector space.*

Let V be a finite dimensional vector space, if ω is a symplectic structure on V, then dim $V = 2n$ and $\omega^n = \omega \wedge \ldots^n \wedge \omega \neq 0$. The linear map $\flat_\omega : V \to V^*$ defined by

$$\flat_\omega(u)(v) = \omega(u, v) \text{ for all } u, v \in V$$

is an isomorphism. Also, if ω is a symplectic structure on V there exists a basis $(\varepsilon^i)_{i=1}^{2n}$ in V^* such that

$$(\omega_{ij}) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

where $\omega = \omega_{ij}\, \varepsilon^i \otimes \varepsilon^j$, 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 1.6.2. *Let W be a subspace of a symplectic vector space (V,ω). Then, the symplectic orthogonal $W^\perp$ of W is the subspace of V given by*

$$W^\perp = \{x \in V \mid \omega(x,y) = 0 \quad \forall y \in W\}.$$

Some useful properties of the symplectic orthogonal are listed below

- $(W^\perp)^\perp = W$,
- $\dim W + \dim W^\perp = \dim V$,
- $W_1 \subset W_2 \Rightarrow W_1^\perp \supset W_2^\perp$,
- $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

Denote by ω_W the 2-form induced by ω on the vector subspace W. In general ω_W is not symplectic anymore and it has kernel,

$$\ker\omega_W = \{x \in W \mid \flat_\omega(x) = 0\} = W \cap W^\perp.$$

Definition 1.6.3. *A vector subspace of a symplectic vector space (V,ω) is said to be*

1. *isotropic if $\omega_W = 0$, i.e., $W \subseteq W^\perp$.*
2. *symplectic if $W \cap W^\perp = \{0\}$.*
3. *coisotropic if $W^\perp$ is isotropic.*
4. *Lagrangian if $W = W^\perp$.*

We remark that a Lagrangian subspace is isotropic and coisotropic at the same time.

1.6.2 Special submanifolds of symplectic manifolds

Recall that a pair (M, ω_M) is called a symplectic manifold if M is a differentiable manifold and there is defined on M a closed nondegenerate 2-form ω_M; that is a 2-form ω_M such that

i) $\mathrm{d}\omega_M = 0$, and

ii) on each tangent space T_xM, $x \in M$, if $\omega_M|_x(X,Y) = 0$ for all $Y \in T_xM$, then $X = 0$.

Definition of symplectic manifold means that the restrictions of ω_M to each $x \in M$ make the tangent space T_xM into a symplectic vector space.

From now on, if M is a symplectic manifold its associated symplectic 2-form will be denoted by ω_M. In addition, if the particular symplectic manifold is the cotangent bundle of an arbitrary manifold Q, that is, $M = T^*Q$; then the associated symplectic 2-form will be denoted by ω_Q. It can be shown (see [16]), that all symplectic manifolds of the same dimension are *locally the same*. This is in sharp contrast to the situation in Riemannian geometry, and indicates that symplectic geometry is essentially a global theory.

Definition 1.6.4. *Given two symplectic manifolds* (M_1, ω_1) *and* (M_2, ω_2)*, let* $\phi : M_1 \to M_2$ *be a smooth map. The map* ϕ *is called symplectic, or a morphism of symplectic manifolds, so long as*

$$\phi^*\omega_2 = \omega_1.$$

Given a symplectic diffeomorphism ϕ, ϕ^{-1} is also symplectic, and ϕ is called *symplectomorphism*. In particular, the cotangent lift of a diffeomorphism is always a symplectomorphism.

Associated with a symplectic manifold (M, ω), there are two canonical *musical isomorphisms*

$$\flat_\omega : \mathfrak{X}(M) \to \Omega^1(M), \quad \sharp_\omega : \Omega^1(M) \to \mathfrak{X}(M)$$

defined as

$$\flat_\omega(X) = i_X\omega \text{ and } \sharp_\omega = \flat_\omega^{-1}.$$

Given a function $f \in C^\infty(M)$, we define the corresponding *Hamiltonian vector field* X_f by

$$\flat_\omega(X_f) = df.$$

The flow of a Hamiltonian vector field leaves the symplectic form invariant, that is, $\mathcal{L}_{X_f}\omega = 0$. Any vector field with this property is called a *locally Hamiltonian vector field*. This terminology follows by the fact that if X is a locally Hamiltonian vector field, the 1-form $i_X\omega$ is closed and by Poincaré's Lemma, it is locally exact, that is, there locally exists a function f such that $i_X\omega = df$.

Now the idea is to take advantage of the previous definitions and to give the notion of an isotropic, coisotropic and Lagrangian submanifold of a symplectic manifold.

Let (M, ω_M) be a symplectic manifold and $i : N \hookrightarrow M$ an immersion. We denote by $\omega_N = i^*\omega_M$ the 2-form induced by ω_M in N, which is in general degenerate. Its kernel, ker $\omega_N = TN \cap TN^\perp$ has constant rank, so that, it defines a completely integrable distribution on N. In fact, since ω_N is closed the result follows.

Definition 1.6.5. *Let* N *be a submanifold of a symplectic manifold* (M, ω_M)*.* N *is said to be isotropic (resp. coisotropic, Lagrangian) at a point* $q \in N$*, if* T_qN *is an isotropic (resp. coisotropic, Lagrangian) subspace of* $(T_qM, \omega_M(q))$*. We say that* N *is an isotropic submanifold (resp. coisotropic, Lagrangian) if it is isotropic (resp. coisotropic, Lagrangian) at every point.*

An alternative definition of Lagrangian submanifold is the following: A submanifold $N \subset M$ is called *Lagrangian* if it is isotropic and there is an isotropic subbundle $E \subset TM\big|_N$ such that $TM\big|_N = TN \oplus E$ (see [177]).

Remark 1.6.6. Notice that $i : N \hookrightarrow M$ is isotropic if and only if $i^*\omega_M = 0$. Also note that if $N \subset M$ is Lagrangian, dim $N = \frac{1}{2}$dim M and $(T_xN)^{\perp} = T_xN$. $\diamond$

This remark provides us with an alternative way to define a Lagrangian submanifold:

Proposition 1.6.7 ([1])**.** *Let (M, ω_M) be a symplectic manifold, $N \subset M$ a submanifold and $i : N \hookrightarrow M$. Then N is Lagrangian if and only if $i^*\omega_M = 0$ and dim $N = \frac{1}{2}$dim M.*

We will focus on the most relevant and know results of the structure of Lagrangian submanifolds on cotangent bundles. As it already know, the cotangent bundle of a manifold M is a symplectic manifold equipped with the closed and non-degenerated (since (T^*M, ω_M) is symplectic) 2-form $\omega_M = -d\lambda$, where λ is the *Liouville 1-form,* that is, the unique 1-form satisfying $\beta^*\lambda = \beta$ for any 1-form β on M. If $(q^1, \ldots, q^m)$ are local coordinates on M and $(q^1, \ldots, q^m, p_1, \ldots, p_m)$ are the induced coordinates on T^*M, one has that

$$\omega_M = dq^i \wedge dp_i.$$

It fact, every symplectic manifold is locally isomorphic to a cotangent bundle by a consequence of the following result

Theorem 1.6.8 (Darboux's theorem)**.** *Let (M, ω_M) be a symplectic manifold of dimension $2m$. Every point $x \in M$ has an open neighborhood U, which is the domain of a chart (U, φ) with local coordinates $x^1, \ldots, x^{2m}$, such that the 2-form ω_M has the local expression*

$$\omega_M = \sum_{i=1}^{m} dx^i \wedge dx^{m+i}.$$

The canonical examples of Lagrangian submanifold on the cotangent bundle which fiber over the base manifold are the following:

Example 1.6.9. Let β be a 1-form over a manifold M and consider the submanifold $\beta(M) \subset T^*M$. Of course, β is an injective immersion and dim $\beta(M) = \frac{1}{2}$dim T^*M.

If $\omega_M = -d\lambda$ denotes the canonical symplectic structure on T^*M, then,

$$\beta^*\omega_M = -d\beta^*\lambda = -d\beta.$$

Therefore, if β is closed, then $\beta(M)$ is isotropic, that is, $\omega_{\beta(M)} = 0$. Hence, if β is a closed 1-form over M then $\beta(M)$ is a Lagrangian submanifold of the symplectic manifold (T^*M, ω_M).

Example 1.6.10. Let (M, ω_M) be a symplectic manifold and $g : M \to M$ a diffeomorphism. Denote by Graph$\,(g)$ the graph of g, that is Graph$\,(g) := \{(x, g(x)),\ x \in M\} \subset M \times M$, and by $pr_i : M \times M \to M$, $i = 0, 1.$, the projections onto the first and second factor, respectively. It is easy to see that $(M \times M, \tilde{\omega}_M)$, where $\tilde{\omega}_M = pr_1^*\,\omega_M - pr_0^*\,\omega_M$, is a symplectic manifold. Let i_g : Graph$\,(g) \hookrightarrow M \times M$ be the inclusion map, then

$$i_g^*\tilde{\omega}_M = (pr_0)^*\left(g^*\omega_M - \omega_M\right).$$

It is quite clear that dim (Graph g) $= \frac{1}{2}$ dim$(M \times M)$. Moreover, if g is a symplectomorphism, then $g^*\omega_M = \omega_M$ and consequently $i_g^*\tilde{\omega}_M = 0$. Finally we can conclude that g is a symplectomorphism if and only if Graph g is a Lagrangian submanifold of $M \times M$.

Let us consider now $M = T^*Q$, the cotangent bundle of a given manifold Q and $\omega_M = \omega_Q$. As we have seen in the previous paragraph, every symplectomorphism $g : T^*Q \to T^*Q$ generates the Lagrangian submanifold Graph $g \subset (T^*Q \times T^*Q, \tilde{\omega}_Q)$, with $\tilde{\omega}_Q = pr_1^* \, \omega_Q - pr_0^* \, \omega_Q$. These Lagrangian submanifolds are generically called canonical relations referring to the map g.

Now we introduce a very important set of Lagrangian submanifolds of the cotangent bundle T^*M.

Theorem 1.6.11 ([117])**.** *Suppose that N is a submanifold of M and $L \in C^\infty(N)$ is a smooth function on N. If π_{T^*M} is the canonical projection of T^*M on M then*

$$\Sigma_L = \{\alpha \in T^*M \mid \pi_{T^*M}(\alpha) = q \in N, \alpha(v) = dL(v) \quad \forall v \in T_qN\} \subset T^*M$$

is a Lagrangian submanifold of T^*M.

Indeed, choosing adapted coordinates to N in such away that $\mu := (q^1, \dots, q^m, p_1, \dots, p_m)$ are local coordinates on T^*M such that the points of N are defined by $q^{(n+i)} = 0$ with $i = 1, \dots, m-n$. The local expression of Σ_L is

$$\Sigma_L = \left\{ \mu \in T^*M \middle| q^{n+i} = 0, \quad p_j = \frac{\partial L}{\partial q^j} \text{ for } i = 1, \dots, m-n, \quad j = 1, \dots, n \right\}.$$

Thus, it follows that dim $\Sigma_L = m = \frac{1}{2}$dim T^*M. Moreover, taking into account the local expression of the canonical symplectic structure ω_M of the cotangent bundle, it is obvious that $\omega_{\Sigma_L} = 0$, where $\omega_{\Sigma_L} = i^*_{\Sigma_L}\omega_M$ with $i_{\Sigma_L} : \Sigma_L \to T^*M$ being the canonical inclusion. Therefore, since Σ_L is isotropic and its dimension is a half of the dimension of the ambient space, Σ_L is a Lagrangian submanifold of the symplectic manifold (T^*M, ω_M) (See [169] for an intrinsic proof).

The importance of this result lies in the fact that Lagrangian submanifolds are associated to the dynamics of Lagrangian and Hamiltonian systems subject or not to constraints.

1.7 Higher-order tangent bundles

In the last decade many papers and books dealing with higher-order derivatives in mechanics has appeared. An extensive analysis of the geometry of higher-order tangent bundles can be found in, for example, [38], [40], [61], [62], [78], [112] and [139]. In this section we recall some basic facts on the higher-order theory following [112].

Let Q be a differentiable 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 (φ, U) of Q such that $q_0 \in U$ and

$$\frac{d^s}{dt^s}\left(\varphi \circ \gamma_1(t)\right)\Big|_{t=0} = \frac{d^s}{dt^s}\left(\varphi \circ \gamma_2(t)\right)\Big|_{t=0},$$

for $s = 0, ..., k$. This is a well defined equivalence relation in $C^\infty(\mathbb{R}, Q)$ and the equivalence class of a curve γ 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, $\tau_{T^{(k)}Q} : T^{(k)}Q \to Q$ where $\tau_{T^{(k)}Q}\left([\gamma]_0^{(k)}\right) = \gamma(0)$ defines a fiber bundle called *tangent bundle of order* k of Q.

We also may define the surjective mappings $\tau_Q^{(l,k)} : T^{(k)}Q \to T^{(l)}Q$, for $l \leq k$, given by $\tau_Q^{(l,k)}\left([\gamma]_0^{(k)}\right) = [\gamma]_0^{(l)}$. It is easy to see that $T^{(1)}Q \equiv TQ$, the tangent bundle of Q, $T^{(0)}Q \equiv Q$ and $\tau_Q^{(0,k)} = \tau_{T^{(k)}Q}$.

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

$$f^{(l,k)}([\gamma]_0^{(k)}) = \frac{d^l}{dt^l}\left(f \circ \gamma(t)\right)\Big|_{t=0}.$$

Of course, these definitions can be applied to functions defined on open sets of Q.

From a local chart (q^i) on a neighborhood U of Q, it is possible to induce local coordinates $(q^{(0)i}, q^{(1)i}, \ldots, q^{(k)i})$ on $T^{(k)}U = (\tau_Q^k)^{-1}(U)$, where $q^{(s)i} = (q^i)^{(s,k)}$ 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

$$X^{(k)}(f^{(l,k)}) = (X(f))^{(l,k)}, \qquad 0 \leq l \leq k,$$

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

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

Now, we consider the *canonical immersion* $j_k : T^{(k)}Q \to T(T^{(k-1)}Q)$ defined as $j_k([\gamma]_0^{(k)}) = [\gamma^{(k-1)}]_0^{(1)}$, where $\gamma^{(k-1)}$ is the lift of the curve γ to $T^{(k-1)}Q$; that is, the curve $\gamma^{(k-1)} : \mathbb{R} \to T^{(k-1)}Q$ is given by $\gamma^{(k-1)}(t) = [\gamma_t]_0^{(k-1)}$ where $\gamma_t(s) = \gamma(t+s)$. In local coordinates

$$j_k(q^{(0)i}, q^{(1)i}, q^{(2)i}, \ldots, q^{(k)i}) = (q^{(0)i}, q^{(1)i}, \ldots, q^{(k-1)i}; q^{(1)i}, q^{(2)i}, \ldots, q^{(k)i}). \tag{1.5}$$

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 $T^{(k+1)}Q$

$$d_T f([\gamma]_0^{k+1}) = j_{k+1}([\gamma]_0^{k+1})(f).$$

1.7.1 The case of Lie groups

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 $g : I \subset \mathbb{R} \to G$ is a curve in $C^{(k)}(\mathbb{R}, G)$ one can consider the application

$$\begin{array}{rccl} \Lambda^{(k)}: & T^{(k)}G & \longrightarrow & G \times k\mathfrak{g} \\ & [g]_0^{(k)} & \longmapsto & \left(g(0), g^{-1}(0)\dot{g}(0), \frac{d}{dt}\Big|_{t=0}(g^{-1}(t)\dot{g}(t)), \ldots, \frac{d^{k-1}}{dt^{k-1}}\Big|_{t=0}(g^{-1}(t)\dot{g}(t))\right). \end{array}$$

Moreover, $\Lambda^{(k)}$ is a diffeomorphism.

We will denote $\xi(t) = g^{-1}(t)\dot{g}(t) \in \mathfrak{g}$. Therefore

$$\Lambda^{(k)}([g]_0^{(k)}) = (g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}) \; ,$$

where

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

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. Observe that $\Lambda^{(1)} : TG \to G \times \mathfrak{g}$ is the standard trivialization of the tangent bundle of a Lie group.

We may also define the surjective mappings $\tau_G^{(l,k)} : T^{(k)}G \to T^{(l)}G$, 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

$$\tau_G^{(l,k)}(g(0), \xi(0), \dot{\xi}(0), \ldots, \xi^{(k-1)}(0)) = (g(0), \xi(0), \dot{\xi}(0), \ldots, \xi^{(l-1)}(0)).$$

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 $j_k : T^{(k)}G \to T(T^{(k-1)}G)$ defined as $j_k([g]_0^{(k)}) = [g^{(k-1)}]_0^{(1)}$, where $g^{(k-1)}$ is the lift of the curve g to $T^{(k-1)}G$; that is, the curve $g^{(k-1)} : \mathbb{R} \to T^{(k-1)}G$ is given by $g^{(k-1)}(t) = [g_t]_0^{(k-1)}$ where $g_t(s) = g(t+s)$. Using the identification given by $\Lambda^{(k)}$ we have that:

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

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

In the same way one can also use the right multiplication to trivialize the space $T^{(k)}G$ as $G \times k\mathfrak{g}$.

1.8 Generalities on Lie algebroids

In this subsection, we introduce some known notions concerning Lie algebroids that are necessary for further developments. Moreover, we illustrate the theory with several examples. We refer the reader to [37, 118] for more information about Lie algebroids and their role in differential geometry.

1.8.1 Lie algebroids and Cartan calculus on Lie algebroids

Definition 1.8.1. *Let E be a vector bundle of rank n over a manifold M of dimension m. An* skew-symmetric algebroid structure *on the vector bundle $\tau_E : E \to M$ is a $\mathbb{R}$-linear bracket $[\![\cdot,\cdot]\!] : \Gamma(\tau_E) \times \Gamma(\tau_E) \to \Gamma(\tau_E)$ on the space $\Gamma(\tau_E)$ the $C^\infty(M)$-module of sections of E, and a vector bundle morphism $\rho : E \to TM$, the* anchor map*, such that:*

1. *$[\![\cdot,\cdot]\!]$ is skew-symmetric, that is,*

$$[\![X,Y]\!] = -[\![Y,X]\!], \quad \textit{for } X, Y \in \Gamma(\tau_E).$$

2. *If we also denote by $\rho : \Gamma(\tau_E) \to \mathfrak{X}(M)$ the homomorphism of $C^\infty(M)$-modules induced by the anchor map then*

$$[\![X, fY]\!] = f[\![X,Y]\!] + \rho(X)(f)Y, \quad \textit{for } X, Y \in \Gamma(\tau_E) \textit{ and } f \in C^\infty(M). \tag{1.6}$$

If the bracket $[\![\cdot,\cdot]\!]$ satisfies the Jacobi identity, that is,

$$[\![X,[\![Y,Z]\!]]\!] + [\![Z,[\![X,Y]\!]]\!] + [\![Y,[\![Z,X]\!]]\!] = 0 \quad \forall X, Y, Z \in \Gamma(\tau_E)$$

we have that the pair $([\![\cdot,\cdot]\!], \rho)$ is a *Lie algebroid structure* on the vector bundle $\tau_E : E \to M$. From now on we will work on Lie algebroids.

In this context, sections of τ_E, play the role of vector fields on M, and the sections of the dual bundle $\tau_{E^*} : E^* \to M$, are like 1-forms on M.

We may consider two type of distinguished functions: given $f \in C^\infty(M)$ one may define a function $\tilde{f}$ on E by $\tilde{f} = f \circ \tau_E$, the *basic functions*. And, given a section θ of the dual bundle $\tau_{E^*} : E^* \to M$, may be regarded as a *lineal function* $\hat{\theta}$ on E as $\hat{\theta}(e) = \langle \theta(\tau_E(e)), e\rangle$ for all $e \in E$. In this sense, $\Gamma(\tau_E)$ is locally generated by the differential of basic and linear functions.

If $X, Y, Z \in \Gamma(\tau_E)$ and $f \in C^\infty(M)$, then using the Jacobi identity we obtain that

$$[\![[\![X,Y]\!], fZ]\!] = f[\![X,[\![Y,Z]\!]]\!] + [\rho(X),\rho(Y)](f)Z. \tag{1.7}$$

Also, from (1.6) it follows that

$$[\![[\![X,Y]\!], fZ]\!] = f[\![[\![X,Y]\!], Z]\!] + \rho[\![X,Y]\!](f)Z. \tag{1.8}$$

Then, using (1.7) and (1.8) and the fact that $[\![\cdot,\cdot]\!]$ is a Lie bracket we conclude that

$$\rho[\![X,Y]\!] = [\rho(X),\rho(Y)],$$

that is, $\rho : \Gamma(\tau_E) \to \mathfrak{X}(M)$ is a homomorphism between the Lie algebras $(\Gamma(\tau_E), [\![\cdot,\cdot]\!])$ and $(\mathfrak{X}, [\cdot,\cdot])$.

The algebra $\bigoplus_k \Gamma(\Lambda^k E^*)$ of multisections of τ_{E^*} plays the role of the algebra of the differential forms and it is possible to define a *differential operator*

Definition 1.8.2. *If $(E, [\![\cdot,\cdot]\!], \rho)$ is a Lie algebroid over M, one can be define the differential of E, $d^E : \Gamma(\bigwedge^k \tau_{E^*}) \to \Gamma(\bigwedge^{k+1} \tau_{E^*})$, as follows;*

$$
\begin{aligned}
d^E\mu(X_0,\dots,X_k) &= \sum_{i=0}^{k}(-1)^i\rho(X_i)(\mu(X_0,\dots,\widehat{X}_i,\dots,X_k)) \\
&+ \sum_{i<j}(-1)^{i+j}\mu([\![X,Y]\!],X_0,\dots,\widehat{X}_i,\dots,\widehat{X}_j,\dots,X_k),
\end{aligned}
$$

for $\mu \in \Gamma(\bigwedge^k \tau_{E^*})$ and $X_0,\dots,X_k \in \Gamma(\tau_E)$.

From the properties of Lie algebroids it follows that d^E is a cohomology operator, that is, $(d^E)^2 = 0$ and $d^E(\alpha\wedge\beta) = d^E\alpha\wedge\beta + (-1)^k\alpha\wedge d^E\beta$, for $\alpha \in \Gamma(\Lambda^k E^*)$ and $\beta \in \Gamma(\Lambda^r E^*)$.

Conversely it is possible to recover the Lie algebroid structure of E from the existence of an exterior differential on $\Gamma(\Lambda^\bullet \tau_{E^*})$. If $f : M \to \mathbb{R}$ is a real smooth function, one can define the anchor map and the Lie bracket as follows (see [110] to the case of skew-symmetric Lie algebroids):

- $d^E f(X) = \rho(X)f$, for $X \in \Gamma(\tau_E)$,
- $i_{[\![X,Y]\!]}\theta = \rho(X)\theta(Y) - \rho(Y)\theta(X) - d^E\theta(X,Y)$ for all $X, Y \in \Gamma(\tau_E)$ and $\theta \in \Gamma(\tau_{E^*})$.

Moreover, from the last equality, the section $\theta \in \Gamma(\tau_{E^*})$ is a *1-cocycle* if and only if $d^E\theta = 0$, or, equivalently,

$$\theta[\![X,Y]\!] = \rho(X)(\theta(Y)) - \rho(Y)(\theta(X)),$$

for all $X, Y \in \Gamma(\tau_E)$.

We may also define the *Lie derivative* with respect to a section $X \in \Gamma(\tau_E)$ as the operator $\mathcal{L}^E_X : \Gamma(\bigwedge^k \tau_{E^*}) \to \Gamma(\bigwedge^k \tau_{E^*})$ given by

$$\mathcal{L}^E_X\theta = i_X \circ d^E\theta + d^E \circ i_X\theta,$$

for $\theta \in \Gamma(\Lambda^k\tau_{E^*})$. One also has the usual identities

- $d^E \circ \mathcal{L}^E_X = \mathcal{L}^E_X \circ d^E$,
- $\mathcal{L}^E_X i_Y - i_X\mathcal{L}^E_Y = i_{[\![X,Y]\!]}$,
- $\mathcal{L}^E_X\mathcal{L}^E_Y - \mathcal{L}^E_Y\mathcal{L}^E_X = \mathcal{L}^E_{[\![X,Y]\!]}$.

We take local coordinates (x^i) on M with $i = 1,\dots,m$ and a local basis $\{e_A\}$ of sections of the vector bundle $\tau_E : E \to M$ with $A = 1,\dots,n$, then we have the corresponding local coordinates on an open subset $\tau_E^{-1}(U)$ of E, (x^i, y^A) (U is an open subset of Q), where $y^A(e)$ is the A-th coordinate of $e \in E$ in the given basis i.e., every $e \in E$ is expressed as $e = y^1 e_1(\tau_E(e)) + \ldots + y^n e_n(\tau_E(e))$.

Such coordinates determine local functions ρ^i_A, $\mathcal{C}^C_{AB}$ on M which contain the local information of the Lie algebroid structure, and accordingly they are called *structure functions of the Lie algebroid.* They are given by

$$\rho(e_A) = \rho^i_A \frac{\partial}{\partial x^i} \quad \text{and} \quad [\![e_A, e_B]\!] = \mathcal{C}^C_{AB} e_C. \tag{1.9}$$

These functions should satisfy the relations

$$\rho^j_A \frac{\partial \rho^i_B}{\partial x^j} - \rho^j_B \frac{\partial \rho^i_A}{\partial x^j} = \rho^i_C \mathcal{C}^C_{AB} \tag{1.10}$$

and

$$\sum_{cyclic(A,B,C)} [\rho^i_A \frac{\partial \mathcal{C}^D_{BC}}{\partial x^i} + \mathcal{C}^D_{AF} \mathcal{C}^F_{BC}] = 0, \tag{1.11}$$

which are usually called *the structure equations.*

If $f \in C^\infty(M)$, we have that

$$d^E f = \frac{\partial f}{\partial x^i} \rho^i_A e^A, \tag{1.12}$$

where $\{e^A\}$ is the dual basis of $\{e_A\}$. On the other hand, if $\theta \in \Gamma(\tau_{E^*})$ and $\theta = \theta_C e^C$ it follows that

$$d^E \theta = (\frac{\partial \theta_C}{\partial x^i} \rho^i_B - \frac{1}{2} \theta_A \mathcal{C}^A_{BC}) e^B \wedge e^C. \tag{1.13}$$

In particular,

$$d^E x^i = \rho^i_A e^A, \quad d^E e^A = -\frac{1}{2} C^A_{BC} e^B \wedge e^C.$$

Examples of Lie algebroids

Example 1.8.3. A *finite dimensional real Lie algebra* $\mathfrak{g}$ where $M = \{m\}$ be a unique point. Then, we consider the vector bundle $\tau_{\mathfrak{g}} : \mathfrak{g} \to M$. The sections of this bundle can be identified with the elements of $\mathfrak{g}$ and therefore we can consider as the Lie bracket the structure of the Lie algebra induced by $\mathfrak{g}$, $[\cdot, \cdot]_{\mathfrak{g}}$. Since $TM = \{0\}$ one may consider the anchor map $\rho \equiv 0$. Thus, $(\mathfrak{g}, [\cdot, \cdot]_{\mathfrak{g}}, 0)$ is a Lie algebroid over a point.

Example 1.8.4. Consider a *tangent bundle* of a manifold M. The sections of the bundle $\tau_{TM} : TM \to M$ are just the set of vector fields on M, the anchor map $\rho : TM \to TM$ is the identity function and the Lie bracket defined on $\Gamma(\tau_{TM})$ is induced by the Lie bracket of vector fields on M.

Example 1.8.5. Let $\phi : M \times G \to M$ be an action of G on the manifold M where G is a Lie group. The induced anti-homomorphism between the Lie algebras $\mathfrak{g}$ and $\mathfrak{X}(M)$ by the action is $\Phi : \mathfrak{g} \to \mathfrak{X}(M)$, $\xi \mapsto \xi_M$, where ξ_M is the infinitesimal generator of the action for $\xi \in \mathfrak{g}$.

The vector bundle $\tau_{M\times\mathfrak{g}} : M \times \mathfrak{g} \to M$ is a Lie algebroid over M. The anchor map is defined by $\rho : M \times \mathfrak{g} \to TM$, $\rho(m, \xi) = -\xi_M(m)$ and the Lie bracket of sections is given by the Lie algebra structure on $\Gamma(\tau_{M\times\mathfrak{g}})$ as

$$[\![\hat{\xi}, \hat{\eta}]\!]_{M\times\mathfrak{g}}(m) = (m, [\xi, \eta]) = \widehat{[\xi, \eta]} \text{ for } m \in M$$

where $\hat{\xi}(m) = (m, \xi)$, $\hat{\eta}(m) = (m, \eta)$ for $\xi, \eta \in \mathfrak{g}$. This Lie algebroid is called *Action Lie algebroid*.

Example 1.8.6. Let G be a Lie group and we assume that G acts free and properly on M. We denote by $\pi : M \to \widehat{M} = M/G$ the associated principal bundle. The tangent lift of the action gives a free and proper action of G on TM and $\widehat{TM} = TM/G$ is a quotient manifold. The quotient vector bundle $\tau_{\widehat{TM}} : \widehat{TM} \to \widehat{M}$ where $\tau_{\widehat{TM}}([v_m]) = \pi(m)$ is a Lie algebroid over $\widehat{M}$. The fiber of $\widehat{TM}$ over a point $\pi(m) \in \widehat{M}$ is isomorphic to T_mM.

The Lie bracket is defined on the space $\Gamma(\tau_{\widehat{TM}})$ which is isomorphic to the Lie subalgebra of G-invariants vector fields, that is,

$$\Gamma(\tau_{\widehat{TM}}) = \mathfrak{X}(M)^G = \{X \in \mathfrak{X}(M) \mid X \text{ is } G\text{-invariant}\}.$$

Thus, the Lie bracket on $\widehat{TM}$ is just the bracket of G-invariant vector fields. The anchor map $\rho : \widehat{TM} \to T\widehat{M}$ is given by $\rho([v_m]) = T_m\pi(v_m)$. Moreover, ρ is a Lie algebra homomorpishm satisfying the compatibility condition since the G-invariant vector fields are π-projectable. This Lie algebroid is called *Lie-Atiyah algebroid* associated with the principal bundle $\pi : M \to \widehat{M}$.

Let $\mathcal{A} : TM \to \mathfrak{g}$ be a principal connection in the principal bundle $\pi : M \to \widehat{M}$ and $B : TM \oplus TM \to \mathfrak{g}$ be the curvature of $\mathcal{A}$. This connection determine an isomorphism $\alpha_{\mathcal{A}}$, between the vector bundles $\widehat{TM} \to \widehat{M}$ and $T\widehat{M} \oplus \widetilde{\mathfrak{g}} \to \widehat{M}$ where $\widetilde{\mathfrak{g}} = (M \times \mathfrak{g})/G$ is the adjoint bundle associated with the principal bundle $\pi : M \to \widehat{M}$ (see [45] for example).

We choose a local trivialization of the principal bundle $\pi : M \to \widehat{M}$ to be $U \times G$, where U is an open subset of $\widehat{M}$. Suppose that e is the identity of G, (x^i) are local coordinates on U and $\{\xi_A\}$ is a basis on $\mathfrak{g}$.

Denote by $\{\overleftarrow{\xi_A}\}$ the corresponding left-invariant vector field on G, that is,

$$\overleftarrow{\xi_A}(g) = (T_eL_g)(\xi_A)$$

for $g \in G$ where $L_g : G \to G$ is the left-translation on G by g. If

$$\mathcal{A}\left(\frac{\partial}{\partial x^i}\Big|_{(x,e)}\right) = \mathcal{A}_i^A(x)\xi_A, \quad \mathcal{B}\left(\frac{\partial}{\partial x^i}\Big|_{(x,e)}, \frac{\partial}{\partial x^j}\Big|_{(x,e)}\right) = \mathcal{B}_{ij}^A(x)\xi_A,$$

for $i, j \in \{1, \ldots, m\}$ and $x \in U$, then the horizontal lift of the vector field $\frac{\partial}{\partial x^i}$ is the vector field on $\pi^{-1}(U) \simeq U \times G$ given by

$$\left(\frac{\partial}{\partial x^i}\right)^h = \frac{\partial}{\partial x^i} - \mathcal{A}_i^A \overleftarrow{\xi}_A.$$

Therefore, the vector fields on $U \times G$

$$\{e_i = \frac{\partial}{\partial x^i} - \mathcal{A}_i^A \overleftarrow{\xi}_A, e_B = \overleftarrow{\xi}_B\}$$

are G-invariant under the action of G over M and they define a local basis $\{\hat{e}_i, \hat{e}_B\}$ on $\Gamma(\widehat{TM}) = \Gamma(\tau_{T\widehat{M}\oplus\tilde{\mathfrak{g}}})$. The corresponding local structure functions of $\tau_{\widehat{TM}} : \widehat{TM} \to \widehat{M}$ are

$$\begin{aligned} \mathcal{C}_{ij}^k &= \mathcal{C}_{iA}^j = -\mathcal{C}_{Ai}^j = \mathcal{C}_{AB}^i = 0, \quad \mathcal{C}_{ij}^A = -\mathcal{B}_{ij}^A, \quad \mathcal{C}_{iA}^C = -\mathcal{C}_{Ai}^C = c_{AB}^C \mathcal{A}_i^B, \\ \mathcal{C}_{AB}^C &= c_{AB}^C, \quad \rho_i^j = \delta_{ij}, \quad \rho_i^A = \rho_A^i = \rho_A^B = 0, \end{aligned}$$

being $\{c_{AB}^C\}$ the constant structure of $\mathfrak{g}$ with respect to the basis $\{\xi_A\}$ (see [111] for more details). That is,

$$[\![\hat{e}_i, \hat{x}_j]\!]_{\widehat{TM}} = -\mathcal{B}_{ij}^C \hat{e}_C, \quad [\![\hat{e}_i, \hat{e}_A]\!]_{\widehat{TM}} = c_{AB}^C \mathcal{A}_i^B \hat{e}_C, \quad [\![\hat{e}_A, \hat{e}_B]\!]_{\widehat{TM}} = c_{AB}^C \hat{e}_C,$$

$$\rho_{\widehat{TM}}(\hat{e}_i) = \frac{\partial}{\partial x^i}, \quad \rho_{\widehat{TM}}(\hat{e}_A) = 0.$$

The basis $\{\hat{e}_i, \hat{e}_B\}$ induce local coordinates $(x^i, y^i, \bar{y}^B)$ on $\widehat{TM} = TM/G$.

1.8.2 Morphisms of Lie algebroids and Lie subalgebroids

Suppose that $(E, [\![\ ,\ ]\!]_E, \rho_E)$ and $(E', [\![\ ,\ ]\!]_{E'}, \rho_{E'})$ are Lie algebroids over M and M', respectively. Then a morphism of vector bundles (F, f) of E on E'

$$\begin{tikzcd} E \arrow[r, "F"] \arrow[d, "\tau_E"'] & E' \arrow[d, "\tau_{E'}"] \\ M \arrow[r, "f"] & M' \end{tikzcd}$$

is a *Lie algebroid morphism* if

$$d^E((F,f)^*\phi') = (F,f)^*(d^{E'}\phi'), \quad \text{for } \phi' \in \Gamma(\wedge^k(E')^*) \quad \forall k; \tag{1.14}$$

where $(F,f)^*\phi'$ is the section of $\Lambda^k E^* \to M$ given by

$$((F,f)^*\phi')_x(e_1, \dots, e_k) = \phi'_{f(x)}(F(e_1), \dots, F(e_k)),$$

for $x \in M$ and $e_1, \dots, e_k \in E_x$. We remark that (1.14) holds if and only if

$$\begin{aligned} d^E(g' \circ f) &= (F,f)^*(d^{E'}g'), \quad \text{for } g' \in C^\infty(M'), \\ d^E((F,f)^*\alpha') &= (F,f)^*(d^{E'}\alpha'), \quad \text{for } \alpha' \in \Gamma((E')^*). \end{aligned}$$

In particular, a Lie algebroid morphism preserves the anchor and the bracket of projectable sections. An equivalent definition of morphism of Lie algebroids could be given in terms of the bracket and the anchor map (see [118]).

If (F, f) is a Lie algebroid morphism, f is an injective immersion and $F|_{E_x} : E_x \to E'_{f(x)}$ is injective, for all $x \in M$, then $(E, [\![\cdot,\cdot]\!]_E, \rho_E)$ is said to be a *Lie subalgebroid* of $(E', [\![\cdot,\cdot]\!]_{E'}, \rho_{E'})$.

An alternative and more natural definition of Lie subalgebroid is given as follows:

Definition 1.8.7. *Let $(E, [\![\cdot,\cdot]\!]_E, \rho_E)$ be a Lie algebroid over M and N is a submanifold of M. A Lie subalgebroid of E over N is a vector subbundle B of E over N*

$$\begin{tikzcd} B \arrow[r, hook, "j"] \arrow[d, "\tau_B"'] & E \arrow[d, "\tau_E"] \\ N \arrow[r, hook, "i"] & M \end{tikzcd}$$

such that $\rho_B = \rho_E \mid_B \colon B \to TN$ is well define and; given $X, Y \in \Gamma(B)$ and $\widetilde{X}, \widetilde{Y} \in \Gamma(E)$ arbitrary extensions of X, Y respectively, we have that $([\![\widetilde{X}, \widetilde{Y}]\!]_E) \mid_N \in \Gamma(B)$.

Examples of Lie subalgebroids

Example 1.8.8. Let E be a Lie algebroid over M. Given a submanifold N of M, if $B = E\mid_N \cap (\rho\mid_N)^{-1}(TN)$ exists as a vector bundle, it will be a Lie subalgebroid of E over N, and will be called *Lie algebroid restriction of E to N* (see [119]).

Example 1.8.9. Let N be a submanifold of M. Then, TN is a Lie subalgebroid of TM.

Now, let $\mathcal{F}$ be a completely integrable distribution on a manifold M. $\mathcal{F}$ equipped with the bracket of vector fields is a Lie algebroid over M since $\tau_E\mid_{\mathcal{F}} \colon \mathcal{F} \to M$ is a vector bundle and if $\mathcal{F}$ is a foliation, $(\Gamma(\mathcal{F}), [\cdot,\cdot])$ is a Lie algebra. The anchor map is the inclusion $i_{\mathcal{F}} : \mathcal{F} \to TM$ ($i_{\mathcal{F}}$ is a Lie algebroid monomorphism).

Moreover, if N is a submanifold of M and $\mathcal{F}_N$ is a foliation on N, then $\mathcal{F}_N$ is a Lie subalgebroid of the Lie algebroid $\tau_{TM} : TM \to M$.

Example 1.8.10. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$ be a Lie subalgebra. If we consider the Lie algebroid induced by $\mathfrak{g}$ and $\mathfrak{h}$ over a point, then $\mathfrak{h}$ is a Lie subalgebroid of $\mathfrak{g}$.

Example 1.8.11. Let $M \times \mathfrak{g} \to M$ be an action Lie algebroid and let N be a submanifold of M. Let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$ such that the infinitesimal generators of the elements of $\mathfrak{h}$ are tangent to N; that is, the application

$$\begin{aligned} \mathfrak{h} &\to \mathfrak{X}(N) \\ \xi &\mapsto \xi_N \end{aligned}$$

is well defined. Thus, the action Lie algebroid $N \times \mathfrak{h} \to N$ is a Lie subalgebroid of $M \times \mathfrak{g} \to M$.

Example 1.8.12. Suppose that the Lie group G acts free and properly on M. Let $\pi : M \to M/G = \widehat{M}$ be the associated $G-$principal bundle. Let N be a $G-$invariant submanifold of M and $\mathcal{F}_N$ be a $G-$invariant foliation over N. We may consider the vector bundle $\widehat{\mathcal{F}_N} = \mathcal{F}_N/G \to N/G = \widehat{N}$ and endow it with a Lie algebroid structure. The sections of $\widehat{\mathcal{F}_N}$ are

$$\Gamma(\widehat{\mathcal{F}}_N) = \{X \in \mathfrak{X}(N) \mid X \text{ is } G\text{-invariant and } X(q) \in \mathcal{F}_N(q), \forall q \in N\}.$$

The standard bracket of vector fields on N induces a Lie algebra structure on $\Gamma(\widehat{\mathcal{F}}_N)$. The anchor map is the canonical inclusion of $\widehat{\mathcal{F}}_N$ on $T\widehat{N}$ and $\widehat{\mathcal{F}}_N$ is a Lie subalgebroid of $\widehat{TM} \to \widehat{M}$.

1.8.3 The prolongation of a Lie algebroid over a smooth map

In this subsection we will recall the definition of the Lie algebroid structure of the prolongation of a Lie algebroid over a smooth map. We will follow [80] and [111].

Let $(E, [\![\cdot,\cdot]\!], \rho)$ be a Lie algebroid of rank n over a manifold M of dimension m and $f : M' \to M$ be a fibration. We consider the subset $\mathcal{T}^f E$ of $E \times TM'$ given by

$$\mathcal{T}^f E := \bigcup_{x' \in M'} E_\rho \times_{Tf} T_{x'}M' = \bigcup_{x' \in M'} \{(b, v') \in E \times T_{x'}M' \mid \rho(b) = (Tf)(v')\},$$

where $Tf : TM' \to TM$ denotes the tangent map to f. We will frequently use the redundant notation (x', b, v') to denote the element $(b, v') \in \mathcal{T}^f E$.

Denoting by $\tau_E^f : \mathcal{T}^f E \to M'$ the map given by

$$\tau_E^f(b, v') = \tau_{M'}(v') = x'$$

for $(b, v') \in \mathcal{T}^f E$ and $\tau_{M'} : TM' \to M'$ the canonical projection. Then if m' is the dimension of M' one may prove that

$$\dim \left(\mathcal{T}^f E\right)_{x'} = n + m' - \dim \left(\rho(E_{f(x')}) + (T_{x'}f)(T_{x'}M')\right)$$

for $x' \in M'$. Thus, if we suppose that there exists $c \in \mathbb{N}$ such that

$$\dim \left(\rho(E_{f(x')}) + (T_{x'}f)(T_{x'}M')\right) = c, \text{ for all } x' \in M', \tag{1.15}$$

then we conclude that $\mathcal{T}^f E$ is a vector bundle over M' with vector bundle projection $\tau_E^f : \mathcal{T}^f E \to M'$. Note that if ρ and Tf are transversal, that is, $\rho(E_{f(x')}) + (T_{x'}f)(T_{x'}M') = T_{f(x')}M$ for all $x' \in M'$. Then it is clear that (1.15) holds.

Next, we will assume that condition (1.15) holds and we will describe the sections of the vector bundle $\tau_E^f : \mathcal{T}^f E \to M'$. We say that a section $\tilde{Y}$ of $\tau_E^f : \mathcal{T}^f E \to M'$ is *projectable* if there exists a section Y of $\tau_E : E \to M$ and a vector field $U \in \mathfrak{X}(M')$ which is f-projectable onto the vector field $\rho(Y)$ and such that $\tilde{Y}(x') = (Y(f(x')), U(x'))$, for all $x' \in M'$. For such projectable section $\tilde{Y}$, we will use the following notation $\tilde{Y} \equiv (Y, U)$. It is easy to prove that one may choose a local basis of projectable sections $\Gamma(\tau_E^f)$ (see [111]).

The Lie bracket of two projectable sections $Z_1 = (Y_1, U_1)$ and $Z_2 = (Y_2, U_2)$ is given by

$$[\![Z_1, Z_2]\!]_f(x') = (x', [\![Y_1, Y_2]\!](q), [U_1, U_2](x')),\ x' \in M',\ q = f(x').$$

Since any section of $\mathcal{T}^f E$ can be locally written as a linear combination of projectable sections, the definition of a Lie bracket for arbitrary sections of $\mathcal{T}^f E$ follows. Therefore $\mathcal{T}^f E$ has a Lie algebroid structure where the anchor of $\mathcal{T}^f E$ is the projection onto the last factor, that is, the map $\rho_f : \mathcal{T}^f E \to TM'$ given by $\rho_f(x', b, v') = v'$.

The Lie algebroid $(\mathcal{T}^f E, [\![\ ,\ ]\!]_f, \rho_f)$ is called the *prolongation* of E *over* f or the $E-$*tangent bundle to* f.

E-tangent bundle to a Lie algebroid E

We consider the prolongation over the canonical projection of the Lie algebroid E over M,

$$\mathcal{T}^{\tau_E}E = \bigcup_{e\in E}(E_\rho \times_{T\tau_E} T_eE) = \bigcup_{e\in E}\{(e', v_e) \in E \times T_eE \mid \rho(e') = (T_e\tau_E)(v_e)\},$$

where $T\tau_E : TE \to TM$ is the tangent map to τ_E.

In fact, $\mathcal{T}^{\tau_E}E$ is a Lie algebroid of rank $2n$ over E where $\tau_E^{(1)} : \mathcal{T}^{\tau_E}E \to E$ is the vector bundle projection, $\tau_E^{(1)}(b, v_e) = \tau_{TE}(v_e) = e$, and the anchor map is $\rho_1 := pr_2 : \mathcal{T}^{\tau_E}E \to TE$; the projection over the second factor. The bracket of sections of this new Lie algebroid will be denoted by $[\![\cdot,\cdot]\!]_{\tau_E^{(1)}}$ (See [134] for more details).

If we now denote by (e, e', v_e) an element $(e', v_e) \in \mathcal{T}^{\tau_E}E$ where $e \in E$ and where v is tangent; we rewrite the definition of the prolongation of the Lie algebroid as the subset of $E \times E \times TE$ by

$$\mathcal{T}^{\tau_E}E = \{(e, e', v_e) \in E \times E \times TE \mid \rho(e') = (T\tau_E)(v_e), v_e \in T_eE \text{ and } \tau_E(e) = \tau_E(e')\}.$$

In this sense, if $(e, e', v_e) \in \mathcal{T}^{\tau_E}E$; then $\rho_1(e, e', v_e) = (e, v_e) \in T_eE$, and $\tau_E^{(1)}(e, e', v_e) = e \in E$.

Let us introduce two canonical operations that we have on a Lie algebroid E. The first one is obtained using the Lie algebroid structure of E and the second one is a consequence of E being a vector bundle.

On one hand, if $f \in C^\infty(M)$ we will denote by f^c the *complete lift* to E of f defined by $f^c(e) = \rho(e)(f)$ for all $e \in E$. Now, let X be a section of E then there exists a unique vector field X^c on E, the *complete lift* of X, satisfying the two following conditions:

1. X^c is τ_E-projectable on $\rho(X)$ and
2. $X^c(\hat{\alpha}) = \widehat{\mathcal{L}_X^E\alpha}$,

for every $\alpha \in \Gamma(\tau_{E^*})$ (see [76]). Here, if $\beta \in \Gamma(\tau_{E^*})$ then $\hat{\beta}$ is the linear function on E defined by

$$\hat{\beta}(e) = \langle \beta(\tau_E(e)), e\rangle, \quad \text{for all } e \in E.$$

We may introduce *the complete lift* $X^{\mathbf{c}}$ of a section $X \in \Gamma(\tau_E)$ as the sections of $\tau_E^{(1)}$: $\mathcal{T}^{\tau_E}E \to E$ given by

$$X^{\mathbf{c}}(e) = (X(\tau_E(e)), X^c(e)) \tag{1.16}$$

for all $e \in E$ (see [134]).

On the other hand, given a section $X \in \Gamma(\tau_E)$ we define the *vertical lift* as the vector field $X^v \in \mathfrak{X}(E)$ given by

$$X^v(e) = X(\tau_E(e))_e^v, \text{ for } e \in E,$$

where ${}_e^v : E_q \to T_eE_q$ for $q = \tau_E(e)$ is the canonical isomorphism between the vector spaces E_q and T_eE_q.

Finally we may introduce the *vertical lift* $X^{\mathbf{v}}$ of a section $X \in \Gamma(\tau_E)$ as a section of $\tau_E^{(1)}$ given by

$$X^{\mathbf{v}}(e) = (0, X^v(e)) \text{ for } e \in E.$$

With these definitions we have the properties (see [76] and [134])

$$[X^c, Y^c] = [\![X, Y]\!]^c, \quad [X^c, Y^v] = [\![X, Y]\!]^v, \quad [X^v, Y^v] = 0 \tag{1.17}$$

for all $X, Y \in \Gamma(\tau_E)$.

If (x^i) are local coordinates on an open subset U of M and $\{e_A\}$ is a basis of sections of τ_E then we have induced coordinates (x^i, y^A) on E. From the basis $\{e_A\}$ we may define a local basis $\{e_A^{(1)}, e_A^{(2)}\}$ of sections of $\tau_E^{(1)}$ (see [111] for more details) given by

$$e_A^{(1)}(e) = \left(e, e_A(\tau_A(e)), \rho_A^i \frac{\partial}{\partial x^i}\Big|_e\right), \quad e_A^{(2)}(e) = \left(e, 0, \frac{\partial}{\partial y^A}\Big|_e\right),$$

for $e \in (\tau_E)^{-1}(U)$ with U an open subset of M.

From this basis one have that the structure of Lie algebroid is determined by

$$\begin{aligned} \rho_1(e_A^{(1)}(e)) &= \left(e, \rho_A^i \tfrac{\partial}{\partial x^i}\Big|_e\right), \qquad \rho_1(e_A^{(2)}(e)) = \left(e, \tfrac{\partial}{\partial y^A}\Big|_e\right) \\ [\![e_A^{(1)}, e_B^{(1)}]\!]_{\tau_E^{(1)}} &= \mathcal{C}_{AB}^C e_C^{(1)}, \\ [\![e_A^{(1)}, e_B^{(2)}]\!]_{\tau_E^{(1)}} &= [\![e_A^{(2)}, e_B^{(2)}]\!]_{\tau_E^{(1)}} = 0, \end{aligned}$$

for all A, B and C; where $\mathcal{C}_{AB}^C$ are the structure functions of E determined by the Lie bracket $[\![\cdot, \cdot]\!]$ with respect to the basis $\{e_A\}$.

Using $\{e_A^{(1)}, e_A^{(2)}\}$ one may introduce local coordinates $(x^i, y^A; z^A, v^A)$ on E. If V is a section of $\tau_E^{(1)}$ then in local coordinates is written as $V(x, y) = (x^i, y^A, z^A(x, y), v^A(x, y))$; and therefore the expression of V in terms of the basis $\{e_A^{(1)}, e_A^{(2)}\}$ is $V = z^A e_A^{(1)} + v^A e_A^{(2)}$ and the vector field $\rho_1(V) \in \mathfrak{X}(E)$ has the expression

$$\rho_1(V) = \rho_A^i z^A(x, y) \frac{\partial}{\partial x^i}\Big|_{(x,y)} + v^A(x, y) \frac{\partial}{\partial y^A}\Big|_{(x,y)}.$$

Finally, if $\{e^A_{(1)}, e^A_{(2)}\}$ denotes the dual basis of $\{e_A^{(1)}, e_A^{(2)}\}$, then we have

$$\begin{aligned} d^{\mathcal{T}^{\tau_E} E} F(x^i, y^A) &= \rho_A^i \frac{\partial F}{\partial x^i} e^A_{(1)} + \frac{\partial F}{\partial y^A} e^A_{(2)}, \\ d^{\mathcal{T}^{\tau_E} E} e^C_{(1)} &= -\frac{1}{2} \mathcal{C}_{AB}^C e^A_{(1)} \wedge e^B_{(1)}, \quad d^{\mathcal{T}^{\tau_E} E} e^C_{(2)} = 0. \end{aligned}$$

Example 1.8.13. In the case of $E = TM$ one may identify $\mathcal{T}^{\tau_E} E$ with TTM with the standard Lie algebroid structure over TM.

Example 1.8.14. Let $\mathfrak{g}$ be a real Lie algebra of finite dimension. Then $\mathfrak{g}$ is a Lie algebroid over a single point $M = \{q\}$. Using that the anchor map of $\mathfrak{g}$ is zero we obtain that

$$\mathcal{T}^{\tau_{\mathfrak{g}}}\mathfrak{g} = \{(\xi_1, \xi_2, v_{\xi_1}) \in \mathfrak{g} \times T\mathfrak{g}\} \simeq \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g} \simeq 3\mathfrak{g}.$$

The vector bundle projection $\tau_{\mathfrak{g}}^{(1)} : \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is given by $\tau_{\mathfrak{g}}^{(1)}(\xi_1, \xi_2, \xi_3) = \xi_1$ with anchor map $\rho_1(\xi_1, \xi_2, \xi_3) = (\xi_1, \xi_3) \simeq v_{\xi_1} \in T_{\xi_1}\mathfrak{g}$.

Let $\{e_A\}$ be a basis of the Lie algebra $\mathfrak{g}$, this basis induces local coordinates y^A on $\mathfrak{g}$, that is, $\xi = y^A e_A$. Also, this basis induces a basis of sections of $\tau_{\mathfrak{g}}^{(1)}$ as

$$e_A^{(1)}(\xi) = (\xi, e_A, 0), \quad e_A^{(2)}(\xi) = \left(\xi, 0, \frac{\partial}{\partial y^A}\right).$$

Moreover

$$\rho_1(e_A^{(1)})(\xi) = (\xi, 0), \quad \rho_1(e_A^{(2)})(\xi) = \left(\xi, \frac{\partial}{\partial y^A}\right).$$

The basis $\{e_A^{(1)}, e_A^{(2)}\}$ induces adapted coordinates (y^A, z^A, v^A) in $\mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}$ and therefore a section $Y \in \Gamma(\tau_{\mathfrak{g}}^{(1)})$ is written as $Y(\xi) = z^A(\xi)e_A^{(1)} + v^A(\xi)e_A^{(2)}$. Thus, the vector field $\rho_1(Y) \in \mathfrak{g}$ has the expression $\rho_1(Y) = v^A(\xi)\frac{\partial}{\partial y^A}\Big|_{\xi}$. Finally, the Lie algebroid structure on $\tau_{\mathfrak{g}}^{(1)}$ is determined by the Lie bracket $[\![(\xi, \tilde{\xi}), (\eta, \tilde{\eta})]\!] = ([\xi, \eta], 0)$.

Example 1.8.15. We consider a Lie algebra $\mathfrak{g}$ acting on a manifold M, that is, we have a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$ mapping every element ξ of $\mathfrak{g}$ to a vector field ξ_M on M. Then we can consider the action Lie algebroid $E = M \times \mathfrak{g}$. Identifying $TE = TM \times T\mathfrak{g} = TM \times \mathfrak{g} \times \mathfrak{g}$, an element of the prolongation Lie algebroid to E over the bundle projection is of the form $(a, b, v_a) = ((x, \xi), (x, \eta), (v_x, \xi, \chi))$ where $x \in M$, $v_x \in T_xM$ and $(\xi, \eta, \chi) \in 3\mathfrak{g}$. The condition $T\tau_{\mathfrak{g}}(v) = \rho(b)$ implies that $v_x = -\eta_M(x)$. Therefore we can identify the prolongation Lie algebroid with $M \times \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}$ with projection onto the first two factors (x, ξ) and anchor map $\rho_1(x, \xi, \eta, \chi) = (-\eta_M(x), \xi, \chi)$. Given a base $\{e_A\}$ of $\mathfrak{g}$ the basis $\{e_A^{(1)}, e_A^{(2)}\}$ of sections of $\mathcal{T}^{\tau_{M\times\mathfrak{g}}}(M \times \mathfrak{g})$ is given by

$$e_A^{(1)}(x, \xi) = (x, \xi, e_A, 0), \quad e_A^{(2)}(x, \xi) = (x, \xi, 0, e_A).$$

Moreover,

$$\rho_1(e_A^{(1)}(x, \xi)) = (x, -(e_A)_M(x), \xi, 0)\,, \quad \rho_2(e_A^{(2)}(x, \xi)) = (x, 0, \xi, e_A).$$

Finally, the Lie bracket of two constant sections is given by $[\![(\xi, \tilde{\xi}), (\eta, \tilde{\eta})]\!] = ([\xi, \eta], 0)$.

Example 1.8.16. Let us describe the E-tangent bundle to E in the case of E being an Atiyah algebroid induced by a trivial principal $G-$bundle $\pi : G \times M \to M$. In such case, by left trivialization we have that the Atiyah algebroid is the vector bundle $\tau_{\mathfrak{g}\times TM} : \mathfrak{g} \times TM \to TM$. If $X \in \mathfrak{X}(M)$ and $\xi \in \mathfrak{g}$ then we may consider the section $X^{\xi} : M \to \mathfrak{g} \times TM$ of the Atiyah algebroid by

$$X^{\xi}(q) = (\xi, X(q)) \text{ for } q \in M.$$

Moreover, in this sense

$$[\![X^{\xi}, Y^{\xi}]\!]_{\mathfrak{g}\times TM} = ([X,Y]_{TM}, [\xi,\eta]_{\mathfrak{g}}), \quad \rho(X^{\xi}) = X.$$

On the other hand, if $(\xi, v_q) \in \mathfrak{g}\times T_qM$, then the fiber of $\mathcal{T}^{\tau_{\mathfrak{g}\times TM}}(\mathfrak{g}\times TM)$ over (ξ, v_q) is

$$\mathcal{T}^{\tau_{\mathfrak{g}\times TM}}_{(\xi,v_q)}(\mathfrak{g}\times TM) = \left\{((\eta,u_q),(\tilde{\eta},X_{v_q})) \in \mathfrak{g}\times T_qM\times\mathfrak{g}\times T_{v_q}(TM) \middle| u_q = T_{v_q}\tau_{\mathfrak{g}\times TM}(X_{v_q})\right\}.$$

This implies that $\mathcal{T}^{\tau_{\mathfrak{g}\times TM}}_{(\xi,v_q)}(\mathfrak{g}\times TM)$ may be identified with the space $(\mathfrak{g}\times\mathfrak{g})\times T_{v_q}(TM)$. Thus, the Lie algebroid $\mathcal{T}^{\tau_{\mathfrak{g}\times TM}}(\mathfrak{g}\times TM)$ may be identified with the vector bundle $\mathfrak{g}\times(\mathfrak{g}\times\mathfrak{g})\times TTM \to \mathfrak{g}\times TM$ whose vector bundle projection is

$$(\xi, ((\eta,\tilde{\eta}), X_{v_q})) \mapsto (\xi, v_q)$$

for $(\xi, ((\eta,\tilde{\eta}), X_{v_q})) \in \mathfrak{g}\times(\mathfrak{g}\times\mathfrak{g})\times TTM$. Therefore, if $(\eta,\tilde{\eta}) \in \mathfrak{g}\times\mathfrak{g}$ and $X \in \mathfrak{X}(TM)$ then one may consider the section $((\eta,\tilde{\eta}), X)$ given by

$$((\eta,\tilde{\eta}), X)(\xi, v_q) = (\xi, ((\eta,\tilde{\eta}), X(v_q))) \text{ for } (\xi, v_q) \in \mathfrak{g}\times T_qM.$$

Moreover,

$$[\![((\eta,\tilde{\eta}), X), ((\xi,\tilde{\xi}), Y)]\!]_{\tau^{(1)}_{\mathfrak{g}\times TM}} = (([\eta,\xi]_{\mathfrak{g}}, 0), [X,Y]_{TM}),$$

and the anchor map $\rho_1 : \mathfrak{g}\times(\mathfrak{g}\times\mathfrak{g})\times TTM \to \mathfrak{g}\times\mathfrak{g}\times TTM$ is defined as

$$\rho_1(\xi, ((\eta,\tilde{\eta}), X)) = ((\xi,\tilde{\eta}), X).$$

E-tangent bundle of the dual bundle of a Lie algebroid

Let $(E, [\![\ ,\ ]\!], \rho)$ be a Lie algebroid of rank n over a manifold of dimension m. Consider the projection of the dual E^* of E over M, $\tau_{E^*} : E^* \to M$, and define the prolongation $\mathcal{T}^{\tau_{E^*}}E$ of E over τ_{E^*}; that is,

$$\mathcal{T}^{\tau_{E^*}}E = \bigcup_{\mu\in E^*}\{(e, v_\mu) \in E\times T_\mu E^* \mid \rho(e) = T\tau_{E^*}(v_\mu)\}.$$

$\mathcal{T}^{\tau_{E^*}}E$ is a Lie algebroid over E^* of rank $2n$ with vector bundle projection $\tau^{(1)}_{E^*} : \mathcal{T}^{\tau_{E^*}}E \to E^*$ given by $\tau^{(1)}_{E^*}(e, v_\mu) = \mu$, for $(e, v_\mu) \in \mathcal{T}^{\tau_{E^*}}E$.

As before, if we now denote by (μ, e, v_μ) an element $(e, v_\mu) \in \mathcal{T}^{\tau_{E^*}}E$ where $\mu \in E^*$, we rewrite the definition of the prolongation Lie algebroid as the subset of $E^*\times E\times TE^*$ by

$$\mathcal{T}^{\tau_{E^*}}E = \{(\mu, e, v_\mu) \in E^*\times E\times TE^* \mid \rho(e) = (T\tau_{E^*})(v_\mu), v_\mu \in T_\mu E^* \text{ and } \tau_{E^*}(\mu) = \tau_E(e)\}.$$

If (x^i) are local coordinates on an open subset U of M, $\{e_A\}$ is a basis of sections of the vector bundle $(\tau_E)^{-1}(U) \to U$ and $\{e^A\}$ is its dual basis, then $\{\tilde{e}^{(1)}_A, \tilde{e}^{(2)}_A\}$ is a basis of sections of the vector bundle $\tau^{(1)}_{E^*}$, where

$$\begin{aligned}
\tilde{e}_A^{(1)}(\mu) &= \left(\mu, e_A(\tau_{E^*}(\mu)), \rho_A^i \frac{\partial}{\partial x^i}\Big|_\mu\right), \\
(\tilde{e}^A)^{(2)}(\mu) &= \left(\mu, 0, \frac{\partial}{\partial p_A}\Big|_\mu\right),
\end{aligned}$$

for $\mu \in (\tau_{E^*})^{-1}(U)$. Here, (x^i, p_A) are the local coordinates on E^* induced by the local coordinates (x^i) and the basis of sections of E^*, $\{e^A\}$.

Using the local basis $\{\tilde{e}_A^{(1)}, (\tilde{e}^A)^{(2)}\}$, one may introduce, in a natural way, local coordinates $(x^i, p_A; z^A, v_A)$ on $\mathcal{T}^{\tau_{E^*}}E$. If ω^* is a point of $\mathcal{T}^{\tau_{E^*}}E$ over $(x,p) \in E^*$, then

$$\omega^*(x,p) = z^A \tilde{e}_A^{(1)}(x,p) + v_A (\tilde{e}^A)^{(2)}(x,p).$$

Denoting by $\rho_{\tau_{E^*}^{(1)}}$ the anchor map of the Lie algebroid $\mathcal{T}^{\tau_{E^*}}E \to E^*$ locally given by $\rho_{\tau_{E^*}^{(1)}}(x^i, p_A, z^A, v_A) = (x^i, p_A, \rho_A^i z^A, v^A)$, we have that

$$\rho_{\tau_{E^*}^{(1)}}(\tilde{e}_A^{(1)})(\mu) = \left(\mu, \rho_A^i \frac{\partial}{\partial x^i}\Big|_\mu\right), \quad \rho_{\tau_{E^*}^{(1)}}((\tilde{e}^A)^{(2)})(\mu) = \left(\mu, \frac{\partial}{\partial p_A}\Big|_\mu\right).$$

Therefore, we have that the corresponding vector field $\rho_{\tau_{E^*}^{(1)}}(V)$ for a section $V = (x^i, p_A, z^A(x,p), v_A(x,p))$ is given by

$$\rho_{\tau_{E^*}^{(1)}}(V) = \rho_A^i z^A \frac{\partial}{\partial x^i}\Big|_\mu + v_A \frac{\partial}{\partial p_A}\Big|_{e^*}.$$

Finally, the structure of the Lie algebroid $(\mathcal{T}^{\tau_{E^*}}E, [\![\cdot,\cdot]\!]_{\tau_{E^*}^{(1)}}, \rho_{\tau_{E^*}^{(1)}})$, is determined by the bracket of sections

$$\begin{aligned}
[\![\tilde{e}_A^{(1)}, \tilde{e}_B^{(1)}]\!]_{\tau_{E^*}^{(1)}} &= \mathcal{C}_{AB}^C \tilde{e}_C^{(1)}, \\
[\![\tilde{e}_A^{(1)}, (\tilde{e}^B)^{(2)}]\!]_{\tau_{E^*}^{(1)}} &= [\![(\tilde{e}^A)^{(2)}, (\tilde{e}^B)^{(2)}]\!]_{\tau_{E^*}^{(1)}} = 0,
\end{aligned}$$

for all A, B and C. Thus, if we denote by $\{\tilde{e}^A_{(1)}, (\tilde{e}_A)_{(2)}\}$ is the dual basis of $\{\tilde{e}_A^{(1)}, (\tilde{e}^A)^{(2)}\}$, then

$$\begin{aligned}
d^{\mathcal{T}^{\tau_{E^*}}E} f(x^i, p_A) &= \rho_A^i \frac{\partial f}{\partial x^i} \tilde{e}^A_{(1)} + \frac{\partial f}{\partial p_A} (\tilde{e}_A)_{(2)}, \\
d^{\mathcal{T}^{\tau_{E^*}}E} \tilde{e}^C_{(1)} &= -\frac{1}{2} \mathcal{C}_{AB}^C \tilde{e}^A_{(1)} \wedge \tilde{e}^B_{(1)}, \qquad d^{\mathcal{T}^{\tau_{E^*}}E} (\tilde{e}_C)_{(2)} = 0,
\end{aligned}$$

for $f \in C^\infty(E^*)$. We refer to [7] and [111] for further details about the Lie algebroid structure of the E-tangent bundle of the dual bundle of a Lie algebroid.

Example 1.8.17. In the case of $E = TM$ one may identify $\mathcal{T}^{\tau_{E^*}}E$ with $T(T^*M)$ with the standard Lie algebroid structure.

Example 1.8.18. Let $\mathfrak{g}$ be a real Lie algebra of finite dimension. Then $\mathfrak{g}$ is a Lie algebroid over a single point. Using that the anchor map is zero we have that $\mathcal{T}^{\tau_{\mathfrak{g}^*}}\mathfrak{g}$ may be identified with the vector bundle $pr_1 : \mathfrak{g}^* \times (\mathfrak{g} \times \mathfrak{g}^*) \to \mathfrak{g}^*$. Under this identification the anchor map is given by

$$\rho_{\tau^{(1)}_{\mathfrak{g}^*}} : \mathfrak{g}^* \times (\mathfrak{g} \times \mathfrak{g}^*) \to T\mathfrak{g}^* \simeq \mathfrak{g}^* \times \mathfrak{g}^*, \quad (\mu, (\xi, \alpha)) \mapsto (\mu, \alpha)$$

and the Lie algebroid of two constant sections $(\xi, \alpha), (\eta, \beta) \in \mathfrak{g} \times \mathfrak{g}^*$ is the constant section $([\xi, \eta], 0)$.

1.8.4 Symplectic Lie algebroids

In this subsection we will recall some results given in [111] about symplectic Lie algebroids.

Definition 1.8.19. *A Lie algebroid $(E, [\![\cdot,\cdot]\!], \rho)$ over a manifold M is said to be symplectic if it admits a symplectic section Ω, that is, Ω is a section of the vector bundle $\bigwedge^2 E^* \to M$ such that:*

- *For all $x \in M$, the 2-form $\Omega_x : E_x \times E_x \to \mathbb{R}$ in the vector space E_x is nondegenerate and*
- *Ω is a 2-cocycle, that is, $d^E\Omega = 0$.*

The canonical symplectic structure of $\mathcal{T}^{\tau_{E^*}}E$

Let $(E, [\![\cdot,\cdot]\!], \rho)$ be a Lie algebroid of rank n over a manifold M of dimension m and $\mathcal{T}^{\tau_{E^*}}E$ be the prolongation of E over the vector bundle projection $\tau_{E^*} : E^* \to M$. We may introduce a canonical section λ_E of $(\mathcal{T}^{\tau_{E^*}}E)^*$ as follows. If $\mu \in E^*$ and (e, v_μ) is a point on the fibre of $\mathcal{T}^{\tau_{E^*}}E$ over μ then

$$\lambda_E(\mu)(e, v_\mu) = \langle \mu, e \rangle. \tag{1.18}$$

λ_E is called the *Liouville section* of $\mathcal{T}^{\tau_{E^*}}E$. Now, in an analogous way that the canonical symplectic form is defined from the Liouville 1-form on the cotangent bundle, we introduce the 2-section Ω_E on $\mathcal{T}^{\tau_{E^*}}E$ as

$$\Omega_E = -d^{\mathcal{T}^{\tau_{E^*}}E}\lambda_E. \tag{1.19}$$

Proposition 1.8.20. *[111] Ω_E is a non-degenerate 2-section of $\mathcal{T}^{\tau_{E^*}}E$ such that $d^{\mathcal{T}^{\tau_{E^*}}E}\Omega_E = 0$.*

Proof. If (x^i) are local coordinates on an open subset U of M, $\{e_A\}$ is a basis of sections of the vector bundle τ_E, (x^i, p_A) are induced local coordinates of E^* on τ_{E^*} and $\{\tilde{e}_A^{(1)}, (\tilde{e}^A)^{(2)}\}$ is the basis of $\tau^{(1)}_{A^*}$ then, using (1.18), it follows that

$$\lambda_E(x^i, p_A) = p_A \tilde{e}^A_{(1)} \tag{1.20}$$

where $\{\tilde{e}^A_{(1)}, (\tilde{e}_A)_{(2)}\}$ is the dual basis of $\{\tilde{e}_A^{(1)}, (\tilde{e}^A)^{(2)}\}$. Thus, from (1.18),(1.19) and (1.20) we obtain that

$$\Omega_E = \tilde{e}^A_{(1)} \wedge (\tilde{e}_A)_{(2)} + \frac{1}{2}\mathcal{C}^C_{AB} p_C \tilde{e}^A_{(1)} \wedge \tilde{e}^B_{(1)}. \tag{1.21}$$

Now, it is straightforward to check that Ω_E is non-degenerate and $d^{\mathcal{T}^{\tau_{E^*}}E}\Omega = 0$. $\square$

Therefore Ω_E is a symplectic 2-section on $\mathcal{T}^{\tau_{E^*}}E$ called *canonical symplectic section* on $\mathcal{T}^{\tau_{E^*}}E$.

Example 1.8.21. If E is the standard Lie algebroid TM then $\lambda_E = \lambda$ and $\Omega_E = \omega_M$ are the usual Liouville 1-form and canonical symplectic 2-form on T^*M, respectively.

Example 1.8.22. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\mathfrak{g}$ is a Lie algebroid over a single point $M = \{q\}$. If $\xi \in \mathfrak{g}$ and $\mu, \alpha \in \mathfrak{g}^*$ then

$$\lambda_{\mathfrak{g}}(\mu)(\xi, \alpha) = \mu(\xi)$$

is the Liouville 1-section on $\mathfrak{g}^* \times (\mathfrak{g} \times \mathfrak{g}^*)$. Thus, the symplectic section $\Omega_{\mathfrak{g}}$ is

$$\Omega_{\mathfrak{g}}(\mu)((\xi, \alpha), (\eta, \beta)) = \langle \mu, [\xi, \eta]\rangle - \langle \alpha, \eta\rangle - \langle \beta, \xi\rangle$$

for $\mu \in \mathfrak{g}^*, (\xi, \alpha), (\eta, \beta) \in \mathfrak{g} \times \mathfrak{g}^*$.

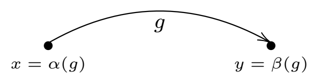
1.9 Lie groupoids

1.9.1 Generalities about Lie groupoids

A groupoid is a small category in which every morphism is an isomorphism (i.e; all morphism is invertible). That is,

Definition 1.9.1. *A groupoid over a set Q, (denoted $G \rightrightarrows Q$) consists of a set of objects Q, a set of morphisms G, and the following structural maps:*

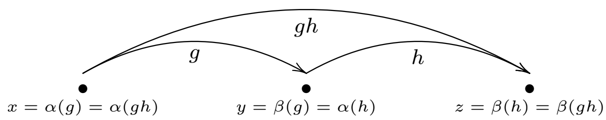
- *a source map $\alpha\colon G \to Q$ and target map $\beta\colon G \to Q$;*



- *an associative multiplication map $m\colon G_2 \to G$, $m(g,h) = gh$, with $\alpha(g) = \alpha(gh)$ and $\beta(h) = \beta(gh)$ where*

$$G_2 = G\ {}_\beta\!\times_\alpha G = \{(g,h) \in G \times G \mid \beta(g) = \alpha(h)\}$$

is call the set of composable pairs; gh is the composite arrow from x to z



- *an identity section of α and β, $\epsilon\colon Q \to G$; such that for all $g \in G$,*

$$\epsilon\left(\alpha(g)\right) g = g = g\epsilon\left(\beta(g)\right);$$

- *an inversion map $i\colon G \to G$, $g \mapsto g^{-1}$, such that for all $g \in G$,*

$$gg^{-1} = \epsilon\left(\alpha(g)\right), \qquad g^{-1}g = \epsilon\left(\beta(g)\right).$$

$$\begin{tikzcd} x = \alpha(g) = \beta(g^{-1}) \arrow[r, bend left, "g"] & y = \beta(g) = \alpha(g^{-1}) \arrow[l, bend left, "g^{-1}"] \end{tikzcd}$$

Remark 1.9.2. Alternatively, a groupoid can be seen as a weak version of a group, where the multiplication will be defined only for elements in $G_2 \subset G \times G$. $\diamond$

Definition 1.9.3. *Given a groupoid $G \rightrightarrows Q$ and an element $g \in G$, define the left translation $\ell_g\colon \alpha^{-1}\left(\beta(g)\right) \to \alpha^{-1}\left(\alpha(g)\right)$ and right translation $r_g\colon \beta^{-1}\left(\alpha(g)\right) \to \beta^{-1}\left(\beta(g)\right)$ by g to be*

$$\ell_g(h) = gh, \qquad r_g(h) = hg.$$

Moreover, $\ell_g^{-1} = \ell_{g^{-1}}$ and $r_g^{-1} = r_{g^{-1}}$.

Now we will focus on a particular class of groupoids, the Lie groupoids which have a differential structure in addition to their algebraic structure.

Definition 1.9.4. *A Lie groupoid is a groupoid $G \rightrightarrows Q$ where*

1. *G and Q are differentiable manifolds,*
2. *α, β are submersions,*
3. *the multiplication map m, is differentiable.*

Remark 1.9.5. In 1.9.4, α and β must be submersions so that G_2 is a differentiable manifold. From the definition it follows that m is a submersion, ϵ is an immersion, and i is a diffeomorphism. $\diamond$

One may introduce the notion of a left (right)-invariant vector field in a Lie groupoid.

Definition 1.9.6. *Given a Lie groupoid $G \rightrightarrows Q$, a vector field $X \in \mathfrak{X}(G)$ is* left-invariant *(respect.*right-invariant*) if X is α-vertical (respect. β-vertical), that is, $T\alpha(X) = 0$ ($T\beta(X) = 0$ respec.) and $\left(T_h\ell_g\right)\left(X(h)\right) = X\left(gh\right)$ ($\left(T_h r_g\right)\left(X(h)\right) = X\left(hg\right)$, respectively) for all $(g,h) \in G_2$ (or $(h,g) \in G_2$, respectively).*

In Lie groups, the infinitesimal version of a Lie group is a Lie algebra, therefore one can think that the infinitesimal version of a Lie groupoid is a Lie algebroid. We will give the definition of *Lie algebroid associated with G.*

Given a Lie groupoid $G \rightrightarrows Q$, consider the following vector bundle

$$AG = V\alpha \to Q$$

where $A_qG = \ker T_{\epsilon(q)}\alpha$, i.e., the tangent space to the α-fiber at the identity section, for $q \in Q$.

There is a bijection between sections $X \in \Gamma(\tau_{AG})$ and left-invariant vector fields $\overleftarrow{X} \in \mathfrak{X}(G)$. Let X be a section of τ_{AG}, then the corresponding left-invariant vector field on G is given by

$$\overleftarrow{X}(g) = \left(T_{\epsilon(\beta(g))}\ell_g\right)\left(X\left(\beta(g)\right)\right). \tag{1.22}$$

The Lie algebroid structure on AG is given by the bracket $[\![\ ,\ ]\!]$ and the anchor map ρ defined as follows

$$\overleftarrow{[\![X,Y]\!]} = \left[\overleftarrow{X}, \overleftarrow{Y}\right], \qquad \rho(X)(q) = \left(T_{\epsilon(q)}\beta\right)\left(X(q)\right),$$

for all $X, Y \in \Gamma(AG)$ and $q \in Q$.

Definition 1.9.7. *Given a Lie groupoid $G \rightrightarrows Q$, the triple $(AG, [\![\cdot,\cdot]\!], \rho)$ is called the Lie algebroid associated to $G \rightrightarrows Q$.*

Remark 1.9.8. Alternatively one can also establish a bijection between sections $X \in \Gamma(\tau_{AG})$ and right-invariant vector fields $\overrightarrow{X} \in \mathfrak{X}(Q)$, by

$$\overrightarrow{X}(g) = -\left(T_{\epsilon(\alpha(g))}\left(r_g \circ i\right)\right)\left(X\left(\alpha(g)\right)\right), \tag{1.23}$$

which yields the Lie bracket relation

$$\overrightarrow{[\![X,Y]\!]} = -\left[\overrightarrow{X}, \overrightarrow{Y}\right], \qquad \left[\overrightarrow{X}, \overleftarrow{Y}\right] = 0.$$

Thus $X \mapsto \overleftarrow{X}$ is a Lie algebra isomorphism, and $X \mapsto \overrightarrow{X}$ is a Lie algebra anti-isomorphism. $\diamond$

Examples of Lie groupoids:

We introduce now some examples of Lie groupoids

- *The pair (or banal) groupoid:* Let Q be a differentiable manifold, and we consider the product manifold $G = Q \times Q$. Then G is a Lie groupoid over Q where the source and target maps α and β are the projections onto the first and second factors respectively, the identity is defined as $\epsilon(q) = (q, q)$ for all $q \in Q$, the multiplication $m((q,s),(s,r)) = (q,r)$ for $(q,s),(s,r) \in Q \times Q$ and the inverse map $i(q,s) = (s,q)$.

Observe that, if q is a point of Q, then

$$V_{\epsilon(q)}\alpha \simeq T_qQ.$$

- *The case of Lie groups:* Let G be a Lie group. G is a Lie groupoid over $\{\mathfrak{e}\}$, the identity element of G. The Lie algebroid associated with G is the Lie algebra $\mathfrak{g} = T_{\mathfrak{e}}G$ of G. The structural maps of the Lie groupoid G are

$$\alpha(g) = \mathfrak{e} \quad \beta(g) = \mathfrak{e}, \quad \epsilon(\mathfrak{e}) = \mathfrak{e}, \quad i(g) = g^{-1}, \quad m(g,h) = gh \quad \text{for } g, h \in G.$$

• *The cotangent groupoid:* Let $G \rightrightarrows Q$ be a Lie groupoid. If A^*G is the dual bundle to AG then the cotangent bundle T^*G is a Lie groupoid over A^*G. The projections $\tilde{\beta}$ and $\tilde{\alpha}$, the partial multiplication $\oplus_{T^*G}$, the identity section $\tilde{\epsilon}$ and the inversion $\tilde{\imath}$ are defined as follows,

$$
\begin{aligned}
&\tilde{\beta}(\mu_g)(X) = \mu_g((T_{\epsilon(\beta(g))}l_g)(X)), \text{ for } \mu_g \in T_g^*G \text{ and } X \in A_{\beta(g)}G,\\
&\tilde{\alpha}(\nu_h)(Y) = \nu_h((T_{\epsilon(\alpha(h))}r_h)(Y - (T_{\epsilon(\alpha(h))}(\epsilon \circ \beta))(Y))),\\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } \nu_h \in T_h^*G \text{ and } Y \in A_{\alpha(h)}G,\\
&(\mu_g \oplus_{T^*G} \nu_h)(T_{(g,h)}m(X_g, Y_h)) = \mu_g(X_g) + \nu_h(Y_h),\\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } (X_g, Y_h) \in T_{(g,h)}G_2,\\
&\tilde{\epsilon}(\mu_x)(X_{\epsilon(x)}) = \mu_x(X_{\epsilon(x)} - (T_{\epsilon(x)}(\epsilon \circ \alpha))(X_{\epsilon(x)}))),\\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } \mu_x \in A_x^*G \text{ and } X_{\epsilon(x)} \in T_{\epsilon(x)}G,\\
&\tilde{\imath}(\mu_g)(X_{g^{-1}}) = -\mu_g((T_{g^{-1}}i)(X_{g^{-1}})), \text{ for } \mu_g \in T_g^*G \text{ and } X_{g^{-1}} \in T_{g^{-1}}G.
\end{aligned}
\tag{1.24}
$$

(for more details, see [60]).

• *Action Lie groupoids.* Let $G \rightrightarrows Q$ be a Lie groupoid and $\pi : P \to Q$ be a smooth map. If $P\, {}_\pi\!\times_\alpha G = \{(p, g) \in P \times G / \pi(p) = \alpha(g)\}$ then a right action of G on π is a smooth map

$$P\, {}_\pi\!\times_\alpha G \to P, \;\; (p, g) \to pg,$$

which satisfies the following relations

$$
\begin{array}{rcll}
\pi(pg) & = & \beta(g), & \text{for } (p, g) \in P\, {}_\pi\!\times_\alpha G,\\
(pg)h & = & p(gh), & \text{for } (p, g) \in P\, {}_\pi\!\times_\alpha G \text{ and } (g, h) \in G_2, \text{ and}\\
p\epsilon(\pi(p)) & = & p, & \text{for } p \in P.
\end{array}
$$

Given such an action one constructs *the action Lie groupoid* $P\, {}_\pi\!\times_\alpha G$ over P by defining

$$
\begin{array}{lcl}
\tilde{\alpha}_\pi : P\, {}_\pi\!\times_\alpha G \longrightarrow P & ; & (p, g) \longrightarrow p,\\
\tilde{\beta}_\pi : P\, {}_\pi\!\times_\alpha G \longrightarrow P & ; & (p, g) \longrightarrow pg,\\
\tilde{\epsilon}_\pi : P \longrightarrow P\, {}_\pi\!\times_\alpha G & ; & p \longrightarrow (p, \epsilon(\pi(p))),\\
\tilde{m}_\pi : (P\, {}_\pi\!\times_\alpha G)_2 \longrightarrow P\, {}_\pi\!\times_\alpha G & ; & ((p, g), (pg, h)) \longrightarrow (p, gh),\\
\tilde{i}_\pi : P\, {}_\pi\!\times_\alpha G \longrightarrow P\, {}_\pi\!\times_\alpha G & ; & (p, g) \longrightarrow (pg, g^{-1}).
\end{array}
$$

Now, if $p \in P$, we consider the map $p\,\cdot : \alpha^{-1}(\pi(p)) \to P$ given by

$$p \cdot g = pg.$$

Then, if $\tau : AG \to Q$ is the Lie algebroid of G, the $\mathbb{R}$-linear map $\Phi : \Gamma(\tau) \to \mathfrak{X}(P)$ defined by

$$\Phi(X)(p) = (T_{\epsilon(\pi(p))}p\,\cdot)(X(\pi(p))), \quad \text{for } X \in \Gamma(\tau) \text{ and } p \in P,$$

induces an action of AG on $\pi : P \to Q$. In addition, the Lie algebroid associated with the Lie groupoid $P\, {}_\pi\!\times_\alpha G \rightrightarrows P$ is the action Lie algebroid (for more details, see [118]).

•*Atiyah or gauge groupoids:* Let $p : Q \to M$ be a principal G-bundle. Then, the free action, $\Phi : G \times Q \to Q$, $(g, q) \to \Phi(g, q) = gq$, of G on Q induces, in a natural way, a free action $\Phi \times \Phi : G \times (Q \times Q) \to Q \times Q$ of G on $Q \times Q$ given by $(\Phi \times \Phi)(g,(q,q')) = (gq, gq')$, for $g \in G$ and $(q,q') \in Q \times Q$. Moreover, one may consider the quotient manifold $(Q \times Q)/G$ and it admits a Lie groupoid structure over M with structural maps given by

$$\begin{array}{lcl}
\tilde{\alpha} : (Q \times Q)/G \longrightarrow M & ; & [(q,q')] \longrightarrow p(q), \\
\tilde{\beta} : (Q \times Q)/G \longrightarrow M & ; & [(q,q')] \longrightarrow p(q'), \\
\tilde{\epsilon} : M \longrightarrow (Q \times Q)/G & ; & x \longrightarrow [(q,q)], \quad \text{if } p(q) = x, \\
\tilde{m} : ((Q \times Q)/G)_2 \longrightarrow (Q \times Q)/G & ; & ([(q,q')],[(gq',q'')]) \longrightarrow [(gq,q'')], \\
\tilde{i} : (Q \times Q)/G \longrightarrow (Q \times Q)/G & ; & [(q,q')] \longrightarrow [(q',q)].
\end{array}$$

This Lie groupoid is called *the Atiyah (gauge) groupoid associated with the principal G-bundle* $p : Q \to M$ (see [124]).

If x is a point of M such that $p(q) = x$, with $q \in Q$, and $p_{Q\times Q} : Q \times Q \to (Q \times Q)/G$ is the canonical projection then it is clear that

$$V_{\tilde{\epsilon}(x)}\tilde{\alpha} = (T_{(q,q)}p_{Q\times Q})(\{0_q\} \times T_qQ).$$

Thus, if $\tau_{TQ/G} : TQ/G \to M$ is the Atiyah algebroid associated with the principal G-bundle $p : G \to M$ then the linear maps

$$(TQ/G)_x \to V_{\tilde{\epsilon}(x)}\tilde{\alpha} \;\; ; \;\; [v_q] \to (T_{(q,q)}p_{Q\times Q})(0_q, v_q), \text{ with } v_q \in T_qQ,$$

induce an isomorphism (over the identity of M) between the Lie algebroids $\tau : A((Q \times Q)/G) \to M$ and $\tau_{TQ/G} : TQ/G \to M$.

1.9.2 Morphism of Lie groupoids and symplectic Lie groupoids

Definition 1.9.9. *Given two Lie groupoids, $G \rightrightarrows Q$ and $G' \rightrightarrows Q'$, a smooth map $\Phi : G \to G'$ is a morphism of Lie groupoids if, for every composable pair $(g,h) \in G_2$, it satisfies $(\Phi(g);\Phi(h)) \in G'_2$ and $\Phi(gh) = \Phi(g)\Phi(h)$.*

A morphism of Lie groupoids $\Phi : G \to G'$ induces a smooth map $\Phi_0 : Q \to Q'$ in such a way

$$\bar{\alpha} \circ \Phi = \Phi_0 \circ \alpha, \quad \bar{\beta} \circ \Phi = \Phi_0 \circ \beta, \quad \Phi \circ \epsilon = \bar{\epsilon} \circ \Phi_0,$$

where $\bar{\alpha}, \bar{\beta}, \bar{\epsilon}$ are some structure maps of the Lie groupoid $G' \rightrightarrows Q'$. That is, the following diagram is commutative,

$$\begin{array}{ccc}
G' & \xrightarrow{\bar{\alpha},\bar{\beta}} & Q' \\
{\scriptstyle\Phi}\big\downarrow & & \big\downarrow{\scriptstyle\Phi_0} \\
G & \xrightarrow{\alpha,\beta} & Q
\end{array}$$

Moreover, Φ induces a morphism $A\Phi : AG \to AG'$ of the corresponding Lie algebroids and

$$\begin{array}{rcl}
\overleftarrow{A\Phi(v)}(g) & = & T\Phi(\overleftarrow{v}(g)), \\
\overrightarrow{A\Phi(w)}(g) & = & T\Phi(\overrightarrow{w}(g)),
\end{array}$$

for all $g \in G$, $v \in A_{\beta(g)}G$ and $w \in A_{\alpha(g)}G$. Moreover, for $X \in \Gamma(AG), \overline{X} \in \Gamma(AG')$ we have that $A\Phi \circ X = \overline{X} \circ \Phi_0$ if and only if $T\Phi \circ \overleftarrow{X} = \overleftarrow{\overline{X}} \circ \Phi$ (or alternatively, $T\Phi \circ \overrightarrow{X} = \overrightarrow{\overline{X}} \circ \Phi$). That is, X and $\overline{X}$ are $A\Phi$-related if and only if their corresponding left-invariant (and right invariant) vector fields are Φ-related (See [118] for more details).

Symplectic groupoids:

Finally, we introduce a subclass of Lie groupoids with an additional structure: symplectic groupoids. They are endowed with a symplectic manifold structure in the following sense,

Definition 1.9.10. *A* symplectic groupoid *is a Lie groupoid $G \rightrightarrows Q$ with a symplectic form ω on G such that the graph of the composition law m is given by*

$$\text{graph}(m) := \{(g, h, r) \in G \times G \times G \mid (g, h) \in G_2 \text{ and } r = m(g, h)\}$$

is a Lagrangian submanifold of $G \times G \times G^-$ with the product symplectic form where the first two factors G being endowed with the symplectic form ω and the third factor G^- being G with the symplectic form $-\omega$.

A typical example of symplectic groupoid is the cotangent groupoid. If $G \rightrightarrows Q$ is a Lie groupoid then the cotangent groupoid $T^*G \rightrightarrows A^*G$ is a symplectic groupoid with the canonical symplectic 2-form ω_G.

1.9.3 The prolongation of a Lie groupoid over a fibration

Given a Lie groupoid $G \rightrightarrows Q$ and a fibration $\pi : P \to Q$, we consider the set

$$\mathcal{P}^\pi G = P_\pi \times_\alpha G_\beta \times_\pi P = \{(p, g, p') \in P \times G \times P / \pi(p) = \alpha(g), \beta(g) = \pi(p')\}.$$

This set has a Lie groupoid structure over P, where the structural maps are given by

$$\begin{aligned}
\alpha^\pi &: \mathcal{P}^\pi G \to P, \quad (p, g, p') \mapsto p;\\
\beta^\pi &: \mathcal{P}^\pi G \to P, \quad (p, g, p') \mapsto p';\\
m^\pi &: (\mathcal{P}^\pi G)_2 \to \mathcal{P}^\pi G, \quad ((p, g, p'), (p', h, p'')) \mapsto (p, gh, p'');\\
\epsilon^\pi &: P \to \mathcal{P}^\pi G, \quad p \mapsto (p, \epsilon(\alpha(p)), p);\\
i^\pi &: \mathcal{P}^\pi G \to \mathcal{P}^\pi G, \quad (p, g, p') \mapsto (p', g^{-1}, p).
\end{aligned}$$

$\mathcal{P}^\pi G$ is called *prolongation of G over $\pi : P \to Q$* (See [118],[124] and [161]).

In what follows we consider the prolongation $\mathcal{P}^\alpha G$ of the Lie groupoid G over its source map $\alpha : G \to Q$, that is, one can consider the subset of $3G := G \times G \times G$,

$$\mathcal{P}^\alpha G = G_\alpha \times_\alpha G_\beta \times_\alpha G = \{(g, h, r) \in 3G / \alpha(g) = \alpha(h), \beta(h) = \alpha(r)\}.$$

$\mathcal{P}^\alpha G$ is a Lie groupoid over G. Observe that $G_2 \subseteq \mathcal{P}^\alpha G$ where the inclusion is given by

$$\begin{aligned}
i_{G_2} : G_2 &\hookrightarrow \mathcal{P}^\alpha G\\
(g, h) &\mapsto (g, g, h).
\end{aligned}$$

Following the constructions given before we can construct the Lie algebroid associated with $\mathcal{P}^\alpha G$. This can be identified with the prolongation $\mathcal{P}^\alpha(AG)$ of AG over $\alpha : G \to Q$; that is,

Definition 1.9.11. *The Lie algebroid associated with a prolongation of a Lie groupoid G over α is given by,*

$$A_g\left(\mathcal{P}^\alpha G\right) = \{(a_{\epsilon(\alpha(g))}; Y_g) \in A_{\alpha(g)}G \times T_gG/(T_g\alpha)(Y_g) = (T_{\epsilon(\alpha(g))}\beta)(a_{\epsilon(\alpha(g))})\}.$$

Remark 1.9.12. If we consider the linear isomorphism

$$\begin{array}{rcl}(\Psi^\alpha)_g : A_g\left(\mathcal{P}^\alpha G\right) & \to & \mathcal{P}^\alpha_g(AG) \subset A_{\alpha(g)}G \times T_gG \\ (0_g, a_{\epsilon(\alpha(g))}, Y_g) & \mapsto & (a_{\epsilon(\alpha(g))}, Y_g) \quad \forall g \in G\end{array}$$

then the mapping $(\Psi^\alpha)_g$, $g \in G$ induces an isomorphism $\Psi^\alpha : A\left(\mathcal{P}^\alpha G\right) \to \mathcal{P}^\alpha(AG)$ between the Lie algebroids $A\left(\mathcal{P}^\alpha G\right)$ and $\mathcal{P}^\alpha(AG)$ (Fore more details see [80] and [124]). $\diamond$

A section Z of $A\left(\mathcal{P}^\alpha G\right)$ can be expressed as

$$Z(g) = (X(\alpha(g)), Y(g))$$

where $X \in \Gamma(\tau_{AG})$ and $Y \in \mathfrak{X}(G)$; such that $T\beta(X) = T\alpha(Y)$.

The corresponding left-invariant and right-invariant vector fields are

$$\overleftarrow{Z}(g, h, r) = (0_g, \overleftarrow{X}(h), Y(r)), \tag{1.25}$$
$$\overrightarrow{Z}(g, h, r) = (-Y(g), \overrightarrow{X}(h), 0_r). \tag{1.26}$$

Chapter 2

Geometrical description of classical mechanics

It is well-known that the velocity phase space of a mechanical system may be identified with the tangent bundle TQ of the n-dimensional configuration space Q. Under this identification, the Lagrangian function is a real C^∞-function L on TQ and the Euler-Lagrange equations are

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) - \frac{\partial L}{\partial q^i} = 0, \quad i = 1, ..., n$$

where $(q^i, \dot{q}^i)$ are local fibred coordinates on TQ, which represent the positions and the velocities of the system, respectively.

If the Lagrangian function is hyperregular one may define the Hamiltonian function $H : T^*Q \longrightarrow \mathbb{R}$ on the phase space of momenta T^*Q and the Euler-Lagrange equations are equivalent to the Hamilton equations for H

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}, \quad i = 1, ..., n.$$

Here, (q^i, p_i) are local fibred coordinates on T^*Q which represent the position and the momenta of the system, respectively.

Solutions of the previous Hamilton equations are just the integral curves of the Hamiltonian vector field X_H on T^*Q which are characterized by the condition

$$i_{X_H}\omega_Q = dH,$$

being ω_Q the canonical symplectic structure of T^*Q (for more details see, for instance, [1, 112]).

Lagrangian (Hamiltonian) mechanics may be also formulated in terms of Lagrangian submanifolds of special symplectic manifolds. Lagrangian submanifolds play an important role in the geometry of several aspects related to classical and quantum mechanics. In [169, 170] Tulczyjew proved that it is possible to interpret the ordinary Lagrangian and Hamiltonian dynamics as Lagrangian submanifolds of convenient special symplectic manifolds. To do that, he introduced canonical isomorphisms which commute the tangent and cotangent functors. This construction is the so-called Tulczyjew triple for classical mechanics. It is possible to interpret first order Lagrangian (and Hamiltonian) dynamics through the double vector bundle structure of the cotangent bundles T^*TQ and T^*T^*Q inducing an anti-symplectomorphism between these double vector bundles.

In this chapter, we explore some geometric techniques to describe the formulation of first-order and higher-order mechanics. We give a brief review of the description of classical mechanics in terms of Lie algebroids following [111] and we will introduce the constraint algorithm for presymplectic Lie algebroids constructed in [87] which generalizes the well-known Gotay-Nester-Hinds algorithm [73]. We will derive first-order and higher-order dynamics respectively in terms of Lagrangian submanifolds using the Tulczyjew triple constructed in [109] and we will construct a double vector bundle anti-symplectomorphism which generalizes to the higher-order case the one given in [77].

2.1 Mechanics on Lie algebroids

The geometric description of the mechanics in terms of Lie algebroids gives a general framework to obtain all the important equations in mechanics (Euler-Lagrange, Euler-Poincaré, Lagrange-Poincaré,...). In this section we will use the notions of Lie algebroid and prolongation of a Lie algebroid described in §1.8.1 to derive the Euler-Lagrange equations and Hamilton equations on Lie algebroids.

In [134] (see also [111]) a geometric formalism for Lagrangian mechanics on Lie algebroids was introduced. It is developed in the prolongation $\mathcal{T}^{\tau_E}E$ of a Lie algebroid E (see §1.8) over the vector bundle projection $\tau_E : E \to M$. The prolongation of the Lie algebroid is playing the same role as TTQ in the standard mechanics. We first derive the canonical geometrical structures defined on $\mathcal{T}^{\tau_E}E$ to derive the Euler-Lagrange equations on Lie algebroids.

Two canonical objects on $\mathcal{T}^{\tau_E}E$ are *the Euler section* Δ and *the vertical endomorphism* S. Δ is the section of $\mathcal{T}^{\tau_E}E \to E$ locally defined by

$$\Delta = y^A e_A^{(2)} \tag{2.1}$$

and S is the section of the vector bundle $(\mathcal{T}^{\tau_E}E) \otimes (\mathcal{T}^{\tau_E}E)^* \to E$ locally characterized by the following conditions

$$Se_A^{(1)} = e_A^{(2)}, \quad Se_A^{(2)} = 0, \quad \text{for all } A. \tag{2.2}$$

Finally, a section ξ of $\mathcal{T}^{\tau_E}E \to E$ is said to be a *second order differential equation* (SODE) on E if $S(\xi) = \Delta$ or, alternatively, $pr_1(\xi(e)) = e$, for all $e \in E$ (for more details, see [111]).

Given a Lagrangian function $L \in C^\infty(E)$ we define the *Cartan 1-section* $\Theta_L \in \Gamma((\mathcal{T}^{\tau_E}E)^*)$, the *Cartan 2-section* $\omega_L \in \Gamma(\wedge^2(\mathcal{T}^{\tau_E}E)^*)$ and the *Lagrangian energy* $E_L \in C^\infty(E)$ as

$$\Theta_L = S^*(d^{\mathcal{T}^{\tau_E}E}L), \qquad \omega_L = -d^{\mathcal{T}^{\tau_E}E}\Theta_L \qquad E_L = \mathcal{L}_\Delta^{\mathcal{T}^{\tau_E}E}L - L.$$

If (x^i, y^A) are local fibred coordinates on E, $(\rho^i_A, \mathcal{C}^C_{AB})$ are the corresponding local structure functions on E and $\{e^{(1)}_A, e^{(2)}_A\}$ the corresponding local basis of sections of $\mathcal{T}^{\tau_E}E$ then

$$\omega_L = \frac{\partial^2 L}{\partial y^A \partial y^B} e^A_{(1)} \wedge e^B_{(2)} + \frac{1}{2}\left(\frac{\partial^2 L}{\partial x^i \partial y^A}\rho^i_B - \frac{\partial^2 L}{\partial x^i \partial y^B}\rho^i_A + \frac{\partial L}{\partial y^A}\mathcal{C}^C_{AB}\right) e^A_{(1)} \wedge e^B_{(1)}, \tag{2.3}$$

$$E_L = \frac{\partial L}{\partial y^A} y^A - L. \tag{2.4}$$

Now, a curve $t \to c(t)$ on E is a solution of the *Euler-Lagrange equations* for L if

- c is *admissible* (that is, $\rho(c(t)) = \dot{m}(t)$, where $m = \tau_E \circ c$) and
- $i_{(c(t),\dot{c}(t))}\omega_L(c(t)) - d^{\mathcal{T}^{\tau_E}E}E_L(c(t)) = 0$, for all t.

If $c(t) = (x^i(t), y^A(t))$ then c is a solution of the Euler-Lagrange equations for L if and only if

$$\dot{x}^i = \rho^i_A y^A, \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^A} + \frac{\partial L}{\partial y^C}\mathcal{C}^C_{AB} y^B - \rho^i_A \frac{\partial L}{\partial x^i} = 0. \tag{2.5}$$

Note that if E is the standard Lie algebroid TQ then the above equations are the classical Euler-Lagrange equations for $L : TQ \to \mathbb{R}$.

On the other hand, the Lagrangian function L is said to be *regular* if ω_L is a symplectic section. In such a case, there exists a unique solution ξ_L verifying

$$i_{\xi_L}\omega_L - d^{\mathcal{T}^{\tau_E}E}E_L = 0\ . \tag{2.6}$$

In addition, one can check that $i_{S\xi_L}\omega_L = i_\Delta \omega_L$ which implies that ξ_L is a SODE section. Thus, the integral curves of ξ_L (that is, the integral curves of the vector field $\rho_1(\xi_L)$) are solutions of the Euler-Lagrange equations for L. ξ_L is called the *Euler-Lagrange section* associated with L.

From (2.3), we deduce that the Lagrangian L is regular if and only if the matrix $(W_{AB}) = \left(\frac{\partial^2 L}{\partial y^A \partial y^B}\right)$ is regular. Moreover, the local expression of ξ_L is

$$\xi_L = y^A e^{(1)}_A + f^A e^{(2)}_A,$$

where the functions f^A satisfy the linear equations

$$\frac{\partial^2 L}{\partial y^B \partial y^A} f^B + \frac{\partial^2 L}{\partial x^i \partial y^A}\rho^i_B y^B + \frac{\partial L}{\partial y^C}\mathcal{C}^C_{AB} y^B - \rho^i_A \frac{\partial L}{\partial x^i} = 0, \ \ \forall\, A.$$

Another possibility is when the matrix $(W_{AB}) = \left(\frac{\partial^2 L}{\partial y^A \partial y^B}\right)$ is non regular. This type of lagrangian is called *singular* or *degenerate lagrangian*. In such a case, ω_L is not a symplectic section and Equation (2.6) has no solution, in general, and even if it exists it will not be unique. In the next section, we will give the extension of the classical Gotay-Nester-Hinds algorithm [73] for presymplectic systems on Lie algebroids given in [87], which in particular will be applied to the case of singular lagrangians in Section 3.1.

For an arbitrary Lagrangian function $L : E \to \mathbb{R}$, we introduce the *Legendre transformation* associated with L as the smooth map $leg_L : E \to E^*$ defined by

$$leg_L(e)(e') = \frac{d}{dt}\Big|_{t=0} L(e + te'),$$

for $e, e' \in E_x$. Its local expression is

$$leg_L(x^i, y^A) = (x^i, \frac{\partial L}{\partial y^A}). \tag{2.7}$$

The Legendre transformation induces a Lie algebroid morphism

$$\mathcal{T} leg_L : \mathcal{T}^{\tau_E} E \to \mathcal{T}^{\tau_{E^*}} E$$

over $leg_L : E \to E^*$ given by

$$(\mathcal{T} leg_L)(e, v) = (e, (T leg_L)(v)),$$

where $Tleg_L : TE \to TE^*$ is the tangent map of $leg_L : E \to E^*$.

We have that (see [111])

$$(\mathcal{T} leg_L, leg_L)^*(\lambda_E) = \Theta_L, \quad (\mathcal{T} leg_L, leg_L)^*(\Omega_E) = \omega_L. \tag{2.8}$$

where λ_E is the Liouville section indroduced in (1.18) and Ω_E is the canonical symplectic section on $\mathcal{T}^{\tau_{E^*}} E$.

On the other hand, from (2.7), it follows that the Lagrangian function L is regular if and only if $leg_L : E \to E^*$ is a local diffeomorphism.

Next, we will assume that L is *hyperregular*, that is, $leg_L : E \to E^*$ is a global diffeomorphism. Then, the pair $(\mathcal{T} leg_L, leg_L)$ is a Lie algebroid isomorphism. Moreover, we may consider the *Hamiltonian function* $H : E^* \to \mathbb{R}$ defined by

$$H = E_L \circ leg_L^{-1}$$

and the *Hamiltonian section* $\xi_H \in \Gamma(\mathcal{T}^{\tau_{E^*}} E)$ which is characterized by the condition

$$i_{\xi_H} \Omega_E = d^{\mathcal{T}^{\tau_{E^*}} E} H.$$

The integral curves of the vector field $\rho_1(\xi_H)$ on E^* satisfy the *Hamilton equations for H*

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \rho^i_A \frac{\partial H}{\partial p_A}, \qquad \frac{\mathrm{d}p_A}{\mathrm{d}t} = -\rho^i_A \frac{\partial H}{\partial x^i} - p_C \mathcal{C}^C_{AB} \frac{\partial H}{\partial p_B}.$$

for $i \in \{1, \dots, m\}$ and $A \in \{1, \dots, n\}$ (see [111]).

In addition, the Euler-Lagrange section ξ_L associated with L and the Hamiltonian section ξ_H are $(\mathcal{T} leg_L, leg_L)$-related, that is,

$$\xi_H \circ leg_L = \mathcal{T} leg_L \circ \xi_L.$$

Thus, if $\gamma : I \to E$ is a solution of the Euler-Lagrange equations associated with L, then $\mu = leg_L \circ \gamma : I \to E^*$ is a solution of the Hamilton equations for H and, conversely, if $\mu : I \to E^*$ is a solution of the Hamilton equations for H then $\gamma = leg_L^{-1} \circ \mu$ is a solution of the Euler-Lagrange equations for L (for more details, see [111]).

Example 2.1.1. Consider the Lie algebroid $E = TQ$, the fiber bundle of a manifold Q of dimension m. If (x^i) are local coordinates on Q, then $\{\dfrac{\partial}{\partial x^i}\}$ is a local basis of $\mathfrak{X}(Q)$ and we have fiber local coordinates $(x^i, \dot{x}^i)$ on TQ. The corresponding structure functions are $\mathcal{C}_{ij}^k = 0$ and $\rho_i^j = \delta_i^j$ for $i, j, k \in \{1, \ldots, m\}$. Therefore given a Lagrangian function $L : TQ \to \mathbb{R}$ the Euler-Lagrange equations associated to L are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i}, \quad i = 1, \ldots, m.$$

Moreover, given a Hamiltonian function $H : T^*Q \to \mathbb{R}$, the Hamilton equations associated to H are

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \quad i = 1, \ldots, m$$

where (x^i, p_i) are local coordinates on T^*Q induced by the local coordinates (x^i) and the local basis $\{dx^i\}$ of T^*Q (see [1] for example).

Example 2.1.2. Consider as a Lie algebroid the finite dimensional Lie algebra $(\mathfrak{g}, [\cdot,\cdot]_{\mathfrak{g}})$ over a point. If e_A is a basis of $\mathfrak{g}$ and $\widetilde{\mathcal{C}}_{AB}^C$ are the structure constants of the Lie algebra, the structures constant of the Lie algebroid $\mathfrak{g}$ with respect to the basis $\{e_A\}$ are $\mathcal{C}_{AB}^C = \widetilde{\mathcal{C}}_{AB}^C$ and $\rho_A^i = 0$. Denote by (y^A) and (μ_A) the local coordinates on $\mathfrak{g}$ and $\mathfrak{g}^*$ respectively, induced by the basis $\{e_A\}$ and its dual basis $\{e^A\}$ respectively. Given a Lagrangian function $L : \mathfrak{g} \to \mathbb{R}$ then the Euler-Lagrange equations for L are just the Euler-Poincaré equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial y^A}\right) = \frac{\partial L}{\partial y^C}\mathcal{C}_{AB}^C y^B.$$

Given a Hamiltonian function $H : \mathfrak{g}^* \to \mathbb{R}$ the Hamilton equations on $\mathfrak{g}^*$ read as the Lie-Poisson equations for H

$$\dot{\mu} = ad^*_{\frac{\partial H}{\partial \mu}} \mu$$

(see [83] for example).

Example 2.1.3. Let G be a Lie group and assume that G acts free and properly on M. We denote by $\pi : M \to \widehat{M} = M/G$ the associated principal bundle. The tangent lift of the action gives a free and proper action of G on TM and $\widehat{TM} = TM/G$ is a quotient manifold. Then we consider the Atiyah algebroid $\widehat{TM}$ over $\widehat{M}$.

According to example 1.8.6, the basis $\{\hat{e}_i, \hat{e}_B\}$ induce local coordinates $(x^i, y^i, \bar{y}^B)$ on $\widehat{TM}$. Given a Lagrangian function $\ell : \widehat{TM} \to \mathbb{R}$ on the Atiyah algebroid $\widehat{TM} \to \widehat{M}$, the Euler-Lagrange equations for ℓ are

$$\begin{aligned}\frac{\partial \ell}{\partial x^j} - \frac{d}{dt}\left(\frac{\partial \ell}{\partial y^j}\right) &= \frac{\partial \ell}{\partial \bar{y}^A}\left(\mathcal{B}^A_{ij} y^i + c^A_{DB}\mathcal{A}^B_j \bar{y}^B\right) \quad \forall j,\\ \frac{d}{dt}\left(\frac{\partial \ell}{\partial \bar{y}^B}\right) &= \frac{\partial \ell}{\partial \bar{y}^A}\left(C^A_{DB}\bar{y}^D - c^A_{DB}\mathcal{A}^D_i y^i\right) \quad \forall B,\end{aligned}$$

which are the Lagrange-Poincaré equations associated to a G-invariant Lagrangian $L : TM \to \mathbb{R}$ (see [45] and [111] for example) where c^C_{AB} are the structure constants of the Lie algebra according to 1.8.6.

2.2 Constraint algorithm for presymplectic Lie algebroids

In this section we introduce the constraint algorithm for presymplectic Lie algebroids given in [87] which generalizes the well-known Gotay-Nester-Hinds algorithm [73]. First we give a review of the Gotay-Nester-Hinds algorithm and then we introduce the construction given in [87] to the case of Lie algebroids.

2.2.1 The Gotay-Nester-Hinds algorithm of constraints

In this subsection we will briefly review the constraint algorithm of constraints for presymplectic systems (see [72] and [73]).

Take the following triple (M, Ω, H) consisting of a smooth manifold M, a closed 2-form Ω and a differentiable function $H : M \to \mathbb{R}$. On M we consider the equation

$$i_X\Omega = dH \tag{2.9}$$

Since we are not assuming that Ω is nondegenerate (that is, Ω is not, in general, symplectic) then Equation (2.9) has no solution in general, or the solutions are not defined everywhere. In the most favorable case, Equation (2.9) admits a global (but not necessarily unique) solution X. In this case, we say that the system admits global dynamics. Otherwise, we select the subset of points of M, where such a solution exists. We denote by M_2 this subset and we will assume that it is a submanifold of $M = M_1$. Then the equations (2.9) admit a solution X defined at all points of M_2, but X need not be tangent to M_2, hence, does not necessarily induce a dynamics on M_2. So we impose an additional tangency condition, and we obtain a new submanifold M_3 along which there exists a solution X, but, however, such X needs to be tangent to M_3. Continuing this process, we obtain a sequence of submanifolds

$$\cdots M_s \hookrightarrow \cdots \hookrightarrow M_2 \hookrightarrow M_1 = M$$

where the general description of M_{l+1} is

$$M_{l+1} = \{p \in M_l \text{ such that there exists } X_p \in T_pM_l \text{ satisfying } i_{X_p}\Omega(p) = dH(p)\}.$$

If the algorithm ends at a final constraint submanifold, in the sense that at some $s \geq 1$ we have $M_{s+1} = M_s$. We will denote this final constraint submanifold by M_f. It may still happen that $\dim M_f = 0$, that is, M_f is a discrete set of points, and in this case the system does not admit a proper dynamics. But, in the case when $\dim M_f > 0$, there exists a well-defined solution X of (2.9) along M_f.

There is another characterization of the submanifolds M_l that we will useful in the sequel. If N is a submanifold of M then we define

$$TN^{\perp} = \{Z \in T_pM,\ p \in N \text{ such that } \Omega(X, Z) = 0 \text{ for all } X \in T_pN\}.$$

Then, at any point $p \in M_l$ there exists $X_p \in T_pM_l$ verifying $i_X\Omega(p) = dH(p)$ if and only if $\langle TM_l^{\perp}, dH\rangle = 0$ (see [72, 73]). Hence, we can define the $l+1$ step of the constraint algorithm as

$$M_{l+1} := \{p \in M_l \text{ such that } \langle TM_l^{\perp}, dH\rangle(p) = 0\}\ .$$

2.2.2 Constraint algorithm for presymplectic Lie algebroids

Let $\tau_E : E \to M$ be a Lie algebroid and suppose that $\Omega \in \Gamma(\wedge^2 E^*)$. Then, we can define the vector bundle morphism $\flat_\Omega : E \to E^*$ (over the identity of M) as follows

$$\flat_\Omega(e) = i(e)\Omega(x), \quad \text{for} \quad e \in E_x.$$

Now, if $x \in M$ and F_x is a subspace of E_x, we may introduce the vector subspace $F_x^{\perp}$ of E_x given by

$$F_x^{\perp} = \{e \in E_x \,|\, \Omega(x)(e, f) = 0, \forall f \in E_x\}.$$

On the other hand, if $\flat_{\Omega_x} = \flat_{\Omega|E_x}$ it is easy to prove that

$$\flat_{\Omega_x}(F_x) \subseteq (F_x^{\perp})^0, \tag{2.10}$$

where $(F_x^{\perp})^0$ is the annihilator of the subspace $F_x^{\perp}$. Moreover, using

$$dimF_x^{\perp} = dimE_x - dimF_x + dim(E_x^{\perp} \cap F_x). \tag{2.11}$$

we obtain that

$$dim(F_x^{\perp})^0 = dimF_x - dim(E_x^{\perp} \cap F_x) = dim(\flat_{\Omega_x}(E_x)).$$

Thus, from (2.10), we deduce that

$$\flat_{\Omega_x}(F_x) = (F_x^{\perp})^{\circ}. \tag{2.12}$$

Next, we will assume that Ω is a presymplectic 2-section ($d^E\Omega = 0$) and that $\alpha \in \Gamma(E^*)$ is a closed 1-section ($d^E\alpha = 0$). Furthermore, we will assume that the kernel of Ω is a vector subbundle of E.

The dynamics of the presymplectic system defined by (Ω, α) is given by a section $X \in \Gamma(E)$ satisfying the dynamical equation

$$i_X\Omega = \alpha\ . \tag{2.13}$$

In general, a section X satisfying (2.13) cannot be found in all points of E. First, we look for the points where (2.13) has sense. We define

$$M_1 = \{x \in M \,|\, \exists e \in E_x : \; i(e)\Omega(x) = \alpha(x)\}$$

From (2.12), it follows that

$$M_1 = \{x \in M \,|\, \alpha(x)(e) = 0, \quad \text{for all} \quad e \in \ker\Omega(x) = E_x^{\perp}\}. \tag{2.14}$$

If M_1 is an embedded submanifold of M, then we deduce that there exists $X : M_1 \to E$ a section of $\tau_E : E \to M$ along M_1 such that (2.13) holds. But $\rho(X)$ is not, in general, tangent to M_1. Thus, we have to restrict to $E_1 = \rho^{-1}(TM_1)$. We remark that, provided that E_1 is a manifold and $\tau_1 = \tau_E \,|_{E_1}: E_1 \to M_1$ is a vector bundle, $\tau_1 : E_1 \to M_1$ is a Lie subalgebroid of $E \to M$.

Now, we must consider the subset M_2 of M_1 defined by

$$\begin{aligned} M_2 &= \{x \in M_1 \,|\, \alpha(x) \in \flat_{\Omega_x}((E_1)_x) = \flat_{\Omega_x}(\rho^{-1}(T_xM_1))\} \\ &= \{x \in M_1 \,|\, \alpha(x)(e) = 0, \;\text{for all } e \in (E_1)_x^{\perp} = (\rho^{-1}(T_xM_1))^{\perp}\}. \end{aligned}$$

If M_2 is an embedded submanifold of M_1, then we deduce that there exists $X : M_2 \to E_1$ a section of $\tau_1 : E_1 \to M_1$ along M_2 such that (2.13) holds. However, $\rho(X)$ is not, in general, tangent to M_2. Therefore, we have that to restrict to $E_2 = \rho^{-1}(TM_2)$. As above, if $\tau_2 = \tau_E \,|_{E_2}: E_2 \to M_2$ is a vector bundle, it follows that $\tau_2 : E_2 \to M_2$ is a Lie subalgebroid of $\tau_1 : E_1 \to M_1$.

Consequently, if we repeat the process, we obtain a sequence of Lie subalgebroids (by assumption)

$$\begin{array}{ccccccccccc} \dots & \hookrightarrow & M_{k+1} & \hookrightarrow & M_k & \hookrightarrow \dots \hookrightarrow & M_2 & \hookrightarrow & M_1 & \hookrightarrow & M_0 = M \\ & & \uparrow \tau_{k+1} & & \uparrow \tau_k & & \uparrow \tau_2 & & \uparrow \tau_1 & & \uparrow \tau_0 = \tau_E \\ \dots & \hookrightarrow & E_{k+1} & \hookrightarrow & E_k & \hookrightarrow \dots \hookrightarrow & E_2 & \hookrightarrow & E_1 & \hookrightarrow & E_0 = E \end{array}$$

where

$$M_{k+1} = \{x \in M_k \,|\, \alpha(x)(e) = 0, \;\text{for all } e \in (\rho^{-1}(T_xM_k))^{\perp}\} \tag{2.15}$$

and

$$E_{k+1} = \rho^{-1}(TM_{k+1}).$$

If there exists $k \in \mathbb{N}$ such that $M_k = M_{k+1}$, then we say that the sequence stabilizes. In such a case, there exists a well-defined (but non necessarily unique) dynamics on the final constraint submanifold $M_f = M_k$. We write

$$M_f = M_{k+1} = M_k, \qquad E_f = E_{k+1} = E_k = \rho^{-1}(TM_k).$$

Then, $\tau_f = \tau_k : E_f = E_k \to M_f = M_k$ is a Lie subalgebroid of $\tau_E : E \longrightarrow M$ (the Lie algebroid restriction of E to E_f). From the construction of the constraint algorithm, we deduce that there exists a section $X \in \Gamma(E_f)$, verifying (2.13). Moreover, if $X \in \Gamma(E_f)$ is a solution of the equation (2.13), then every arbitrary solution is of the form $X' = X + Y$, where $Y \in \Gamma(E_f)$ and $Y(x) \in \ker \Omega(x)$, for all $x \in M_f$. In addition, if we denote by Ω_f and α_f the restriction of Ω and α, respectively, to the Lie algebroid $E_f \longrightarrow M_f$, we have that Ω_f is a presymplectic 2-section and then any $X \in \Gamma(E_f)$ verifying Equation (2.13) also satisfies

$$i_X \Omega_f = \alpha_f \tag{2.16}$$

but, in principle, there are solutions of (2.16) which are not solutions of (2.13) since $\ker \Omega \cap E_f \subset \ker \Omega_f$.

Remark 2.2.1. Note that one can generalize the previous procedure to the general setting of *implicit differential equations on a Lie algebroid*. More precisely, let $\tau_E : E \to M$ be a Lie algebroid and $S \subset E$ be a submanifold of E (not necessarily a vector subbundle). Then, the corresponding sequence of submanifolds of E is

$$\begin{aligned}
S_0 &= S \\
S_1 &= S_0 \cap \rho^{-1}\big(T\tau_E(S_0)\big) \\
&\vdots \\
S_{k+1} &= S_k \cap \rho^{-1}\big(T\tau_E(S_k)\big) \\
&\vdots
\end{aligned}$$

In our case, $S_k = \rho^{-1}(TM_k)$ (equivalently, $M_k = \tau_E(S_k)$). $\diamond$

2.3 The Tulczyjew's triple

In this section we summarize a classical result due to W.M. Tulczyjew showing a natural identification of T^*TQ and TT^*Q as symplectic manifolds where Q is any smooth manifold. This construction plays a key role in Lagrangian and Hamiltonian mechanics.

In [169, 170], Tulczyjew established two identifications, the first one between TT^*Q and T^*TQ (useful to describe Lagrangian mechanics) and the second one between TT^*Q and T^*T^*Q (useful to describe Hamiltonian mechanics). The Tulczyjew map α_Q is an isomorphism between TT^*Q and T^*TQ. Beside, it is also a symplectomorphism between these double vector bundles (see [71, 153] for further details) as symplectic manifolds, i.e. $(TT^*Q\,, \omega_Q^c)$, where ω_Q^c is the complete lift of ω_Q, and (T^*TQ, ω_{TQ}) where ω_{TQ} is the symplectic structure on T^*TQ.

Before giving the full picture, we begin with two basic definitions. The *canonical involution* (see [71] for further details) of TTQ is the smooth map $\kappa_Q : TTQ \to TTQ$ given by

$$\kappa_Q\left(\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}}{\mathrm{d}t}\chi(s,t)\Big|_{t=0}\right)\Big|_{s=0}\right) := \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\chi}(s,t)\Big|_{t=0}\right)\Big|_{s=0},$$

where $\chi : \mathbb{R}^2 \to Q$ and $\tilde{\chi}(s,t) := \chi(t,s)$. If (q^i) are local coordinates for Q, (q^i, v^i) are the induced coordinates for TQ and $(q^i, v^i, \dot{q}^i, \dot{v}^i)$ are the corresponding local coordinates on TTQ, then the canonical involution can be locally defined by $\kappa_Q(q^i, v^i, \dot{q}^i, \dot{v}^i) = (q^i, \dot{q}^i, v^i, \dot{v}^i)$.

The relation among the canonical involution and tangent bundles is expressed in the following diagram

$$\begin{tikzcd}
TTQ \arrow[r, "\kappa_Q"'] \arrow[d, "\tau_{TTQ}"'] & TTQ \arrow[d, "T\tau_{TQ}"] \\
TQ \arrow[r, "\mathrm{Id}"'] & TQ,
\end{tikzcd}$$

The *tangent pairing* between TT^*Q and TTQ is the fibred map $\langle \cdot, \cdot \rangle^T : TT^*Q \times_Q TTQ \to \mathbb{R}$ given by

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)\Big|_{t=0}, \frac{\mathrm{d}}{\mathrm{d}t}\delta(t)\Big|_{t=0} \right\rangle^T := \frac{\mathrm{d}}{\mathrm{d}t}\langle \gamma(t), \delta(t)\rangle^T \Big|_{t=0},$$

where $\gamma : \mathbb{R} \to T^*Q$ and $\delta : \mathbb{R} \to TQ$ are such that $\pi_{T^*Q} \circ \gamma \equiv \tau_{TQ} \circ \delta$.

Definition 2.3.1. *The Tulczyjew's diffeomorphism α_Q is the map $\alpha_Q : TT^*Q \to T^*TQ$ given by*

$$\langle \alpha_Q(V), W \rangle := \langle V, \kappa_Q(W) \rangle^T, \quad V \in TT^*Q, \;\; W \in TTQ.$$

Locally

$$\alpha_Q(q^i, p_i, \dot{q}^i, \dot{p}_i) = (q^i, \dot{q}^i, \dot{p}_i, p_i).$$

This map is a symplectomorphism when we consider on TT^*Q the symplectic structure given by the complete lift ω_Q^c of the canonical symplectic form ω_Q on T^*Q.

In the following diagram we show the relationship among the double vector bundles and the α_Q-Tulczyjew's isomorphism:

$$\begin{tikzcd}
& TT^*Q \arrow[rr, "\alpha_Q"] \arrow[dl, "\tau_{TT^*Q}"'] \arrow[dr, "T\pi_{T^*Q}"] & & T^*TQ \arrow[dl, "\pi_{T^*TQ}"] \arrow[dr, "T^*\tau_{TQ}"] & \\
T^*Q \arrow[drr, "\pi_{T^*Q}"'] & & TQ \arrow[d, "\tau_{TQ}"] & & T^*Q \arrow[dll, "\pi_{T^*Q}"] \\
& & Q & &
\end{tikzcd}$$

The definition of $T^*\tau_{TQ}$ is given in the following remark.

Remark 2.3.2. Given a tangent bundle $\tau_{TN} : TN \to N$, for each $y \in T_xN$ we can define

$$\mathcal{V}_y\tau_{TN} = \ker\{T_y\tau_{TN} : T_yTN \to T_xN\}, \quad \tau_{TN}(y) = x.$$

Summing over all y we obtain a vector bundle $\mathcal{V}\tau_{TN}$ of rank n over TN. Any element $u \in T_xN$ determines a vertical vector at any point y in the fibre over x, called its vertical lift to y, denoted by $u^\vee(y)$. It is the tangent vector at $t = 0$ to the curve $y + t\,u$. If X is a vector field on N, we may define its vertical lift as $X^\vee(y) = (X(\tau_{TN}(y)))^\vee$. Locally, if $X = X^i \frac{\partial}{\partial x^i}$ in a neighborhood U with local coordinates x^i, then $X^\vee$ is locally given by

$$X^\vee = X^i \frac{\partial}{\partial v^i},$$

with respect to induced coordinates (x^i, v^i) on TU.

Now, we define $T^*\tau_{TQ} : T^*TQ \to T^*Q$ by $\langle T^*\tau_{TQ}(\alpha_u), w\rangle = \langle \alpha_u, w_u^\vee\rangle$; $u, w \in T_qQ$, $\alpha_u \in T_u^*TQ$ and $w_u^\vee \in T_uTQ$. $\diamond$

Definition 2.3.3. *The Tulczyjew's isomorphism β_Q is the map $\beta_Q : TT^*Q \to T^*T^*Q$ defined by the isomorphism of vector bundles*

$$\beta_Q(V) := i_V\omega_Q, \quad V \in TT^*Q,$$

where ω_Q is the canonical symplectic form on T^*Q.

Locally,

$$\beta_Q(q^i, p_i, \dot q^i, \dot p_i) = (q^i, \dot q^i, \dot p_i, p_i).$$

This map is an anti-symplectomorphism when on T^*T^*Q we consider the canonical symplectic structure ω_{T^*Q}.

By means of the Tulczyjew's isomorphisms α_Q and β_Q, the double vector bundle TT^*Q may be endowed with two (a priori) different symplectic structures. Let ω_{TQ} and ω_{T^*Q} be the symplectic structures corresponding to T^*TQ and T^*T^*Q respectively. Therefore, $\omega_{\alpha_Q} := \alpha_Q^*\,\omega_{TQ}$ and $\omega_{\beta_Q} := \beta_Q^*\,\omega_{T^*Q}$ define symplectic structures on TT^*Q which turn out to be the same, more precisely: $\omega_{\alpha_Q} = -\omega_{\beta_Q}$. As mentioned before, there exists a third canonical symplectic structure on TT^*Q which comes from the complete lift of the canonical symplectic form ω_Q, denoted by ω_Q^c and which coincides with the previous ones, that is $\omega_Q^c = \omega_{\alpha_Q}$. In coordinates:

$$\Theta_{\alpha_Q} = \alpha_Q^*\,\Theta_{TQ} = \dot p_i \mathrm{d}q^i + p_i \mathrm{d}\dot q^i \ \text{ and } \ \Theta_{\beta_Q} = \beta_Q^*\,\Theta_{T^*Q} = -\dot p_i \mathrm{d}q^i + \dot q^i \mathrm{d}p_i,$$

where Θ_{TQ} and Θ_{T^*Q} are the Liouville 1-forms on TQ and T^*Q, respectively.

$$\begin{tikzcd}
T^*TQ \arrow[dr, "\pi_{T^*TQ}"] & & TT^*Q \arrow[ll, "\alpha_Q"'] \arrow[rr, "\beta_Q"] \arrow[dl, "T\pi_{T^*Q}"'] \arrow[dr, "\tau_{TT^*Q}"] & & T^*T^*Q \arrow[dl, "\pi_{T^*T^*Q}"] \\
& TQ \arrow[rr, "leg_L"] \arrow[dr, "\tau_{TQ}"'] & & T^*Q \arrow[dl, "\pi_{T^*Q}"] & \\
& & Q & &
\end{tikzcd}$$

2.3.1 Implicit description of mechanics

In middle seventies, W.M. Tulczyjew [169, 170] introduced the notion of special symplectic manifold, which is a symplectic manifold symplectomorphic to a cotangent bundle. Using this notion, Tulczyjew gave a nice interpretation of Lagrangian and Hamiltonian dynamics as Lagrangian submanifolds of convenient special symplectic manifolds. Thus, in order to describe this interpretation we are going to use the notion of Lagrangian submanifold introduced in §1.6.1 and that of Tulczyjew's isomorphisms introduced in the previous subsection §2.3.

Theorem 2.3.4. *Given a Hamiltonian function $H : T^*Q \to \mathbb{R}$, consider the associated Hamiltonian vector field $X_H \in \mathfrak{X}(T^*Q)$. The following assertions hold:*

1. *The image of X_H is a Lagrangian submanifold S_{X_H} of $(TT^*Q, \omega_{\beta_Q})$.*
2. *The image of $\mathrm{d}H$ is a Lagrangian submanifold S_H of (T^*T^*Q, ω_{T^*Q}).*
3. *The isomorphism β_Q maps one into each other, i.e. $\beta_Q(S_{X_H}) = S_H$.*

Lemma 2.3.5. *Given a Lagrangian function $L : TQ \to \mathbb{R}$, then $S_L = ImdL$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}).*

Proposition 2.3.6. *Given a hyper-regular Lagrangian function $L : TQ \to \mathbb{R}$, consider the associated Hamiltonian $H = E_L \circ leg_L^{-1}$. Then $\alpha_Q^{-1}(S_L) = S_{X_H} = \beta_Q^{-1}(S_H)$.*

The results 2.3.4, 2.3.5, 2.3.6 are summarized in the following diagram

$$
\begin{tikzcd}
 & & N_L \arrow[dr] & & S_{X_H} \arrow[dl] & & \\
S_L \arrow[r, hook] & T^*TQ \arrow[dr, "\pi_{T^*TQ}"] & & TT^*Q \arrow[ll, "\alpha_Q"'] \arrow[rr, "\beta_Q"] \arrow[dl, "T\pi_{T^*Q}"] \arrow[dr, "\tau_{TT^*Q}"] & & T^*T^*Q \arrow[dl, "\pi_{T^*T^*Q}"'] & S_H \arrow[l, hook'] \\
 & & TQ \arrow[ul, bend left, "dL"] \arrow[rr, "leg_L"] \arrow[dr, "\tau_{TQ}"'] & & T^*Q \arrow[ur, bend right, "dH"'] \arrow[dl, "\pi_{T^*Q}"] & & \\
 & & & Q & & &
\end{tikzcd}
$$

In proposition 2.3.6 we derive a Lagrangian submanifold of TT^*Q with a Lagrangian as starting point. To extract the integrable part of the corresponding equations of motion that this submanifold implies, it is just necessary to use the constraint integrability algorithm developed in [122].

Observe that the three spaces $T(T^*Q)$, $T^*(T^*Q)$ and $T^*(TQ)$ involved in the Tulczyjew triple are symplectic manifolds; the two maps α_Q and β_Q involved in the construction are a symplectomorphism and an anti-symplectomorphism, respectively; and the dynamical equations (Euler-Lagrange and Hamilton equations) are the local equations defining the Lagrangian submanifolds $N_L = \alpha_Q^{-1}(S_L)$ and S_{X_H}. Moreover, the Lagrangian and Hamiltonian functions are not involved in the definition of the triple. In this sense, the triple is canonical. Finally, we would like to point out that the construction can be applied to an arbitrary Lagrangian function, not necessarily a regular Lagrangian.

2.4 Dynamics generated by Lagrangian submanifolds

In this section we will give an alternative definition of the Tulczyjew's diffeomorphism α_Q defining it as the composition of two anti-symplectomorphism.

It is well known (see [77]) that the cotangent bundles T^*TQ and T^*T^*Q are examples of double vector bundles

$$\begin{tikzcd} T^*T^*Q \arrow[r, "T^*\pi_{T^*Q}"] \arrow[d, "\pi_{T^*T^*Q}"'] & TQ \arrow[d, "\tau_{TQ}"] \\ T^*Q \arrow[r, "\pi_{T^*Q}"] & Q \end{tikzcd} \qquad \begin{tikzcd} T^*TQ \arrow[r, "T^*\tau_{TQ}"] \arrow[d, "\pi_{T^*TQ}"'] & T^*Q \arrow[d, "\pi_{T^*Q}"] \\ TQ \arrow[r, "\tau_{TQ}"] & Q \end{tikzcd}$$

The concept of a double vector bundle is due to J. Pradines [152, 153], see also [101, 118]. In particular, all the arrows correspond to vector bundle structures and all pairs of vertical and horizontal arrows are vector bundle morphisms. The above double vector bundles are canonically isomorphic with the vector bundle isomorphism

$$\mathcal{R} : T^*TQ \to T^*T^*Q \tag{2.17}$$

over the identity of T^*Q being an anti-symplectomorphism and also an isomorphism of double vector bundles (see [77, 101]). It is completely determined by the condition

$$\langle \mathcal{R}(\alpha_u), W_{T^*\tau_{TQ}(\alpha_u)} \rangle = -\langle \alpha_u, \widetilde{W}_u \rangle + \langle W_{T^*\tau_{TQ}(\alpha_u)}, \widetilde{W}_u \rangle^T$$

for all $\alpha_u \in T_u^*TQ$, $\widetilde{W}_u \in T_uTQ$ and $W_{T^*\tau_{TQ}(\alpha_u)} \in TT^*Q$ satisfying

$$T\tau_{TQ}(\widetilde{W}_u) = T\pi_{T^*Q}(W_{T^*\tau_{TQ}(\alpha_u)}).$$

Locally, if we denote $(q^i, \dot{q}^i, p_i, \tilde{p}_i)$ local coordinates on $T^*(TQ)$ the anti-symplectomorphism $\mathcal{R}$ is given by

$$\mathcal{R}(q^i, \dot{q}^i, p_i, \tilde{p}_i) = (q^i, \tilde{p}_i, -p_i, \dot{q}^i).$$

The following diagram summarizes the situation

$$\begin{tikzcd}
& T^*TQ \arrow[rrr, "\mathcal{R}"] \arrow[dr, "\pi_{T^*TQ}"] \arrow[ddl, "T^*\tau_{TQ}"'] & & & T^*T^*Q \arrow[dr, "T^*\pi_{T^*Q}"] \arrow[ddl, "\pi_{T^*T^*Q}"'] & \\
& & TQ \arrow[rrr, "id"] \arrow[ddl, "\tau_{TQ}"'] & & & TQ \arrow[ddl, "\tau_{TQ}"'] \\
T^*Q \arrow[rrr, "id"] \arrow[dr, "\pi_{T^*Q}"] & & & T^*Q \arrow[dr, "\pi_{T^*Q}"] & & \\
& Q \arrow[rrr, "id"] & & & Q &
\end{tikzcd}$$

We define the symplectomorphism $\widetilde{\mathcal{R}} : T^*TQ \to TT^*Q$ given by

$$\widetilde{\mathcal{R}} = \beta_Q^{-1} \circ \mathcal{R},$$

where β_Q is the Tulzcyjew's symplectomorphism. In local coordinates $\widetilde{\mathcal{R}}$ is of the form

$$\widetilde{\mathcal{R}}(q^i, \dot{q}^i, p_i, \tilde{p}_i) = (q^i, \tilde{p}_i, \dot{q}^i, p_i).$$

Observe that $\widetilde{\mathcal{R}} = \alpha_Q^{-1}$. Therefore, an alternative definition of Tulczyjew's diffeomorphism is given by $\alpha_Q = \mathcal{R}^{-1} \circ \beta_Q : TT^*Q \to T^*TQ$. From now on, we will use α_Q^{-1} instead of $\widetilde{\mathcal{R}}$.

Now, let N be the Whitney sum between the tangent bundle and the cotangent bundle of a manifold Q, that is, $N = TQ \oplus T^*Q = TQ \times_Q T^*Q$. Observe that N is a submanifold of $TQ \times T^*Q$. Let f be the pairing between vectors and covectors on Q, $f = \langle \cdot, \cdot \rangle : N \to \mathbb{R}$. Applying the Tulczyjew's construction (1.6.11) we construct a Lagrangian submanifold $\Sigma_{f,N}$ of $T^*(TQ \times T^*Q) \simeq T^*TQ \times T^*T^*Q$ as

$$\Sigma_{f,N} = \{\mu \in T^*(TQ \times T^*Q) \mid i_N^*\mu = df\} \subset T^*(TQ \times T^*Q),$$

where $i_N : N \hookrightarrow TQ \times T^*Q$ denotes the canonical inclusion.

The relationship among these spaces is summarized in the following diagram:

$$\begin{tikzcd}
\Sigma_{f,N} \arrow[r, hook, "i_{\Sigma_{f,N}}"] \arrow[d, "(\pi_{T^*N})|_{\Sigma_{f,N}}"'] & T^*(TQ \times T^*Q) \arrow[d, "\pi_{T^*(TQ \times T^*Q)}"] \\
N \arrow[r, hook, "i_N"] & TQ \times T^*Q
\end{tikzcd}$$

Locally, $\Sigma_{f,N}$ is characterized by

$$\Sigma_{f,N} = \{((q^i, \dot{q}^i, p_i, \tilde{p}_i), (q^i, \tilde{p}_i, -p_i, \dot{q}^i)) \in T^*TQ \times T^*T^*Q\}.$$

Observe that $\Sigma_{f,N}$ is the graph of $\mathcal{R}$ which is a Lagrangian submanifold of $T^*TQ \times T^*T^*Q$ where the symplectic form is

$$\tilde{\omega} = pr_0^*\omega_{TQ} - pr_1^*\omega_{T^*Q}.$$

Here, ω_{TQ} denotes the canonical symplectic 2-form on T^*TQ, ω_{T^*Q} is the canonical symplectic 2-form on T^*T^*Q, and pr_0, pr_1 are the projections of $T^*TQ \times T^*T^*Q$ onto the first and second factors, respectively.

We introduce now the dynamics through a Lagrangian $L : TQ \to \mathbb{R}$. This Lagrangian defines the *phase dynamics* $N_L = \alpha_Q^{-1}(\operatorname{Im} dL) \subset TT^*Q$ which can be understood as an implicit differential equation on T^*Q, and solutions of it are curves $\gamma : I \subset \mathbb{R} \to T^*Q$ satisfying $\widetilde{\gamma} \in N_L$, where $\widetilde{\gamma}$ is the tangent prolongation of γ, that is, $\widetilde{\gamma} : I \subset \mathbb{R} \to TT^*Q$ given by $\widetilde{\gamma}(t) = (\gamma(t), \dot{\gamma}(t))$. Since,

$$N_L = \left\{ \widetilde{\gamma} = (q^i, p_i, \dot{q}^i, \dot{p}_i) \in TT^*Q \middle| p_i = \frac{\partial L}{\partial \dot{q}^i}, \dot{p}_i = \frac{\partial L}{\partial q^i} \right\}$$

then $\gamma(t)$ verifies the *Euler-Lagrange equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{\partial L}{\partial q^i}.$$

Remark 2.4.1. Observe that alternatively we can introduce the dynamics in the following way: Let us consider the Legendre transformation $leg_L : TQ \to T^*Q$. Given a curve $\sigma : I \subset \mathbb{R} \to Q$ on the base manifold Q, the dynamics can be expressed as

$$\frac{d}{dt}\left(leg_L(\sigma(t), \dot{\sigma}(t))\right) = \alpha_Q^{-1}(dL(\sigma(t), \dot{\sigma}(t))). \tag{2.18}$$

That is, a curve $\sigma : I \subset \mathbb{R} \to Q$ is a solution of the Euler-Lagrange equations if, and only if, it satisfies the equation (2.18)

◇

2.5 Higher-order mechanical systems

The aim of this section is to build up the Lagrangian and Hamiltonian formalism for systems involving higher-order derivatives using a generalization of Tulczyjew's construction for higher-order tangent bundles.

A *Lagrangian of order k* is a real smooth function $L : T^{(k)}Q \to \mathbb{R}$. That is, a real function which depends on higher-order derivatives, where k denotes the order of the derivative. We want to generalize to this case, the description of first-order Lagrangian dynamics given in Section 2.3, following the results given in [109].

2.5.1 Higher-order mechanical systems: variational description

Given two points $x, y \in T^{(k-1)}Q$ we define the infinite-dimensional manifold $C^{2k}(x, y)$ of $2k$-differentiable curves which connect x and y as

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

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

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

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

$$\begin{array}{rccl} \mathcal{J}: & C^{2k}(x,y) & \longrightarrow & \mathbb{R} \\ & c & \longmapsto & \int_0^T L(c^{(k)}(t))\ dt\ . \end{array}$$

Definition 2.5.1. *(Hamilton's principle) A curve $c \in C^{2k}(x,y)$ 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 $\mathcal{J}$.*

Let us introduce a theorem (see [112]) which describes the *higher-order Euler-Lagrange equations* and some geometrical structures in higher-order mechanics

Theorem 2.5.2. *Let $L: T^{(k)}Q \to \mathbb{R}$ be a higher-order Lagrangian and*

$$\mathcal{J}(c)=\int_0^1 L(c^{(k)}(t))dt$$

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

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

and an unique 1−form Θ_L on $T^{(2k-1)}Q$ such that for all variations of the form $\delta c_s \in T_cC^{2k}(x,y)$ with fixed endpoints we have that

$$\frac{d}{ds}\mathcal{J}(c_s(t))\Big|_{s=0}=\int_0^1 \mathcal{E}L(c^{(2k)}(t))\cdot\delta c(t)\ dt+\Big(\Theta_L(c^{(2k-1)}(t))\cdot\delta^{(2k-1)}c(t)\Big)\Big|_0^1.$$

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

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

$$\Theta_L=\sum_{l=0}^{k-1}\hat{p}_{(l)i}\,\mathbf{d}q^{(l)i},$$

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

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

The equations of motion are called higher-order Euler-Lagrange, and are written as

$$\sum_{l=0}^{k}(-1)^l\frac{d^l}{dt^l}\left(\frac{\partial L}{\partial q^{(l)i}}\right)=0. \tag{2.19}$$

Here $c_s \in C^{2k}(x,y)$ is a family of curves with $c_0 = c$ and $s \in (-b,b) \subset \mathbb{R}$, $\delta c^i = \frac{d}{ds} c_s^{(i)}\Big|_{s=0}$ and $\delta^l c^i = \frac{d^l}{dt^l}\delta c^i$.

Therefore, from Theorem (2.5.2) it is possible to define the 2-form $\Omega_L = -d\Theta_L$. In local coordinates, we have that

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

and it is easy to see [112] that Ω_L is symplectic if and only if,

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

The higher-order Lagrangian is said to be *regular* if Ω_L is symplectic.

In the following we assume that the Lagrangian L is regular. Take now the restriction $\mathcal{J}_L$ of the action functional $\mathcal{J}$ to the subspace C_L of solutions of Euler-Lagrange equations. This space can be identified with the space of initial conditions $T^{(2k-1)}Q$ of the Euler-Lagrange equations. Therefore is easy to show that

$$\mathbf{d}\mathcal{J}_L = F_t^*\Theta_L - \Theta_L$$

where F_t is the flow of the Euler-Lagrange vector field X_L, defined on $T^{(2k-1)}Q$ by $\mathcal{E}L \circ X_L = 0$. Since $\mathbf{d}^2 = 0$ 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

$$0 = i_{\xi_Q^{(2k-1)}}\mathbf{d}\mathcal{J}_L = i_{\xi_Q^{(2k-1)}}(F_t^*\Theta_L - \Theta_L) = F_t^*(i_{\xi_Q^{(2k-1)}}\Theta_L) - i_{\xi_Q^{(2k-1)}}\Theta_L,$$

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

Finally, let us recall that the *Legendre transformation* is locally defined to be the map $Leg_L : T^{(2k-1)}Q \to T^*(T^{(k-1)}Q)$ locally given by

$$Leg_L(q^{(0)i}, q^{(1)i}, \ldots, q^{(2k-1)i}) = (q^{(0)i}, q^{(1)i}, \ldots, q^{(k-1)i}, \hat{p}_{(0)i}, \ldots, \hat{p}_{(k-1)i}). \tag{2.20}$$

If L is regular then Leg_L is a local diffeomorphism, and conversely. Observe that, when $k=1$, $Leg_L = leg_L$.

2.5.2 Geometrical description of higher-order mechanics

In this subsection we study a natural application of the Tulczyjew triple to obtain the higher-order Euler-Lagrange equations following some results of M. de León and E. Lacomba [109].

Consider the canonical immersion $j_k : T^{(k)}Q \to TT^{(k-1)}Q$ defined in (1.5). Then, if $x \in T^{(k)}Q$, then $j_k^* : T^*_{j_k(x)}(TT^{(k-1)}Q) \to T^*_x(T^{(k)}Q)$ is given by

$$j_k^*\mu = \mu \circ Tj_k, \quad \forall \mu \in T^*_{j_k(x)}TT^{(k-1)}Q.$$

By Tulczyjew's Theorem (1.6.11) we can construct a Lagrangian submanifold of $T^*(TT^{(k-1)}Q)$ induced by a higher-order Lagrangian $L : T^{(k)}Q \to \mathbb{R}$,

$$\Sigma_L = \{\mu \in T^*(TT^{(k-1)}Q) \Big| j_k^*\mu = dL\} \subset T^*(TT^{(k-1)}Q) \tag{2.21}$$

fibering onto $j_k(T^{(k)}Q)$.

Now, we will use the construction given in Section 2.3 for first order systems to the particular case when our tangent vector bundle is $TT^{(k-1)}Q$.

Definition 2.5.3. *The Tulczyjew's isomorphism $\beta_{T^{k-1}Q}$ is the map $\beta_{T^{k-1}Q} : TT^*T^{(k-1)}Q \to T^*T^*T^{(k-1)}Q$ defined by*

$$\beta_{T^{k-1}Q}(V) := i_V\omega_{T^{k-1}Q},$$

with $V \in TT^*T^{(k-1)}Q$ and where $\omega_{T^{k-1}Q}$ is the canonical symplectic form of $T^*T^{(k-1)}Q$.

As we have said, this map is an anti-symplectomorphism when we consider $T^*T^*T^{(k-1)}Q$ with the canonical symplectic structure $\omega^c_{T^{(k-1)}Q} = d_T\omega_{T^{(k-1)}Q}$.

Definition 2.5.4. *The higher-order Tulczyjew's diffeomorphism $A_{T^{k-1}Q}$ is the map $A_{T^{k-1}Q} : T(T^*T^{(k-1)}Q) \to T^*(TT^{k-1}Q)$ given by*

$$\langle A_{T^{k-1}Q}(V), W\rangle := \langle V, \kappa_{T^{(k-1)}Q}(W)\rangle^T, \quad V \in T(T^*T^{(k-1)}Q),\ \ W \in T^{(k)}(TQ).$$

Locally, if we denote by

$$(\mathbf{q}^{k-1}, \mathbf{p}_{k-1}, \dot{\mathbf{q}}^{k-1}, \dot{\mathbf{p}}_{k-1}) = (q^{(0)}, \ldots, q^{(k-1)}, p_0, \ldots, p_{k-1}, \dot{q}^{(0)}, \ldots, \dot{q}^{(k-1)}, \dot{p}_0, \ldots, \dot{p}_{k-1})$$

the local coordinates on $TT^*(T^{(k-1)})Q$, we have that

$$\begin{aligned} A_{T^{k-1}Q}(\mathbf{q}^{k-1}, \mathbf{p}_{k-1}, \dot{\mathbf{q}}^{k-1}, \dot{\mathbf{p}}_{k-1}) &= (\mathbf{q}^{k-1}, \dot{\mathbf{q}}^{k-1}, \dot{\mathbf{p}}_{k-1}, \mathbf{p}_{k-1}) \\ \beta_{T^{k-1}Q}(\mathbf{q}^{k-1}, \mathbf{p}_{k-1}, \dot{\mathbf{q}}^{k-1}, \dot{\mathbf{p}}_{k-1}) &= (\mathbf{q}^{k-1}, \mathbf{p}_{k-1}, -\dot{\mathbf{p}}_{k-1}, \dot{\mathbf{q}}^{k-1}). \end{aligned}$$

This map is a symplectomorphism when we consider on $TT^*T^{(k-1)}Q$ the symplectic structure given by the complete lift $\omega^c_{T^{(k-1)}Q}$ of the canonical symplectic structure $\omega_{T^{(k-1)}Q}$ on $T^*TT^{(k-1)}Q$.

In the following diagram we show the different relationships among the double vector bundles and the $A_{T^{k-1}Q}$–Tulczyjew's isomorphism:

$$\begin{tikzcd}
 & T(T^*T^{(k-1)}Q) \arrow[rr, "A_{T^{k-1}Q}"] \arrow[dl, "\tau_{TT^*T^{k-1}Q}"'] \arrow[dr, "T\pi_{T^*T^{k-1}Q}"] & & T^*(TT^{(k-1)}Q) \arrow[dl, "\pi_{T^*TT^{k-1}Q}"'] \arrow[dr, "T^*\tau_{TT^{k-1}Q}"] & \\
T^*T^{(k-1)}Q \arrow[drr, "\pi_{T^*T^{k-1}Q}"'] & & TT^{(k-1)}Q \arrow[d, "\tau_{TT^{k-1}Q}"'] & & T^*T^{(k-1)}Q \arrow[dll, "\pi_{T^*T^{k-1}Q}"] \\
 & & T^{(k-1)}Q & &
\end{tikzcd}$$

The Lagrangian submanifold Σ_L is locally parametrized by the $2k$ points $(q^{(0)}, \ldots, q^{(k)}, \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-2)})$ and is immersed in $T^*TT^{(k-1)}Q$ as

$$\left\{ \left(q^{(0)}, \ldots, q^{(k-1)}; q^{(1)}, \ldots, q^{(k)}; \frac{\partial L}{\partial q^{(0)}}, \frac{\partial L}{\partial q^{(1)}} - \tilde{p}^{(0)}, \ldots, \frac{\partial L}{\partial q^{(k-1)}} - \tilde{p}^{(k-2)}; \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-2)}, \frac{\partial L}{\partial q^{(k)}} \right) \right\}$$

Therefore, taking this into account, the Lagrangian dynamics is given by the Lagrangian submanifold $N_L = A^{-1}_{T^{k-1}Q}(\Sigma_L)$ of $TT^*T^{(k-1)}Q$. Locally, N_L is given by the elements in $TT^*T^{(k-1)}Q$ which have the form

$$\left(q^{(0)}, \ldots, q^{(k-1)}; \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-2)}, \frac{\partial L}{\partial q^{(k)}}; q^{(1)}, \ldots, q^{(k)}; \frac{\partial L}{\partial q^{(0)}}, \frac{\partial L}{\partial q^{(1)}} - \tilde{p}^{(0)}, \ldots, \frac{\partial L}{\partial q^{(k-1)}} - \tilde{p}^{(k-2)} \right).$$

N_L determines the following set of differential equations:

$$\frac{d}{dt}\tilde{p}^{(0)} = \frac{\partial L}{\partial q^{(0)}}, \tag{2.22}$$

$$\frac{d}{dt}\tilde{p}^{(i)} + \tilde{p}^{(i-1)} = \frac{\partial L}{\partial q^{(i)}}, \tag{2.23}$$

$$\frac{\partial L}{\partial q^{(k-1)}} - \tilde{p}^{(k-2)} = \frac{d}{dt}\left(\frac{\partial L}{\partial q^{(k)}}\right), \tag{2.24}$$

where $1 \leq i \leq k-2$.

Differentiating respect to the time equation (2.24), and replacing into equation (2.23) for $j = k-1$ we obtain that

$$\frac{d^2}{dt^2}\left(\frac{\partial L}{\partial q^{(k)}}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial q^{(k-1)}}\right) - \frac{\partial L}{\partial q^{(k-2)}} - \tilde{p}^{k-3}.$$

Differentiating with respect to the time the last equality and replacing into (2.23) when $i = k-3$ we have

$$\frac{d^3}{dt^3}\left(\frac{\partial L}{\partial q^{(k)}}\right) = \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial q^{(k-1)}}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial q^{(k-2)}}\right) + \frac{\partial L}{\partial q^{(k-3)}} - \tilde{p}^{k-4}.$$

Proceeding as before, differentiating $k-4$ with respect to time and replacing into equation (2.23), we obtain

$$\frac{d^k}{dt^k}\left(\frac{\partial L}{\partial q^{(k)}}\right) = \frac{d^{k-1}}{dt^{k-1}}\left(\frac{\partial L}{\partial q^{(k-1)}}\right) - \frac{d^{k-2}}{dt^{k-2}}\left(\frac{\partial L}{\partial q^{(k-2)}}\right) + \ldots - \frac{d}{dt}\left(\frac{\partial L}{\partial q^{(1)}}\right) + \frac{d}{dt}\tilde{p}^{0}.$$

Finally using (2.22) we obtain the equation

$$\sum_{j=0}^{k}(-1)^j \frac{d^j}{dt^j}\left(\frac{\partial L}{\partial q^{(j)}}\right) = 0,$$

which is exactly the higher-order Euler-Lagrange equation (2.19) for L (see [109]).

Solutions of this system are curves in the Lagrangian submanifold Σ_L, $\mu : I \subset \mathbb{R} \to \Sigma_L$ satisfying

$$\pi_{T^*TT^{k-1}Q} |_{\Sigma_L} (\mu(t)) = \gamma^{(k)}(t),$$

where $\gamma^{(k)}(t)$ is the k-lift of a curve $\gamma : I \to Q$ and $\pi_{T^*TT^{k-1}Q} |_{\Sigma_L}: \Sigma_L \to TT^{(k-1)}Q$ is the canonical projection from Σ_L to $TT^{(k-1)}Q$.

Definition 2.5.5. *We define the higher-order Legendre transformation* $\mathbb{F}L : \Sigma_L \to T^*T^{(k-1)}Q$ *as the mapping* $\mathbb{F}L = \tau_{TT^*T^{k-1}Q} \circ (A_{T^{k-1}Q})^{-1} |_{\Sigma_L}$.

Locally, this map is given by

$$\mathbb{F}L(q^{(0)}, \ldots, q^{(k)}, \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-2)}) = \left(q^{(0)}, \ldots, q^{(k-1)}, \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-2)}, \frac{\partial L}{\partial q^{(k)}} \right).$$

A higher-order Lagrangian system is regular if, and only if, $\mathbb{F}L$ *is a local diffeomorphism.*

Observe that a higher-order Lagrangian system is regular if and only if $\left(\frac{\partial^2 L}{\partial q^{(k)i} \partial q^{(k)j}} \right)$ is a nondegenerate matrix. In such a case, since $\tilde{p}^{(k-1)} = \frac{\partial L}{\partial q^{(k)}}(q^{(0)}, \ldots, q^{(k)})$, applying the implicit function theorem, we can define $q^{(k)}$ as a function f depending of $q^{(0)}, \ldots, q^{(k-1)}, \tilde{p}^{(k-1)}$; that is,

$$q^{(k)} = f(q^{(0)}, \ldots, q^{(k-1)}, \tilde{p}^{(k-1)}). \tag{2.25}$$

Using the Legendre transformation we can state in an alternatively way the solutions of the higher-order Lagrangian system as the curves on Σ_L, $\mu : I \subset \mathbb{R} \to \Sigma_L$ satisfying

$$A^{-1}_{T^{k-1}Q}(\mu(t)) = \frac{d}{dt}\mathbb{F}L(\mu(t)),$$

where μ satisfies $\pi_{T^*TT^{k-1}Q} |_{\Sigma_L} (\mu(t)) = \gamma^{(k)}(t)$, where $\gamma^{(k)}(t)$ is the k-lift of a curve $\gamma : I \to Q$.

Finally, we would like to point out that when the higher-order Lagrangian system is regular we can establish the Hamiltonian formalism defining a Hamiltonian function $H : T^*T^{(k-1)}Q \to \mathbb{R}$ as

$$H(q^{(0)}, \ldots, q^{(k-1)}, \tilde{p}^{(0)}, \ldots, \tilde{p}^{(k-1)}) = \sum_{i=0}^{k-1} \tilde{p}^{(i)} q^{(i)} - L(q^{(0)}, \ldots, q^{(k)})$$

where $q^{(k)}$ is given implicitly by (2.25).

The corresponding Hamiltonian vector field X_H is determined by $i_{X_H}\omega_{T^{k-1}Q} = dH$. In this particular case we have that

$$Im(X_H) = X_H(T^*T^{(k-1)}Q) = \beta^{-1}_{T^{k-1}Q}(dH(T^*T^{(k-1)}Q)) = A^{-1}_{T^{k-1}Q}(\Sigma_L).$$

In the singular case, the submanifold Im (dH) is not transversal with respect to $\pi_{T^*T^*T^{(k-1)}Q}$, then is necessary to apply the integrability algorithm to find, if it exists, a subset where there are consistent solutions of the dynamics (see [72] and [73], for example).

Finally, to end this chapter we will give an alternative characterization of the dynamics in the Lagrangian submanifols Σ_L in terms of the solution of the higher-order Euler-Lagrange equations.

Theorem 2.5.6. *The curve $q(t) \in Q$ is a solution of the higher-order Euler-Lagrange equations for $L : T^{(k)}Q \to \mathbb{R}$ if and only if*

$$A_{T^{(k-1)}Q}\left(\frac{d}{dt}(Leg_L(q(t), \dot{q}(t), \ddot{q}(t), \dots, q^{(2k-1)}))\right) \in \Sigma_L$$

where $Leg_L : T^{(2k-1)}Q \to T^*(T^{(k-1)}Q)$ was defined in (2.20).

Proof. Given the curve $q(t) \in Q$ we consider its tangent lift to $T^{(2k-1)}Q$, that is $\sigma(t) = (q(t), \dot{q}(t), \ddot{q}(t), \dots, q^{(2k-1)}(t))$ Therefore $Leg_L(\sigma(t)) = (q^{(0)}(t), \dots, q^{(k-1)}(t), \hat{p}_0(t), \dots, \hat{p}_{k-1}(t))$ where $\hat{p}_r$ with $0 \leq r \leq k-1$ denotes the Jacobi-Osrtogradski momenta defined in (2.5.2).

Now, using the definition of $A_{T^{k-1}Q}$ we have that

$$A_{T^{k-1}Q}\left(\frac{d}{dt}(Leg_L(\sigma(t)))\right) = \left(q^{(0)}(t), \dots, q^{(k-1)}(t); q^{(1)}(t), \dots, q^{(k)}(t), \frac{d}{dt}\hat{p}_0, \frac{d}{dt}\hat{p}_1, \dots, \frac{d}{dt}\hat{p}_{k-1}, \hat{p}^{(0)}(t), \dots, \hat{p}^{(k-1)}(t)\right)$$

Observe that the condition $A_{T^{k-1}Q}\left(\frac{d}{dt}(Leg_L(\sigma(t)))\right) \in \Sigma_L$ implies the equations

$$\frac{d}{dt}\hat{p}_{(0)} = \frac{\partial L}{\partial q^{(0)}}, \tag{2.26}$$

$$\frac{d}{dt}\hat{p}_{(j)} + \hat{p}_{(j-1)} = \frac{\partial L}{\partial q^{(i)}} \text{ where } 1 \leq j \leq k-2, \tag{2.27}$$

$$\hat{p}_{(k-1)} = \frac{\partial L}{\partial q^{(k)}}, \tag{2.28}$$

which implies the higher-order Euler-Lagrange equations

$$\sum_{j=0}^{k}(-1)^j \frac{d^j}{dt^j}\left(\frac{\partial L}{\partial q^{(j)}}\right) = 0.$$

□

Remark 2.5.7. We could also develop the case of higher-order Lagrangian systems with higher-order constrains using a similar formalism.

In this case, consider a submanifold $\mathcal{M} \subset T^{(k)}Q$ and a Lagrangian function $L : \mathcal{M} \to \mathbb{R}$. Observe that $\mathcal{M} \hookrightarrow T^{(k)}Q \hookrightarrow TT^{(k-1)}Q$. Denote by $i_{\mathcal{M}} : \mathcal{M} \hookrightarrow TT^{(k-1)}Q$ the composition of both inclusions. Now construct

$$\Sigma_L = \left\{\mu \in T^*TT^{(k-1)}Q \mid i^*_{\mathcal{M}}\mu = dL\right\}$$

and proceed as before. ◇

Chapter 3

Skinner and Rusk formalism for higher-order mechanical systems

In the present chapter we consider the Skinner and Rusk formalism for mechanical systems [167] [168]. R. Skinner and R. Rusk considered a geometric framework where the velocities and the momenta are independent coordinates. To do this, they considered the dynamics on the Whitney sum of TQ (velocity phase space) and T^*Q (the phase space).

Given a Lagrangian function $L : TQ \to \mathbb{R}$ one considers the bundle $TQ \oplus T^*Q$ with canonical projections $pr_1 : TQ \oplus T^*Q \to TQ$ and $pr_2 : TQ \oplus T^*Q \to T^*Q$ onto the first and second factors. We then define a function $H : TQ \oplus T^*Q \longrightarrow \mathbb{R}$ by $H(X_p, \alpha_p) = \alpha_p(X_p) - L(X_p)$. In bundle coordinates (q^i, v^i, p_i), H is given by $H(q^i, v^i, p_i) = v^i p_i - L(q^i, v^i)$, and it is sometimes refereed as the *Pontryagin Hamiltonian* or *generalized energy* (see [180] for example). We can also define a 2-form Ω on $TQ \oplus T^*Q$ by $\Omega = pr_2^*(\omega_Q)$, where ω_Q denotes the canonical symplectic 2-form on T^*Q. Then, one discuss the presymplectic system $(TQ \oplus T^*Q, \Omega, dH)$ and obtain the corresponding sequence of constraint submanifolds, which, of course, have a close relation with those obtained by Gotay and Nester (and extended in the framework of Lie algebroids by Iglesias, Marrero, Martín de Diego and Sosa) on the Lagrangian and Hamiltonian side. It should be noticed that this algorithm includes the SODE condition just from the very beginning.

In this chapter we will consider higher-order mechanics from the point of view of the Skinner and Rusk formalism to obtain higher-order Euler-Lagrange equations, higher-order Euler-Poincaré equations and higher-order Lagrange-Poincaré equations. Also, we will study the case of systems with higher-order constraints. The extension of this theory to the natural setting of Lie algebroids will be also developed.

One of the main objectives in the chapter is to characterize geometrically the equations of motion of an optimal control problem for an underactuated mechanical system. In this last system, the trajectories are "parameterized" by the admissible controls and the necessary conditions for extremals in the optimal control problem are expressed using a "pseudo-Hamiltonian formulation" based on the Pontryagin maximun principle or an appropriate variational setting using some smoothness conditions [3]. Many of the concrete examples under study have additional geometric properties, as for instance, the configuration space is not only a differentiable manifold but it also has a compatible structure of group, that is, the configuration space is a Lie group. In this chapter, we will take advantage of this property to give an intrinsic expression of the equations of motion for higher-order mechanical systems and for optimal control problems with symmetries.

3.1 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 [167] and [168]. We consider a dynamical system of n degrees of freedom modeled by a configuration space Q of dimension n. The behavior of this system is described by the Lagrangian $L \in C^\infty(TQ)$ which contains the information of the dynamics associated with the system.

Consider the following phase space,

$$TQ \times_Q T^*Q \simeq TQ \oplus T^*Q,$$

that is, the Whitney sum of the velocity phase space and phase space, also called *Pontryagin bundle*. This space is endowed with two canonical projections, $pr_1 : TQ \times_Q T^*Q \to TQ$ and $pr_2 : TQ \times_Q T^*Q \to T^*Q$. We denote by W this Whitney sum, $W = TQ \oplus T^*Q$. Using the canonical projections of the tangent bundle and cotangent bundle over the manifold Q we can construct the following diagram which illustrates the situation,

$$\begin{tikzcd}
 & W = TQ \times_Q T^*Q \arrow[dl, "pr_1"'] \arrow[dr, "pr_2"] & \\
TQ \arrow[dr, "\tau_Q"] & & T^*Q \arrow[dl, "\pi_Q"] \\
 & Q &
\end{tikzcd}$$

Figure 3.1: Skinner and Rusk formalism

If (U, φ) is a local chart of Q, and $\varphi = (q^i), i = 1, \ldots, n$; we can induce natural coordinates on TQ and T^*Q in $\tau_Q^{-1}(U)$ and $\pi_Q^{-1}(U)$ respectively. These coordinates are denoted by (q^i, v^i) and (q^i, p_i) respectively. Therefore, (q^i, v^i, p_i) are natural coordinates in W. Observe that $dim(W) = 3n$.

Let λ be the Liouville one-form of the cotangent bundle and $\omega_Q = -d\lambda$ the canonical symplectic 2-form on T^*Q. We define the 2-form Ω on W as

$$\Omega = pr_2^*(\omega_Q).$$

Observe that Ω is a closed 2-form, but nevertheless, this form is always degenerate and therefore is a *presymplectic form*. Using the expression in local coordinates (q^i, v^i, p_i) in W,

$\omega_Q = dq^i \wedge dp_i$ and taking account that $pr_2(q^i, v^i, p_i) = (q^i, p_i)$ we have that $\Omega = dq^i \wedge dp_i$. From this local expression, is clear that $\left\{\frac{\partial}{\partial v^i}\right\}$ is a local basis of $\ker\,\Omega$, that is,

$$\ker\,\Omega = \mathrm{span}\left\langle \frac{\partial}{\partial v^i} \right\rangle$$

and therefore the 2-form Ω is degenerate.

Then, we have a presymplectic manifold (W, Ω) and our objective is to obtain a presymplectic Hamiltonian system in order to deduce the equations of motion following the procedure given in [10], [41] and [154]. Nevertheless, in this formalism, we suppose that the information of the dynamics is specified by a Lagrangian $L \in C^\infty(TQ)$.

To define a Hamiltonian function, first consider the function $C \in C^\infty(W)$, defined canonically in the following way: if $p \in Q$, $v_p \in T_pQ$ is a tangent vector to Q at p and $\alpha_p \in T^*_pQ$ is a covector on p, we define C as

$$\begin{aligned} C: \quad & TQ \times_Q T^*Q \to \mathbb{R} \\ & (v_p, \alpha_p) \mapsto \langle \alpha_p, v_p \rangle. \end{aligned}$$

In local coordinates, $C(q^i, v^i, p_i) = v^i p_i$.

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

$$H(q^i, p_i, v^i) = C - pr_1^*(L) = p_i v^i - L(q^i, v^i)$$

and therefore we have a presymplectic Hamiltonian system (W, Ω, H). The presymplectic algorithm given in Section (2.2) can be applied and the equations of motion are given by the solutions of the following equation

$$i_X \Omega = dH.$$

3.2 Skinner-Rusk formalism for higher order mechanical systems

In this section we will consider higher-order mechanics from the point of view of the Skinner and Rusk formalism and we will analyze the case when this system is also subject to higher-order constraints (see [56] for the first order case).

Let us consider the *higher-order Pontryagin bundle* W_0, that is, the Whitney sum

$$W_0 := T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$$

with the canonical projections

$$\begin{aligned} pr_1 &: W_0 \longrightarrow T^{(k)}Q, \\ pr_2 &: W_0 \longrightarrow T^*(T^{(k-1)}Q). \end{aligned}$$

We construct on W_0 the presymplectic 2-form

$$\Omega_{W_0} = pr_2^*(\omega_{T^{(k-1)}Q}),$$

where $\omega_{T^{(k-1)}Q}$ is the canonical symplectic form on $T^*(T^{(k-1)}Q)$. Also we define the function $H_{W_0} : W_0 \to \mathbb{R}$ given by

$$H_{W_0}(p, \alpha) = \langle \alpha, j_k(p) \rangle - L(p)$$

where $(p, \alpha) \in W_0$. Here $\langle \cdot, \cdot \rangle$ denotes the natural pairing between vectors and covectors on $T^{(k-1)}Q$ (observe that $j_k(p) \in TT^{(k-1)}Q$ was defined in (1.5)).

We will see that the dynamics of the higher-order problem is *intrinsically characterized* as the solutions of the presymplectic hamiltonian equations

$$i_X \Omega_{W_0} = dH_{W_0}\,. \tag{3.1}$$

Observe that locally

$$\ker \Omega_{W_0} = \text{span}\, \left\langle \mathcal{V}_i = \frac{\partial}{\partial q^{(k)i}} \right\rangle .$$

Taking local coordinates $(q^{(0)i}, q^{(1)i}, \dots, q^{(k-1)i}; p_i^{(0)}, \dots, p_i^{(k-1)}, q^{(k)i})$ on W_0, then the local expressions of the presymplectic 2-form Ω_{W_0} and the hamiltonian H_{W_0} are

$$\begin{aligned}
\Omega_{W_0} &= \sum_{r=0}^{k-1} dq^{(r)i} \wedge dp_i^{(r)}, \\
H_{W_0} &= \sum_{r=0}^{k-1} q^{(r+1)i} p_i^{(r)} - L(q^{(0)i}, q^{(1)i}, \dots, q^{(k)i}).
\end{aligned}$$

Consider a vector field X on W_0 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 $\Omega_{W_0}(X, \cdot) = dH_{W_0}(\cdot)$: Given $V \in W_0$,

$$V = \sum_{r=0}^{k} v^{(r)i} \frac{\partial}{\partial q^{(r)i}} + \sum_{r=0}^{k-1} \widetilde{v}_i^{(r)} \frac{\partial}{\partial p_i^{(r)}},$$

we have that

$$\begin{aligned}
\Omega_{W_0}(X, V) &= \sum_{r=0}^{k-1} dq^{(r)i} \wedge dp_i^{(r)}(X, V) \\
&= \sum_{r=0}^{k-1} \left[dq^{(r)i}(X) dp_i^{(r)}(V) - dq^{(r)i}(V) dp_i^{(r)}(X) \right] = \sum_{r=0}^{k-1} \left[X^{(r)i} \widetilde{v}_i^{(r)} - v^{(r)i} Y_i^{(r)} \right].
\end{aligned}$$

Therefore, since

$$dH_{W_0}(V) = \sum_{r=0}^{k} \frac{\partial H_{W_0}}{\partial q^{(r)i}} v^{(r)i} + \sum_{r=0}^{k-1} \frac{\partial H_{W_0}}{\partial p_i^{(r)}} \widetilde{v}_i^{(r)}$$

from equation (3.1) we obtain that

$$\sum_{r=0}^{k-1}\left[X^{(r)i}\widetilde{v}_i^{(r)} - v^{(r)i}Y_i^{(r)}\right] = \sum_{r=0}^{k}\frac{\partial H_{W_0}}{\partial q^{(r)i}}v^{(r)i} + \sum_{r=0}^{k-1}\frac{\partial H_{W_0}}{\partial p_i^{(r)}}\widetilde{v}_i^{(r)}. \tag{3.2}$$

From the other hand, differentiating H_{W_0} we have:

$$\begin{aligned}\frac{\partial H_{W_0}}{\partial q^{(0)i}} &= -\frac{\partial L}{\partial q^{(0)i}}\\ \frac{\partial H_{W_0}}{\partial q^{(r)i}} &= p_i^{(r-1)} - \frac{\partial L}{\partial q^{(r)i}}; \quad r = 1,\dots,k-1.\end{aligned}$$

As (3.2) holds for every vector field V in W_0, we have that

$$-Y_i^{(r)} = \frac{\partial H_{W_0}}{\partial q^{(r)i}}, \quad r = 0,\dots,k-1.$$

Therefore,

$$\begin{aligned}-Y_i^{(0)} &= -\frac{\partial L}{\partial q^{(0)i}}\\ -Y_i^{(r)} &= p_i^{(r-1)} - \frac{\partial L}{\partial q^{(r)i}}; \quad r = 1,\dots,k-1\\ X_i^{(r)} &= \frac{\partial H_{W_0}}{\partial p_i^{(r)}} = q^{(r+1)i}; \quad r = 0,\dots,k-1\\ 0 &= \frac{\partial H_{W_0}}{\partial q^{(k)i}} = p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}}.\end{aligned}$$

The solutions of equation (3.1) are defined on the first constraint submanifold given by the set of points $x \in W_0$ such that $dH_{W_0}(x)(Z) = 0$, for all $Z \in \ker \Omega_{W_0}(x)$. Locally these restrictions are defined from the following relations

$$\varphi_i^1 = p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}} = 0, \quad i = 1,\dots,n\,. \tag{3.3}$$

The equations $\varphi_i^1 = 0$ (primary relations) determine the set of points W_1 of W_0 where (3.1) 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 Section 2.2 and [73]).

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

$$\frac{dq^{(r)i}}{dt} = q^{(r+1)i}, \qquad r = 0,\dots,k-1, \tag{3.4}$$

$$\frac{dp_i^{(0)}}{dt} = \frac{\partial L}{\partial q^{(0)i}} \tag{3.5}$$

$$\frac{dp_i^{(r)}}{dt} = -p_i^{(r-1)} + \frac{\partial L}{\partial q^{(r)i}}, \qquad r = 1,\dots,k-1 \tag{3.6}$$

and the constraint equation (3.3).

Differentiating with respect to time the equations φ_i^1, substituting into (3.6) and proceeding further, we find the equations of motion for the higher-order variational problem, analyzed in the last chapter, i.e.

$$\sum_{r=0}^{k}(-1)^r \frac{d^r}{dt^r}\left(\frac{\partial L}{\partial q^{(r)i}}\right) = 0 \,. \tag{3.7}$$

The solution of equations (3.1) 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 (assuming that all the subsets generated by the algorithm are submanifolds, see Section 2.2)

$$\cdots \hookrightarrow W_k \hookrightarrow \cdots \hookrightarrow W_2 \hookrightarrow W_1 \hookrightarrow W_0 \,.$$

Algebraically, these constraint submanifolds can be described as

$$W_i = \left\{ x \in T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q) \;\middle|\; dH_{W_0}(x)(v) = 0 \;\; \forall v \in (T_xW_{i-1})^{\perp} \right\} \quad i \geq 1 \,, \tag{3.8}$$

where $(T_xW_{i-1})^{\perp} = \left\{ v \in T_xW_0 \;\middle|\; \Omega_{W_0}(x)(u,v) = 0 \;\; \forall\, u \in T_xW_{i-1} \right\}$.

If this constraint algorithm stabilizes, i.e., there exists a positive integer $k \in \mathbb{N}$ such that $W_{k+1} = W_k$ and $\dim W_k \geq 1$, then we will have at least a well defined solution X on $W_f = W_k$ such that

$$(i_X\Omega_{W_0} = dH_{W_0})_{|W_f} \,.$$

If the bilinear form defined by $\left(\frac{\partial^2 L}{\partial q^{(k)i}\partial q^{(k)j}}\right)$ is nondegenerate we have that the final constraint submanifold is the first one, i.e., $W_f = W_1$. Observe that the dimension of W_1 is even, $\dim\, W_1 = 2kn$. In what follows, we will investigate when this constraint submanifold is equipped with a symplectic 2-form in order to determine a unique solution X. More precisely, if we denote by Ω_{W_1} the restriction of the presymplectic 2-form Ω_{W_0} to W_1, then we have the following result:

Theorem 3.2.1. (W_1, Ω_{W_1}) *is a symplectic manifold if and only if the bilinear form defined by*

$$\left(\frac{\partial^2 L}{\partial q^{(k)i}\partial q^{(k)j}}\right) \tag{3.9}$$

is nondegenerate.

Proof. Ω_{W_1} is symplectic if and only if $T_xW_1 \cap (T_xW_1)^{\perp} = 0$ for all $x \in W_1$. This condition is satisfied if and only if the matrix $d\varphi_1(\frac{\partial}{\partial q^{(k)j}})$ is regular, that is,

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

□

3.2.1 Skinner and Rusk formalism for higher-order constrained mechanical systems

Now, we will consider higher-order mechanical systems subject to higher-order constraints. Let us consider a submanifold $\mathcal{M}$ of $T^{(k)}Q$ locally determined by the vanishing of the constraints functions $\Phi^\alpha : T^{(k)}Q \to \mathbb{R}$, $1 \leq \alpha \leq m$. We will develop a geometric characterization of higher-order constrained variational problems using, as an essential tool, the Skinner and Rusk formulation.

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

$$\left(\frac{\partial(\Phi^1, ..., \Phi^m)}{\partial(q_1^{(k)}, ..., q_n^{(k)})} \right)$$

has constant rank equal to m at all points of $\mathcal{M}$.

Let us take the submanifold $\overline{W}_0 = pr_1^{-1}(\mathcal{M}) = \mathcal{M} \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$ and the restrictions to $\overline{W}_0$ of the canonical projections pr_1 and pr_2

$$\begin{aligned} \pi_1 &: \overline{W}_0 \to \mathcal{M}, \\ \pi_2 &: \overline{W}_0 \to T^*(T^{(k-1)}Q) . \end{aligned}$$

We consider on $\overline{W}_0$ the presymplectic 2-form

$$\Omega_{\overline{W}_0} = \pi_2^*(\omega_{T^{(k-1)}Q})$$

and the function $H_{\overline{W}_0} : \overline{W}_0 \to \mathbb{R}$ given by

$$H_{\overline{W}_0}(p, \alpha) = \langle \alpha, j_k(p) \rangle - L(p)$$

where $(p, \alpha) \in \overline{W}_0$.

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

$$i_X \Omega_{\overline{W}_0} = dH_{\overline{W}_0} . \tag{3.10}$$

Definition 3.2.2. *The presymplectic Hamiltonian system $(\overline{W}_0, \Omega_{\overline{W}_0}, H_{\overline{W}_0})$ will be called a 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 $\overline{W}_0$.

Let us consider $\Omega = pr_2{}^*(\omega_{T^{(k-1)}Q})$ and $H : T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q) \to \mathbb{R}$ given by

$$H(\mu) = \langle pr_2(\mu), pr_1(\mu) \rangle - L(pr_1(\mu)).$$

Then, it is easy to show that equations (3.10) are equivalent to (see [59])

$$\left\{ \begin{array}{rcl} i_X\Omega - dH & \in & (T\overline{W}_0)^0 \\ X & \in & T\overline{W}_0 \, , \end{array} \right. \tag{3.11}$$

where $(T\overline{W}_0)^0$ is the annihilator of $T\overline{W}_0$ locally spanned by $\{d\Phi^\alpha\}$, where $\Phi^\alpha : \overline{W}_0 \to \mathbb{R}$ denote the constraints $\Phi^\alpha = \Phi^\alpha \circ pr_1$ (for notational simplicity, we do not distinguish the notation between constraints on $\mathcal{M}$ and constraints on $\overline{W}_0$).

The solutions of Equation (3.10) are defined on the first constraint submanifold given by the set of points $x \in \overline{W}_0$ such that $(dH + \lambda_\alpha d\Phi^\alpha)(x)(Z) = 0$, for all $Z \in \ker \Omega(x)$. 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 $\overline{W}_1$ (primary constraint submanifold) of $\overline{W}_0$ where (3.10) has a solution.

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

$$\begin{aligned} \Phi^\alpha &= 0 \quad \alpha = 1, \ldots, m \quad \text{(constraints determining } \mathcal{M}\text{)}, && (3.12)\\ \varphi_i^1 &= 0 \quad i = 1, \ldots, n \quad \text{(primary relations)}. && (3.13) \end{aligned}$$

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

$$\begin{aligned} \frac{d}{dt} q^{(r)i} &= q^{(r+1)i}, \qquad r = 0, \ldots, k-1, && (3.14)\\ -\frac{d}{dt} p_i^{(0)} &= -\frac{\partial L}{\partial q^{(0)i}} + \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial q^{(0)i}} && (3.15)\\ -\frac{d}{dt} p_i^{(r)} &= p_i^{(r-1)} - \frac{\partial L}{\partial q^{(r)i}} + \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial q^{(r)i}}, \qquad r = 1, \ldots, k-1 && (3.16) \end{aligned}$$

and the relations (3.12) and (3.13).

Differentiating with respect to time the equations φ_i^1, substituting into (3.16) and proceeding further, we find the equations of motion for the higher-order constrained variational problem

$$\sum_{r=0}^{k} (-1)^r \frac{d^r}{dt^r} \left(\frac{\partial L}{\partial q^{(r)i}} - \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial q^{(r)i}} \right) = 0 \,. \tag{3.17}$$

The solution of equation (3.10) on $\overline{W}_1$ may not be tangent to $\overline{W}_1$. In such a case, we have to restrict $\overline{W}_1$ to the submanifold $\overline{W}_2$ where there exists at least a solution tangent to $\overline{W}_1$. Proceeding further, we obtain a sequence of submanifolds

$$\cdots \hookrightarrow \overline{W}_k \hookrightarrow \cdots \hookrightarrow \overline{W}_2 \hookrightarrow \overline{W}_1 \hookrightarrow \overline{W}_0 \,.$$

Algebraically, these constraint submanifolds can be described as

$$\overline{W}_i = \left\{ x \in \mathcal{M} \times_{T^{(k-1)}Q} T^* \left(T^{(k-1)}Q \right) \;\middle|\; dH_{\overline{W}_0}(x)(v) = 0 \;\; \forall v \in \left(T_x \overline{W}_{i-1} \right)^\perp \right\} \quad i \geq 1 \,, \tag{3.18}$$

where $\left(T_x\overline{W}_{i-1}\right)^{\perp} = \left\{v \in T_x\overline{W}_0 \;\middle|\; \Omega_{\overline{W}_0}(x)(u,v) = 0 \;\; \forall\, u \in T_x\overline{W}_{i-1}\right\}$.

If this constraint algorithm stabilizes, i.e., there exists a positive integer $k \in \mathbb{N}$ such that $\overline{W}_{k+1} = \overline{W}_k$ and $\dim \overline{W}_k \geq 1$, then we will have at least a well defined solution X on $\overline{W}_f = \overline{W}_k$ such that

$$\left(i_X\Omega_{\overline{W}_0} = dH_{\overline{W}_0}\right)|_{\overline{W}_f} \ .$$

Denote by $\Omega_{\overline{W}_1}$, the pullback of the presymplectic 2-form $\Omega_{\overline{W}_0}$ to $\overline{W}_1$. In order to establish a necessary and sufficient condition for the symplecticity of the 2-form $\Omega_{\overline{W}_1}$, we define the extended Lagrangian

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

Theorem 3.2.3. *For any choice of coordinates* $(q^{(0)i}, q^{(1)i}, \ldots, q^{(k-1)i}; p_i^{(0)}, \ldots, p_i^{(k-1)}, q^{(k)i})$ *in* $T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$, *we have that* $(\overline{W}_1, \Omega_{\overline{W}_1})$ *is a symplectic manifold if and only if*

$$\det\begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial q^{(k)i}\partial q^{(k)j}} & -\frac{\partial \Phi^\alpha}{\partial q^{(k)i}} \\ \frac{\partial \Phi^\beta}{\partial q^{(k)j}} & \mathbf{0} \end{pmatrix} = det\begin{pmatrix} \frac{\partial^2 L}{\partial q^{(k)i}\partial q^{(k)j}} - \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial q^{(k)i}\partial q^{(k)j}} & -\frac{\partial \Phi^\alpha}{\partial q^{(k)i}} \\ \frac{\partial \Phi^\beta}{\partial q^{(k)j}} & \mathbf{0} \end{pmatrix} \neq 0 \qquad (3.19)$$

The proof of this theorem follows the same lines that the one in Proposition (3.6.2).

Remark 3.2.4. Observe that if the determinant of the matrix in Theorem 3.2.3 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 λ_α and higher-order velocities $q^{(k)i}$ in terms of coordinates $(q^{(0)i}, \ldots, q^{(k-1)i}, p_i^{(0)}, \ldots, p_i^{(k-1)})$, i.e.,

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

Thus we can consider $(q^{(0)i}, q^{(1)i}, \ldots, q^{(k-1)i}, p_i^{(0)}, \ldots, p_i^{(k-1)})$ as local coordinates in $\overline{W}_1$. In this case,

$$\Omega_{\overline{W}_1} = \sum_{r=0}^{k-1} dq^{(r)i} \wedge dp_i^{(r)}$$

which is obviously symplectic. ◇

3.3 Skinner and Rusk formalism for higher-order problems on Lie groups

Now, we will give an adaptation of the Skinner-Rusk algorithm to the case of higher-order theories on Lie groups (see [15] and [50] for the standard case). Using the results given in Section (1.7.1) and Appendix A we have the identifications

$$\begin{aligned} T^{(k)}G &\equiv G \times k\mathfrak{g} \ , \\ T^*T^{(k-1)}G &\equiv G \times (k-1)\mathfrak{g} \times k\mathfrak{g}^* \ . \end{aligned}$$

For developing our geometric formalism for higher-order variational 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 $\omega_{G\times(k-1)\mathfrak{g}}$ after left-trivialization. Denote by $\boldsymbol{\xi}\in(k-1)\mathfrak{g}$ and $\boldsymbol{\alpha}\in k\mathfrak{g}^*$ with components $\boldsymbol{\xi}=(\xi^{(0)},\dots,\xi^{(k-2)})$ and $\boldsymbol{\alpha}=(\alpha_0,\dots,\alpha_{k-1})$. Then, after a technical computation (see Appendix A) we deduce that

$$\begin{aligned}
(\theta_{G\times(k-1)\mathfrak{g}})_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}(\boldsymbol{\xi}_1,\boldsymbol{\nu}^1) &= \langle\boldsymbol{\alpha},\boldsymbol{\xi}_1\rangle\,,\\
(\omega_{G\times(k-1)\mathfrak{g}})_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}\big((\boldsymbol{\xi}_1,\boldsymbol{\nu}^1),(\boldsymbol{\xi}_2,\boldsymbol{\nu}^2)\big) &= -\langle\boldsymbol{\nu}^1,\boldsymbol{\xi}_2\rangle+\langle\boldsymbol{\nu}^2,\boldsymbol{\xi}_1\rangle+\langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle\\
&= -\sum_{i=0}^{k-1}\Big[\langle\nu^1_{(i)},\xi_2^{(i)}\rangle+\langle\nu^2_{(i)},\xi_1^{(i)}\rangle\Big]+\langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle,
\end{aligned}$$

where $\boldsymbol{\xi}_a\in k\mathfrak{g}$ and $\boldsymbol{\nu}^a\in k\mathfrak{g}^*$, $a=1,2$ with components $\boldsymbol{\xi}_a=(\xi_a^{(i)})_{0\leq i\leq k-1}$ and $\boldsymbol{\nu}^a=(\nu^a_{(i)})_{0\leq i\leq k-1}$ where each component $\xi_a^{(i)}\in\mathfrak{g}$ and $\nu^a_{(i)}\in\mathfrak{g}^*$. Observe that α_0 comes from the identification $T^*G=G\times\mathfrak{g}^*$.

Consider the *higher-order Pontryagin 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

$$\begin{aligned}
pr_1(g,\boldsymbol{\xi},\xi^{(k-1)},\boldsymbol{\alpha}) &= (g,\boldsymbol{\xi},\xi^{(k-1)})\\
pr_2(g,\boldsymbol{\xi},\xi^{(k-1)},\boldsymbol{\alpha}) &= (g,\boldsymbol{\xi},\boldsymbol{\alpha})
\end{aligned}$$

where, as usual, $\boldsymbol{\xi}=(\xi^{(0)},\dots,\xi^{(k-2)})\in(k-1)\mathfrak{g}$ and $\boldsymbol{\alpha}=(\alpha_0,\dots,\alpha_{k-1})\in k\mathfrak{g}^*$.

The following diagram summarizes the structure of the higher-order Pontryagin bundle:

$$\begin{tikzcd}
& W_0=G\times k\mathfrak{g}\times k\mathfrak{g}^* \arrow[dl, "pr_1"'] \arrow[dr, "pr_2"] & \\
G\times k\mathfrak{g} \arrow[dr, "\tau_G^{(k-1,k)}"'] & & G\times(k-1)\mathfrak{g}\times k\mathfrak{g}^* \arrow[dl, "\pi_{G\times(k-1)\mathfrak{g}}"] \\
& G\times(k-1)\mathfrak{g} &
\end{tikzcd}$$

For developing the Skinner and Rusk formalism it is only necessary to construct the presymplectic 2-form Ω_{W_0} by $\Omega_{W_0}=pr_2^*\omega_{G\times(k-1)\mathfrak{g}}$ and the Hamiltonian function $H:W_0\to\mathbb{R}$ by

$$H(g,\boldsymbol{\xi},\xi^{(k-1)},\boldsymbol{\alpha})=\sum_{i=0}^{k-1}\langle\alpha_i,\xi^{(i)}\rangle-L(g,\boldsymbol{\xi},\xi^{(k-1)})\,. \tag{3.20}$$

Therefore

$$\begin{aligned}
(\Omega_{W_0})_{(g,\boldsymbol{\xi},\xi^{(k-1)},\boldsymbol{\alpha})}\Big((\boldsymbol{\xi}_1,\xi_1^{(k)},\boldsymbol{\nu}^1),(\boldsymbol{\xi}_2,\xi_2^{(k)},\boldsymbol{\nu}^2)\Big) &= -\langle\boldsymbol{\nu}^1,\boldsymbol{\xi}_2\rangle+\langle\boldsymbol{\nu}^2,\boldsymbol{\xi}_1\rangle\\
+\langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle &= -\sum_{i=0}^{k-1}\Big[\langle\nu^1_{(i)},\xi_2^{(i)}\rangle-\langle\nu^2_{(i)},\xi_1^{(i)}\rangle\Big]+\langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle\,,
\end{aligned}$$

where $\boldsymbol{\xi}_a \in k\mathfrak{g}$, $\boldsymbol{\nu}^a \in k\mathfrak{g}^*$, and $\xi_a^{(k)} \in \mathfrak{g}$, $a = 1, 2$. As a consequence of the definition of the presymplectic 2-form Ω_{W_0}, the variable $\xi_a^{(k)}$ does not appear on the right-hand side of the previous expression. Moreover,

$$\begin{aligned} dH_{(g,\boldsymbol{\xi},\xi^{(k-1)},\boldsymbol{\alpha})}(\boldsymbol{\xi}_2, \xi_2^{(k)}, \boldsymbol{\nu}^2) \; = \; & \langle -\pounds_g^* \left(\frac{\delta L}{\delta g}(g, \boldsymbol{\xi}, \xi^{(k-1)}) \right), \xi_2^{(0)} \rangle \\ & + \sum_{i=0}^{k-2} \langle \alpha_i - \frac{\delta L}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \xi^{(k-1)}), \xi_2^{(i+1)} \rangle \\ & + \langle \boldsymbol{\nu}^2, \boldsymbol{\xi} \rangle \; . \end{aligned}$$

Therefore, the *intrinsic equations of motion* of a higher-order problem on Lie groups are now

$$i_X \Omega_{W_0} = dH \; . \tag{3.21}$$

If we look for a solution $X(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}) = (\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1)$ of Equation (3.21) we deduce:

$$\begin{aligned} \xi_1^{(i)} &= \xi^{(i)}, \quad 0 \leq i \leq k-1 \; , \\ \nu_{(0)}^1 &= \pounds_g^* \left(\frac{\delta L}{\delta g}(g, \boldsymbol{\xi}, \xi^{(k-1)}) \right) + ad_{\xi_1^{(0)}} \alpha_0 \; , \\ \nu_{(i+1)}^1 &= \frac{\delta L}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \xi^{(k-1)}) - \alpha_i, \quad 0 \leq i \leq k-2 \; , \end{aligned}$$

and the constraint functions

$$\alpha_{k-1} - \frac{\delta L}{\delta \xi^{(k-1)}}(g, \boldsymbol{\xi}, \xi^{(k-1)}) = 0 \; .$$

Observe that the coefficients ξ_1^k are still undetermined.

An integral curve of X, that is a curve of the type

$$t \longrightarrow (g(t), \xi(t), \ldots, \xi^{(k-1)}(t), \alpha_0(t), \ldots, \alpha_{k-1}(t)) \; ,$$

must satisfy the following system of differential-algebraic equations (DAEs):

$$\dot{g} = g\xi \; , \tag{3.22}$$

$$\frac{d\xi^{(i-1)}}{dt} = \xi^{(i)}, \quad 1 \leq i \leq k-1 \; , \tag{3.23}$$

$$\frac{d\alpha_0}{dt} = \pounds_g^* \left(\frac{\delta L}{\delta g}(g, \boldsymbol{\xi}, \xi^{(k-1)}) \right) + ad_\xi^* \alpha_0 \; , \tag{3.24}$$

$$\frac{d\alpha_{i+1}}{dt} = \frac{\delta L}{\delta \xi^{(i)}}(g, \boldsymbol{\xi}, \xi^{(k-1)}) - \alpha_i, \quad 0 \leq i \leq k-2 \; , \tag{3.25}$$

$$\alpha_{k-1} = \frac{\delta L}{\delta \xi^{(k-1)}}(g, \boldsymbol{\xi}, \xi^{(k-1)}) \; . \tag{3.26}$$

If $k \geq 2$, combining Equation (3.26) with Equation (3.25) for $i = k-2$, we obtain

$$\frac{d}{dt} \frac{\delta L}{\delta \xi^{(k-1)}} = \frac{\delta L}{\delta \xi^{(k-2)}} - \alpha_{k-2} \; .$$

Proceeding successively, now with $i = k-3$ and ending with $i = 0$ we obtain the following relation:

$$\alpha_0 = \sum_{i=0}^{k-1} (-1)^i \frac{d^i}{dt^i} \frac{\delta L}{\delta \xi^{(i)}} .$$

This last expression is also valid for $k \geq 1$. Substituting in Equation (3.24) we finally deduce the k^{th}*-order trivialized Euler-Lagrange equations*:

$$\left(\frac{d}{dt} - ad_\xi^* \right) \sum_{i=0}^{k-1} (-1)^i \frac{d^i}{dt^i} \frac{\delta L}{\delta \xi^{(i)}} = \pounds_g^* \left(\frac{\delta L}{\delta g} \right) . \tag{3.27}$$

Of course if the Lagrangian $L : T^{(k)}G \equiv G \times k\mathfrak{g} \longrightarrow \mathbb{R}$ is left-invariant, that is

$$L(g, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}) = L(h, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}) ,$$

for all $g, h \in G$, then defining the reduced Lagrangian $\ell : k\mathfrak{g} \longrightarrow \mathbb{R}$ by

$$\ell(\xi, \dot{\xi}, \ldots, \xi^{(k-1)}) = L(e, \xi, \dot{\xi}, \ldots, \xi^{(k-1)}) ,$$

we write Equations (3.27) as

$$\left(\frac{d}{dt} - ad_\xi^* \right) \sum_{i=0}^{k-1} (-1)^i \frac{d^i}{dt^i} \frac{\delta \ell}{\delta \xi^{(i)}} = 0 , \tag{3.28}$$

which are the k^{th}*-order Euler-Poincaré equations*. These equations are exactly the same that the ones derived by Gay-Balmaz *et al* in [68] using variational techniques. Our derivation allows us to identify automatically the geometric preservation properties of the system, for instance, preservation of the Hamiltonian or (pre-) symplecticity of the flow.

The constraint algorithm

From the presymplectic character of Ω_{W_0} and the expression of $H : W_0 \to \mathbb{R}$ we have that (3.21) has not solution on W_0 then it is necessary to identify the unique maximal submanifold W_f along which (3.21) possesses tangent solutions on W_f and therefore, the existence of solutions is guaranteed. This final constraint submanifold W_f is detected using the constraint algorithm (2.2.2). 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

$$W_i = \left\{ x \in G \times k\mathfrak{g} \times k\mathfrak{g}^* \;\middle|\; dH(x)(\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1) = 0 \; \forall (\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1) \in (T_x W_{i-1})^\perp \right\} \tag{3.29}$$

with $i \geq 1$ and where

$$(T_x W_{i-1})^\perp = \Big\{ (\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1) \in (k+1)\mathfrak{g} \times k\mathfrak{g}^* \;\Big|\; \Omega_{W_0}(x)((\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1), (\boldsymbol{\xi}_2, \xi_2^{(k)}, \boldsymbol{\nu}^2)) = 0 \; \forall \, (\boldsymbol{\xi}_2, \xi_2^{(k)}, \boldsymbol{\nu}^2) \in T_x W_{i-1} \Big\} .$$

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 $\dim W_k \geq 1$, then we will have at least a well defined solution X on $W_f = W_k$ such that

$$(i_X \Omega_{W_0} = dH)_{|W_f} \ .$$

From these definitions, we deduce that the first constraint submanifold W_1 is defined by the vanishing of the constraint functions

$$\varphi = \alpha_{k-1} - \frac{\delta L}{\delta \xi^{(k-1)}} = 0 \ . \tag{3.30}$$

The existence of solutions satisfying the constraints Equations (3.30) implies using Equations (3.24), (3.25) the following compatibility conditions: if $k > 2$,

$$\frac{\delta L}{\delta \xi^{(k-2)}} - \alpha_{k-2} = \frac{\delta^2 L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} \xi_1^{(k)} + \sum_{i=0}^{k-2} \frac{\delta^2 L}{\delta \xi^{(k-1)} \delta \xi^{(i)}} \xi^{i+1} + \pounds_g^* \left(\frac{\delta^2 L}{\delta \xi^{(k-1)} \delta g} \right) \xi \ ,$$

and, in the particular case $k = 1$, we deduce the equation

$$\pounds_g^* \left(\frac{\delta L}{\delta g} \right) + ad_\xi^* \alpha_0 = \frac{\delta^2 L}{\delta \xi^2} \xi^{(1)} + \pounds_g^* \left(\frac{\delta^2 L}{\delta \xi \delta g} \right) \xi \ .$$

In both cases, these equations impose restrictions over the remainder coefficients $\xi_1^{(k)}$ of the vector field X.

If the bilinear form $\frac{\delta^2 L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ defined by

$$\frac{\delta^2 L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} (g, \boldsymbol{\xi}, \xi^{(k-1)})(\xi, \tilde{\xi}) = \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} L(g, \boldsymbol{\xi}, \xi^{(k-1)} + t\xi + s\tilde{\xi})$$

is nondegenerate, we have a special case when the constraint algorithm finishes at the first step W_1. More precisely, if we denote by Ω_{W_1} the restriction of the presymplectic 2-form Ω to W_1, then we have the following result:

Theorem 3.3.1. (W_1, Ω_{W_1}) *is a symplectic manifold if and only if*

$$\frac{\delta^2 L}{\delta \xi^{(k-1)} \delta \xi^{(k-1)}} \tag{3.31}$$

is nondegenerate.

Proof. For all $x = (g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}) \in W_1$, defined by the vanishing of the constraints (3.30), that is, $\varphi(x) = 0$, we have that

$$T_x \varphi : T_x W_0 \longrightarrow T_0 \mathfrak{g}^* \equiv \mathfrak{g}^*$$

is obviously surjective ($\varphi : W_0 \to \mathfrak{g}^*$ is a submersion).

Now, assume that Ω_{W_1} is symplectic, then for all $x \in W_1$, $(T_xW_1)^\perp \cap T_xW_1 = 0$. Since $\ker \Omega_x \subseteq T_xW_1^\perp$, therefore $\ker \Omega_x \cap T_xW_1 = 0$. Now observing that any element in $V_x \in \ker \Omega_x$ is written as

$$V_x = (\mathbf{0}, \xi_1^{(k)}, \mathbf{0}) \in \ker \Omega_x \subset k\mathfrak{g} \times \mathfrak{g} \times k\mathfrak{g}^* ,$$

we deduce that $T_x\varphi_{|\ker \Omega_x} : \ker \Omega_x \equiv \mathfrak{g} \longrightarrow \mathfrak{g}^*$ is an isomorphism and in consequence, the bilinear form (3.31) is nondegenerate.

Conversely, since we know that the nondegeneracy of (3.31) is equivalent to say that $T_x\varphi_{|\ker \Omega_x}$ is an isomorphism, then we deduce that

$$T_xW_0 = T_xW_1 \oplus \ker \Omega_x . \tag{3.32}$$

Observe that if $Z_x \in T_xW_0$, we can write $Z_x = (Z_x - V_x) + V_x$ where V_x is the unique element of $\ker \Omega_x$ such that $T_x\varphi(V_x) = T_x\varphi(Z_x)$. Then, $Z_x - V_x \in T_xW_1$. From Equation (3.32) we deduce that $T_xW_1^\perp = \ker \Omega_x$, for all $x \in W_1$ since

$$\dim(T_xW_1)^\perp = \dim T_xW_0 - \dim T_xW_1 + \dim(T_xW_1 \cap \ker \Omega_x) .$$

In consequence, $0 = \ker \Omega_x \cap T_xW_1 = (T_xW_1)^\perp \cap T_xW_1 = 0$ and Ω_{W_1} is a symplectic 2-form.

□

3.3.1 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, $L_{\mathcal{M}} : \mathcal{M} \to \mathbb{R}$, is an extension of the previous framework. First, it is necessary to note that for constrained system, we understand a variational problem subject to constraints (vakonomic mechanics), being this analysis completely different in the case of nonholonomic constraints (see [17, 138, 56]). The case of nonholonomic mechanical systems will be studied in Chapter 4.

Given the pair $(\mathcal{M}, L_{\mathcal{M}})$ we can define the space

$$\overline{W}_0 = \mathcal{M} \times k\mathfrak{g}^* .$$

Taking the inclusion $i_{\overline{W}_0} : \overline{W}_0 \hookrightarrow G \times k\mathfrak{g} \times k\mathfrak{g}^*$, then we can construct the following presymplectic form

$$\Omega_{\overline{W}_0} = (pr_2 \circ i_{\overline{W}_0})^* \Omega_{G\times(k-1)\mathfrak{g}\times k\mathfrak{g}^*} ,$$

and the function $\bar{H} : \overline{W}_0 \to \mathbb{R}$ defined by

$$\bar{H}(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}) = \sum_{i=0}^{k-1} \langle \alpha_i, \xi^{(i)} \rangle - L_{\mathcal{M}}(g, \boldsymbol{\xi}, \xi^{(k-1)}) ,$$

where $(g, \boldsymbol{\xi}, \xi^{(k-1)}) \in \mathcal{M}$.

With these two elements it is possible to write the following presymplectic system:

$$i_X \Omega_{\overline{W}_0} = d\bar{H} \; . \tag{3.33}$$

This then justifies the use of the following terminology.

Definition 3.3.2. *The presymplectic Hamiltonian system $(\overline{W}_0, \Omega_{\overline{W}_0}, \bar{H})$ will be called 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 $L : G \times k\mathfrak{g} \to \mathbb{R}$ of $L_{\mathcal{M}} : \mathcal{M} \to \mathbb{R}$. The main idea is to take into account that Equation (3.33) is equivalent to

$$\left\{ \begin{array}{rcl} i_X \Omega_{W_0} - dH & \in & \operatorname{ann} T\overline{W}_0 \; , \\ X & \in & T\overline{W}_0 \; , \end{array} \right.$$

where ann denotes the annihilator of a distribution and H is the function defined in (3.20).

Assuming that $\mathcal{M}$ is determined by the vanishing of m-independent constraints

$$\Phi^A(g, \boldsymbol{\xi}, \xi^{(k-1)}) = 0, \; 1 \leq A \leq m \; ,$$

then, locally, $\operatorname{ann} T\overline{W}_0 = \operatorname{span} \{d\Phi^A\}$, and therefore the previous equations are rewritten as

$$\left\{ \begin{array}{rcl} i_X \Omega_{W_0} - dH & = & \lambda_A d\Phi^A \, , \\ X(\Phi^A) & = & 0 \; , \end{array} \right.$$

where λ_A are Lagrange multipliers to be determined.

If $X(g, \boldsymbol{\xi}, \xi^{(k-1)}, \boldsymbol{\alpha}) = (\boldsymbol{\xi}_1, \xi_1^{(k)}, \boldsymbol{\nu}^1)$ then, as in the previous subsection, we obtain the following prescription about these coefficients:

$$\begin{array}{rcl} \xi_1^{(i)} & = & \xi^{(i)}, \quad 0 \leq i \leq k-1 \, , \\ \nu^1_{(0)} & = & \pounds_g^* \left(\dfrac{\delta L}{\delta g} - \lambda_A \dfrac{\delta \Phi^A}{\delta g} \right) + ad_{\xi_1^{(0)}} \alpha_0 \; , \\ \nu^1_{(i+1)} & = & \dfrac{\delta L}{\delta \xi^{(i)}} - \lambda_A \dfrac{\delta \Phi^A}{\delta \xi^{(i)}} - \alpha_i, \quad 0 \leq i \leq k-2 \; , \\ 0 & = & \pounds_g^* \left(\dfrac{\delta \Phi^A}{\delta g} \right) \xi + \displaystyle\sum_{i=1}^{k-2} \dfrac{\delta \Phi^A}{\delta \xi^{(i)}} \xi^{(i+1)} + \dfrac{\delta \Phi^A}{\delta \xi^{(k-1)}} \xi_1^{(k)} \; , \quad 1 \leq A \leq m \, , \end{array}$$

and the algebraic equations:

$$\begin{array}{rcl} \alpha_{k-1} - \dfrac{\delta L}{\delta \xi^{(k-1)}} + \lambda_A \dfrac{\delta \Phi^A}{\delta \xi^{(k-1)}} & = & 0 \; , \\ \Phi^A & = & 0 \; . \end{array}$$

The integral curves of X satisfy the system of differential-algebraic equations with additional unknowns (λ_A):

$$\begin{aligned}
\dot{g} &= g\xi \, , \\
\frac{d\xi^{(i-1)}}{dt} &= \xi^{(i)}, \quad 1 \leq i \leq k-1 \, , \\
\frac{d\alpha_0}{dt} &= \pounds_g^* \left(\frac{\delta L}{\delta g} - \lambda_A \frac{\delta \Phi^A}{\delta g} \right) + ad_\xi^* \alpha_0 \, , \\
\frac{d\alpha_{i+1}}{dt} &= \frac{\delta L}{\delta \xi^{(i)}} - \lambda_A \frac{\delta \Phi^A}{\delta \xi^{(i)}} - \alpha_i \, , \\
0 &= \pounds_g^* \left(\frac{\delta \Phi^A}{\delta g} \right) \xi + \sum_{i=1}^{k-2} \frac{\delta \Phi^A}{\delta \xi^{(i)}} \xi^{(i+1)} + \frac{\delta \Phi^A}{\delta \xi^{(k-1)}} \xi_1^{(k-1)} \\
\alpha_{k-1} &= \frac{\delta L}{\delta \xi^{(k-1)}} - \lambda_A \frac{\delta \Phi^A}{\delta \xi^{(k-1)}} \, , \\
\Phi^A &= 0 \, .
\end{aligned}$$

As a consequence we finally obtain the k^{th}*-order trivialized constrained Euler-Lagrange equations*,

$$\left(\frac{d}{dt} - ad_\xi^* \right) \sum_{i=0}^{k-1} (-1)^i \frac{d^i}{dt^i} \left[\frac{\delta L}{\delta \xi^{(i)}} - \lambda_A \frac{\delta \Phi^A}{\delta \xi^{(i)}} \right] = \pounds_g^* \left(\frac{\delta L}{\delta g} - \lambda_A \frac{\delta \Phi^A}{\delta g} \right) \, . \tag{3.34}$$

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 $\ell : k\mathfrak{g} \longrightarrow \mathbb{R}$ and the reduced constraints $\phi^A : k\mathfrak{g} \to \mathbb{R}$ we write Equations (3.34) as

$$\left(\frac{d}{dt} - ad_\xi^* \right) \sum_{i=0}^{k-1} (-1)^i \frac{d^i}{dt^i} \left[\frac{\delta \ell}{\delta \xi^{(i)}} - \lambda_A \frac{\delta \phi^A}{\delta \xi^{(i)}} \right] = 0 \, . \tag{3.35}$$

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 (3.33).

Observe that the first constraint submanifold $\overline{W}_1$ is determined by the conditions

$$\begin{aligned}
\varphi_\lambda = \alpha_{k-1} - \frac{\delta L}{\delta \xi^{(k-1)}} + \lambda_A \frac{\delta \Phi^A}{\delta \xi^{(k-1)}} &= 0 \, , && (3.36) \\
\Phi^A &= 0 \, . && (3.37)
\end{aligned}$$

If we denote by $\Omega_{\overline{W}_1}$ the pullback of the presymplectic 2-form $\Omega_{\overline{W}_0}$ to $\overline{W}_1$, then we deduce the following theorem

Theorem 3.3.3. $(\overline{W}_1, \Omega_{\overline{W}_1})$ *is a symplectic manifold if and only if*

$$\begin{pmatrix} \dfrac{\delta^2 L}{\delta\xi^{(k-1)}\delta\xi^{(k-1)}} + \lambda_A \dfrac{\delta^2 \Phi^A}{\delta\xi^{(k-1)}\delta\xi^{(k-1)}} & \dfrac{\delta\Phi^A}{\delta\xi^{(k-1)}} \\ \dfrac{\delta\Phi^A}{\delta\xi^{(k-1)}} & \mathbf{0} \end{pmatrix} \tag{3.38}$$

is nondegenerate, considered as a bilinear form on the vector space $\mathfrak{g} \times \mathbb{R}^m$.

Proof. Consider the following extended manifold $W_0 \times \mathbb{R}^m$ with the induced presymplectic 2-form $\Omega_{W_0\times\mathbb{R}^m} = (\mathrm{Pr}_1)^*\Omega_{W_0}$ where Pr_1 is the projection of $W_0 \times \mathbb{R}^m$ onto the first factor. It is clear that

$$\ker \Omega_{W_0\times\mathbb{R}^m} = \mathrm{span}\ \{(V, \mathbf{0}), (\mathbf{0}, \frac{\partial}{\partial\lambda_A})\}\ ,$$

where $V \in \ker \Omega_{W_0}$ and (λ_A) are coordinates in $\mathbb{R}^m$. We have a new presymplectic Hamiltonian system given by the triple $(W_0 \times \mathbb{R}^m, \Omega_{W_0\times\mathbb{R}^m}, H_{W_0\times\mathbb{R}^m})$ where $H_{W_0\times\mathbb{R}^m} = H + \lambda_A\Phi^A$. Applying the constraint algorithm to this presymplectic system, we observe that (3.36) and (3.37) are exactly the primary constraints and determine a primary constraint submanifold $\widetilde{W}_1$ of $W_0 \times \mathbb{R}^m$. We construct the map

$$\begin{array}{rccc} (\varphi_\lambda, \Phi^A): & W_0 \times \mathbb{R}^m & \longrightarrow & \mathfrak{g}^* \times \mathbb{R}^m \\ & (x, \lambda_A) & \longmapsto & (\varphi_\lambda(x, \lambda_A), \Phi^A(x)). \end{array}$$

This map is a submersion and therefore, applying similar arguments to Theorem 3.2.1, we deduce that condition of nondegeneracy (3.38) is equivalent to the symplecticity of $(\widetilde{W}_1, \Omega_{\widetilde{W}_1})$ where $\Omega_{\widetilde{W}_1}$ is the pullback of $\Omega_{W_0\times\mathbb{R}^m}$ to $\widetilde{W}_1$. The proof is finished observing, that under these regularity conditions, we have that, via Pr_1, $\widetilde{W}_1$ and $\overline{W}_1$ are diffeomorphic and $(\mathrm{Pr}_1)^*_{|\widetilde{W}_1}\Omega_{\overline{W}_1} = \Omega_{\widetilde{W}_1}$.

□

3.4 Extension to trivial principal bundles

In this section we show how to combine the results given in the previous sections and as a consequence to obtain the Skinner-Rusk formalism for higher-order mechanical systems whose configuration space is a trivial principal bundle.

3.4.1 Unconstrained problem

As in the previous section, consider the higher-order Pontryagin bundle

$$\begin{aligned} W &= T^{(k)}(M \times G) \times_{T^{(k-1)}(M\times G)} T^*(T^{(k-1)}(M \times G)) \\ &\simeq (T^{(k)}M \times_{T^{(k-1)}M} T^*(T^{(k-1)}M)) \times (G \times k\mathfrak{g} \times k\mathfrak{g}^*) =: W_M \times W_G \end{aligned}$$

where M is a m-dimensional smooth manifold and G is a finite dimensional Lie group.

Let $(q^{(r)i}, q^{(k)i}, p_i^{(r)})$, where $0 \leqslant r \leqslant k-1$ and $1 \leqslant i \leqslant m$, be a set of local coordinates in W_M and $(g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha})$, where $\boldsymbol{\xi} = (\xi^0, \dots, \xi^{k-2}) \in (k-1)\mathfrak{g}$ and $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{k-1}) \in k\mathfrak{g}^*$, a set of local coordinates in W_G. Then, the induced natural coordinates in $W = W_M \times W_G$ are $(q^{(r)i}, q^{(k)i}, p_i^{(r)}, g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha})$. Using these coordinates, we denote the projections of W, W_M, W_G onto their factors in the following way

$$\mathrm{pr}_1(q^{(r)i}, q^{(k)i}, p_i^{(r)}, g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}) = (q^{(r)i}, q^{(k)i}, p_i^{(r)}), \quad \mathrm{pr}_2(q^{(r)i}, q^{(k)i}, p_i^{(r)}, g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}) = (g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}),$$
$$\widetilde{\mathrm{pr}}_1(q^{(r)i}, q^{(k)i}, p_i^{(r)}) = (q^{(r)i}, q^{(k)i}), \quad \widetilde{\mathrm{pr}}_2(q^{(r)i}, q^{(k)i}, p_i^{(r)}) = (q^{(r)i}, p_i^{(r)}),$$
$$\overline{\mathrm{pr}}_1(g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}) = (g, \boldsymbol{\xi}, \xi^{k-1}), \quad \overline{\mathrm{pr}}_2(g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}) = (g, \boldsymbol{\xi}, \boldsymbol{\alpha}).$$

The bundle W is endowed with some canonical geometric structures. As in the previous sections, let $\omega_{T^{(k-1)}Q}$ and $\omega_{G\times(k-1)\mathfrak{g}}$ be the canonical symplectic forms on $T^*(T^{(k-1)}M)$ and $T^*(T^{(k-1)}G) \simeq G \times (k-1)\mathfrak{g} \times k\mathfrak{g}^*$, respectively. Then, we can consider the presymplectic forms $\Omega_M = \widetilde{\mathrm{pr}}_2^* \omega_{T^{(k-1)}Q}$ and $\Omega_G = \overline{\mathrm{pr}}_2^* \omega_{G\times(k-1)\mathfrak{g}}$. We define the following presymplectic form in W

$$\Omega = \mathrm{pr}_1^* \Omega_M + \mathrm{pr}_2^* \Omega_G, \tag{3.39}$$

where a local basis for $\ker \Omega$ is

$$\ker \Omega = \left\langle \frac{\partial}{\partial q^{(k)i}}, \frac{\partial}{\partial \xi^{k-1}} \right\rangle. \tag{3.40}$$

Now, given a Lagrangian function $L \in \mathrm{C}^\infty(T^{(k)}(M \times G)) = \mathrm{C}^\infty(T^{(k)}M \times G \times k\mathfrak{g})$, we can define the *Hamiltonian function* $H \in \mathrm{C}^\infty(W)$ locally as

$$H(q^{(r)i}, q^{(k)i}, p_i^{(r)}, g, \boldsymbol{\xi}, \xi^{k-1}, \boldsymbol{\alpha}) = \sum_{r=0}^{k-1} \left(p_i^{(r)} q^{(r+1)i} + \langle \alpha_i, \xi^i \rangle \right) - L(q^{(r)i}, q^{(k)i}, g, \boldsymbol{\xi}, \xi^{k-1}). \tag{3.41}$$

The equations of motion for a presymplectic Hamiltonian system (W, Ω, H) are given by

$$i_X \Omega = \mathrm{d}H \ , \quad \text{for } X \in \mathfrak{X}(W). \tag{3.42}$$

As in the previous sections, the first constraint submanifold $W_1 \hookrightarrow W$ is locally defined by the constraints

$$p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}} = 0 \quad ; \quad \alpha_{k-1} - \frac{\partial L}{\partial \xi^{k-1}} = 0 \, .$$

Let $X \in \mathfrak{X}(W)$ be a generic vector field locally given by

$$X = F^{(r)i} \frac{\partial}{\partial q^{(r)i}} + F^{(k)i} \frac{\partial}{\partial q^{(k)i}} + G_i^{(r)} \frac{\partial}{\partial p_i^{(r)}} + \xi_1^0 \frac{\partial}{\partial g} + \xi_1^{i+1} \frac{\partial}{\partial \xi^i} + \nu_i^1 \frac{\partial}{\partial \alpha_i} \, . \tag{3.43}$$

Then, from (3.42) we have the following system of equations

$$F^{(r)i} = q^{(r+1)i}\,, \tag{3.44}$$

$$G_i^{(0)} = \frac{\partial L}{\partial q^{(0)i}} \quad , \quad G_i^{(r)} = \frac{\partial L}{\partial q^{(r)i}} - p_i^{(r-1)}\,, \tag{3.45}$$

$$p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}} = 0\,, \tag{3.46}$$

$$\xi_1^i = \xi^i\,, \tag{3.47}$$

$$\nu_0^1 = \pounds_g^* \frac{\partial L}{\partial g} + ad^*_{\xi_1^0}\alpha_0 \quad , \quad \nu_{i+1}^1 = \frac{\partial L}{\partial \xi^i} - \alpha_i\,, \tag{3.48}$$

$$\alpha_{k-1} - \frac{\partial L}{\partial \xi^{k-1}} = 0\,. \tag{3.49}$$

Observe that the coefficients $F^{(k)i}$ and ξ_1^k are yet to be determined and that the tangency condition for a vector field X solution to equation (3.42) along W_1 gives the following equations

$$\begin{aligned}
\frac{\partial L}{\partial q^{(k-1)i}} - p_i^{(k-2)} &= q^{(r+1)j}\frac{\partial^2 L}{\partial q^{(r)j}\partial q^{(k)i}} + F^{(k)j}\frac{\partial^2 L}{\partial q^{(k)j}\partial q^{(k)i}} \\
&\quad + \pounds_g^* \frac{\partial^2 L}{\partial g \partial q^{(k)i}}\xi^0 + \frac{\partial^2 L}{\partial \xi^i \partial q^{(k)i}}\xi^{i+1} + \frac{\partial^2 L}{\partial \xi^{k-1}\partial q^{(k)i}}\xi_1^k\,, \\
\frac{\partial L}{\partial \xi^{k-2}} - \alpha_{k-2} &= q^{(r+1)j}\frac{\partial^2 L}{\partial q^{(r)j}\partial \xi^{k-1}} + F^{(k)j}\frac{\partial^2 L}{\partial q^{(k)j}\partial \xi^{k-1}} \\
&\quad + \pounds_g^* \frac{\partial^2 L}{\partial g \partial \xi^{k-1}}\xi^0 + \frac{\partial^2 L}{\partial \xi^i \partial \xi^{k-1}}\xi^{i+1} + \frac{\partial^2 L}{\partial \xi^{k-1}\partial \xi^{k-1}}\xi_1^k\,.
\end{aligned} \tag{3.50}$$

These equations enable us to determinate the remaining coefficients $F^{(k)i}$ and ξ_1^k of the vector field X. Observe that if the Hessian matrix of L with respect to the highest-order "velocities", $q^{(k)i}$ and ξ^{k-1}, is invertible, that is,

$$\det\begin{pmatrix} \dfrac{\partial^2 L}{\partial q^{(k)j}\partial q^{(k)i}} & \dfrac{\partial^2 L}{\partial q^{(k)i}\partial \xi^{k-1}} \\ \dfrac{\partial^2 L}{\partial \xi^{k-1}\partial q^{(k)i}} & \dfrac{\partial^2 L}{\partial \xi^{k-1}\partial \xi^{k-1}} \end{pmatrix}(p) \neq 0\,, \quad \text{for every } p \in T^kM \times G \times k\mathfrak{g}\,,$$

then the previous system of equations has an unique solution for $F^{(k)i}$ and ξ_1^k, obtaining a unique vector field $X \in \mathfrak{X}(W)$ solution to the equation (3.42). In particular, the constraint algorithm finishes at the first step. Otherwise, new constraints may arise from equations (3.50), and the algorithm continues if necessary.

Now, let $\gamma\colon \mathbb{R} \to W$ be an integral curve of X locally given by

$$\gamma(t) = (q^{(r)i}(t), q^{(k)i}(t), p_i^{(r)}(t), g(t), \xi^i(t), \alpha_i(t))\,. \tag{3.51}$$

γ must be satisfy

$$\dot{q}^{(r)i} = q^{(r+1)i}\,, \tag{3.52}$$

$$\dot{p}^0_i = \frac{\partial L}{\partial q^{(0)i}} \quad , \quad \dot{p}^{(r)}_i = \frac{\partial L}{\partial q^{(r)i}} - p^{(r-1)}_i\,, \tag{3.53}$$

$$\dot{g} = g\xi^0 \quad , \quad \dot{\xi}^{i-1} = \xi^i\,, \tag{3.54}$$

$$\dot{\alpha}_0 = \pounds^*_g \frac{\partial L}{\partial g} + ad^*_{\xi^0_1}\alpha_0 \quad , \quad \dot{\alpha}_{i+1} = \frac{\partial L}{\partial \xi^i} - \alpha_i\,, \tag{3.55}$$

in addition to the restrictions (3.46) and (3.49). Now, using equations (3.46) in combination with equations (3.53) we obtain the kth-order Euler-Lagrange equations

$$\sum_{r=0}^{k} (-1)^r \frac{d^r}{dt^r}\frac{\partial L}{\partial q^{(r)i}} = 0\,. \tag{3.56}$$

On the other hand, using equations (3.49) in combination with equations (3.55) we obtain the kth-order trivialized Euler-Lagrange equations

$$\left(\frac{d}{dt} - ad^*_{\xi^0}\right)\sum_{i=0}^{k-1}(-1)^i \frac{d^i}{dt^i}\frac{\partial L}{\partial \xi^i} = \pounds^*_g\left(\frac{\partial L}{\partial g}\right)\,. \tag{3.57}$$

Therefore, a dynamical trajectory $\gamma\colon \mathbb{R} \to W$ of the system must satisfy the following local equations

$$\sum_{r=0}^{k} (-1)^r \frac{d^r}{dt^r}\frac{\partial L}{\partial q^{(r)i}} = 0 \quad , \quad \left(\frac{d}{dt} - ad^*_{\xi^0}\right)\sum_{i=0}^{k-1}(-1)^i \frac{d^i}{dt^i}\frac{\partial L}{\partial \xi^i} = \pounds^*_g\left(\frac{\partial L}{\partial g}\right)\,.$$

Finally, if the Lagrangian function $L \in C^\infty(T^{(k)}M \times G \times k\mathfrak{g})$ is left-invariant, that is,

$$L(q^{(r)i}, q^{(k)i}, g, \xi^i) = L(q^A_i, q^A_k, h, \xi^i)\,,$$

for all $g, h \in G$, then we can define the reduced Lagrangian $\ell \in C^\infty(T^{(k)}M \times k\mathfrak{g})$ by

$$\ell(q^{(r)i}, q^{(k)i}, \xi^i) = L(q^A_i, q^A_k, e, \xi^i)\,,$$

and therefore equations (3.57) become the kth order Euler-Poincaré equations

$$\left(\frac{d}{dt} - ad^*_{\xi^0}\right)\sum_{i=0}^{k-1}(-1)^i \frac{d}{dt^i}\frac{\partial \ell}{\partial \xi^i} = 0\,. \tag{3.58}$$

Observe that equations (3.56) remain the same with the reduced Lagrangian function, just replacing L by ℓ.

3.4.2 Constrained problem

Now, as in the previous sections, we assume that the dynamic of the system is constrained. Let $i_{\mathcal{M}} \colon \mathcal{M} \hookrightarrow T^{(k)}Q$ be the constraint submanifold, with $\operatorname{codim} \mathcal{M} = m$, and $L_{\mathcal{M}} \in C^\infty(\mathcal{M})$ the Lagrangian function describing the dynamics of the constrained dynamical system.

Consider the submanifold $\overline{W} = \mathcal{M} \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$ of $T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$ with canonical inclusion $i_{\overline{W}} \colon \overline{W} \hookrightarrow T^{(k)}Q \times_{T^{(k-1)}Q} T^*(T^{(k-1)}Q)$ and natural projection $\mathrm{pr}_{\mathcal{M}} \colon \overline{W} \to \mathcal{M}$. If we take $Q = M \times G$, then $\overline{W} = \mathcal{M} \times_{T^{(k-1)}M} T^*(T^{(k-1)}M) \times k\mathfrak{g}^*$.

Now, using the results given in Section 3.4.1, we can define a closed 2-form in $\overline{W}$ as $\overline{\Omega} = i_{\overline{W}}^* \Omega$, where Ω is the presymplectic form defined in (3.39), and a Hamiltonian function $\overline{H}$ in $\overline{W}$ from the canonical pairing and $\mathrm{pr}_{\mathcal{M}}^* L_{\mathcal{M}}$. With these elements we can state the dynamical equation for the constrained problem, which is

$$i_X \overline{\Omega} = \mathrm{d}\overline{H} \,. \tag{3.59}$$

We adopt an "extrinsic point of view", that is, we will work in the bundle W, and then require the solutions to lie in the submanifold $\overline{W} \hookrightarrow W$.

In order to do this, we must construct a Hamiltonian function $H \in C^\infty(W)$ using the Lagrangian function $L_{\mathcal{M}} \in C^\infty(\mathcal{M})$ containing the dynamical information of the system. Hence, let $L \in C^\infty(T^{(k)}(M \times G))$ be an arbitrary extension of $L_{\mathcal{M}}$, and let H be the Hamiltonian function defined in (3.41) using this arbitrary extension of the Lagrangian function $L_{\mathcal{M}}$.

The equations of motion for the constrained dynamical system are determined by

$$i_X \Omega - \mathrm{d}H \in \operatorname{ann}(T\overline{W}) \ , \quad \text{for } X \in \mathfrak{X}(W) \text{ tangent to } \overline{W} \,. \tag{3.60}$$

Let $\Phi^A \in C^\infty(T^{(k)}(M \times G))$, $1 \leqslant A \leqslant m$, be local functions defining the submanifold $\mathcal{M} \hookrightarrow T^{(k)}(M \times G)$. With some abuse of notation, we also denote by Φ^A the pull-back of the constraint functions to W. Then, the annihilator of $T\overline{W}$ is locally given by $\operatorname{ann}(T\overline{W}) = \left\langle \mathrm{d}\Phi^A \right\rangle$. Therefore, the equation defining the submanifold W_1 may be written locally as $i_Y \, \mathrm{d}H = \lambda_A \mathrm{d}\Phi^A$, $\forall \, Y \in \ker \Omega$, where λ_A, $1 \leqslant A \leqslant m$ are the Lagrange multipliers. The equations defining locally the submanifold W_1 are

$$p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}} + \lambda_A \frac{\partial \Phi^A}{\partial q^{(k)i}} = 0 \quad ; \quad \alpha_{k-1} - \frac{\partial L}{\partial \xi^{k-1}} + \lambda_A \frac{\partial \Phi^A}{\partial \xi^{k-1}} = 0 \quad ; \quad \Phi^A = 0 \,.$$

Now, let us compute the local expression of equation (3.60). If we assume that $\mathcal{M}$ is determined by the vanishing of the m functions Φ^A, then equation (3.60) may be rewritten as $i_X \, \Omega - \mathrm{d}H = \lambda_A \mathrm{d}\Phi^A$, where λ_A are Lagrange multipliers to be determined. Then, bearing in mind the local expression of $\mathrm{d}H$ and Ω, taking a generic vector field locally given by (3.43) we obtain the following system of equations

$$F^{(r)i} = q^{(r+1)i}, \tag{3.61}$$

$$G_i^{(0)} = \frac{\partial L}{\partial q^{(0)i}} - \lambda_A \frac{\partial \Phi^A}{\partial q^{(0)i}} \quad , \quad G_i^{(r)} = \frac{\partial L}{\partial q^{(r)i}} - \lambda_A \frac{\partial \Phi^A}{\partial q^{(r)i}} - p_i^{(r-1)} , \tag{3.62}$$

$$p_i^{(k-1)} - \frac{\partial L}{\partial q^{(k)i}} + \lambda_A \frac{\partial \Phi^A}{\partial q^{(k)i}} = 0 , \tag{3.63}$$

$$\xi_1^i = \xi^i , \tag{3.64}$$

$$\nu_0^1 = \pounds_g^* \left(\frac{\partial L}{\partial g} - \lambda_A \frac{\partial \Phi^A}{\partial g} \right) + ad^*_{\xi_1^0} \alpha_0 \quad , \quad \nu_{i+1}^1 = \frac{\partial L}{\partial \xi^i} - \lambda_A \frac{\partial \Phi^A}{\partial \xi^i} - \alpha_i , \tag{3.65}$$

$$\alpha_{k-1} - \frac{\partial L}{\partial \xi^{k-1}} + \lambda_A \frac{\partial \Phi^A}{\partial \xi^{k-1}} = 0 , \tag{3.66}$$

$$\Phi^A(q^{(r)i}, q^{(k)i}, g, \boldsymbol{\xi}, \xi^{k-1}) = 0 . \tag{3.67}$$

If we denote by Ω_{W_1} the pullback of the presymplectic 2-form Ω to W_1, then we deduce the following theorem.

Theorem 3.4.1. (W_1, Ω_{W_1}) *is a symplectic manifold if and only if the bilinear form defined as*

$$\begin{pmatrix} \dfrac{\partial^2 L}{\partial q^{(k)j} \partial q^{(k)i}} - \lambda_A \dfrac{\partial^2 \Phi^A}{\partial q^{(k)j} \partial q^{(k)i}} & \dfrac{\partial^2 L}{\partial q^{(k)i} \partial \xi^{k-1}} - \lambda_A \dfrac{\partial^2 \Phi^A}{\partial q^{(k)i} \partial \xi^{k-1}} & -\left(\dfrac{\partial \Phi^A}{\partial q^{(k)i}} \right)^T \\ \dfrac{\partial^2 L}{\partial \xi^{k-1} \partial q^{(k)i}} - \lambda_A \dfrac{\partial^2 \Phi^A}{\partial \xi^{k-1} \partial q^{(k)i}} & \dfrac{\partial^2 L}{\partial \xi^{k-1} \partial \xi^{k-1}} - \lambda_A \dfrac{\partial^2 \Phi^A}{\partial \xi^{k-1} \partial \xi^{k-1}} & -\left(\dfrac{\partial \Phi^A}{\partial \xi^{k-1}} \right)^T \\ \dfrac{\partial \Phi^A}{\partial q^{(k)i}} & \dfrac{\partial \Phi^A}{\partial \xi^{k-1}} & \mathbf{0} \end{pmatrix}$$

is nondegenerate along W_1.

Now, let $\gamma \colon \mathbb{R} \to W$ be an integral curve of X locally given by (3.51). Then the condition $X \circ \gamma = \dot{\gamma}$ gives the following system of differential equations for the component functions of γ

$$\dot{q}^{(r)i} = q^{(r+1)i}, \tag{3.68}$$

$$\dot{p}_i^{(0)} = \frac{\partial L}{\partial q^{(0)i}} - \lambda_A \frac{\partial \Phi^A}{\partial q^{(0)i}} \quad , \quad \dot{p}_i^{(r)} = \frac{\partial L}{\partial q^{(r)i}} - \lambda_A \frac{\partial \Phi^A}{\partial q^{(r)i}} - p_i^{(r-1)}, \tag{3.69}$$

$$\dot{g} = g\xi^0 \quad , \quad \dot{\xi}^{i-1} = \xi^i , \tag{3.70}$$

$$\dot{\alpha}_0 = \pounds_g^* \left(\frac{\partial L}{\partial g} - \lambda_A \frac{\partial \Phi^A}{\partial g} \right) + ad^*_{\xi_1^0} \alpha_0 \quad , \quad \dot{\alpha}_{i+1} = \frac{\partial L}{\partial \xi^i} - \lambda_A \frac{\partial \Phi^A}{\partial \xi^i} - \alpha_i , \tag{3.71}$$

in addition to equations (3.63), (3.66) and (3.67). Now, using equations (3.63) in combination with (3.69) we obtain the kth order constrained Euler-Lagrange equations

$$\sum_{r=0}^{k} (-1)^r \frac{d^r}{dt^r} \left(\frac{\partial L}{\partial q^{(r)i}} - \lambda_A \frac{\partial \Phi^A}{\partial q^{(r)i}} \right) = 0 . \tag{3.72}$$

On the other hand, using equations (3.66) in combination with (3.69) we obtain the kth order trivialized constrained Euler-Lagrange equation

$$\left(\frac{d}{dt} - ad^*_{\xi^0}\right)\sum_{i=0}^{k-1}(-1)^i\frac{d^i}{dt^i}\left(\frac{\partial L}{\partial \xi^i} - \lambda_A\frac{\partial \Phi^A}{\partial \xi^i}\right) = \pounds_g^*\left(\frac{\partial L}{\partial g} - \lambda_A\frac{\partial \Phi^A}{\partial g}\right). \tag{3.73}$$

Therefore, a dynamical trajectory $\gamma\colon \mathbb{R} \to W$ of the system must satisfy the equations (3.72) and (3.73), in addition to $\Phi^A(q^{(r)i}(t), q^{(k)i}(t), g(t), \xi^i(t)) = 0$.

Finally, if both the extended Lagrangian function $L \in C^\infty(T^{(k)}M \times G \times k\mathfrak{g})$ and the constraint functions $\Phi^A \in C^\infty(T^{(k)}M \times G \times k\mathfrak{g})$ are left-invariant, then we can define the reduced Lagrangian function $\ell \in C^\infty(T^{(k)}M \times k\mathfrak{g})$ and the reduced constraint functions $\phi^A \in C^\infty(T^{(k)}M \times k\mathfrak{g})$ as

$$\ell(q^{(r)i}, q^{(k)i}, \xi^i) = L(q^{(r)i}, q^{(k)i}, e, \xi^i) \quad , \quad \phi^A(q^{(r)i}, q^{(k)i}, \xi^i) = \Phi^A(q^{(r)i}, q^{(k)i}, e, \xi^i)\,,$$

and then equations (3.73) become

$$\left(\frac{d}{dt} - ad^*_{\xi^0}\right)\sum_{i=0}^{k-1}(-1)^i\frac{d^i}{dt^i}\left(\frac{\partial \ell}{\partial \xi^i} - \lambda_A\frac{\partial \phi^A}{\partial \xi^i}\right) = 0\,.$$

Note that equations (3.72) remain the same, just replacing L by ℓ and Φ^A by ϕ^A.

3.5 Extension to Lie algebroids

In this section we will develop a geometrical description for second-order mechanics on Lie algebroids in the Skinner and Rusk formalism, given a general geometric framework for the previous results in this chapter and using strongly the results given in [87].

First, we will review the description of vakonomics mechanics on Lie algebroids given by Iglesias, Marrero, Martín de Diego and Sosa in [87]. After it we will introduce the notion of admissible elements on a Lie algebroid and we will particularize the previous construction to the case when the Lie algebroid is the prolongation of a Lie algebroid and the constraint submanifold is the set of admissible elements. Then we will obtain the second-order Skinner and Rusk formulation on Lie algebroids.

3.5.1 Vakonomic mechanics on Lie algebroids

Let $\tau_{\widetilde{E}} : \widetilde{E} \to Q$ be a Lie algebroid of rank n over a manifold Q of dimension m with anchor map $\rho : \widetilde{E} \to TQ$ and $L : \widetilde{E} \to \mathbb{R}$ be a Lagrangian function on $\widetilde{E}$. Moreover, let $\mathcal{M} \subset \widetilde{E}$ be an embedded submanifold of dimension $n + m - \bar{m}$ such that $\tau_{\mathcal{M}} = \tau_{\widetilde{E}}\big|_{\mathcal{M}} : \mathcal{M} \to Q$ is a surjective submersion.

Suppose that e is a point of $\mathcal{M}$ with $\tau_{\mathcal{M}}(e) = x \in Q$, (x^i) are local coordinates on an open subset U of Q, $x \in U$, and $\{e_A\}$ is a local basis of $\Gamma(\widetilde{E})$ on U. Denote by (x^i, y^A) the corresponding local coordinates for $\widetilde{E}$ on the open subset $\tau_{\widetilde{E}}^{-1}(U)$. Assume that

$$\mathcal{M} \cap \tau_{\widetilde{E}}^{-1}(U) \equiv \{(x^i, y^A) \in \tau_{\widetilde{E}}^{-1}(U) \,|\, \Phi^\alpha(x^i, y^A) = 0,\ \alpha = 1, \dots, \bar{m}\}$$

where Φ^α are the local independent constraint functions for the submanifold $\mathcal{M}$.

We will suppose, without loss of generality, that the $(\bar{m} \times n)$-matrix

$$\left(\frac{\partial \Phi^\alpha}{\partial y^B}\Big|_e \right)_{\alpha=1,\ldots,\bar{m};B=1,\ldots,n}$$

is of maximal rank.

Now, using the implicit function theorem, we obtain that there exists an open subset $\widetilde{V}$ of $(\tau_{\widetilde{E}})^{-1}(U)$, an open subset $W \subseteq \mathbb{R}^{m+n-\bar{m}}$ and smooth real functions $\Psi^\alpha : W \to \mathbb{R}$, $\alpha = 1, \ldots, \bar{m}$, such that

$$\mathcal{M} \cap \widetilde{V} \equiv \{(x^i, y^A) \in \widetilde{V} \,|\, y^\alpha = \Psi^\alpha(x^i, y^a), \text{ with } \alpha = 1, \ldots, \bar{m} \text{ and } \bar{m}+1 \leq a \leq n\}.$$

Consequently, (x^i, y^a) are local coordinates on $\mathcal{M}$ and we will denote by $\tilde{L}$ the restriction of L to $\mathcal{M}$.

Consider the Whitney sum of $\widetilde{E}^*$ and $\widetilde{E}$, that is, $W = \widetilde{E} \oplus \widetilde{E}^*$, and the canonical projections $pr_1 : \widetilde{E} \oplus \widetilde{E}^* \longrightarrow \widetilde{E}$ and $pr_2 : \widetilde{E} \oplus \widetilde{E}^* \longrightarrow \widetilde{E}^*$. Now, let W_0 be the submanifold $W_0 = pr_1^{-1}(\mathcal{M}) = \mathcal{M} \times_Q \widetilde{E}^*$ and the restrictions $\pi_1 = pr_1|_{W_0}$ and $\pi_2 = pr_2|_{W_0}$. Also denote by $\nu : W_0 \longrightarrow Q$ the canonical projection of W_0 over the base manifold.

Next, we consider the prolongation of the Lie algebroid $\widetilde{E}$ over $\tau_{\widetilde{E}^*} : \widetilde{E}^* \to Q$ (respectively, $\nu : W_0 \to Q$). We will denote this Lie algebroid by $\mathcal{T}^{\tau_{\widetilde{E}^*}}\widetilde{E}$ (respectively, $\mathcal{T}^\nu \widetilde{E}$). Moreover, we can prolong $\pi_2 : W_0 \to \widetilde{E}^*$ to a morphism of Lie algebroids $\mathcal{T}\pi_2 : \mathcal{T}^\nu \widetilde{E} \to \mathcal{T}^{\tau_{\widetilde{E}^*}}\widetilde{E}$ defined by $\mathcal{T}\pi_2 = (Id, T\pi_2)$.

If (x^i, p_A) are the local coordinates on $\widetilde{E}^*$ associated with the local basis $\{e^A\}$ of $\Gamma(\widetilde{E}^*)$, then (x^i, p_A, y^a) are local coordinates on W_0 and we may consider the local basis $\{\widetilde{e}_A^{(1)}, (\widetilde{e}^A)^{(2)}, e_a^{(2)}\}$ of $\Gamma(\mathcal{T}^\nu \widetilde{E})$ defined by

$$\begin{aligned}
\widetilde{e}_A^{(1)}(\check{e}, e^*) &= \left(e_A(x), \rho_A^i \frac{\partial}{\partial x^i}\Big|_{(\check{e},e^*)} \right), \\
(\widetilde{e}^A)^{(2)}(\check{e}, e^*) &= \left(0, \frac{\partial}{\partial p_A}\Big|_{(\check{e},e^*)} \right), \\
e_a^{(2)}(\check{e}, e^*) &= \left(0, \frac{\partial}{\partial y^a}\Big|_{(\check{e},e^*)} \right),
\end{aligned}$$

where $(\check{e}, e^*) \in W_0$ and $\nu(\check{e}, e^*) = x$. If $([\![\cdot,\cdot]\!]^\nu, \rho^\nu)$ is the Lie algebroid structure on $\mathcal{T}^\nu \widetilde{E}$, we have that

$$[\![\widetilde{e}_A^{(1)}, \widetilde{e}_B^{(1)}]\!]^\nu = \mathcal{C}_{AB}^C \widetilde{e}_C^{(1)},$$

and the rest of the fundamental Lie brackets are zero. Moreover,

$$\rho^\nu(\widetilde{e}_A^{(1)}) = \rho_A^i \frac{\partial}{\partial x^i}, \quad \rho^\nu((\widetilde{e}^A)^{(2)}) = \frac{\partial}{\partial p_A}, \quad \rho^\nu(e_a^{(2)}) = \frac{\partial}{\partial y^a}.$$

The Pontryagin Hamiltonian H_{W_0} is a function defined on $W_0 = \mathcal{M} \times_Q \widetilde{E}^*$ given by

$$H_{W_0}(\check{e}, e^*) = \langle e^*, \check{e} \rangle - \tilde{L}(\check{e}),$$

or, in local coordinates,

$$H_{W_0}(x^i, p_A, y^a) = p_a y^a + p_\alpha \Psi^\alpha(x^i, y^a) - \tilde{L}(x^i, y^a)\,. \tag{3.74}$$

Moreover, one can consider the presymplectic 2-section $\Omega_0 = (\mathcal{T}\pi_2, \pi_2)^*\Omega_{\widetilde{E}}$, where $\Omega_{\widetilde{E}}$ is the canonical symplectic section on $\mathcal{T}^{\tau_{\widetilde{E}^*}}\widetilde{E}$ defined in Equation (1.19). In local coordinates,

$$\Omega_0 = \widetilde{e}^A_{(1)} \wedge \widetilde{e}^{(2)}_A + \frac{1}{2}\mathcal{C}^C_{AB} p_C \widetilde{e}^A_{(1)} \wedge \widetilde{e}^B_{(1)}, \tag{3.75}$$

where $\{\widetilde{e}^A_{(1)}, \widetilde{e}^{(2)}_A, e^a_{(2)}\}$ denotes the dual basis of $\{\widetilde{e}^{(1)}_A, (\widetilde{e}^A)^{(2)}, e^{(2)}_a\}$.

Therefore, we have the triple $(\mathcal{T}^\nu\widetilde{E}, \Omega_0, d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0})$ as a presymplectic hamiltonian system.

Definition 3.5.1. *The vakonomic problem on Lie algebroids consists on finding the solutions for the equation*

$$i_X\Omega_0 = d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0}; \tag{3.76}$$

that is, to solve the constraint algorithm for $(\mathcal{T}^\nu\widetilde{E}, \Omega_0, d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0})$.

In local coordinates, we have that

$$d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0} = \left(p_\alpha\frac{\partial\Psi^\alpha}{\partial x^i} - \frac{\partial\tilde{L}}{\partial x^i}\right)\rho^i_A\widetilde{e}^A_{(1)} + \Psi^\alpha\widetilde{e}^{(2)}_\alpha + y^a\widetilde{e}^{(2)}_a + \left(p_a + p_\alpha\frac{\partial\Psi^\alpha}{\partial y^a} - \frac{\partial\tilde{L}}{\partial y^a}\right)e^a_{(2)}.$$

If we apply the constraint algorithm,

$$W_1 = \{w \in \mathcal{M}\times_Q \widetilde{E}^* \,|\, d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0}(w)(Y) = 0, \quad \forall Y \in \ker\ \Omega_0(w)\}.$$

Since $\ker\Omega_0 = \operatorname{span}\{e^{(2)}_a\}$, we get that W_1 is locally characterized by the equations

$$\varphi_a = d^{\mathcal{T}^\nu\widetilde{E}}H_{W_0}(e^{(2)}_a) = p_a + p_\alpha\frac{\partial\Psi^\alpha}{\partial y^a} - \frac{\partial\tilde{L}}{\partial y^a} = 0,$$

or

$$p_a = \frac{\partial\tilde{L}}{\partial y^a} - p_\alpha\frac{\partial\Psi^\alpha}{\partial y^a}, \quad \bar{m}+1 \leq a \leq n.$$

Let us also look for the expression of X satisfying Eq. (3.76). A direct computation shows that

$$X = y^a\widetilde{e}^{(1)}_a + \Psi^\alpha\widetilde{e}^{(1)}_\alpha + \Big[\Big(\frac{\partial\tilde{L}}{\partial x^i} - p_\alpha\frac{\partial\Psi^\alpha}{\partial x^i}\Big)\rho^i_A - y^a\mathcal{C}^B_{Aa}p_B - \Psi^\alpha\mathcal{C}^B_{A\alpha}p_B\Big](\widetilde{e}^A)^{(2)} + \Upsilon^a e^{(2)}_a.$$

Therefore, the vakonomic equations are

$$\begin{cases} \dot{x}^i = y^a\rho^i_a + \Psi^\alpha\rho^i_\alpha, \\ \dot{p}_\alpha = \Big(\dfrac{\partial\tilde{L}}{\partial x^i} - p_\beta\dfrac{\partial\Psi^\beta}{\partial x^i}\Big)\rho^i_\alpha - y^a\mathcal{C}^B_{\alpha a}p_B - \Psi^\beta\mathcal{C}^B_{\alpha\beta}p_B, \\ \dfrac{d}{dt}\left(\dfrac{\partial\tilde{L}}{\partial y^a} - \rho_\alpha\dfrac{\partial\Psi^\alpha}{\partial y^a}\right) = \Big(\dfrac{\partial\tilde{L}}{\partial x^i} - p_\alpha\dfrac{\partial\Psi^\alpha}{\partial x^i}\Big)\rho^i_a - y^b\mathcal{C}^B_{ab}p_B - \Psi^\alpha\mathcal{C}^B_{a\alpha}p_B. \end{cases}$$

Of course, we know that there exist sections X of $\mathcal{T}^{\nu}\widetilde{E}$ along W_1 satisfying (3.76), but they may not be sections of $(\rho^{\nu})^{-1}(TW_1) = \mathcal{T}^{\nu_1}\widetilde{E}$, in general (here $\nu_1 : W_1 \to Q$). Then, following the procedure detailed in Section 2.2.2, we obtain a sequence of embedded submanifolds

$$\ldots \hookrightarrow W_{k+1} \hookrightarrow W_k \hookrightarrow \ldots \hookrightarrow W_2 \hookrightarrow W_1 \hookrightarrow W_0 = \mathcal{M} \times_Q \widetilde{E}^*.$$

If the algorithm stabilizes, then we find a final constraint submanifold W_f on which at least a section $X \in \Gamma(\mathcal{T}^{\nu_f}E)$ verifies

$$(i_X\Omega_0 = d^{\mathcal{T}^{\nu}\widetilde{E}} H_{W_0})\big|_{W_f}$$

where $\nu_f : W_f \to Q$.

One of the most important cases is when $W_f = W_1$. The authors of [87] have analyzed this case with the following result: Consider the restriction Ω_1 of Ω_0 to $\mathcal{T}^{\nu_1}\widetilde{E}$;

Proposition 3.5.2. *Ω_1 is a symplectic section of the Lie algebroid $\mathcal{T}^{\nu_1}\widetilde{E}$ if and only if for any system of coordinates (x^i, p_A, y^a) on W_0 we have that*

$$\det\left(\frac{\partial^2 \tilde{L}}{\partial y^a \partial y^b} - p_\alpha \frac{\partial^2 \Psi^\alpha}{\partial y^a \partial y^b}\right) \neq 0, \text{ for all point in } W_1.$$

3.5.2 Prolongation of a Lie algebroid over a smooth map (cont'd)

This subsection is devoted to study some additional properties and characterizations about the prolongation of a Lie algebroid over a smooth map (see subsection 1.8.3).

Let $\widetilde{E}$ be a Lie algebroid over Q with fiber bundle projection $\tau_{\widetilde{E}} : \widetilde{E} \to Q$ and anchor map $\rho : \widetilde{E} \to TQ$.

Definition 3.5.3. *A tangent vector v at the point $e \in \widetilde{E}$ is called admissible if $\rho(e) = T_e\tau_{\widetilde{E}}(v)$; and a curve on $\widetilde{E}$ is admissible if its tangent vectors are admissible. The set of admissible elements on $\widetilde{E}$ will be denote $E^{(2)}$.*

Notice that v is admissible if and only if (e, e, v) is in $\mathcal{T}^{\tau_{\widetilde{E}}}\widetilde{E}$. We will consider $E^{(2)}$ as the subset of the prolongation of $\widetilde{E}$ over $\tau_{\widetilde{E}}$, that is, $E^{(2)} \subset \widetilde{E}_\rho \times_{T\tau_{\widetilde{E}}} T\widetilde{E}$ is given by

$$E^{(2)} = \{(e, v_e) \in \widetilde{E} \times T\widetilde{E} \mid \rho(e) = T\tau_{\widetilde{E}}(v_e)\}.$$

Other authors call this set $\mathrm{Adm}(\widetilde{E})$ (see [43] and [134]).

We consider $E^{(2)}$ as the substitute of the second order tangent bundle in classical mechanics. If (x^i) are local coordinates on Q and $\{e_A\}$ is a basis of sections of $\widetilde{E}$ then we denote (x^i, y^A) the corresponding local coordinates on $\widetilde{E}$ and $(x^i, y^A; z^A, v^A)$ local coordinates on $\mathcal{T}^{\tau_{\widetilde{E}}}\widetilde{E}$ induced by the basis of sections $\{e_A^{(1)}, e_A^{(2)}\}$ of $\mathcal{T}^{\tau_{\widetilde{E}}}\widetilde{E}$ (see subsection 1.8.3). Therefore, the set $E^{(2)}$ is locally characterized by the condition $\{(x^i, y^A; z^A, v^A) \in \mathcal{T}^{\tau_{\widetilde{E}}}E \mid y^A = z^A\}$, that is $(x^i, y^A, v^A) := (x^i, y^A, \dot{y}^A)$ are local coordinates on $E^{(2)}$.

We denote the canonical inclusion of $E^{(2)}$ on the prolongation of $\widetilde{E}$ over $\tau_{\widetilde{E}}$ as

$$\begin{array}{rcl} i_{E^{(2)}} : E^{(2)} & \hookrightarrow & \mathcal{T}^{\tau_{\widetilde{E}}}\widetilde{E}, \\ (x^i, y^A, \dot{y}^A) & \mapsto & (x^i, y^A, y^A, \dot{y}^A). \end{array}$$

Example 3.5.4. Let M be a differentiable manifold of dimension m, if (x^i) are local coordinates on M, then $\{\frac{\partial}{\partial x^i}\}$ is a local basis of $\mathfrak{X}(M)$ and then we have fiber local coordinates $(x^i, \dot{x}^i)$ on TM. The corresponding local structure functions of the Lie algebroid $\tau_{TM} : TM \to M$ are

$$\mathcal{C}_{ij}^k = 0 \text{ and } \rho_i^j = \delta_i^j, \text{ for } i, j, k \in \{1, \ldots, m\}.$$

In this case, we have seen that the prolongation Lie algebroid over τ_{TM} is just the tangent bundle $T(TM)$ where the Lie algebroid structure of the vector bundle $T(TM) \to TM$ is as we have described above as the tangent bundle of a manifold.

The set of admissible elements is given by

$$E^{(2)} = \{(x^i, v^i, \dot{x}^i, w^i) \in T(TM) \mid \dot{x}^i = v^i\}$$

and observe that this subset is just the second-order tangent bundle of a manifold M, that is, $E^{(2)} = T^{(2)}M$.

Now, let $\tau_E : E \to M$ be a Lie algebroid with anchor map $\rho : E \to TM$ and let $\mathcal{T}^{\tau_E}E$ be the $E-$tangent bundle to E. Now we will define the bundle $\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E)$ over $\mathcal{T}^{\tau_E}E$. This bundle plays the role of $\tau_{T(TM)} : T(TTM) \to T(TM)$ in ordinary Lagrangian Mechanics.

In what follows we will describe the Lie algebroid structure of the E-tangent bundle to the prolongation Lie algebroid over $\tau_E : E \to Q$.

As we know from subsection (1.8.3), the basis of sections $\{e_A\}$ of E induces a local basis of the sections of $\mathcal{T}^{\tau_E}E$ given by

$$e_A^{(1)}(e) = \left(e, e_A(\tau_E(e)), \rho_A^i \frac{\partial}{\partial x^i}\Big|_e\right), \quad e_A^{(2)}(e) = \left(e, 0, \frac{\partial}{\partial y^A}\Big|_e\right),$$

for $e \in E$. From this basis we can induce local coordinates $(x^i, y^A; z^A, v^A)$ on $\mathcal{T}^{\tau_E}E$. Now, from this basis, we can induce a local basis of sections of $\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E)$ in the following way: consider an element $(e, v_b) \in \mathcal{T}^{\tau_E}E$, then define the components of the basis $\{e_A^{(1,1)}, e_A^{(2,1)}, e_A^{(1,2)}, e_A^{(2,2)}\}$ as

$$\begin{aligned}
e_A^{(1,1)}(e, v_b) &= \left((e, v_b), e_A^{(1)}(e), \rho_A^i \frac{\partial}{\partial x^i}\Big|_{(e,v_b)}\right),\\
e_A^{(2,1)}(e, v_b) &= \left((e, v_b), e_A^{(2)}(e), \frac{\partial}{\partial y^A}\Big|_{(e,v_b)}\right),\\
e_A^{(1,2)}(e, v_b) &= \left((e, v_b), 0, \frac{\partial}{\partial z^A}\Big|_{(e,v_b)}\right),\\
e_A^{(2,2)}(e, v_b) &= \left((e, v_b), 0, \frac{\partial}{\partial v^A}\Big|_{(e,v_b)}\right).
\end{aligned}$$

The basis $\{e_A^{(1,1)}, e_A^{(2,1)}, e_A^{(1,2)}, e_A^{(2,2)}\}$ induces local coordinates $(x^i, y^A, z^A, v^A, b^A, c^A, d^A, w^A)$ on $\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E)$. If we denote by $(\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E), [\![\cdot,\cdot]\!]_{\tau_E^{(2)}}, \rho_2)$ the Lie algebroid structure of $\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E)$, it is characterized by

$$
\begin{aligned}
\rho_2(e_A^{(1,1)})(e,v_b) &= \left((e,v_b),\rho_A^i\frac{\partial}{\partial x^i}\Big|_{(e,v_b)}\right),\\
\rho_2(e_A^{(2,1)})(e,v_b) &= \left((e,v_b),\frac{\partial}{\partial y^A}\Big|_{(e,v_b)}\right),\\
\rho_2(e_A^{(1,2)})(e,v_b) &= \left((e,v_b),\frac{\partial}{\partial z^A}\Big|_{(e,v_b)}\right),\\
\rho_2(e_A^{(2,2)})(e,v_b) &= \left((e,v_b),\frac{\partial}{\partial v^A}\Big|_{(e,v_b)}\right),\\
[\![e_A^{(1,1)},e_B^{(1,1)}]\!]_{\tau_E^{(2)}} &= \mathcal{C}_{AB}^C e_C^{(1,1)},\\
[\![e_A^{(1,1)},e_B^{(1,2)}]\!]_{\tau_E^{(2)}} &= [\![e_A^{(1,2)},e_B^{(1,2)}]\!]_{\tau_E^{(2)}}=0,\\
[\![e_A^{(1,1)},e_A^{(2,2)}]\!]_{\tau_E^{(2)}} &= [\![e_A^{(2,1)},e_A^{(2,2)}]\!]_{\tau_E^{(2)}}=[\![e_A^{(1,2)},e_A^{(2,1)}]\!]_{\tau_E^{(2,1)}}=[\![e_A^{(1,1)},e_B^{(2,1)}]\!]_{\tau_E^{(2)}}=0.
\end{aligned}
$$

for all A,B and C where $\mathcal{C}_{AB}^C$ are the structure constants of E.

In the same way, from the basis $\{\widetilde{e}_A^{(1)},(\widetilde{e}^A)^{(2)}\}$ of sections of $\mathcal{T}^{\tau_{E^*}}E$ given by

$$
\begin{aligned}
\widetilde{e}_A^{(1)}(e^*) &= \left(e^*,e_A(\tau_{E^*}(e^*)),\rho_A^i\frac{\partial}{\partial x^i}\Big|_{e^*}\right),\\
(\widetilde{e}^A)^{(2)}(e^*) &= \left(e^*,0,\frac{\partial}{\partial p_A}\Big|_{e^*}\right),
\end{aligned}
$$

where $e^*\in E$, we construct the basis of sections of $\mathcal{T}^{\tau_{(\mathcal{T}^{\tau_E}E)^*}}\mathcal{T}^{\tau_E}E$, denoted $\{\widetilde{e}_A^{(1,1)},(\widetilde{e}^A)^{(2,1)},\widetilde{e}_A^{(1,2)},(\widetilde{e}^A)^{(2,2)}\}$. In what follows $(x^i,y^A,p_A,\bar{p}_A)$ denotes local coordinates on $\mathcal{T}^{\tau_{E^*}}E$ induced by the basis $\{\widetilde{e}_A^{(1)},(\widetilde{e}^A)^{(2)}\}$.

This basis is given by

$$
\begin{aligned}
\widetilde{e}_A^{(1,1)}(\alpha^*) &= \left(\alpha^*,e_A^{(1)}(\tau_{(\mathcal{T}^{\tau_E}E)^*}(\alpha^*)),\rho_A^i\frac{\partial}{\partial x^i}\Big|_{\alpha^*}\right),\\
(\widetilde{e}_A)^{(2,1)}(\alpha^*) &= \left(\alpha^*,e_A^{(2)}(\tau_{(\mathcal{T}^{\tau_E}E)^*}(\alpha^*)),\frac{\partial}{\partial y^A}\Big|_{\alpha^*}\right),\\
(\widetilde{e}^A)^{(1,2)}(\alpha^*) &= \left(\alpha^*,0,\frac{\partial}{\partial p_A}\Big|_{\alpha^*}\right),\\
(\widetilde{e}^A)^{(2,2)}(\alpha^*) &= \left(\alpha^*,0,\frac{\partial}{\partial \bar{p}_A}\Big|_{\alpha^*}\right).
\end{aligned}
$$

where $\alpha^*\in(\mathcal{T}^{\tau_E}E)^*$ and $\tau_{(\mathcal{T}^{\tau_E}E)^*}:(\mathcal{T}^{\tau_E}E)^*\to E$ is the vector bundle projection.

The Lie algebroid structure $(\mathcal{T}^{\tau_{(\mathcal{T}^{\tau_E}E)^*}}(\mathcal{T}^{\tau_E}E);[\![\cdot,\cdot]\!]_2,\rho_2)$ is given by

$$
\begin{aligned}
\rho_2(\widetilde{e}_A^{(1,1)}(\alpha^*)) &= \left(\alpha^*,\rho_A^i\frac{\partial}{\partial x^i}\Big|_{\alpha^*}\right),\quad \rho_2((\widetilde{e}^A)^{(2,1)}(\alpha^*))=\left(\alpha^*,\frac{\partial}{\partial y^A}\Big|_{\alpha^*}\right),\\
\rho_2(\widetilde{e}_A^{(1,2)}(\alpha^*)) &= \left(\alpha^*,\frac{\partial}{\partial p_A}\Big|_{\alpha^*}\right),\quad \rho_2((\widetilde{e}^A)^{(2,2)}(\alpha^*))=\left(\alpha^*,\frac{\partial}{\partial \bar{p}_A}\Big|_{\alpha^*}\right),
\end{aligned}
$$

where the unique non-zero Lie bracket is $[\![\widetilde{e}_A^{(1,1)}, \widetilde{e}_B^{(1,1)}]\!]_2 = \mathcal{C}_{AB}^{C}\widetilde{e}_C^{(1,1)}$. This basis induces local coordinates $(x^i, y^A, p_A, \bar{p}_A, q^A, \bar{q}^A; l_A, \bar{l}_A)$ on $\mathcal{T}^{\tau_{(\mathcal{T}^{\tau_E}E)^*}}\mathcal{T}^{\tau_E}E$.

3.5.3 Second-order unconstrained problem on Lie algebroids

Consider the Whitney sum of $(\mathcal{T}^{\tau_E}E)^*$ and $\mathcal{T}^{\tau_E}E$, $W = \mathcal{T}^{\tau_E}E \times_E (\mathcal{T}^{\tau_E}E)^*$ and its canonical projections $pr_1 : W \to \mathcal{T}^{\tau_E}E$ and $pr_2 : W \to (\mathcal{T}^{\tau_E}E)^*$. Now, let W_0 be the submanifold $W_0 = pr_1^{-1}(E^{(2)}) = E^{(2)} \times_E (\mathcal{T}^{\tau_E}E)^*$ and the restrictions $\pi_1 = pr_1 \mid_{W_0}$ and $\pi_2 = pr_2 \mid_{W_0}$. Also we denote by $\nu : W_0 \to E$ the canonical projection. The diagram in Figure 3.2 illustrates the situation.

$$\begin{tikzcd}
 & W_0 = E^{(2)} \times_E (\mathcal{T}^{\tau_{E^*}}E)^* \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] \arrow[dd, "\nu"] & \\
E^{(2)} \arrow[dr, "\tau_E^{(2,1)}"'] & & (\mathcal{T}^{\tau_E}E)^* \arrow[dl, "\tau_{(\mathcal{T}^{\tau_E}E)^*}"] \\
 & E &
\end{tikzcd}$$

Figure 3.2: Second order Skinner and Rusk formalism on Lie algebroids

Consider the prolongations of $\mathcal{T}^{\tau_E}E$ by $\tau_{(\mathcal{T}^{\tau_E}E)^*}$ and by ν, respectively. We will denote these Lie algebroids by $\mathcal{T}^{\tau_{(\mathcal{T}^{\tau_E}E)^*}}(\mathcal{T}^{\tau_E}E)$ and $\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E$ respectively. Moreover, we can prolong $\pi_2 : W_0 \to (\mathcal{T}^{\tau_E}E)^*$ to a morphism of Lie algebroids $\mathcal{T}\pi_2 : \mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E \to \mathcal{T}^{\tau_{(\mathcal{T}^{\tau_E}E)^*}}(\mathcal{T}^{\tau_{E^*}}E)$ defined by $\mathcal{T}\pi_2 = (Id, T\pi_2)$.

We denote by $(x^i, y^A, p_A, \bar{p}_A)$ local coordinates on $(\mathcal{T}^{\tau_E}E)^*$ induced by $\{e^A_{(1)}, e^A_{(2)}\}$, the dual basis of the basis $\{e_A^{(1)}, e_A^{(2)}\}$, a basis of $\mathcal{T}^{\tau_E}E$. Then, $(x^i, y^A, p_A, \bar{p}_A, z^A)$ are local coordinates in W_0 and we may consider the local basis $\{\widetilde{e}_A^{(1,1)}, \widetilde{e}_A^{(2,1)}, (\widetilde{e}^A)^{(1,2)}, (\widetilde{e}^A)^{(2,2)}, \check{e}_A^{(1,2)}\}$ of $\Gamma(\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E)$ defined by

$$\begin{aligned}
\widetilde{e}_A^{(1,1)}(\breve{\alpha}, \alpha^*) &= \left((\breve{\alpha}, \alpha^*), e_A^{(1)}(\tau_{(\mathcal{T}^{\tau_E}E)^*}(\alpha^*)), \rho_A^i \frac{\partial}{\partial x^i}\Big|_{(\breve{\alpha}, \alpha^*)}\right), \\
\widetilde{e}_A^{(2,1)}(\breve{\alpha}, \alpha^*) &= \left((\breve{\alpha}, \alpha^*), e_A^{(2)}(\tau_{(\mathcal{T}^{\tau_E}E)^*}(\alpha^*)), \frac{\partial}{\partial y^A}\Big|_{(\breve{\alpha}, \alpha^*)}\right), \\
(\widetilde{e}^A)^{(1,2)}(\breve{\alpha}, \alpha^*) &= \left((\breve{\alpha}, \alpha^*), 0, \frac{\partial}{\partial p_A}\Big|_{(\breve{\alpha}, \alpha^*)}\right), \\
(\widetilde{e}^A)^{(2,2)}(\breve{\alpha}, \alpha^*) &= \left((\breve{\alpha}, \alpha^*), 0, \frac{\partial}{\partial \bar{p}_A}\Big|_{(\breve{\alpha}, \alpha^*)}\right), \\
\check{e}_A^{(1,2)}(\breve{\alpha}, \alpha^*) &= \left((\breve{\alpha}, \alpha^*), 0, \frac{\partial}{\partial z^A}\Big|_{(\breve{\alpha}, \alpha^*)}\right)
\end{aligned}$$

for $\alpha^* \in (\mathcal{T}^{\tau_E}E)^*$, $\breve{\alpha} \in E^{(2)}$, $(\breve{\alpha}, \alpha^*) \in W_0$, and $\tau_{(\mathcal{T}^{\tau_E}E)^*} : (\mathcal{T}^{\tau_E}E)^* \to E$ is the canonical projection.

If $(\llbracket\cdot,\cdot\rrbracket^{\nu},\rho^{\nu})$ is the Lie algebroid structure on $\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E$, we have that

$$\llbracket\widetilde{e}_A^{(1,1)},\widetilde{e}_B^{(1,1)}\rrbracket^{\nu}=\mathcal{C}_{AB}^{C}\widetilde{e}_C^{(1,1)},$$

and the rest of the fundamental Lie brackets are zero. Moreover,

$$\begin{aligned}
\rho^{\nu}(\widetilde{e}_A^{(1,1)}(\breve{\alpha},\alpha^*)) &= \left((\breve{\alpha},\alpha^*),\rho^i_A\frac{\partial}{\partial x^i}\Big|_{(\breve{\alpha},\alpha^*)}\right), \quad \rho^{\nu}(\widetilde{e}_A^{(2,1)}(\breve{\alpha},\alpha^*))=\left((\breve{\alpha},\alpha^*),\frac{\partial}{\partial y^A}\Big|_{(\breve{\alpha},\alpha^*)}\right),\\
\rho^{\nu}((\widetilde{e}^A)^{(1,2)}(\breve{\alpha},\alpha^*)) &= \left((\breve{\alpha},\alpha^*),\frac{\partial}{\partial p_A}\Big|_{(\breve{\alpha},\alpha^*)}\right), \quad \rho^{\nu}((\widetilde{e}^A)^{(2,2)}(\breve{\alpha},\alpha^*))=\left((\breve{\alpha},\alpha^*),\frac{\partial}{\partial \bar{p}_A}\Big|_{(\breve{\alpha},\alpha^*)}\right),\\
\rho^{\nu}(\breve{e}_A^{(1,2)}(\breve{\alpha},\alpha^*)) &= \left((\breve{\alpha},\alpha^*),\frac{\partial}{\partial z^A}\Big|_{(\breve{\alpha},\alpha^*)}\right).
\end{aligned}$$

The Pontryagin Hamiltonian H_{W_0} is a function in W_0 given by

$$H_{W_0}(\breve{\alpha},\alpha^*)=\langle\alpha^*,\breve{\alpha}\rangle-L(\breve{\alpha}),$$

or in local coordinates

$$H_{W_0}(x^i,y^A,p_A,\bar{p}_A,z^A)=\bar{p}_Az^A+p_Ay^A-L(x^i,y^A,z^A).$$

Moreover, one can consider the presymplectic 2-section $\Omega_0=(\mathcal{T}\pi_2,\pi_2)^*\Omega_E$, where Ω_E is the canonical symplectic section on $\mathcal{T}^{\tau_{E^*}}E$. In local coordinates,

$$\Omega_0=\widetilde{e}^A_{(1,1)}\wedge(\widetilde{e}_A)^{(1,2)}+\widetilde{e}^A_{(2,1)}\wedge(\widetilde{e}_A)^{(2,2)}+\frac{1}{2}\widetilde{\mathcal{C}}^C_{AB}p_C\widetilde{e}^A_{(1,1)}\wedge\widetilde{e}^B_{(1,1)}.$$

Here $\{\widetilde{e}^A_{(1,1)},\widetilde{e}^A_{(2,1)},(\widetilde{e}_A)^{(1,2)},(\widetilde{e}_A)^{(2,2)},\breve{e}^A_{(1,2)}\}$ denotes the dual basis of $\{\widetilde{e}_A^{(1,1)},\widetilde{e}_A^{(2,1)},(\widetilde{e}^A)^{(1,2)},(\widetilde{e}^A)^{(2,2)},\breve{e}_A^{(1,2)}\}$

Therefore, we have the triple $\left(\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E,\Omega_0,d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E}H_{W_0}\right)$ as a presymplectic Hamiltonian system.

The second-order problem on the Lie algebroid $\tau_E:E\to M$ consists on finding the solutions of the equation

$$i_X\Omega_0=d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E}H_{W_0},$$

that is, to solve the constraint algorithm for $\left(\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E,\Omega_0,d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E}H_{W_0}\right)$.

In local coordinates we have

$$d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E}H_{W_0}=-\rho^i_A\frac{\partial L}{\partial x^i}\widetilde{e}^A_{(1,1)}+\left(p_A-\frac{\partial L}{\partial y^A}\right)\widetilde{e}^A_{(2,1)}+\left(\bar{p}_A-\frac{\partial L}{\partial z^A}\right)\breve{e}^A_{(2,1)}+z^A(\widetilde{e}_A)^{(2,2)}+y^A(\widetilde{e}_A)^{(1,2)}.$$

If we apply the constraint algorithm, since $\ker\Omega_0=\operatorname{span}\{\breve{e}_A^{(2,1)}\}$ the first constraint submanifold W_1 is locally characterized by the equation

$$\varphi_A=d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E}H_{W_0}(\breve{e}_A^{(2,1)})=\bar{p}_A-\frac{\partial L}{\partial z^A}=0,$$

or

$$\bar{p}_A = \frac{\partial L}{\partial z^A}.$$

Looking for the expression of X satisfying the equation for the second-order problem we have that the second-order equations are

$$\begin{aligned}
\dot{x}^i &= \rho^i_A y^A, \\
\dot{p}_A &= \rho^i_A \frac{\partial L}{\partial x^i} + \mathcal{C}^C_{AB} p_C y^B, \\
\dot{\bar{p}}_A &= -p_A + \frac{\partial L}{\partial y^A}, \\
\bar{p}_A &= \frac{\partial L}{\partial z^A}.
\end{aligned}$$

After some straightforward computations the last equations are equivalent to the following equations:

$$0 = \frac{d^2}{dt^2}\frac{\partial L}{\partial z^A} + \mathcal{C}^C_{AB} y^B \frac{d}{dt}\left(\frac{\partial L}{\partial z^A}\right) - \frac{d}{dt}\frac{\partial L}{\partial y^A} - \mathcal{C}^C_{AB} y^B \left(\frac{\partial L}{\partial y^A}\right) + \rho^i_A \frac{\partial L}{\partial x^i}. \tag{3.77}$$

As in the previous section it is possible to apply the constraint algorithm (2.2.2) 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 $i_X\Omega_0 = d^{\mathcal{T}^{\nu}\mathcal{T}^{\tau_E}E} H_{W_0}$.

Observe that the first constraint submanifold W_1 is determined by the conditions

$$\varphi_A = \bar{p}_A - \frac{\partial L}{\partial z^A} = 0.$$

If we denote by Ω_{W_1} the pullback of the presymplectic 2-section Ω_{W_0} to W_1, then we deduce the result which is the same than the theorem given in [87] explained in section 3.5.1 to the case when the $M = E^{(2)}$.

Proposition 3.5.5. *Ω_{W_1} is a symplectic section of the Lie algebroid $\mathcal{T}^{\nu_1}\mathcal{T}^{\tau_E}E$ if and only if*

$$\left(\frac{\partial^2 L}{\partial z^A \partial z^B}\right)$$

is nondegenerate along W_1, where $\nu_1 = \nu\,|_{W_1} \colon W_1 \to E$.

Remark 3.5.6. Observe that we can particularize the equations (3.77) to the case of Atiyah algebroids to obtain the second-order Lagrange-Poincaré equations.

Let G be a Lie group and we assume that G acts free and properly on M. We denote by $\pi : M \to \widehat{M} = M/G$ the associated principal bundle. The tangent lift of the action gives a free and proper action of G on TM and $\widehat{TM} = TM/G$ is a quotient manifold. Then we consider the Atiyah algebroid $\widehat{TM}$ over $\widehat{M}$.

According to example 1.8.6, the basis $\{\hat{e}_i, \hat{e}_B\}$ induce local coordinates $(x^i, y^i, \bar{y}^B)$. From this basis one can induces a basis of the prolongation Lie algebroid, namely $\{\hat{e}^{(1)}_i, \hat{e}^{(1)}_B\}$. This basis induce adapted coordinates $(x^i, y^i, \bar{y}^B, \dot{y}^i, \dot{\bar{y}}^B)$ on $\widehat{T^{(2)}M} = (T^{(2)}M)/G$.

Given a Lagrangian function $\ell : \widehat{T^{(2)}M} \to \mathbb{R}$ over the set of admissible elements of the Atiyah algebroid $\widehat{TTM} \to \widehat{TM}$, where $\widehat{TTM} = (TTM)/G$, the Euler-Lagrange equations for ℓ are

$$\begin{aligned}\frac{\partial \ell}{\partial x^j} - \frac{d}{dt}\left(\frac{\partial \ell}{\partial y^j}\right) + \frac{d^2}{dt^2}\left(\frac{\partial \ell}{\partial \dot{y}^j}\right) &= \left(\frac{d}{dt}\left(\frac{\partial \ell}{\partial \dot{\bar{y}}^A}\right) - \frac{\partial \ell}{\partial \bar{y}^A}\right)\left(\mathcal{B}^A_{ij} y^i + c^A_{DB}\mathcal{A}^B_j \bar{y}^B\right) \quad \forall j,\\ \frac{d^2}{dt^2}\left(\frac{\partial \ell}{\partial \dot{\bar{y}}^B}\right) - \frac{d}{dt}\left(\frac{\partial \ell}{\partial \bar{y}^B}\right) &= \left(\frac{d}{dt}\left(\frac{\partial \ell}{\partial \dot{\bar{y}}^A}\right) - \frac{\partial \ell}{\partial \bar{y}^A}\right)\left(C^A_{DB}\bar{y}^D - c^A_{DB}\mathcal{A}^D_i y^i\right) \quad \forall B,\end{aligned}$$

which are the second-order Lagrange-Poincaré equations associated to a G-invariant Lagrangian $L : T^{(2)}M \to \mathbb{R}$ (see [70] and for example) where c^C_{AB} are the structure constants of the Lie algebra according to 1.8.6. ◇

3.5.4 Second-order constrained problem on Lie algebroids

Now, we will consider second-order mechanical systems subject to second-order constraints. Let $\mathcal{M} \subset E^{(2)}$ be an embedded submanifold of dimension $n + m - \bar{m}$ (locally determined by the vanishing of the constraint functions $\Phi^\alpha : \mathcal{M} \to \mathbb{R}$, $\alpha = 1, \ldots, m$) such that the bundle projection $\tau^{(2,1)}_E |_{\mathcal{M}}: \mathcal{M} \to E$ is a surjective submersion.

We will suppose that the $(\bar{m} \times n)-$matrix $\left(\frac{\partial \Phi^\alpha}{\partial z^B}\right)$ with $\alpha = 1, \ldots, \bar{m}$ and $B = 1, \ldots, n$ is of maximal rank. Then, we will use the following notation $z^A = (z^\alpha, z^a)$ for $1 \leq A \leq n$, $1 \leq \alpha \leq \bar{m}$ and $\bar{m} + 1 \leq a \leq n$. Therefore, using the implicit function theorem we can write

$$z^\alpha = \Psi^\alpha(x^i, y^A, z^a).$$

Consequently we can consider local coordinates on $\mathcal{M}$ by (x^i, y^A, z^a) and we will denote by $\widetilde{L}$ the restriction of L to $\mathcal{M}$.

Proposition 3.5.7 ([118]). *Let $(E, [\![\ ,\ ]\!], \rho)$ be a Lie algebroid over a manifold M with projection $\tau_E : E \to M$ and anchor map with constant rank. Consider a submanifold N of M. If $\tau_E\big|_{\rho^{-1}(TN)} : \rho^{-1}(TN) \to M$ is a vector subbundle, then $\rho^{-1}(TN)$ is a Lie algebroid over N.*

Let us take the submanifold $\overline{W}_0 = pr_1^{-1}(\mathcal{M}) = \mathcal{M} \times_E (\mathcal{T}^{\tau_E}E)^*$ and the restrictions of $\overline{W}_0$ of the canonical projections π_1 and π_2 given by $\pi_1 = pr_1 |_{\overline{W}_0}$ and $\pi_2 = pr_1 |_{\overline{W}_0}$. We will denote local coordinates on $\overline{W}_0$ by $(x^i, y^A, p_A, \bar{p}_A, z^a)$.

Therefore, proceeding as in the unconstrained case one can construct the presymplectic Hamiltonian system $(\overline{W}_0, \Omega_{\overline{W}_0}, H_{\overline{W}_0})$, where $\Omega_{\overline{W}_0}$ is the presymplectic 2-section on $\overline{W}_0$ and the Hamiltonian function $H : \overline{W}_0 \to \mathbb{R}$ is locally given by

$$H_{\overline{W}_0}(x^i, y^A, p_A, \bar{p}_A, z^a) = p_A y^A + \bar{p}_a z^a + \bar{p}_\alpha \Psi^\alpha(x^i, y^A, z^a) - \widetilde{L}(x^i, y^A, z^a).$$

With these two elements it is possible to write the following presymplectic system

$$i_X \Omega_{\overline{W}_0} = d^{(\rho^\nu)^{-1}(T\overline{W}_0)} H_{\overline{W}_0}, \tag{3.78}$$

where $(\rho^\nu)^{-1}(TW_0)$ denotes the Lie subalgebroid of $\mathcal{T}^\nu\mathcal{T}^{\tau_E}E$ over $\overline{W}_0 \subset W_0$.

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 $L : E^{(2)} \to \mathbb{R}$ of $L_{\mathcal{M}} : \mathcal{M} \to \mathbb{R}$. The main idea is to take into account that Equation (3.78) is equivalent to

$$\left\{\begin{array}{rcl} i_X\Omega_{W_0} - d^{\mathcal{T}^\nu\mathcal{T}^{\tau_E}E}H & \in & \text{ann}\,(\rho^\nu)^{-1}(T_x\overline{W}_0)\,, \\ X & \in & (\rho^\nu)^{-1}(T_x\overline{W}_0)\ \text{ and } x \in \overline{W}_0, \end{array}\right.$$

where $H : W_0 \to \mathbb{R}$ is the function defined in the last section and ann denotes the set of sections $\widetilde{X} \in \Gamma((\mathcal{T}^\nu\mathcal{T}^{\tau_E}E)^*)$ such that $\langle \widetilde{X}, Y\rangle = 0$ for all $Y \in (\rho^\nu)^{-1}(TW_0)$.

Assuming that $\mathcal{M}$ is determined by the vanishing of $\overline{m}$-independent constraints

$$\Phi^\alpha(x^i, y^A, z^a) = 0,\ 1 \leq \alpha \leq \overline{m}\,,$$

then, locally, ann $(\rho^\nu)^{-1}(T\overline{W}_0) = \text{span}\,\{d^{\mathcal{T}^\nu\mathcal{T}^{\tau_E}E}\Phi^\alpha\}$, and therefore the previous equations are rewritten as

$$\left\{\begin{array}{lcl} i_X\Omega_{W_0} - d^{\mathcal{T}^\nu\mathcal{T}^{\tau_E}E}H & = & \lambda_\alpha d^{\mathcal{T}^\nu\mathcal{T}^{\tau_E}E}\Phi^\alpha\,, \\ X(x) \in (\rho^\nu)^{-1}(T_x\overline{W}_0) & & \text{for all } x \in \overline{W}_0\,, \end{array}\right.$$

where λ_α are Lagrange multipliers to be determined.

Proceeding as in the last section one can obtain the following system of equations for $\widetilde{L} = L + \lambda_\alpha\Phi^\alpha$

$$\begin{array}{rcl} 0 & = & \dfrac{d^2}{dt^2}\dfrac{\partial\widetilde{L}}{\partial z^A} + \mathcal{C}^C_{AB}y^B\dfrac{d}{dt}\left(\dfrac{\partial\widetilde{L}}{\partial z^A}\right) - \dfrac{d}{dt}\dfrac{\partial\widetilde{L}}{\partial y^A} - \mathcal{C}^C_{AB}y^B\left(\dfrac{\partial\widetilde{L}}{\partial y^A}\right) + \rho^i_A\dfrac{\partial\widetilde{L}}{\partial x^i} \\ 0 & = & \Phi^\alpha(x^i, y^A, z^A). \end{array} \tag{3.79}$$

Here the first constraint submanifold $\overline{W}_1$ is determined by the condition

$$\begin{array}{rcl} 0 & = & \bar{p}_A - \dfrac{\partial L}{\partial z^A} + \lambda_\alpha\dfrac{\partial\Phi^\alpha}{\partial z^A} \\ 0 & = & \Phi^\alpha(x^i, y^A, z^A). \end{array}$$

If we denote by $\Omega_{\overline{W}_1}$ the pullback of the presymplectic section $\Omega_{\overline{W}_0}$ to $\overline{W}_1$, then we can deduce that $\Omega_{\overline{W}_1}$ is a symplectic section if and only if

$$\begin{pmatrix} \dfrac{\partial^2 L}{\partial z^A\partial z^B} + \lambda_\alpha\dfrac{\partial^2\Phi^\alpha}{\partial z^A\partial z^B} & \dfrac{\partial\Phi^\alpha}{\partial z^A} \\ \dfrac{\partial\Phi^\alpha}{\partial z^B} & \mathbf{0} \end{pmatrix} \tag{3.80}$$

is nondegenerate.

3.6 Optimal control of underactuated mechanical systems

This section is devoted to the so-called underactuated mechanical control 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 [8, 162]).

After introducing the Skinner and Rusk formalism for higher-order mechanical systems with higher-order constraints in the last sections, we may turn to the geometric framework for optimal control of underactuated mechanical systems.

Definition 3.6.1. *A control system is called underactuated if the number of control inputs is less than the dimension of the configuration space.*

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. Underactuated systems include spacecrafts, underwater vehicles, mobile robots, helicopters, wheeled vehicles, underactuated manipulators...

There are many papers in which optimal control problems are addressed using geometric techniques (see, for instance, [18, 90, 91, 165] and references therein). In this section we introduce an optimization strategy in an underactuated mechanical system, that is, we are interested in 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 mechanical systems strongly inspired on the Skinner and Rusk formulation. Since the controlled Euler-Lagrange equations are a set of second-order ordinary differential equations we will need to implement the higher-order version of this classical Skinner and Rusk formalism. 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 such as the symplecticity and the preservation of the hamiltonian.

3.6.1 Optimal control of mechanical systems defined on TQ

Let Q be the configuration manifold, where (q^A) are local coordinates with $A = 1, ..., n$. We consider a mechanical system described by a regular Lagrangian $L : TQ \to \mathbb{R}$. The induced local coordinates on TQ are just $(q^A, \dot{q}^A)$. Additionally, there are *control parameters* in our picture. To define these *control parameters* we introduce the *control manifold* $U \subset \mathbb{R}^l$ $(l \leq n)$ for a given interval $I = [0, T]$. The *control path space* is defined by

$$\mathcal{P}(U) = \mathcal{P}([0, T], U) = \{u : [0, T] \to U \,|\, u \in L^\infty\},$$

where $u(t) \in U$ is the *control parameter* (see [146]).

Also, we interpret a *control force* as a parameter-dependent force, that is a parameter-dependent fiber-preserving map $f(u) : TQ \to T^*Q$ over the identity Id_Q, which can be written in local coordinates as

$$f(u) : (q^A, \dot{q}^A) \mapsto (q^A, f(u)(q^A, \dot{q}^A)).$$

We shall assume that all the control systems in this work are controllable, that is, for any two points q_0 and q_T in the configuration space Q, there exists a control parameter $u(t)$ defined on some interval $[0, T] \subset \mathbb{R}$ such that the system with initial condition q_0 reaches the point q_T in time T (see [34] for details).

Consider initially the class of underactuated Lagrangian control system (superarticulated mechanical system following the nomenclature by [8]) where the configuration space Q is the cartesian product of two differentiable manifolds, $Q = Q_1 \times Q_2$. Denote by $(q^A) = (q^a, q^\alpha)$, $1 \leq A \leq n$, local coordinates on Q where (q^a), $1 \leq a \leq r$ and (q^α), $r + 1 \leq \alpha \leq n$, are local coordinates on Q_1 and Q_2, respectively.

Given a Lagrangian $L : TQ \equiv TQ_1 \times TQ_2 \to \mathbb{R}$, we assume that the controlled external forces can be applied only to the coordinates (q^a). Thus, the equations of motion are given by

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a} &= u^a, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} &= 0\,, \end{aligned} \tag{3.81}$$

where $a = 1, \ldots, r$, and $\alpha = r + 1, \ldots, n$.

We study the optimal control problem that consists on finding a trajectory $(q^A(t), u^a(t))$ of state variables and control inputs satisfying equations (6.29) from given initial and final conditions, $(q^A(t_0), \dot{q}^A(t_0))$, $(q^A(t_f), \dot{q}^A(t_f))$ respectively, minimizing the cost functional

$$\mathcal{A} = \int_{t_0}^{t_f} C(q^A, \dot{q}^A, u^a)\, dt,$$

where $C : T^{(2)}Q \times U \to \mathbb{R}$ is the *cost function* (continuously differentiable).

This optimal control problem is equivalent to the following constrained variational problem:

Extremize

$$\widetilde{\mathcal{A}} = \int_{t_0}^{t_f} \widetilde{L}(q^A(t), \dot{q}^A(t), \ddot{q}^A(t))\, dt$$

subject to the second order constraints given by

$$\Phi^\alpha(q^A(t), \dot{q}^A(t), \ddot{q}^A(t)) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} = 0\,,$$

and the boundary conditions, where $\widetilde{L} : T^{(2)}Q \to \mathbb{R}$ is defined as

$$\widetilde{L}(q^A(t), \dot{q}^A(t), \ddot{q}^A(t)) = C\left(q^A, \dot{q}^A, \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a}\right).$$

Now, according to the formulation given in Section 3.2.1, 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 3.2.1 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 Φ^α, we will see how to define local coordinates on $\mathcal{M}$.

From the constraint equations we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} = 0 \Longleftrightarrow \frac{\partial^2 L}{\partial \dot q^\beta \partial \dot q^\alpha}\ddot q^\beta = F_\alpha(q^A, \dot q^A, \ddot q^a).$$

Let us assume that the matrix $(W_{\alpha\beta}) = \left(\frac{\partial^2 L}{\partial \dot q^\alpha \partial \dot q^\beta}\right)$ is non-singular and denote by $(W^{\alpha\beta})$ its inverse. Thus,

$$\ddot q^\alpha = W^{\alpha\beta} F_\alpha(q^A, \dot q^A, \ddot q^a) = G^\alpha(q^A, \dot q^A, \ddot q^a).$$

Therefore, we can consider $(q^A, \dot q^A, \ddot q^a)$ as a system of local coordinates on $\mathcal{M}$. The canonical inclusion $i_{\mathcal{M}} : \mathcal{M} \hookrightarrow TTQ$ can be written as

$$\begin{array}{ccc} \mathcal{M} & \overset{i_{\mathcal{M}}}{\to} & TTQ \\ (q^A, \dot q^A, \ddot q^a) & \mapsto & (q^A, \dot q^A, \ddot q^a, G^\alpha(q^A, \dot q^A, \ddot q^a)) \,. \end{array}$$

Define the restricted lagrangian $\widetilde{L}_{\mathcal{M}} : \mathcal{M} \to \mathbb{R}$.

$$\begin{tikzcd}
 & W_0 = \mathcal{M} \times_{TQ} T^*(TQ) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] \arrow[dd, "\pi_{W_0}"] & \\
\mathcal{M} \arrow[dr, "(\tau_{TQ})|_{\mathcal{M}}"] & & T^*TQ \arrow[dl, "\pi_{T^*Q}"] \\
 & TQ &
\end{tikzcd}$$

Figure 3.3: Second order Skinner and Rusk formalism

We will consider $W_0 = \mathcal{M} \times_{TQ} T^*(TQ)$ whose coordinates are $(q^A, \dot q^A; p_A^0, p_A^1, \ddot q^a)$.

Let us define the 2-form $\Omega_{W_0} = \pi_1^*(\omega_{TQ})$ on W_0 and $H_{W_0}(\alpha_x, v_x) = \langle \alpha_x, i_{\mathcal{M}}(v_x)\rangle - \widetilde{L}_{\mathcal{M}}(v_x)$ where $x \in TQ$, $v_x \in \mathcal{M}_x = ((\tau_Q^{(1,2)})|_{\mathcal{M}})^{-1}(x)$ and $\alpha_x \in T_x^*TQ$. In local coordinates,

$$\begin{array}{rcl} \Omega_{W_0} & = & dq^A \wedge dp_A^0 + d\dot q^A \wedge dp_A^1, \\ H_{W_0} & = & p_A^0 \dot q^A + p_a^1 \ddot q^a + p_\alpha^1 G^\alpha(q^A, \dot q^A, \ddot q^a) - \widetilde{L}_{\mathcal{M}}(q^A, \dot q^A, \ddot q^a). \end{array}$$

The dynamics of this variational constrained problem is determined by the solution of the equation

$$i_X \Omega_{W_0} = dH_{W_0}. \tag{3.82}$$

(see for example [10], [15],[36], [155] and [173] for example).

It is clear that Ω_{W_0} is a presymplectic form on W_0 and locally

$$\ker \Omega_{W_0} = \text{span}\left\langle \frac{\partial}{\partial \ddot q^a} \right\rangle.$$

Following the constraints algorithm (2.2.2) we obtain the primary constraints

$$dH_{W_0}\left(\frac{\partial}{\partial \ddot{q}^a}\right) = 0\,.$$

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 $4n$ with local coordinates $(q^A, \dot{q}^A, \ddot{q}^a, p_A^0, p_\alpha^1)$.

Consider a solution curve $(q^A(t), \dot{q}^A(t), \ddot{q}^a(t), p_A^0(t), p_A^1(t))$ of Equation (3.82). Then, this curve satisfies the following system of differential equations

$$\frac{dq^A}{dt} = \dot{q}^A\,, \qquad \frac{d^2q^a}{dt^2} = \ddot{q}^a\,, \tag{3.83}$$

$$\frac{d^2q^\alpha}{dt^2} = G^\alpha(q^A, \frac{dq^A}{dt}, \frac{d^2q^a}{dt^2})\,, \tag{3.84}$$

$$\frac{dp_A^0}{dt} = -p_\alpha^1 \frac{\partial G^\alpha}{\partial q^A} + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^A}\,, \tag{3.85}$$

$$\frac{dp_A^1}{dt} = -p_A^0 - p_\alpha^1 \frac{\partial G^\alpha}{\partial \dot{q}^A} + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^A}\,, \tag{3.86}$$

$$p_a^1 = -p_\alpha^1 \frac{\partial G^\alpha}{\partial \ddot{q}^a} + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a}\,. \tag{3.87}$$

From Equations (3.86) and (3.87) we deduce

$$\frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial \ddot{q}^a}\right) = -p_a^0 - p_\alpha^1 \frac{\partial G^\alpha}{\partial \dot{q}^a} + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^a}\,.$$

Differentiating with respect to time, replacing in the previous equality and using (3.85) we obtain the following system of 4-order differential equations

$$\frac{d^2}{dt^2}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial \ddot{q}^a}\right) - \frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial \dot{q}^a}\right) + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial q^a} = 0\,. \tag{3.88}$$

Also, using (3.85) and (3.86) we deduce

$$\frac{d^2 p_\alpha^1}{dt^2} = \frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^\alpha} - p_\beta^1 \frac{\partial G^\beta}{\partial \dot{q}^\alpha}\right) - \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^\alpha} - p_\beta^1 \frac{\partial G^\beta}{\partial q^\alpha}\right)\,. \tag{3.89}$$

If we solve the implicit system of differential equations given by (3.88) and (3.89) then from Equations (3.86) and (3.87) we deduce that the values of p_a^0 and p_α^0 are

$$p_a^0 = \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial \dot{q}^a} - \frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a} - p_\alpha^1 \frac{\partial G^\alpha}{\partial \ddot{q}^a}\right)\,, \tag{3.90}$$

$$p_\alpha^0 = \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^\alpha} - p_\beta^1 \frac{\partial G^\beta}{\partial \dot{q}^\alpha} - \frac{dp_\alpha^1}{dt}\,. \tag{3.91}$$

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 (3.88), (3.89) and (3.84) without explicitly calculate the values $p^0_a(t)$.

Now, we are interested in the geometric properties of the dynamics. First, consider the submanifold W_1 of W_0 determined by

$$W_1 = \{x \in \mathcal{M} \times_{TQ} T^*TQ \mid dH_{W_0}(x)(V) = 0 \ \forall\ V \in \ker \Omega(x)\}$$

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

$$\varphi^1_a = p^1_a + p^1_\alpha \frac{\partial G^\alpha}{\partial \ddot{q}^a} - \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a} = 0 \; .$$

Therefore, we can consider local coordinates $(q^A, \dot{q}^A, \ddot{q}^a, p^0_A, p^1_\alpha)$ on W_1.

Theorem 3.6.2. (W_1, Ω_{W_1}) *is symplectic if and only if for any choice of local coordinates* $(q^A, \dot{q}^A, \ddot{q}^a, p^0_A, p^1_A)$ *on* W_0*,*

$$\det(\mathcal{R}_{ab}) = \det \left(\frac{\partial^2 \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a \partial \ddot{q}^b} - p^1_\alpha \frac{\partial^2 G^\alpha}{\partial \ddot{q}^a \partial \ddot{q}^b} \right)_{(n-r)\times(n-r)} \neq 0 \quad \textit{along } W_1 \; . \tag{3.92}$$

Proof. Let us recall that Ω_{W_1} is symplectic if and only if $T_xW_1 \cap (T_xW_1)^\perp = 0 \quad \forall x \in W_1$, where

$$(T_xW_1)^\perp = \{v \in T_x(T^*TQ) \times_{TQ} \mathcal{M} \; / \; \Omega_{W_0}(x)(v, w) = 0, \text{for all } w \in T_xW_1\} \, .$$

Suppose that (W_1, Ω_{W_1}) is symplectic and that

$$\lambda^a \mathcal{R}_{ab}(x) = 0 \text{ for some } \lambda^a \in \mathbb{R} \text{ and } x \in W_1 \; .$$

Hence

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

Therefore, $\lambda^b \left. \frac{\partial}{\partial \ddot{q}^b} \right|_x \in T_xW_1$ but it is also in $T_xW_1^\perp$. This implies that $\lambda_b = 0$ for all b and that the matrix $(\mathcal{R}_{ab})$ is regular.

Now, suppose that the matrix $(\mathcal{R}_{ab})$ is regular. Since

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

then, $\left. \frac{\partial}{\partial \ddot{q}^b} \right|_x \notin T_xW_1$ and, in consequence,

$$T_xW_1 \oplus \text{span} \left\{ \left. \frac{\partial}{\partial \ddot{q}^b} \right|_x \right\} = T_xW_0.$$

Now, let $Z \in T_xW_1 \cap (T_xW_1)^{\perp}$ with $x \in W_1$. It follows that

$$0 = i_Z\Omega_{W_0}(x)\left(\left.\frac{\partial}{\partial \ddot{q}^a}\right|_x\right), \text{ for all } a \text{ and } i_Z\Omega_{W_0}(x)(\bar{Z}) = 0, \text{ for all } \bar{Z} \in T_xW_1 .$$

Then, $Z \in \ker \Omega_{W_0}(x)$. This implies that

$$Z = \lambda_b \left.\frac{\partial}{\partial \ddot{q}^b}\right|_x$$

Since $Z \in T_xW_1$ then

$$0 = d\varphi_a(x)(Z) = d\varphi_a(x)\left(\lambda_b \left.\frac{\partial}{\partial \ddot{q}^b}\right|_x\right) = \lambda_b \mathcal{R}_{ab}$$

and, consequently, $\lambda_b = 0$, for all b, and $Z = 0$.

□

In the case when the matrix (3.92) is regular then the equations (3.88), (3.89) and (3.84) can be written as an explicit system of differential equations of the form

$$\frac{d^4q^a}{dt^4} = \Gamma^a\left(q^A, \frac{dq^A}{dt}, \frac{d^2q^a}{dt^2}, \frac{d^3q^a}{dt^2}, p^1_\alpha, \frac{dp^1_\alpha}{dt}\right) \tag{3.93}$$

$$\frac{d^2q^\alpha}{dt^2} = G^\alpha(q^A, \frac{dq^A}{dt}, \frac{d^2q^a}{dt^2}) \tag{3.94}$$

$$\frac{d^2p^1_\alpha}{dt^2} = \frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{q}^\alpha} - p^1_\beta \frac{\partial G^\beta}{\partial \dot{q}^\alpha}\right) - \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial q^\alpha} - p^1_\beta \frac{\partial G^\beta}{\partial q^\alpha}\right) . \tag{3.95}$$

Remark 3.6.3. The Pontryagin's maximun principle [157] 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 (3.93), (3.94) and (3.95), 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:

$$i_X\Omega_{W_1} = dH_{W_1} .$$

This is the case of a regular optimal control problem [12]. From (3.93), (3.94) and (3.95) we obtain a unique curve $(q^A(t))$ (fixed appropriate initial conditions) which determine the controls from

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a} = u^a .$$

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 3.6.4. Now, we will analyze an alternative characterization of the condition (3.92) and its relationship with the matrix condition that appears in Theorem 3.2.1. Using the chain rule for

$$\widetilde{L}_{\mathcal{M}}(q^A, \dot{q}^A, \ddot{q}^a) = \widetilde{L}\left(q^A, \dot{q}^A, \ddot{q}^a, G^\alpha(q^A, \dot{q}^A, \ddot{q}^a)\right)$$

we have that

$$\begin{aligned}
\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a} &= \frac{\partial \widetilde{L}}{\partial \ddot{q}^a} + \frac{\partial \widetilde{L}}{\partial \ddot{q}^\alpha}\frac{\partial G^\alpha}{\partial \ddot{q}^a}\\
\frac{\partial^2 \widetilde{L}_{\mathcal{M}}}{\partial \ddot{q}^a \partial \ddot{q}^b} &= \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^a \partial \ddot{q}^b} + \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^a \partial \ddot{q}^\beta}\frac{\partial G^\beta}{\partial \ddot{q}^b} + \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^\alpha \partial \ddot{q}^b}\frac{\partial G^\alpha}{\partial \ddot{q}^a} + \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^\alpha \partial \ddot{q}^\beta}\frac{\partial G^\alpha}{\partial \ddot{q}^a}\frac{\partial G^\beta}{\partial \ddot{q}^b} + \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^\alpha}\frac{\partial^2 G^\alpha}{\partial \ddot{q}^a \partial \ddot{q}^b}\,.
\end{aligned}$$

Define $W_{ij} = \left(\frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^i \partial \ddot{q}^j}\right)$, where $\Phi^\alpha = \ddot{q}^\alpha - G^\alpha$. Then we can write (3.92) as

$$\mathcal{R}_{ab} = W_{ab} - W_{a\beta}\frac{\partial \Phi^\beta}{\partial \ddot{q}^b} - W_{\alpha b}\frac{\partial \Phi^\alpha}{\partial \ddot{q}^a} + W_{\alpha\beta}\frac{\partial \Phi^\alpha}{\partial \ddot{q}^a}\frac{\partial \Phi^\beta}{\partial \ddot{q}^b} + \left(p^1_\alpha - \frac{\partial \widetilde{L}}{\partial \dot{q}^\alpha}\right)\frac{\partial^2 \Phi^\alpha}{\partial \ddot{q}^a \partial \ddot{q}^b}\,.$$

Consider now the extended lagrangian $\mathcal{L} = \widetilde{L} - \lambda_\alpha \Phi^\alpha$ where $\lambda_\alpha = \frac{\partial \widetilde{L}}{\partial \dot{q}^\alpha} - p^1_\alpha$.

Then, the matrix $(\overline{W}_{ij}) = \left(\frac{\partial^2 \mathcal{L}}{\partial \ddot{q}^i \partial \ddot{q}^j}\right)$ is equal, along W_1, to

$$\overline{W}_{ij} = \begin{pmatrix} \overline{W}_{ab} & W_{a\beta} \\ W_{\alpha b} & W_{\alpha\beta} \end{pmatrix} \tag{3.96}$$

where $\overline{W}_{ab} = \frac{\partial^2 \widetilde{L}}{\partial \ddot{q}^a \partial \ddot{q}^b} - \lambda_\alpha \frac{\partial^2 \Phi^\alpha}{\partial \ddot{q}^a \partial \ddot{q}^b}$.

The elements of the matrix (3.92) are given by

$$\mathcal{R}_{ab} = \overline{W}_{ab} - \overline{W}_{a\beta}\frac{\partial \Phi^\beta}{\partial \ddot{q}^b} - \overline{W}_{\alpha b}\frac{\partial \Phi^\alpha}{\partial \ddot{q}^a} + \overline{W}_{\alpha\beta}\frac{\partial \Phi^\alpha}{\partial \ddot{q}^a}\frac{\partial \Phi^\beta}{\partial \ddot{q}^b}\,. \tag{3.97}$$

Now, using elemental linear algebra the matrix (3.97) is regular if and only if the matrix of elements (3.96) is regular.

◇

Remark 3.6.5. Condition (3.92) 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 most 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 system determined by the control equations $\ddot{x} = u_1$, $\ddot{y} = u_2$ and cost function $C(x, y, \dot{x}, \dot{y}, u_1, u_2) = \frac{1}{2}(u_1^2 + 2u_1u_2 + u_2^2)$. If we apply our techniques we deduce that W_1 is determined by the constraints

$$p^1_x - (\ddot{x} + \ddot{y}) = 0\,, \qquad p^1_y - (\ddot{x} + \ddot{y}) = 0\,.$$

But the solution of the dynamics is only consistently defined on the submanifold W_2 of W_1 determined by the additional (secondary) constraint

$$p_x^0 - p_y^0 = 0 \ .$$

$\diamond$

Example 3.6.6 (Optimal control of a cart with a pendulum)**.** A Cart-Pole System consists of a cart and an inverted pendulum on it (see [17] and references therein). The coordinate x denotes the position of the cart on the x-axis and θ 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

$$\mathbb{I} = \begin{pmatrix} M+m & ml\cos\theta \\ ml\cos\theta & ml^2 \end{pmatrix},$$

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 $V(\theta) = mgl\cos(\theta)$.

The Lagrangian of the system (kinetic energy minus potential energy) is given by

$$L(q,\dot{q}) = L(x,\theta,\dot{x},\dot{\theta}) = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x}^2 + 2\dot{x}l\dot{\theta}\cos\theta + l^2\dot{\theta}^2) - mgl\cos\theta - mg\widetilde{h} \ ,$$

where $\widetilde{h}$ is the car height.

The controller can apply a force u, the control input, parallel to the track remaining the joint angle θ unactuated. Therefore, the equations of motion of the controlled system are

$$\begin{array}{rcl} (M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + ml\ddot{\theta}\cos\theta & = & u \\ \ddot{x}\cos\theta + l\ddot{\theta} - g\sin\theta & = & 0 \end{array}$$

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

$$\mathcal{A} = \frac{1}{2}\int_0^T u^2 dt.$$

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

$$\widetilde{L}(x,\theta,\dot{x},\dot{\theta},\ddot{x},\ddot{\theta}) = \frac{1}{2}\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x}\right)^2 = \frac{1}{2}\left[(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + ml\ddot{\theta}\cos\theta\right]^2 .$$

We rewrite the second-order constraint as

$$\ddot{\theta} = \frac{g\sin\theta - \ddot{x}\cos\theta}{l} .$$

Thus, the submanifold $\mathcal{M}$ of $T^{(2)}(\mathbb{R}\times\mathbb{S}^1)$ is given by

$$\mathcal{M} = \left\{(x,\theta,\dot{x},\dot{\theta},\ddot{x},\ddot{\theta}) \mid \ddot{x}\cos\theta + l\ddot{\theta} - g\sin\theta = 0\right\} .$$

Let us consider the submanifold $W_0 = T^*(T(\mathbb{R}\times\mathbb{S}^1)) \times_{T(\mathbb{R}\times\mathbb{S}^1)} \mathcal{M}$ with induced coordinates $(x,\theta,\dot{x},\dot{\theta};p_x^0,p_\theta^0,p_x^1,p_\theta^1,\ddot{x})$.

Now, we consider the restriction of $\widetilde{L}$ to $\mathcal{M}$ given by

$$\begin{aligned}\widetilde{L}|_{\mathcal{M}} &= \frac{1}{2}\left[(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + ml\cos\theta(\frac{g\sin\theta - \ddot{x}\cos\theta}{l})\right]^2 \\ &= \frac{1}{2}\left[(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + mg\cos\theta\sin\theta - m\ddot{x}\cos^2\theta\right]^2 .\end{aligned}$$

For simplicity, denote by

$$G^\theta = \frac{g\sin\theta - \ddot{x}\cos\theta}{l}.$$

Now, the presymplectic 2-form Ω_{W_0}, the Hamiltonian H_{W_0} and the primary constraint φ_x^1 are, respectively

$$\begin{aligned}\Omega_{W_0} &= dx\wedge dp_x^0 + d\theta\wedge dp_\theta^0 + d\dot{x}\wedge dp_x^1 + d\dot{\theta}\wedge dp_\theta^1 ,\\ H_{W_0} &= p_x^0\dot{x} + p_\theta^0\dot{\theta} + p_x^1\ddot{x} + p_\theta^1\left[\frac{g\sin\theta - \ddot{x}\cos\theta}{l}\right]\\ &\quad -\frac{1}{2}\left[(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + mg\cos\theta\sin\theta - m\ddot{x}\cos^2\theta\right]^2 ,\\ \varphi_x^1 &= \frac{\partial\widetilde{H}}{\partial\ddot{x}} = p_x^1 + p_\theta^1\frac{\partial G^\theta}{\partial\ddot{x}} - \frac{\partial\widetilde{L}_{\mathcal{M}}}{\partial\ddot{x}} = 0,\end{aligned}$$

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 3.6.2 we deduce that the 2-form Ω_{W_1}, restriction of Ω_{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 $X \in \mathfrak{X}(W_1)$ which satisfies $i_X\Omega_{W_1} = dH_{W_1}$. 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 $F_t : W_1 \to W_1$ of the vector field X then we have that $F_t^*\Omega_{W_1} = \Omega_{W_1}$. Obviously, the Hamiltonian function

$$
\begin{aligned}
\widetilde{H}\big|_{W_1} &= p_x^0\dot{x} + p_\theta^0\dot{\theta} + \left[-p_\theta^1\frac{\partial G^\theta}{\partial \ddot{x}} + \frac{\partial \widetilde{L}_N}{\partial \ddot{x}}\right]\ddot{x} + p_\theta^1\left[\frac{g\sin\theta - \ddot{x}\cos\theta}{l}\right] \\
&- \frac{1}{2}\left[(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + mg\cos\theta\sin\theta - m\ddot{x}\cos^2\theta\right]^2
\end{aligned}
$$

is preserved by the solution of the optimal control problem, that is $\widetilde{H}\big|_{W_1} \circ F_t = \widetilde{H}\big|_{W_1}$. 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

$$
\begin{aligned}
\frac{d^4x}{dt^4} &= -\frac{1}{[(M+m) - m\cos^2\theta]^2}\Big\{[4m\dot{\theta}\cos\theta\sin\theta]\times \\
&\quad [(M+m)\dddot{x} - ml\dot{\theta}^3\cos\theta - 2m\sin\theta\dot{\theta}(g\sin\theta - \ddot{x}\cos\theta) + mg\dot{\theta}\cos(2\theta)] \\
&+ \ 2m[(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + mg\cos\theta\sin\theta - m\ddot{x}\cos^2\theta]\times \\
&\quad \times[\dot{\theta}^2\cos(frm{-}e\theta) + \cos\theta\sin\theta\ddot{\theta}] \\
&+ \ \frac{1}{l}\left(\frac{d^2}{dt^2}p_\theta 1\cos\theta - 2\frac{d}{dt}p_\theta^1\dot{\theta}\sin\theta - p_\theta 1(\dot{\theta}^2\cos\theta + \ddot{\theta}\sin\theta)\right)\Big\} \\
&+ \ \frac{1}{[(M+m) - m\cos^2\theta]}\Big\{\frac{2m\sin\theta}{l}(g\dot{\theta}\cos\theta - \dddot{x}\cos\theta + \ddot{x}\dot{\theta}\sin\theta)^2 \\
&+ \ 2m\dot{\theta}\sin\theta(g\dot{\theta}\cos\theta - \dddot{x}\cos\theta + \ddot{x}\dot{\theta}\sin\theta) + 4m\dot{\theta}^2\cos\theta(q\sin\theta - \ddot{x}\cos\theta) \\
&- \ ml\dot{\theta}^4\sin\theta - 4mg\sin\theta\cos\theta\dot{\theta}^2 + \frac{mg}{l}\cos(2\theta)(g\sin\theta - \ddot{x}\cos\theta) \\
&- \ 4m\dddot{x}\dot{\theta}\cos\theta + 2m\ddot{x}\dot{\theta}^2\cos(2\theta) + \frac{2m}{l}\ddot{x}\cos\theta\sin\theta(g\sin\theta - \ddot{x}\cos\theta)\Big\} \\
\frac{d^2\theta}{dt^2} &= \frac{1}{l}(g\sin\theta - \ddot{x}\cos\theta)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2p_\theta^1}{dt^2} &= \Big\{(M+m)\dddot{x} - ml(2\dot{\theta}\sin\theta(\frac{g\sin\theta-\ddot{x}\cos\theta}{l}) + \dot{\theta}^3\cos\theta) + mg\dot{\theta}\cos(2\theta) \\
&\quad -m\dddot{x}\cos^2\theta + 2m\ddot{x}\dot{\theta}\cos\theta\sin\theta\Big\} \\
&\quad \times(-2ml\dot{\theta}\sin\theta - mg\cos(2\theta) + ml\dot{\theta}^2\cos\theta - 2m\ddot{x}\cos\theta\sin\theta) \\
&+ \Big\{(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + mg\cos\theta\sin\theta - m\ddot{x}\cos\theta\Big\}\times \\
&\quad \times\left(-2ml\left(\frac{(g\sin\theta-\ddot{x}\cos\theta)}{l}\sin\theta + \dot{\theta}^2\cos\theta\right) + 2m\dot{\theta}\cos\theta(g\sin\theta - \ddot{x}cos\theta)\right. \\
&\quad \left. -ml\dot{\theta}^3\sin\theta - 2m\dddot{x}\cos\theta\sin\theta - 2m\ddot{x}\dot{\theta} + 2mg\dot{\theta}\sin\theta\cos\theta(1+m\ddot{x})\right) \\
&+ \frac{1}{l}\Big[-p_\theta^0 + (M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + m\cos\theta(g\sin\theta - \ddot{x}\cos\theta)\Big](-2ml\sin\theta) \\
&\quad +p_\theta 1(\dddot{x}\sin\theta + \ddot{x}\dot{\theta}\cos\theta - g\dot{\theta}\sin\theta)\Big].
\end{aligned}
$$

3.6.2 Underactuated control systems on Lie groups

Now we consider the optimal control of a mechanical system on a finite dimensional Lie group. The goal is to move the system within the time interval $I = [0, T]$, under the influence of control forces f with chosen control parameter $u(t)$, from its current state to a desired state in an optimal way, e.g. by minimizing distance, control effort, or time, which will be representated by a suitable cost function.

Let the configuration space be a n-dimensional Lie group G with Lie algebra $\mathfrak{g}$. We assume that the controlled equations are trivialized where $L : G\times\mathfrak{g}\to\mathbb{R}$. These equations are

$$
\frac{d}{dt}\left(\frac{\delta L}{\delta\xi}\right) - ad_\xi^*\left(\frac{\delta L}{\delta\xi}\right) - \pounds_g^*\frac{\partial L}{\partial g} = u_a e^a, \tag{3.98}
$$

where we are assuming that $\{e^a\}$ are independent elements on $\mathfrak{g}^*$ and (u_a) are the control parameters. Complete it to a basis $\{e^a, e^\alpha\}$ of the vector space $\mathfrak{g}^*$. Take its dual basis $\{e_A\} = \{e_a, e_\alpha\}$ on $\mathfrak{g}$ with bracket relations:

$$
[e_A, e_B] = \mathcal{C}_{AB}^C e_C
$$

The basis $\{e_A\} = \{e_a, e_\alpha\}$ induces coordinates $(y^a, y^\alpha) = (y^A)$ on $\mathfrak{g}$, that is, if $e\in\mathfrak{g}$ then $e = y^A e_A = y^a e_a + y^\alpha e_\alpha$. In $\mathfrak{g}^*$, we have induced coordinates (p_a, p_α) for the previous fixed basis $\{e^A\}$.

In these coordinates, the equations of motion are rewritten as

$$
\begin{aligned}
\frac{d}{dt}\left(\frac{\partial L}{\partial y^a}\right) - \mathcal{C}_{Aa}^B y^A\frac{\partial L}{\partial y^B} - \langle\pounds_g^*\frac{\delta L}{\delta g}, e_a\rangle &= u_a\,, \\
\frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) - \mathcal{C}_{A\alpha}^B y^A\frac{\partial L}{\partial y^B} - \langle\pounds_g^*\frac{\delta L}{\delta g}, e_\alpha\rangle &= 0\,.
\end{aligned}
$$

With these equations we can study the optimal control problem that consists on finding trajectories $(g(t), u^a(t))$ of the state variables and control inputs satisfying the previous equations from given initial and final conditions $(g(t_0), y^A(t_0))$ and $(g(t_f), y^A(t_f))$, respectively, and extremizing the functional

$$\mathcal{J} = \int_{t_0}^{t_f} C(g(t), y^A(t), u^a(t))\, dt. \tag{3.99}$$

The proposed optimal control problem is equivalent to a variational problem with second order constraints, determined by the Lagrangian $\widetilde{L} : G \times 2\mathfrak{g} \to \mathbb{R}$ given, in the selected coordinates, by

$$\widetilde{L}(g, y^A, \dot{y}^A) = C\left(g, y^A, \frac{d}{dt}\left(\frac{\partial L}{\partial y^a}\right) - \mathcal{C}^B_{Aa} y^A \frac{\partial L}{\partial y^B} - \langle \pounds^*_g \frac{\delta L}{\delta g}, e_a\rangle\right),$$

subjected to the second-order constraints

$$\Phi^A(g, y^A, \dot{y}^A) = \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) - \mathcal{C}^B_{A\alpha} y^A \frac{\partial L}{\partial y^B} - \langle \pounds^*_g \frac{\delta L}{\delta g}, e_\alpha\rangle = 0\,, \tag{3.100}$$

which determine the submanifold $\mathcal{M}$ of $G \times 2\mathfrak{g}$ with projections described on the following diagram

$$\begin{tikzcd}
& \overline{W}_0 = \mathcal{M} \times 2\mathfrak{g}^* \arrow[dl, "pr_1 \circ i_{\overline{W}_0}"'] \arrow[dr, "pr_2 \circ i_{\overline{W}_0}"] & \\
\mathcal{M} \arrow[dr, "(\tau_G^{(1,2)})|_M"'] & & G \times \mathfrak{g} \times 2\mathfrak{g}^* \arrow[dl, "\pi_{G\times\mathfrak{g}}"] \\
& G \times \mathfrak{g} &
\end{tikzcd}$$

Figure 3.4: Skinner and Rusk formalism

Here $i_{\overline{W}_0} : \overline{W}_0 \hookrightarrow G \times 2\mathfrak{g} \times 2\mathfrak{g}^*$ denotes the canonical inclusion and $\tau_G^{(1,2)} : G \times 2\mathfrak{g} \to G \times \mathfrak{g}$ was defined in section 1.7.

Observe that from the constraint equations we have that

$$\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\dot{y}^\beta + \frac{\partial^2 L}{\partial y^\alpha \partial y^b}\dot{y}^b - \mathcal{C}^B_{A\alpha} y^A \frac{\partial L}{\partial y^B} - \langle \pounds^*_g \frac{\delta L}{\delta g}, e_\alpha\rangle \quad = \quad 0\,.$$

Therefore, assuming that the matrix $(W_{\alpha\beta}) = \left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$ is regular we can write the constraint equations as

$$\begin{aligned}
\dot{y}^\beta &= -W^{\beta\alpha}\left(\frac{\partial^2 L}{\partial y^\alpha \partial y^b}\dot{y}^b - \mathcal{C}^B_{A\alpha} y^A \frac{\partial L}{\partial y^B} - \langle \pounds^*_g \frac{\delta L}{\delta g}, e_\alpha\rangle\right) \\
&= G^\beta(g, y^A, \dot{y}^a)
\end{aligned}$$

where $W^{\beta\alpha} = (W_{\beta\alpha})^{-1}$.

This means that we can identify $T\mathcal{M} \equiv G \times \text{span}\ \{(e_A, \mathbf{0}, \mathbf{0}), (\mathbf{0}, e_A, \mathbf{0}), (\mathbf{0}, \mathbf{0}, e_a)\}$ where $(e_A, \mathbf{0}, \mathbf{0}), (\mathbf{0}, e_A, \mathbf{0}), (\mathbf{0}, \mathbf{0}, e_a) \in 3\mathfrak{g}$.

Therefore, we can choose coordinates $(g, y^A, \dot{y}^a)$ on $\mathcal{M}$. This choice allows us to consider an "intrinsic point view", that is, to work directly on $\overline{W}_0 = \mathcal{M} \times 2\mathfrak{g}^*$ avoiding the use of Lagrange multipliers.

Define the restricted Lagrangian $\widetilde{L}_{\mathcal{M}}$ by $\widetilde{L}_{\mathcal{M}} = \widetilde{L}\ |_{\mathcal{M}}\colon \mathcal{M} \to \mathbb{R}$ and induced coordinates on $\overline{W}_0$ are $\gamma = (g, y^A, \dot{y}^a, p_A, \tilde{p}_A)$. Consider the presymplectic 2-form on $\overline{W}_0$, $\Omega_{\overline{W}_0} = (pr_2 \circ i_{\overline{W}_0})^*(\omega_{G\times\mathfrak{g}})$.

Using the notation $(e_A)_0 = (e_A, \mathbf{0}, \mathbf{0}; \mathbf{0}, \mathbf{0}) \in 3\mathfrak{g} \times 2\mathfrak{g}^*$ and, in the same way $(e_A)_1 = (\mathbf{0}, e_A, \mathbf{0}; \mathbf{0}, \mathbf{0})$, $(e_a)_2 = (\mathbf{0}, \mathbf{0}, e_a; \mathbf{0}, \mathbf{0})$, $(e^A)_3 = (\mathbf{0}, \mathbf{0}, \mathbf{0}; e^A, \mathbf{0})$ and $(e^A)_4 = (\mathbf{0}, \mathbf{0}, \mathbf{0}; \mathbf{0}, e^A)$ then the unique nonvanishing elements on the expression of $\Omega_{\overline{W}_0}$ are:

$$
\begin{array}{rcl}
(\Omega_{\overline{W}_0})_\gamma((e_A)_0, (e_B)_0) & = & \mathcal{C}^C_{AB} p_C\,, \\
(\Omega_{\overline{W}_0})_\gamma((e_A)_0, (e^B)_3) & = & -(\Omega_{\overline{W}_0})_\gamma((e^A)_3, (e_B)_0) = \delta^B_A\,, \\
(\Omega_{\overline{W}_0})_\gamma((e_A)_1, (e^B)_4) & = & -(\Omega_{\overline{W}_0})_\gamma((e^A)^4, (e_B)_1) = \delta^B_A\,.
\end{array}
$$

Taking the dual basis $(e^A)_0 = (e^A, \mathbf{0}, \mathbf{0}; \mathbf{0}, \mathbf{0}) \in 3\mathfrak{g}^* \times 2\mathfrak{g}$ and, in the same way $(e^A)_1 = (\mathbf{0}, e^A, \mathbf{0}; \mathbf{0}, \mathbf{0})$, $(e^a)_2 = (\mathbf{0}, e^a, \mathbf{0}; \mathbf{0}, \mathbf{0})$, $(e_A)_3 = (\mathbf{0}, \mathbf{0}, \mathbf{0}; e_A, \mathbf{0})$ and $(e_A)_4 = (\mathbf{0}, \mathbf{0}, \mathbf{0}; \mathbf{0}, e_A)$ we deduce that

$$
(\Omega_{\overline{W}_0}) = (e^A)_0 \wedge (e_A)_3 + (e^A)_1 \wedge (e_A)_4 + \frac{1}{2}\mathcal{C}^C_{AB} p_C (e^A)_0 \wedge (e^B)_0.
$$

Moreover we define the Hamiltonian $\bar{H} : \overline{W}_0 \to \mathbb{R}$ by

$$
\bar{H} = y^A p_A + \dot{y}^a \tilde{p}_a + G^\alpha(g, y^A, \dot{y}^a)\tilde{p}_\alpha - \widetilde{L}_{\mathcal{M}}(g, y^A, \dot{y}^a),
$$

and, in consequence,

$$
\begin{array}{rcl}
d\bar{H} & = & -\langle \mathcal{L}^*_g \left(\dfrac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g} + \tilde{p}_\beta \dfrac{\delta G^\beta}{\delta g} \right), e_A \rangle (e^A)_0 + \left(p_A - \dfrac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^A} + \tilde{p}_\beta \dfrac{\partial G^\beta}{\partial y^A} \right) (e^A)_1 \\
& & + \left(\tilde{p}_a - \dfrac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^a} + \tilde{p}_\beta \dfrac{\partial G^\beta}{\partial \dot{y}^a} \right) (e^a)_2 + y^A (e_A)_3 + \dot{y}^a (e_a)_4 + G^\alpha (e_\alpha)_4\,.
\end{array}
$$

The conditions for the integral curves $t \to (g(t), y^A(t), \dot{y}^a(t), p_\alpha(t), \tilde{p}_\alpha(t))$ of a vector field X

satisfying equations $i_X\Omega_{\overline{W}_0} = d\bar{H}$ are

$$
\begin{aligned}
\frac{dg}{dt} &= g(y^A(t)e_A)\\
\frac{dy^a}{dt} &= \dot{y}^a\\
\frac{dy^\alpha}{dt} &= G^\alpha(g, y^A, \dot{y}^a)\\
\frac{dp_A}{dt} &= \langle \pounds_g^* \left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g} - \widetilde{p}_\beta \frac{\delta G^\beta}{\delta g} \right), e_A \rangle + \mathcal{C}_{AB}^C p_C y^B\\
\frac{d\widetilde{p}_A}{dt} &= -p_A + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^A} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^A}\\
\widetilde{p}_a &= \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^a} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial \dot{y}^a} = \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^a} + W^{\beta\alpha} \widetilde{p}_\beta \frac{\partial^2 L}{\partial y^\alpha \partial y^a}
\end{aligned}
$$

To shorten the number of unknown variables involved in the previous set of equations, we can write them using as variables $(g(t), y^A, \dot{y}^a, \widetilde{p}_\alpha)$

$$
\begin{aligned}
\frac{dg}{dt} &= g(y^A(t)e_A)\\
\frac{dy^\alpha}{dt} &= G^\alpha(g, y^A, \dot{y}^a)\\
0 &= \frac{d^2}{dt^2}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^a} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial \dot{y}^a} \right] - \mathcal{C}_{Aa}^b y^A \left(\frac{d}{dt}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^b} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial \dot{y}^b} \right] \right)\\
&\quad - \frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^a} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^a} \right) + \mathcal{C}_{Aa}^C y^A \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^C} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^C} \right)\\
&\quad + \langle \pounds_g^* \left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g} - \widetilde{p}_\beta \frac{\delta G^\beta}{\delta g} \right), e_a \rangle - \mathcal{C}_{Aa}^\gamma y^A \frac{d\widetilde{p}_\gamma}{dt}\\
0 &= \frac{d^2 \widetilde{p}_\alpha}{dt^2} + \mathcal{C}_{A\alpha}^\beta y^A \frac{d\widetilde{p}_\beta}{dt} - \mathcal{C}_{A\alpha}^C y^A \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^C} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^C} \right]\\
&\quad - \frac{d}{dt}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^\alpha} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^\alpha} \right] + \langle \pounds_g^* \left(\frac{\delta \widetilde{L}_{\mathcal{M}}}{\delta g} - \widetilde{p}_\beta \frac{\delta G^\beta}{\delta g} \right), e_\alpha \rangle\\
&\quad + \mathcal{C}_{A\alpha}^b y^A \left(\frac{d}{dt}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial \dot{y}^b} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial \dot{y}^b} \right] \right) - \mathcal{C}_{A\alpha}^b y^A \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^b} - \widetilde{p}_\beta \frac{\partial G^\beta}{\partial y^b} \right]
\end{aligned}
$$

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 an explicit system of third-order differential equations. This regularity assumption is equivalent to the condition that the constrain algorithm stops at the first constraint submanifold $\overline{W}_1$ (see Theorem (3.2.1)).

Example 3.6.7 (optimal control of an underactuated rigid body)**.** We consider the motion of a rigid body where the configuration space is the Lie group $G = SO(3)$ (see [26, 105]). Therefore, $TSO(3) \simeq SO(3) \times \mathfrak{so}(3)$, where $\mathfrak{so}(3) \equiv \mathbb{R}^3$ is the Lie algebra of the Lie group $SO(3)$ (here we are using the well know isomorphism between the Lie algebra $\mathfrak{so}(3)$ and $\mathbb{R}^3$, see [81] and [130] for example). The Lagrangian function for this system is given by $L : SO(3) \times \mathfrak{so}(3) \to \mathbb{R}$,

$$L(R, \Omega_1, \Omega_2, \Omega_3) = \frac{1}{2}\left(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2\right).$$

Now, denote by $t \to R(t) \in SO(3)$ 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 equation:

$$\dot{R}(t) = R(t)\begin{pmatrix} 0 & -\Omega_3(t) & \Omega_2(t) \\ \Omega_3(t) & 0 & -\Omega_1(t) \\ -\Omega_2(t) & \Omega_1(t) & 0 \end{pmatrix} = R(t)\left(\Omega_1(t)E_1 + \Omega_2(t)E_2 + \Omega_3(t)E_3\right)$$

where

$$E_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad E_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad E_3 := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and the equations for the angular velocities Ω_i with $i = 1, 2, 3$:

$$\begin{aligned} I_1\dot{\Omega}_1(t) &= (I_2 - I_3)\Omega_2(t)\Omega_3(t) + u_1(t) \\ I_2\dot{\Omega}_2(t) &= (I_3 - I_1)\Omega_3(t)\Omega_1(t) + u_2(t) \\ I_3\dot{\Omega}_3(t) &= (I_1 - I_2)\Omega_1(t)\Omega_2(t) \end{aligned}$$

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 $(R(t_0), \Omega(t_0))$, $(R(t_f), \Omega(t_f))$ respectively and minimizing the cost functional

$$\mathcal{A} = \int_0^T \mathcal{C}(\Omega, u_1, u_2)dt = \int_0^T \left[c_1(u_1^2 + u_2^2) + c_2(\Omega_1^2 + \Omega_2^2 + \Omega_3^2)\right] dt,$$

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,

$$\min \widetilde{\mathcal{J}} = \int_0^T \widetilde{L}(\Omega, \dot{\Omega})dt$$

subject to the constraint $I_3\dot{\Omega}_3 - (I_1 - I_2)\Omega_1\Omega_2 = 0$, where

$$\widetilde{L}(\Omega, \dot{\Omega}) = \mathcal{C}\left(\Omega, I_1\dot{\Omega}_1 - (I_2 - I_3)\Omega_2\Omega_3, I_2\dot{\Omega}_2 - (I_3 - I_1)\Omega_3\Omega_1\right) .$$

Thus, the submanifold $\mathcal{M}$ of $G \times 2\mathfrak{so}(3)$, is given by

$$\mathcal{M} := \{(R, \Omega, \dot{\Omega}) \mid \dot{\Omega}_3 = \left(\frac{I_1 - I_2}{I_3}\right)\Omega_1\Omega_2\}.$$

We consider the submanifold $\overline{W}_0 = \mathcal{M} \times 2\mathfrak{so}^*(3)$ with induced coordinates

$$(g, \Omega_1, \Omega_2, \Omega_3, \dot{\Omega}_1, \dot{\Omega}_2, p_1, p_2, p_3, \tilde{p}_1, \tilde{p}_2, \tilde{p}_3).$$

Now, we consider the restriction of the Lagrangian $\widetilde{L}$ to the constraint submanifold, $\widetilde{L}_{\mathcal{M}}$ given by

$$\tilde{L}_{\mathcal{M}} = c_1\left[\left(I_1\dot{\Omega}_1 - (I_2 - I_3)\Omega_2\Omega_3\right)^2 + \left(I_2\dot{\Omega}_2 - (I_3 - I_1)\Omega_3\Omega_1\right)^2\right] + c_2\left(\Omega_1^2 + \Omega_2^2 + \Omega_3^2\right) .$$

For simplicity we denote by $G^3 = \frac{I_1 - I_2}{I_3}\Omega_1\Omega_2$. Therefore, we can write the equations of motion of the optimal control problem 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 problem are:

$$\begin{aligned}
\Omega_2(t)\frac{d\tilde{p}_3}{dt} - 2\left(c_2\frac{d\Omega_1}{dt} + c_1\Omega_3(t)\frac{d^2\Omega_2}{dt^2} - c_1\frac{d^3\Omega_1}{dt^3}\right) &= 0,\\
-\Omega_1(t)\frac{d\tilde{p}_3}{dt} - 2\left(c_2\frac{d\Omega_2}{dt} - c_1\Omega_3(t)\frac{d^2\Omega_1}{dt^2} - c_1\frac{d^3\Omega_2}{dt^3}\right) &= 0,\\
\frac{d^2\tilde{p}_3}{dt^2} - 2c_2\frac{d\Omega_3}{dt} - 2c_1\Omega_2(t)\frac{d^2\Omega_1}{dt^2} + 2c_1\Omega_1(t)\frac{d^2\Omega_2}{dt^2} &= 0,\\
\frac{d\Omega_3}{dt} &= 0.
\end{aligned}$$

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 (c_1) and also is trying to minimize the overall angular velocity (c_2). In Figures (3.5) and (3.6) we compare their behavior in two particular cases: $c_1 = 1/2$, $c_2 = 1/2$; and $c_1 = 1/2$, $c_2 = 0$.

The most typical case is of course the problem of minimize the total fuel expenditure, that is $c_1 = 1$ and $c_2 = 0$. Now, the explicit system of differential equations is

$$\begin{aligned}
\Omega_2(t)\frac{d\tilde{p}_3}{dt} - 2\left(\frac{d^2\Omega_2}{dt^2}\Omega_3(t) - \frac{d^3\Omega_1}{dt^3}\right) &= 0,\\
-\Omega_1(t)\frac{d\tilde{p}_3}{dt} + 2\left(\frac{d^2\Omega_1}{dt^2}\Omega_3(t) + \frac{d^3\Omega_2}{dt^3}\right) &= 0,\\
\frac{d^2\tilde{p}_3}{dt^2} - 2\left(\Omega_2(t)\frac{d^2\Omega_1}{dt^2} - \Omega_1(t)\frac{d^2\Omega_2}{dt^2}\right) &= 0,\\
\frac{d\Omega_3}{dt} &= 0.
\end{aligned}$$

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 $R(t_0)$ and $R(t_f)$.

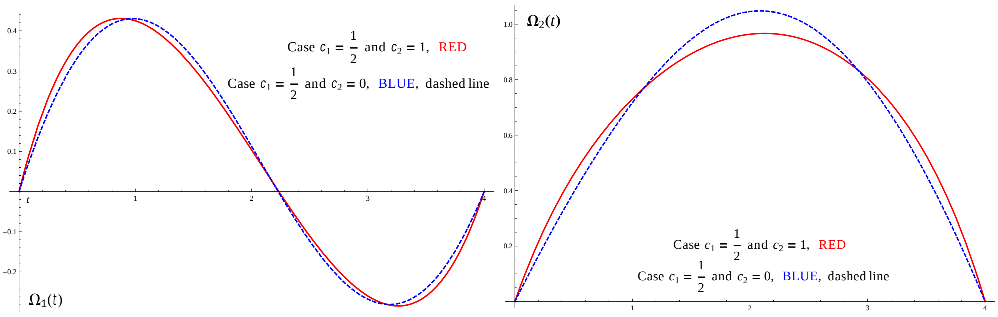


Figure 3.5: Angular velocity values for initial conditions satisfying $\Omega_i(0) = \Omega_i(4) = 0$, $i = 1, 2$ and fixed values of $R(0)$ and $R(4)$.

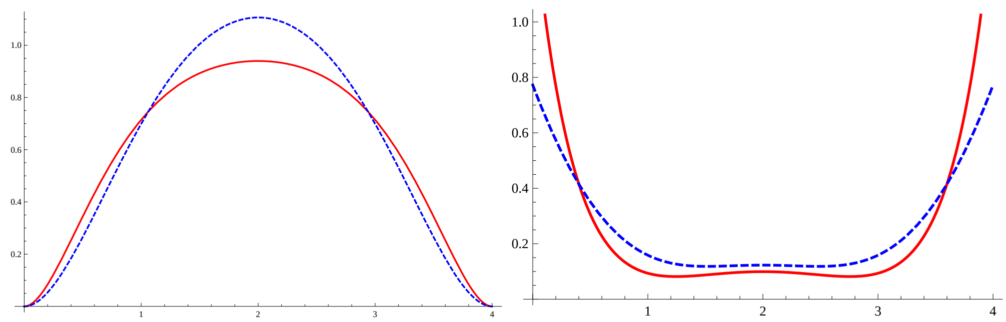


Figure 3.6: Comparison of the functions $1/2\left(\Omega_1^2(t) + \Omega_2^2(t) + \Omega_3^2(t)\right)$ (left) and $1/2\left(u_1^2(t) + u_2^2(t)\right)$ (right) in both cases

Remark 3.6.8. Also, it is possible to consider a slightly more complicated example where the moments of inertia are not equal. The resulting system of equations are computed numerically using Mathematica. We include the equations for different moments of inertia and it is clear that it is necessary to develop numerical methods preserving the geometric structure for this mechanical control systems.

$$
\begin{aligned}
0 \quad = \quad & \Omega_2\tilde{p}_3' + \frac{(I_1 - I_2)\,\Omega_2\tilde{p}_3'}{I_3} - 2c_1\Omega_1' + \Omega_3\Bigg(\frac{\tilde{p}_3\,(-I_1 + I_2)\,\Omega_1}{I_3} + 2c_1\Omega_2 \\
& +2\,(-I_2 + I_3)\,c_2\Omega_3\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\Bigg) + \frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_2'}{I_3} \\
& -\Omega_2\left(2c_1\Omega_3 + c_2\left(2\,(-I_2 + I_3)\,\Omega_2\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\right.\right. \\
& \left.\left. +2\,(I_1 - I_3)\,\Omega_1\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\right)\right) \\
& -2\,(I_1 - I_3)\,c_2\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\Omega_3' \\
& -2I_2c_2\Omega_3\left((I_1 - I_3)\left(\Omega_3\Omega_1' + \Omega_1\Omega_3'\right) + I_2\Omega_2''\right) \\
& -2\,(I_1 - I_3)\,c_2\Omega_3\left((I_1 - I_3)\left(\Omega_3\Omega_1' + \Omega_1\Omega_3'\right) + I_2\Omega_2''\right) \\
& +2I_1c_2\Big(-\,(I_2 - I_3)\left(2\Omega_2'\Omega_3' + \Omega_3\Omega_2'' + \Omega_2\Omega_3''\right) + I_1\Omega_1^{(3)}\Big) \\
0 \quad = \quad & -\Omega_1\tilde{p}_3' + \frac{(I_1 - I_2)\,\Omega_1\tilde{p}_3'}{I_3} + \frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_1'}{I_3} - 2c_1\Omega_2' \\
& -\Omega_3\left(2c_1\Omega_1 - \frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_2}{I_3} + 2\,(I_1 - I_3)\,c_2\Omega_3\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\right) \\
& +\Omega_1\left(2c_1\Omega_3 + c_2\left(2\,(-I_2 + I_3)\,\Omega_2\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\right.\right. \\
& \left.\left. +2\,(I_1 - I_3)\,\Omega_1\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\right)\right) \\
& -2\,(-I_2 + I_3)\,c_2\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\Omega_3' \\
& +2I_1c_2\Omega_3\left((-I_2 + I_3)\,\Omega_3\Omega_2' + (-I_2 + I_3)\,\Omega_2\Omega_3' + I_1\Omega_1''\right) \\
& -2\,(-I_2 + I_3)\,c_2\Omega_3\left((-I_2 + I_3)\,\Omega_3\Omega_2' + (-I_2 + I_3)\,\Omega_2\Omega_3' + I_1\Omega_1''\right) \\
& +2I_2c_2\Big(2\,(I_1 - I_3)\,\Omega_1'\Omega_3' + (I_1 - I_3)\,\Omega_3\Omega_1'' + (I_1 - I_3)\,\Omega_1\Omega_3'' + I_2{\Omega_2}^{(3)}\Big),
\end{aligned}
$$

$$
\begin{aligned}
0 \quad = \quad & \Omega_1\left(-\frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_1}{I_3} + 2c_1\Omega_2 + 2\,(-I_2 + I_3)\,c_2\Omega_3\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\right) \\
& -\Omega_2\left(2c_1\Omega_1 - \frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_2}{I_3} + 2\,(I_1 - I_3)\,c_2\Omega_3\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\right) \\
& -2c_1\Omega_3' + \tilde{p}_3' - 2c_2\left((-I_2 + I_3)\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\Omega_2'\right. \\
& +(I_1 - I_3)\,\Omega_1'\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right) + (-I_2 + I_3)\,\Omega_2\left((-I_2 + I_3)\,\Omega_3\Omega_2'\right. \\
& \left. +(-I_2 + I_3)\,\Omega_2\Omega_3' + I_1\Omega_1''\right) + (I_1 - I_3)\,\Omega_1\left((I_1 - I_3)\,\Omega_3\Omega_1'\right. \\
& \left.\left. +\ (I_1 - I_3)\,\Omega_1\Omega_3' + I_2\Omega_2''\right)\right) \\
0 \quad = \quad & \Omega_1\left(-\frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_1}{I_3} + 2c_1\Omega_2 + 2\,(-I_2 + I_3)\,c_2\Omega_3\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\right) \\
& -\Omega_2\left(2c_1\Omega_1 - \frac{\tilde{p}_3\,(I_1 - I_2)\,\Omega_2}{I_3} + 2\,(I_1 - I_3)\,c_2\Omega_3\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right)\right) \\
& -2c_1\Omega_3' + \tilde{p}_3' - 2c_2\left((-I_2 + I_3)\left((-I_2 + I_3)\,\Omega_2\Omega_3 + I_1\Omega_1'\right)\Omega_2'\right. \\
& +(I_1 - I_3)\,\Omega_1'\left((I_1 - I_3)\,\Omega_1\Omega_3 + I_2\Omega_2'\right) + (-I_2 + I_3)\,\Omega_2\left((-I_2 + I_3)\,\Omega_3\Omega_2'\right. \\
& \left.\left. +(-I_2 + I_3)\,\Omega_2\Omega_3' + I_1\Omega_1''\right) + (I_1 - I_3)\,\Omega_1\left((I_1 - I_3)\,\Omega_3\Omega_1' + (I_1 - I_3)\,\Omega_1\Omega_3' + I_2\Omega_2''\right)\right)
\end{aligned}
$$

$\diamond$

The case $c_1 = 0$ and $c_2 = 1$, that is, we only try to minimize the overall angular velocity (see [163] for the fully-actuated case) is singular. We obtain the following system

$$\begin{aligned}
\Omega_2(t)\frac{d\tilde{p}_3}{dt} - 2\frac{d\Omega_1}{dt} &= 0,\\
-\Omega_1(t)\frac{d\tilde{p}_3}{dt} - 2\frac{d\Omega_2}{dt} &= 0,\\
\frac{d^2\tilde{p}_3}{dt^2} - 2\frac{d\Omega_3}{dt} &= 0,\\
\frac{d\Omega_3}{dt} &= 0\,.
\end{aligned}$$

Observe that in this case it is not possible to impose arbitrary boundary conditions $(R(t_0), \Omega(t_0))$ and $(R(t_f), \Omega(t_f))$ although it is always possible to find a trajectory verifying initial and final attitude conditions $R(t_0)$ and $R(t_f)$. It is interesting to observe that in this particular case the algorithm does not stop at the first step. It is easy to show that

$$\begin{aligned}
\overline{W}_1 &= \{(g, \Omega_1, \Omega_2, \Omega_3, \dot{\Omega}_1, \dot{\Omega}_2, p_1, p_2, p_3, \tilde{p}_1, \tilde{p}_2, \tilde{p}_3) \in \overline{W}_0\ |\tilde{p}_1 = 0, \tilde{p}_2 = 0\}\\
\overline{W}_2 &= \{(g, \Omega_1, \Omega_2, \Omega_3, \dot{\Omega}_1, \dot{\Omega}_2, p_1, p_2, p_3, 0, 0, \tilde{p}_3) \in \overline{W}_1 \mid p_1 = \frac{1}{2}\Omega_1, p_2 = \frac{1}{2}\Omega_2\}\\
\overline{W}_3 &= \{(g, \Omega_1, \Omega_2, \Omega_3, \dot{\Omega}_1, \dot{\Omega}_2, \frac{1}{2}\Omega_1, \frac{1}{2}\Omega_2, p_3, 0, 0, \tilde{p}_3) \in \overline{W}_2 \mid \frac{1}{2}\dot{\Omega}_1 = p_3\Omega_2 - \frac{1}{2}\Omega_2\Omega_3,\\
&\quad \frac{1}{2}\dot{\Omega}_2 = -p_3\Omega_1 + \frac{1}{2}\Omega_1\Omega_3\}.
\end{aligned}$$

Remark 3.6.9. For numerical simulation of underactuated optimal control systems is usually difficult to satisfy numerically the second-order constraints (3.100) then it is useful to combine our analysis with other techniques for underactuated systems. For instance, assuming that our underactuated system of ODEs (3.98) posses the property of differential flatness (often referred as flatness, see [67, 143]), with $z(t) = F(g(t), y^i(t), \dot{y}^i(t), u^a(t))$ as a flat output, such that the states and inputs may be determined from equations of the form

$$(g(t), y^i(t), u^a(t)) \quad = \quad \phi(z(t), \ldots, z^{(\beta)}(t))$$

Then substituting in the cost functional (3.99) we obtain the new functional

$$\mathcal{J}_z = \int_{t_0}^{t_f} L(\phi(z(t), \ldots, z^{(\beta)}(t))\, dt = \int_{t_0}^{t_f} \tilde{L}(z(t), \ldots, z^{(\beta)}(t))\, dt$$

now *without additional constraints*. Consequently, the number of variables in the optimal control problem will be reduced to expedite real-time computation.

◇

3.6.3 Optimal control of underactuated mechanical systems on a principal bundle

Let us consider the configuration space Q of the system as a trivial principal bundle, that is, $Q = M \times G$, where M is an m-dimensional smooth manifold and G a finite dimensional

Lie group. Let $L \in C^\infty(TM \times \mathfrak{g})$ be a left-trivialized Lagrangian function, where $\mathfrak{g}$ is the Lie algebra of G.

The Euler-Lagrange equations with controls are

$$\begin{aligned} &\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = u_a \mu^a_A(q)\,, \quad (1 \leqslant A \leqslant m) \\ &\frac{d}{dt}\left(\frac{\partial L}{\partial \xi}\right) - ad^*_\xi\left(\frac{\partial L}{\partial \xi}\right) = u_a \eta^a(q)\,, \end{aligned} \tag{3.101}$$

where $\mathcal{B}^a = \{(\mu^a, \eta^a)\}$, $\mu^a(q) \in T^*_qM$, $\eta^a(q) \in \mathfrak{g}^*$, $a = 1, \ldots, r$, is a set of independent sections of the bundle $\pi\colon T^*M \times \mathfrak{g}^* \to M$, and u_a are admissible control parameters.

We complete $\mathcal{B}^a$ to a basis $\{\mathcal{B}^a, \mathcal{B}^\alpha\}$ of $\Gamma(\pi)$, and let us consider its dual basis $\{\mathcal{B}_a, \mathcal{B}_\alpha\}$, that is, a basis of $\Gamma(\tau)$, where $\tau\colon TM \times \mathfrak{g} \to M$. Observe that $\Gamma(\tau) = \mathfrak{X}(M) \times C^\infty(M, \mathfrak{g})$ (see [111] for details). This basis induces coordinates $(q^A, \dot{q}^A, \xi^a, \xi^\alpha)$ on $TM \times \mathfrak{g}$.

If we denote $\mathcal{B}_a = \{(X_a, \Xi_a)\} \subset \Gamma(\tau)$ and $\mathcal{B}_\alpha = \{X_\alpha, \Xi_\alpha\} \subset \Gamma(\tau)$, where $X_a, X_\alpha \in \mathfrak{X}(M)$ are locally given by $X_a(q) = X^A_a(q) \left.\dfrac{\partial}{\partial q^A}\right|_q$, $X_\alpha(q) = X^A_\alpha(q) \left.\dfrac{\partial}{\partial q^A}\right|_q$, and $\Xi_a(q), \Xi_\alpha(q) \in \mathfrak{g}$, with $q \in M$, then equations (3.101) can be rewritten as

$$\begin{aligned} &\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A}\right) X^A_a(q) + \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \xi}\right) - ad^*_\xi \frac{\partial L}{\partial \xi}\right) \Xi_a(q) = u_a\,, \\ &\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A}\right) X^A_\alpha(q) + \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \xi}\right) - ad^*_\xi \frac{\partial L}{\partial \xi}\right) \Xi_\alpha(q) = 0\,. \end{aligned} \tag{3.102}$$

The optimal control problem consists on finding a trajectory $(q(t), \dot{q}(t), \xi(t), u(t))$ of the state variables and control inputs solving equation (3.102) given initial and final boundary conditions $(q(0), \dot{q}(0), \xi(0))$ and $(q(T), \dot{q}(T), \xi(T))$, respectively, and minimizing the following cost functional

$$\mathcal{A}(q, \dot{q}, \xi, u) = \int_0^T C(q(t), \dot{q}(t), \xi(t), u(t))dt\,,$$

where $C\colon (TM \times \mathfrak{g}) \times U \to \mathbb{R}$ is a cost function.

This optimal control problem is equivalent to solving the following second-order variational problem with second-order constraints

$$\begin{cases} \min \widetilde{L}(q^A, \dot{q}^A, \ddot{q}^A, \xi^i, \dot{\xi}^i), \\ \text{subject to } \Phi^\alpha(q^A, \dot{q}^A, \ddot{q}^A, \xi^i, \dot{\xi}^i)\,,\ \alpha = 1, \ldots, m \end{cases}$$

where $\widetilde{L}, \Phi^\alpha \in \mathrm{C}^\infty(T^{(2)}M \times 2\mathfrak{g})$ are given by

$$\widetilde{L}(q^A, \dot{q}^A, \ddot{q}^A, \xi^i, \dot{\xi}^i) = C\left(q^A, \dot{q}^A, \xi^i, F_a(q^A, \dot{q}^A, \ddot{q}^A, \xi^i, \dot{\xi}^i)\right)\,,$$

where C is the cost function and

$$F_a(q^A, \dot{q}^A, \ddot{q}^A, \xi^i, \dot{\xi}^i) = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A}\right) X^A_a(q) + \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \xi}\right) - \left(ad^*_\xi \frac{\partial L}{\partial \xi}\right)\right) \Xi_a(q)\,.$$

The Lagrangian $\widetilde{L}$ is subjected to the second-order constraints:

$$\Phi^{\alpha}(q^A,\dot{q}^A,\ddot{q}^A,\xi^i,\dot{\xi}^i)=\left(\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^A}\right)-\frac{\partial L}{\partial q^A}\right)X^A_{\alpha}(q)+\left(\frac{d}{dt}\left(\frac{\partial L}{\partial\xi}\right)-\left(ad^*_{\xi}\frac{\partial L}{\partial\xi}\right)\right)\Xi_{\alpha}(q)\,.$$

Example 3.6.10 (Optimal control of an underactuated vehicle)**.** Consider a rigid body moving in the special Euclidean group of the plane $SE(2)$ with a thruster to adjust its pose. The configuration of this system is determined by a tuple (x,y,θ,γ), where (x,y) is the position of the center of mass, θ is the orientation of the blimp with respect to a fixed basis, and γ the orientation of the thrust with respect to a body basis. Therefore, the configuration manifold is $Q=SE(2)\times\mathbb{S}^1$ (see [58] and references therein), where (x,y,θ) are the local coordinates of $SE(2)$ and γ is the coordinate corresponding to the configuration on $\mathbb{S}^1$.

The Lagrangian of this system is given by its kinetic energy

$$L(x,y,\theta,\gamma,\dot{x},\dot{y},\dot{\theta},\dot{\gamma})=\frac{1}{2}m(\dot{x}^2+\dot{y}^2)+\frac{1}{2}J_1\dot{\theta}^2+\frac{1}{2}J_2(\dot{\theta}+\dot{\gamma})^2\,,$$

and the input forces are

$$F^1=\cos(\theta+\gamma)\,\mathrm{d}x+\sin(\theta+\gamma)\,\mathrm{d}y-p\sin\gamma\mathrm{d}\theta;\quad F^2=\mathrm{d}\gamma,$$

where the control forces that we consider are applied to a point on the body with distance $p>0$ from the center of mass (m is the mass of the rigid body), along the body x-axis.

The system is invariant under the left multiplication of the Lie group $G=SE(2)$:

$$\begin{array}{rcl}\Phi\colon SE(2)\times SE(2)\times\mathbb{S}^1 & \longrightarrow & SE(2)\times\mathbb{S}^1\\ ((a,b,\alpha),(x,y,\theta,\gamma)) & \longmapsto & (x\cos\alpha-y\sin\alpha+a,x\sin\alpha+y\cos\alpha+b,\theta+\alpha,\gamma).\end{array}$$

A basis of the Lie algebra $\mathfrak{se}(2)\simeq\mathbb{R}^3$ of $SE(2)$ is given by

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

From this basis we have that

$$[e_1,e_2]=e_3,\quad [e_1,e_3]=-e_2,\quad [e_2,e_3]=0.$$

Thus, we can write down the structure constants as

$$\mathcal{C}^2_{31}=\mathcal{C}^1_{23}=-1,\mathcal{C}^2_{13}=\mathcal{C}^1_{32}=1\,,$$

and all others are zero. An element $\xi\in\mathfrak{se}(2)$ is of the form $\xi=\xi_1\,e_1+\xi_2\,e_2+\xi_3\,e_3$; therefore the reduced Lagrangian $\ell\colon T\mathbb{S}^1\times\mathfrak{se}(2)\to\mathbb{R}$ is given by

$$\ell(\gamma,\dot{\gamma},\xi)=\frac{1}{2}m(\xi_1^2+\xi_2^2)+\frac{J_1+J_2}{2}\xi_3^2+J_2\xi_3\dot{\gamma}+\frac{J_2}{2}\dot{\gamma}^2\,.$$

Then the reduced Euler-Lagrange equations with controls are given by

$$
\begin{aligned}
m\dot{\xi}_1 &= u_1 \cos\gamma\,,\\
m\dot{\xi}_2 + (J_1+J_2)\xi_1\xi_3 + J_2\xi_1\dot{\gamma} - m\xi_1\xi_3 &= u_1 \sin\gamma\,,\\
(J_1+J_2)\dot{\xi}_3 + J_2\ddot{\gamma} - m\xi_2(\xi_1+\xi_3) &= -u_1 p \sin\gamma\,,\\
J_2(\dot{\xi}_3+\ddot{\gamma}) &= u_2\,.
\end{aligned}
$$

On the other hand, choosing the adapted basis $\{\mathcal{B}_a, \mathcal{B}_\alpha\}$ the modified equations of motion (3.102) read in this case as

$$
\begin{aligned}
m(\cos\gamma\dot{\xi}_1 + \sin\gamma(\dot{\xi}_2 - \xi_1\xi_3)) + (J_1+J_2)\xi_1\xi_3\sin\gamma + J_2\xi_1\dot{\gamma}\sin\gamma &= u_1\,,\\
m(\cos\gamma(\dot{\xi}_2 - \xi_1\xi_3) - \sin\gamma\dot{\xi}_1) + \xi_1\xi_3(J_1+J_2)\cos\gamma + J_2\xi_1\dot{\gamma}\cos\gamma &= 0\,,\\
\frac{J_1+J_2}{p}(\dot{\xi}_3 + p\xi_1\xi_3) + \frac{J_2}{p}(\ddot{\gamma} + p\xi_1\dot{\gamma}) + m\left(\dot{\xi}_2 - \xi_1\xi_3 - \frac{\xi_2\xi_1+\xi_3\xi_2}{p}\right) &= 0\,,\\
J_2(\dot{\xi}_3+\ddot{\gamma}) &= u_2\,.
\end{aligned}
$$

Now, we can study the optimal control problem that consists, as mentioned before, on finding a trajectory of state variables and control inputs satisfying the previous equations from given initial and final boundary conditions $(\gamma(0), \dot{\gamma}(0), \xi(0))$, $(\gamma(T), \dot{\gamma}(T), \xi(T))$ respectively, and extremizing the cost functional

$$
\int_0^T (\rho_1 u_1^2 + \rho_2 u_2^2)\,\mathrm{d}t\,,
$$

where ρ_1 and ρ_2 are non-zero constants.

The related optimal control problem is equivalent to the second-order Lagrangian problem with second-order constraints defined as follows. Extremize

$$
\widetilde{\mathcal{A}} = \int_0^T \widetilde{L}(\xi, \dot{\xi}, \gamma, \dot{\gamma}, \ddot{\gamma})\mathrm{d}t,
$$

subject to the second-order constraints given by the functions

$$
\Phi^1 = m(\cos\gamma(\dot{\xi}_2 - \xi_1\xi_3) - \sin\gamma\dot{\xi}_1) + \xi_1\xi_3(J_1+J_2)\cos\gamma + J_2\xi_1\dot{\gamma}\cos\gamma\,, \tag{3.103a}
$$

$$
\Phi^2 = \frac{J_1+J_2}{p}(\dot{\xi}_3 + p\xi_1\xi_3) + \frac{J_2}{p}(\ddot{\gamma} + p\xi_1\dot{\gamma}) + m\left(\dot{\xi}_2 - \xi_1\xi_3 - \frac{\xi_2\xi_1+\xi_3\xi_2}{p}\right)\,. \tag{3.103b}
$$

Here, $\widetilde{L}\colon T^{(2)}\mathbb{S}^1 \times 2\mathfrak{se}(2) \to \mathbb{R}$ is defined by

$$
\begin{aligned}
\widetilde{L}(\gamma, \dot{\gamma}, \ddot{\gamma}, \xi, \dot{\xi}) = \rho_1 \Big(&m(\cos\gamma\dot{\xi}_1 + \sin\gamma(\dot{\xi}_2 - \xi_1\xi_3)) + (J_1+J_2)\xi_1\xi_3\sin\gamma + J_2\xi_1\dot{\gamma}\sin\gamma\Big)^2\\
&+ \rho_2 J_2^2(\dot{\xi}_3+\ddot{\gamma})^2\,,
\end{aligned} \tag{3.104}
$$

which basically is the cost function $C = \rho_1 u_1^2 + \rho_2 u_2^2$ in terms of the new variables.

The submanifold $\mathcal{M}$ of $T^{(2)}\mathbb{S}^1 \times SE(2) \times 2\mathfrak{se}(2)$ is given by the constraints

$$\ddot{\gamma} = -\frac{mp}{J_2}\left(\dot{\xi}_2 - \xi_1\xi_3 - \xi_2\frac{\xi_1+\xi_3}{p}\right) - \frac{J_1+J_2}{J_2}(\dot{\xi}_3 + p\xi_1\xi_3) - p\xi_1\dot{\gamma}\,,$$
$$\dot{\xi}_1 = \frac{1}{\tan\gamma}\left(\frac{J_1+J_2}{m}\xi_1\xi_3 + \frac{J_2}{m}\xi_1\dot{\gamma} + \dot{\xi}_2 - \xi_1\xi_3\right).$$

We consider the submanifold $\overline{W} = \mathcal{M} \times T^*(T\mathbb{S}^1) \times 2\mathfrak{se}(2)^*$ with induced coordinates

$$(\gamma, \dot{\gamma}, g, \xi_1, \xi_2, \xi_3, \dot{\xi}_2, \dot{\xi}_3, \eta_1, \eta_2, p_1, p_2, p_3, \tilde{p}_1, \tilde{p}_2, \tilde{p}_3)\,,$$

and the restriction $\widetilde{L}_{\mathcal{M}}$ of $\widetilde{L}$ given by

$$\widetilde{L}_{\mathcal{M}} = \rho_1\left[\frac{m\cos\gamma}{\tan\gamma}\left(\frac{J_1+J_2}{m}\xi_1\xi_3 + \frac{J_2}{m}\xi_1\dot{\gamma} + \dot{\xi}_2 - \xi_1\xi_3\right) + (J_1+J_2)\xi_1\xi_3\sin\gamma + J_2\xi_1\dot{\gamma}\sin\gamma \right.$$
$$\left. + \sin\gamma(\dot{\xi}_2 - \xi_1\xi_3)\right]^2 + \rho_2 J_2^2\left[\dot{\xi}_3 - \frac{mp}{J_2}\left(\dot{\xi}_2 - \xi_1\xi_3 - \xi_2\frac{\xi_1+\xi_3}{p}\right) - \frac{J_1+J_2}{J_2}(\dot{\xi}_3 + p\xi_1\xi_3) - p\xi_1\dot{\gamma}\right]^2.$$

Observe that we use the intrinsic formulation in the submanifold $\mathcal{M}$ because the constraints enable us to write the variables $\ddot{\gamma}$ and $\dot{\xi}_1$ in terms of the others, and thus it is easy to determine a subset of intrinsic coordinates.

For simplicity, we consider the particular case $J_1 = J_2 = 1$ and $m = p = 1$ then the equations of motion of the optimal control problem are

$$\dot{\xi}_i = \frac{d}{dt}\xi_i \quad , \quad \dot{\gamma} = \frac{d}{dt}\gamma \quad , \quad \ddot{\gamma} = \frac{d}{dt}\dot{\gamma}\,, \quad i = 1,2,3.$$
$$\dot{\eta}_1 = 2\rho_1\left(\left(\frac{\cos\gamma}{\tan\gamma} + \sin\gamma\right)\cdot\mathcal{A}\right)\left(\cos\gamma - \frac{\sin\gamma}{\tan\gamma} - \frac{1}{\cos^2\gamma + \tan^2\gamma}\right) + \lambda_1\sin\gamma\mathcal{A}\left(1 + \frac{1}{\tan\gamma}\right),$$
$$\dot{\eta}_2 = 2\xi_1\rho_1\left(\left(\frac{\cos\gamma}{\tan\gamma} + \sin\gamma\right)^2\cdot\mathcal{A}\right) - \xi_1(\lambda_1\cos\gamma + \lambda_2 - 2\mathcal{B}\rho_2) - \eta_1\,,$$
$$\dot{\tilde{p}}_1 = 2\rho_1\left(\left(\frac{\cos\gamma}{\tan\gamma} + \sin\gamma\right)^2\cdot\mathcal{A}\right)(\mathcal{B} - \lambda_2)(\xi_3 + \dot{\gamma} - \xi_2) - \lambda_1\cos\gamma(\dot{\gamma} + \xi_3)\,,$$
$$\dot{\tilde{p}}_2 = (\lambda_2 - 2\mathcal{B}\rho_2)(\xi_1 + \xi_3)\,,$$
$$\dot{\tilde{p}}_3 = 2\xi_1\rho_1\left(\left(\frac{\cos\gamma}{\tan\gamma} + \sin\gamma\right)^2\cdot\mathcal{A}\right) - \lambda_1\cos\gamma\xi_1 - (\xi_1 - \xi_2)(\lambda_2 + 2\rho_2\mathcal{B}),$$
$$\dot{p}_i = ad^*_{\xi}p_i\,, \quad i = 1,2,3.$$

where

$$\xi = (\xi_1, \xi_2, \xi_3) \quad ; \quad \mathcal{A} = \xi_1\xi_3 + \xi_1\dot{\gamma} + \dot{\xi}_2 \quad ; \quad \mathcal{B} = \dot{\xi}_3 + \xi_1\xi_3 + \dot{\xi}_2 - \xi_2\xi_1 - \xi_2\xi_3 + \xi_1\dot{\gamma}$$

and the coadjoint operator is just the cross product, $ad^*_{\xi}p = \xi \times p$ using the identification of $\mathfrak{se}(2)$ with $\mathbb{R}^3$ (see [81] for example).

In all cases we additionally have the reconstruction equation

$$\dot{g}(t) = g(t)(\xi_1(t)e_1 + \xi_2(t)e_2 + \xi_3(t)e_3)$$

with boundary conditions $g(t_0)$ and $g(t_f)$, where $g(t) = (x(t), y(t), \theta(t))$.

Finally, the regularity condition is given by the matrix

$$A = \begin{pmatrix} 2\rho_2 & 0 & 0 & 2\rho_2 & 0 & 1 \\ 0 & 2\rho_1\cos^2\gamma & 2\rho_1\sin\gamma\cos\gamma & 0 & -\sin\gamma & 0 \\ 0 & 2\rho_1\sin\gamma\cos\gamma & 2\rho_1\sin^2\gamma & 0 & \cos\gamma & 1 \\ 2\rho_2 & 0 & 0 & 2\rho_2 & 0 & 2 \\ 0 & -\sin\gamma & \cos\gamma & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 0 & 0 \end{pmatrix},$$

whose determinant is

$$\det A = 4\rho_1\rho_2\sin^4\gamma + 4\rho_1\rho_2\cos^4\gamma + 8\rho_1\rho_2\sin^2\gamma\cos^2\gamma = 4\rho_1\rho_2(\sin^2\gamma + \cos^2\gamma)^2 = 4\rho_1\rho_2 \neq 0\,.$$

Therefore the algorithm stabilizes at the first constraint submanifold W_1. Moreover, there exists an unique solution of the dynamics, the vector field $X \in \mathfrak{X}(W_1)$ which satisfies $i_X\,\Omega_{W_1} = \mathrm{d}H_{W_1}$. In consequence, we have a unique control input which extremizes (minimizes) the objective function $\mathcal{A}$. If we take the flow $F_t\colon W_1 \to W_1$ of the solution vector field X then we have that $F_t^*\Omega_{W_1} = \Omega_{W_1}$.

3.6.4 Optimal control of underactuated systems on Lie algebroids

In the general situation, the dynamics is specified fixed a Lagrangian $L : E \to \mathbb{R}$ where $(E, [\![\cdot,\cdot]\!], \rho)$ is a Lie algebroid over a manifold Q with fiber bundle projection $\tau_E : E \to Q$.

If we take local coordinates (x^i) on Q and a local basis $\{e_A\}$ of sections of E, then we have the corresponding local coordinates (x^i, y^A) on E. Such coordinates determine the local structures functions ρ^i_A and $\mathcal{C}^C_{AB}$ and then the Euler-Lagrange equations on Lie algebroids can be written as (see §2.1)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial y^A}\right) - \rho^i_A\frac{\partial L}{\partial x^i} + \mathcal{C}^C_{AB}y^B\frac{\partial L}{\partial y^A} = 0.$$

These equations are precisely the components of the *Euler-Lagrange operator* $\mathcal{E}L : E^{(2)} \to E^*$, which locally reads

$$\mathcal{E}L = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial y^A}\right) - \rho^i_A\frac{\partial L}{\partial x^i} + \mathcal{C}^C_{AB}y^B\frac{\partial L}{\partial y^A}\right)e^A,$$

where $\{e^A\}$ is the dual basis of $\{e_A\}$ (see [57]). In terms of the Euler-Lagrange operator, the equations of motion just read

$$\mathcal{E}L = 0.$$

Now, we add controls in our picture. Assume that the controlled Euler-Lagrange equations are

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial y^A}\right)-\rho^i_A\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{AB}y^B\frac{\partial L}{\partial y^A}\right)e^A=u_ae^a, \tag{3.105}$$

where we are denoting as $\{e^A\}=\{e^a,e^\alpha\}$ the dual basis of $\{e_A\}$ and u_a are admissible control parameters. Using the basis of sections of E, equations (3.105) can be rewritten as

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial y^a}\right)-\rho^i_a\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{aB}y^B\frac{\partial L}{\partial y^C} &= u_a,\\ \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right)-\rho^i_\alpha\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{\alpha B}y^B\frac{\partial L}{\partial y^C} &= 0\end{aligned} \tag{3.106}$$

The optimal control problem consists on finding an admissible trajectory $\gamma(t)=(x^i(t),y^A(t),u(t))$ of the state variables and control inputs given initial and final boundary conditions $(x^i(0),y^A(0))$ and $(x^i(T),y^A(T))$, respectively, solving the controlled Euler-Lagrange equations (3.106) and minimizing

$$\mathcal{A}(x^i,y^A,u_a)=\int_0^T C(x^i,y^A,u_a)dt\,,$$

where $C\colon E\times U\to\mathbb{R}$ denotes the cost function.

To Solve this optimal control problem is equivalent to solve the following second-order problem:

$$\begin{aligned}&\min\ \widetilde{L}(x^i(t),y^A(t),z^A(t))\\ &\text{subject to } \Phi^\alpha(x^i(t),y^A(t),z^A(t))\,,\ \alpha=1,\ldots,m\end{aligned}$$

where $\widetilde{L},\Phi^\alpha\in C^\infty(E^{(2)})$. Here

$$\widetilde{L}(x^i(t),y^A(t),z^A(t))=C\left(x^i(t),y^A(t),F_a(x^i(t),y^A(t),z^A(t))\right)\,,$$

where C is the cost function and

$$F_a(x^i(t),y^A(t),z^A(t))=\frac{d}{dt}\left(\frac{\partial L}{\partial y^a}\right)-\rho^i_a\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{aB}y^B\frac{\partial L}{\partial y^C}\,.$$

The Lagrangian $\widetilde{L}$ is subjected to the second-order constraints:

$$\Phi^\alpha(x^i(t),y^A(t),z^A(t))=\frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right)-\rho^i_\alpha\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{\alpha B}y^B\frac{\partial L}{\partial y^C}\,,$$

which determines a submanifold $\mathcal{M}$ of $E^{(2)}$. Observe that from the constraint equations we have that

$$\frac{\partial^2 L}{\partial y^\alpha\partial y^\beta}z^\beta+\frac{\partial^2 L}{\partial y^\alpha\partial y^a}z^a-\rho^i_\alpha\frac{\partial L}{\partial x^i}+\mathcal{C}^C_{\alpha B}y^B\frac{\partial L}{\partial y^C}=0.$$

Therefore, assuming that the matrix $W_{\alpha\beta} = \left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$ is regular, we can write the equations as

$$\begin{aligned} z^\alpha &= -W^{\alpha\beta}\left(\frac{\partial^2 L}{\partial y^\beta \partial y^a} z^a - \rho^i_\beta \frac{\partial L}{\partial x^i} + \mathcal{C}^C_{\beta B} y^B \frac{\partial L}{\partial y^C}\right) \\ &= G^\alpha(x^i, y^A, z^a) \end{aligned}$$

where $W^{\alpha\beta} = (W_{\alpha\beta})^{-1}$.

Therefore, we can choose coordinates (x^i, y^A, z^a) on $\mathcal{M}$. This choose allows us to consider an intrinsic point of view, that is, to work directly on $\overline{W} = \mathcal{M} \times (\mathcal{T}^{\tau_E} E)^*$ avoiding the use of the Lagrange multipliers.

Define the restricted Lagrangian $\widetilde{L}_{\mathcal{M}}$ by $\widetilde{L}\big|_{\mathcal{M}} : \mathcal{M} \to \mathbb{R}$ and take induced coordinates $(x^i, y^A, z^a, p_A, \overline{p}_A)$ on $\overline{W}$. Applying the same procedure than in section 3.5.4 we derive the following system of equations

$$\begin{aligned} \dot{x}^i &= \rho^i_A y^A, \\ \frac{dy^a}{dt} &= z^a, \\ \frac{dy^\alpha}{dt} &= G^\alpha(x^i, y^A, z^a), \\ \frac{dp_A}{dt} &= \rho^i_A \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial x^i} - \bar{p}_\beta \frac{\partial G^\beta}{\partial x^i}\right) + \mathcal{C}^C_{AB} p_C y^B, \\ \frac{d\bar{p}_A}{dt} &= -p_A + \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^A} - \bar{p}_\beta \frac{\partial G^\beta}{\partial y^A}, \\ \bar{p}_a &= \frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial z^a} - \bar{p}_\beta \frac{\partial G^\beta}{\partial z^a}. \end{aligned}$$

To shorten the number of unknown variables involved in the previous set of equations, we can write them using as variables $(x^i, y^A, z^a, \overline{p}_\alpha)$

$$\begin{aligned} \dot{x}^i &= \rho^i_A y^A, \\ \frac{dy^\alpha}{dt} &= G^\alpha(x^i, y^A, z^a), \\ 0 &= \frac{d^2}{dt^2}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial z^a} - \overline{p}_\beta \frac{\partial G^\beta}{\partial z^a}\right) - \mathcal{C}^b_{Aa} y^A \left(\frac{d}{dt}\left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial z^b} - \overline{p}_\beta \frac{\partial G^\beta}{\partial z^b}\right]\right) \\ &\quad -\frac{d}{dt}\left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^a} - \overline{p}_\beta \frac{\partial G^\beta}{\partial y^a}\right) + \mathcal{C}^C_{Aa} y^A \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^C} - \overline{p}_\beta \frac{\partial G^\beta}{\partial y^C}\right) \\ &\quad +\rho^i_a \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial x^i} - \overline{p}_\beta \frac{\partial G^\beta}{\partial x^i}\right) - \mathcal{C}^\gamma_{Aa} y^A \frac{d\overline{p}_\gamma}{dt}. \end{aligned}$$

$$
\begin{aligned}
0 \quad = \quad & \frac{d^2\overline{p}_\alpha}{dt^2} + \mathcal{C}^\beta_{A\alpha} y^A \frac{d\overline{p}_\beta}{dt} - \mathcal{C}^C_{A\alpha} y^A \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^C} - \overline{p}_\beta \frac{\partial G^\beta}{\partial y^C} \right] \\
& - \frac{d}{dt} \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^\alpha} - \overline{p}_\beta \frac{\partial G^\beta}{\partial y^\alpha} \right] + \rho^i_\alpha \left(\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial x^i} - \overline{p}_\beta \frac{\partial G^\beta}{\partial x^i} \right) \\
& + \mathcal{C}^b_{A\alpha} y^A \left(\frac{d}{dt} \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial z^b} - \overline{p}_\beta \frac{\partial G^\beta}{\partial z^b} \right] \right) - \mathcal{C}^b_{A\alpha} y^A \left[\frac{\partial \widetilde{L}_{\mathcal{M}}}{\partial y^b} - \overline{p}_\beta \frac{\partial G^\beta}{\partial y^b} \right]
\end{aligned}
$$

If the matrix

$$
\left(\frac{\partial^2 \widetilde{L}_{\mathcal{M}}}{\partial z^a \partial z^b} \right)
$$

is regular then we can write the previous equations as an explicit system of third-order differential equations. This regularity assumption is equivalent to the condition that the constrain algorithm stops at the first constraint submanifold.

Chapter 4

Optimal control of nonholonomic mechanical systems

Many important problems in robotics, the dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics, which typically means mechanical systems with rolling constraints. Some of the important issues are trajectory tracking, dynamic stability and feedback stabilization, bifurcation and control.

As is well known, the application of tools from modern differential geometry in the fields of mechanics and control theory has caused an important progress in these research areas. For example, the study of the geometrical formulation of the nonholonomic equations of motion has led to a better comprehension of locomotion generation, controllability, motion planning, and trajectory tracking, raising new interesting questions in these subjects (see [17], [19], [20], [21], [25], [31], [34], [94], [100], [113], [141], [149] and references therein). On the other hand, there are by now many papers in which optimal control problems are addressed using geometric techniques (references [21], [90], [91], and [165] are good examples). In this context, we present a geometrical formulation of the dynamics of higher-order mechanical systems with nonholonomic constraints as a higher-order constrained systems.

In this chapter we will study optimal control problems of mechanical systems subject to nonholonomic constraints. Of a grat interest in the present chapter are the recent developments that utilize a geometric approach and in particular the theory of Lagrangian submanifolds and Lie algebroids. The class of nonholonomic systems that we study in this chapter includes, in particular, any wheeled-type vehicle, such as robots on wheels and or tracks. The fact that most of these robotic systems apply torques and forces internal to the system, which makes these system move in an undulatory fashion (see [149] and references therein for more on undulatory locomotion), without the application of any external forces, makes the system under-actuated. Hence, including under-actuated systems in our study is crucial in covering a wide range of robotic applications.

In our framework we have implicitly a reduction process, that is; after the geometric procedure applied thorough in this chapter to describe the dynamical equations for the optimal control problem we can reduce the degrees of freedom of the configuration space where is defined the lagrangian which describes the control problem. We will see how this framework can be easily extended when instead of working on TQ we consider an arbitrary Lie algebroid.

4.1 Nonholonomic mechanical systems

We shall start with a configuration space Q, which is an n-dimensional differentiable manifold with local coordinates (q^i), $1 \leq i \leq n = \dim Q$. Linear constraints on the velocities are locally given by equations of the form

$$\phi^a(q^i, \dot{q}^i) = \mu_i^a(q)\dot{q}^i = 0, \quad 1 \leq a \leq m,$$

depending, in general, on their configuration coordinates and their velocities. From an intrinsic point of view, the linear constraints are defined by a regular distribution $\mathcal{D}$ on Q of constant rank $n - m$ such that the annihilator of $\mathcal{D}$ is locally given at each point of Q by

$$\mathcal{D}_q^o = \operatorname{span}\left\{\mu^a(q) = \mu_i^a dq^i\, ; 1 \leq a \leq m\right\}$$

where the one-forms μ^a are independent at each point of Q.

The various kinds of constraints we are concerned with typically are divided in two types: *holonomic* and *nonholonomic*, depending on whether the constraint is derived from a constraint in the configuration space or not. Therefore, the dimension of the space of configurations is reduced by holonomic constraints but not by nonholonomic constraints. Thus, holonomic constraints allow a reduction in the number of coordinates of the configuration space needed to formulate a given problem (see [144]).

We will restrict ourselves to the case of nonholonomic constraints. In this case, the constraints are given by a nonintegrable distribution $\mathcal{D}$. In addition to these constraints, we need to specify the dynamical evolution of the system, usually by fixing a Lagrangian function $L\colon TQ \to \mathbb{R}$. In mechanics, the central concepts permitting the extension of mechanics from the Newtonian point of view to the Lagrangian one are the notions of virtual displacements and virtual work; these concepts were formulated in the developments of mechanics, in their application to statics. In nonholonomic dynamics, the procedure is given by the *Lagrange–d'Alembert principle*. This principle allows us to determine the set of possible values of the constraint forces from the set $\mathcal{D}$ of admissible kinematic states alone. The resulting equations of motion are

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\right]\delta q^i = 0,$$

where δq^i denotes the virtual displacements verifying

$$\mu_i^a \delta q^i = 0$$

(for the sake of simplicity, we will assume that the system is not subject to non-conservative forces). This must be supplemented by the constraint equations. By using the Lagrange multiplier rule, we obtain

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a.$$

The term on the right hand side represents the constraint force or reaction force induced by the constraints.

The functions λ_a are Lagrange multipliers which, after being computed using the constraint equations, allow us to obtain a set of second order differential equations.

Now we restrict ourselves to the case of nonholonomic mechanical systems where the Lagrangian is of mechanical type

$$L(v_q) = \frac{1}{2}\mathcal{G}(v_q, v_q) - V(q), \quad v_q \in T_qQ.$$

Here $\mathcal{G}$ denotes a Riemannian metric on the configuration space Q and $V : Q \to \mathbb{R}$ is a potential function. Locally, the metric is determined by the matrix $M = (\mathcal{G}_{ij})_{1\leq i,j\leq n}$ where $\mathcal{G}_{ij} = \mathcal{G}(\partial/\partial q^i, \partial/\partial q^j)$.

We denote by $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ the canonical projection of $\mathcal{D}$ over Q and $\Gamma(\tau_{\mathcal{D}})$ the set of sections of τ_D, which in this case is just the set of vector fields $\mathfrak{X}(Q)$ taking values on $\mathcal{D}$. If $X, Y \in \mathfrak{X}(Q)$, then $[X, Y]$ denotes the standard Lie bracket of vector fields.

Definition 4.1.1. *A nonholonomic mechanical system on a manifold Q is given by the triple $(\mathcal{G}, V, \mathcal{D})$ where $\mathcal{G}$ is a Riemannian metric on Q, specifying the kinetic energy of the system, $V : Q \to \mathbb{R}$ is a smooth function representing the potential energy and $\mathcal{D}$ a non-integrable distribution on Q representing the nonholonomic constraints.*

Remark 4.1.2. Given $X, Y \in \Gamma(\tau_{\mathcal{D}})$ that is, $X(x) \in \mathcal{D}_x$ and $Y(x) \in \mathcal{D}_x$ for all $x \in Q$, then it may happen that $[X, Y] \notin \Gamma(\tau_{\mathcal{D}})$ since $\mathcal{D}$ is nonintegrable. ⋄

We want to obtain a bracket defined for sections of $\mathcal{D}$. Using the Riemannian metric $\mathcal{G}$ we can construct two complementary projectors

$$\mathcal{P}\colon TQ \to \mathcal{D},$$
$$\mathcal{Q}\colon TQ \to \mathcal{D}^{\perp},$$

with respect to the tangent bundle orthogonal decomposition $\mathcal{D} \oplus \mathcal{D}^{\perp} = TQ$.

Therefore, given $X, Y \in \Gamma(\tau_{\mathcal{D}})$ we define the *nonholonomic bracket* $[\![\cdot,\cdot]\!] : \Gamma(\tau_{\mathcal{D}}) \times \Gamma(\tau_{\mathcal{D}}) \to \Gamma(\tau_{\mathcal{D}})$ as

$$[\![X, Y]\!] := \mathcal{P}[X, Y]$$

(see [9],[11],[75]). It is clear that this Lie bracket verifies the usual properties of a Lie bracket except the Jacobi identity.

Definition 4.1.3. *Consider the restriction of the Riemannian metric $\mathcal{G}$ to the distribution $\mathcal{D}$*

$$\mathcal{G}^{\mathcal{D}} : \mathcal{D} \times_Q \mathcal{D} \to \mathbb{R}$$

and define the Levi-Civita connection

$$\nabla^{\mathcal{G}^{\mathcal{D}}} : \Gamma(\tau_{\mathcal{D}}) \times \Gamma(\tau_{\mathcal{D}}) \to \Gamma(\tau_{\mathcal{D}})$$

determined by the following two properties:

1. $[\![X,Y]\!] = \nabla^{\mathcal{G}^{\mathcal{D}}}_X Y - \nabla^{\mathcal{G}^{\mathcal{D}}}_Y X$, *(Symmetry)*

2. $X(\mathcal{G}^{\mathcal{D}}(Y,Z)) = \mathcal{G}^{\mathcal{D}}(\nabla^{\mathcal{G}^{\mathcal{D}}}_X Y, Z) + \mathcal{G}^{\mathcal{D}}(Y, \nabla^{\mathcal{G}^{\mathcal{D}}}_X Z)$ *(Metricity).*

Let (q^i) be coordinates on Q and $\{e_A\}$ vector fields on $\Gamma(\tau_D)$ such that

$$\mathcal{D}_x = span\{e_A(x)\}, \quad x \in U \subset Q.$$

Then, we determine the *Christoffel symbols* Γ^A_{BC} of the connection $\nabla^{\mathcal{G}^{\mathcal{D}}}$ by

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{e_B} e_C = \Gamma^A_{BC}(q) e_A.$$

Definition 4.1.4. *A curve $\gamma : I \subset \mathbb{R} \to \mathcal{D}$ is admissible if there exists a curve $\sigma : I \subset \mathbb{R} \to Q$ projecting γ over Q, that is, $\tau_{\mathcal{D}} \circ \gamma = \sigma$; such that*

$$\gamma(t) = \frac{d}{dt}\sigma(t).$$

Given local coordinates on Q, (q^i) $i = 1, \ldots, n$; and $\{e_A\}$ sections on $\Gamma(\tau_{\mathcal{D}})$ such that $e_A = \rho^i_A(q)\dfrac{\partial}{\partial q^i}$ we introduce induced coordinates (q^i, y^A) on $\mathcal{D}$ where, if $e \in \mathcal{D}_x$ then $e = y^A e_A(x)$. Therefore, $\gamma(t) = (q^i(t), y^A(t))$ is admissible if

$$\dot{q}^i = \rho^i_A y^A.$$

Consider the restricted Lagrangian function $\ell : \mathcal{D} \to \mathbb{R}$,

$$\ell(v) = \frac{1}{2}\mathcal{G}^{\mathcal{D}}(v,v) - V(\tau_D(v)), \text{ with } v \in \mathcal{D}.$$

Definition 4.1.5 ([11])**.** *A solution of the nonholonomic problem is an admissible curve $\gamma : I \to \mathcal{D}$ such that*

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))) = 0.$$

Here the section $grad_{\mathcal{G}^{\mathcal{D}}} V \in \Gamma(\tau_{\mathcal{D}})$ is characterized by

$$\mathcal{G}^{\mathcal{D}}(grad_{\mathcal{G}^{\mathcal{D}}} V, X) = \rho(X)V, \quad \text{for every } X \in \Gamma(\tau_{\mathcal{D}}).$$

These equations are equivalent to the *nonholonomic equations*. Locally, are given by

$$\begin{aligned} \dot{q}^i &= \rho^i_A y^A, \\ \dot{y}^C &= -\Gamma^C_{AB} y^A y^B - (\mathcal{G}^{\mathcal{D}})^{CB} \rho^i_B \frac{\partial V}{\partial q^i}, \end{aligned}$$

where $(\mathcal{G}^{\mathcal{D}})^{AB}$ denotes the coefficients of the inverse matrix of $(\mathcal{G}^{\mathcal{D}})_{AB}$ where $\mathcal{G}^{\mathcal{D}}(e_A, e_B) = (\mathcal{G}^{\mathcal{D}})_{AB}$.

Remark 4.1.6. Observe that the last equations only depend of the coordinates (q^i, y^A) on $\mathcal{D}$. Therefore the nonholonomic equations are free of Lagrange multipliers. These equations are equivalent to the well known *Hamel equations* (see [137] for example, and reference therein). $\diamond$

4.2 Optimal control of nonholonomic mechanical systems

The purpose of this section is to study optimal control problems for a nonholonomic mechanical systems.

Definition 4.2.1. *A solution of a fully actuated nonholonomic problem is an admissible curve $\gamma : I \to \mathcal{D}$ such that*

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}}V(\tau_{\mathcal{D}}(\gamma(t))) = u^A(t)e_A(\tau_{\mathcal{D}}(\gamma(t)),$$

where u^A are the control inputs.

Locally, the last equations are written as

$$\begin{aligned} \dot{q}^i &= \rho^i_A y^A \\ \dot{y}^C &= -\Gamma^C_{AB}y^A y^B - (\mathcal{G}^{\mathcal{D}})^{CB}\rho^i_B \frac{\partial V}{\partial q^i} + u^C. \end{aligned}$$

Given a cost function

$$\begin{aligned} C \ : \ & U \times \mathcal{D} \to \mathbb{R} \\ & (u^A, q^i, y^A) \mapsto C(q^i, y^A, u^A) \end{aligned}$$

the *optimal control problem* consists on finding an admissible curve $\gamma : I \to \mathcal{D}$ solution of the fully actuated nonholonomic problem given initial and final boundary conditions on $\mathcal{D}$ and minimizing the functional

$$\mathcal{J}(\gamma(t), u(t)) := \int_0^T C(\gamma(t), u(t))dt$$

where γ is an admissible curve.

We define the submanifold $\mathcal{D}^{(2)}$ of $T\mathcal{D}$ by

$$\mathcal{D}^{(2)} := \{v \in T\mathcal{D} \mid v = \dot{\gamma}(0) \text{ where } \gamma : I \to \mathcal{D} \text{ is admissible}\}, \tag{4.1}$$

and we can choose coordinates $(x^i, y^A, \dot{y}^A)$ on $\mathcal{D}^{(2)}$ where the inclusion on $T\mathcal{D}$, $i_{\mathcal{D}^{(2)}} : \mathcal{D}^{(2)} \hookrightarrow T\mathcal{D}$, is given by

$$i_{\mathcal{D}^{(2)}}(q^i, y^A, \dot{y}^A) = (q^i, y^A, \rho^i_A y^A, \dot{y}^A).$$

Therefore, $\mathcal{D}^{(2)}$ is locally described by the constraints on $T\mathcal{D}$

$$\dot{q}^i - \rho^i_A y^A = 0.$$

Observe now that our optimal control problem is alternatively determined by a function $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ where

$$\mathcal{L}(q^i, y^A, \dot{y}^C) = C\left(q^i, y^A, \dot{y}^C + \Gamma^C_{AB}y^A y^B + (\mathcal{G}^{\mathcal{D}})^{CB}\rho^i_B \frac{\partial L}{\partial q^i}\right). \tag{4.2}$$

The following diagram summarizes the situation:

$$\begin{tikzcd}
\mathcal{D}^{(2)} \arrow[rr, hook, "i_{D^{(2)}}"] \arrow[dr, "\tau^{(2,1)}_{\mathcal{D}}"] \arrow[dd, "\mathcal{L}"'] & & T\mathcal{D} \arrow[dl, "\tau_{T\mathcal{D}}"'] \arrow[dr, "T\tau_{\mathcal{D}}"] & \\
& \mathcal{D} \arrow[rr, hook, "j"] \arrow[dr, "\tau_{\mathcal{D}}"'] & & TQ \arrow[dl, "\tau_{TQ}"] \\
\mathbb{R} & & Q &
\end{tikzcd}$$

Here $j : \mathcal{D} \to TQ$ is the canonical inclusion from $\mathcal{D}$ to TQ, $\tau^{(2,1)}_{\mathcal{D}} : \mathcal{D}^{(2)} \to \mathcal{D}$ and $\tau_{T\mathcal{D}} : T\mathcal{D} \to \mathcal{D}$ are the projections locally given by $\tau^{(2,1)}_{\mathcal{D}}(q^i, y^A, \dot{y}^A) = (q^i, y^A)$ and $\tau_{T\mathcal{D}}(q^i, y^A, v^i, \dot{y}^A) = (q^i, y^A)$, respectively. Finally, $T\tau_{\mathcal{D}} : T\mathcal{D} \to TQ$ is locally given by the mapping $(q^i, y^A, \dot{q}^i, \dot{y}^A) \mapsto (q^i, \dot{q}^i)$.

To derive the equations of motion for $\mathcal{L}$ we can use standard variational calculus for systems with constraints defining the extended Lagrangian $\widetilde{\mathcal{L}}$ (see § 4.3 for an intrinsic approach)

$$\widetilde{\mathcal{L}} = \mathcal{L} + \lambda_i(\dot{q}^i - \rho^i_A y^A)$$

and therefore the equations of motion are

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{q}^i}\right) - \frac{\partial \widetilde{\mathcal{L}}}{\partial q^i} &= \dot{\lambda}_i - \frac{\partial \mathcal{L}}{\partial q^i} + \lambda_j \frac{\partial \rho^j_A}{\partial q^i} y^A = 0, \\
\frac{d}{dt}\left(\frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{y}^A}\right) - \frac{\partial \widetilde{\mathcal{L}}}{\partial y^A} &= \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^A}\right) - \frac{\partial \mathcal{L}}{\partial y^A} + \rho^i_A \lambda_i = 0, \\
\dot{q}^i &= \rho^i_A y^A.
\end{aligned} \tag{4.3}$$

Remark 4.2.2. Our initial idea for work in $\mathcal{D}$ is given by the fact that we consider $\mathcal{D}$ as the velocity phase space of a nonholonomic system. Of course to obtain the equations of motion of a nonholonomic system is necessary more information from TQ, but it is obtained using the projection of the standard Lie bracket.

Also, we want to point out that it is possible to obtain the corresponding Levi-Civita connection in a similar way that the one given in [24]. We think that our framework clarifies the situation where the dynamics is defined and simplify the use of Lagrange multipliers techniques. $\diamond$

4.2.1 An illustrative example: the vertical coin

We want to study the optimal control of the vertical coin where we assume that the coin can not fall sideways (see [24]). The scenario for this system is the following: m denotes the mass of the coin, $J > 0$ is the inertia about the vertical axis. The position of the point of contact between the coin and the plane is denoted by $(x, y) \in \mathbb{R}^2$ and the heading direction is denoted by θ. Therefore, the configuration space is the Lie group $SE(2)$ and the configuration is given by $q = (x, y, \theta) \in SE(2)$.

The control inputs are denoted by u_1 and u_2. The first one corresponds to applied a perpendicular force to the center of mass of the coin and the second ones, is the torque applied about the vertical axis.

The constraint is given by the no slipping condition and is expressed in differential form by

$$\omega = \sin\theta dx - \cos\theta dy.$$

Therefore the constraint distribution $\mathcal{D}$ is given by

$$\mathcal{D} = \left\{ \frac{1}{J}\frac{\partial}{\partial\theta}, \frac{\cos\theta}{m}\frac{\partial}{\partial x} + \frac{\sin\theta}{m}\frac{\partial}{\partial y} \right\}.$$

That is, the distribution is given by the span of the vector fields

$$\begin{aligned} X_1(q) &= \frac{1}{J}\frac{\partial}{\partial\theta}, \\ X_2(q) &= \frac{\cos\theta}{m}\frac{\partial}{\partial x} + \frac{\sin\theta}{m}\frac{\partial}{\partial y}. \end{aligned}$$

The Lagrangian is metric on Q where the matrix associated with the metric $\mathcal{G}$ is

$$\mathcal{G} = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J \end{pmatrix}.$$

Then the Lagrangian $L : TSE(2) \to \mathbb{R}$ is given by

$$L(q, \dot{q}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{J}{2}\dot{\theta}^2.$$

The projection map $\mathcal{P} : TQ \to \mathcal{D}$ is

$$\mathcal{P}(q, \dot{q}) = \cos^2\theta dx \otimes \frac{\partial}{\partial x} + \cos\theta\sin\theta dx \otimes \frac{\partial}{\partial y} + \cos\theta\sin\theta dy \otimes \frac{\partial}{\partial x} + \sin^2\theta dy \otimes \frac{\partial}{\partial y} + d\theta \otimes \frac{\partial}{\partial\theta}.$$

Obviously the projection map $\mathcal{P}$ satisfies $\mathcal{P}(X_1) = X_1$ and $\mathcal{P}(X_2) = X_2$.

Let $q = (x, y, \theta)$ be coordinates on the base manifold $SE(2)$ and take the basis $\{X_1, X_2\}$ of sections of $SE(2)$. This basis induce adapted coordinates $(x, y, \theta, y^1, y^2) \in \mathcal{D}$. The nonholonomic bracket is given by

$$[\![\, , \,]\!] = \mathcal{P}([\cdot, \cdot]).$$

Observe now,

$$[\![X_1, X_2]\!] = \mathcal{P}[X_1, X_2] = \mathcal{P}\left(-\frac{1}{Jm}\sin\theta\frac{\partial}{\partial x} + \frac{\cos\theta}{Jm}\frac{\partial}{\partial y} \right) = 0.$$

Then, this implies that all the structure functions are zero, that is, $\mathcal{C}^A_{BC} = 0$ for all $1 \leq A, B, C \leq 2$.

We introduce the fiber map $\rho : \mathcal{D} \to TSE(2)$ locally given by

$$\begin{aligned} \rho_1^1 &= 0, \quad \rho_1^2 = 0, \quad \rho_1^3 = \frac{1}{J}, \\ \rho_2^1 &= \frac{\cos\theta}{m}, \quad \rho_2^2 = \frac{\sin\theta}{m}, \quad \rho_2^3 = 0. \end{aligned}$$

The conditions to be an admissible curve are

$$\begin{aligned} \dot{x} &= \rho_1^1 y^1 + \rho_2^1 y^2 = \frac{\cos\theta}{m} y^2, \\ \dot{y} &= \rho_1^2 y^1 + \rho_2^2 y^2 = \frac{\sin\theta}{m} y^2, \\ \dot{\theta} &= \rho_1^3 y^1 + \rho_2^3 y^2 = \frac{1}{J} y^1. \end{aligned}$$

The restricted Lagrangian function in these new adapted coordinates is rewritten as

$$\ell(x, y, \theta, y^1, y^2) = \frac{1}{2m}(y^2)^2 + \frac{1}{J}(y^1)^2.$$

The Euler-Lagrange equations, together the admissibility conditions, for this lagrangian are

$$\begin{aligned} \frac{\dot{y}_1}{J} &= 0, \quad \frac{\dot{y}_2}{m} = 0, \\ \dot{x} &= \frac{\cos\theta}{m} y_2, \quad \dot{y} = \frac{\sin\theta}{m} y_2, \quad \dot{\theta} = \frac{1}{J} y_1. \end{aligned}$$

Now, we add controls in our picture. Therefore the controlled Euler-Lagrange equations are now

$$\frac{\dot{y}_1}{J} = u_2, \quad \frac{\dot{y}_2}{m} = u_1,$$

together with

$$\dot{x} = \frac{\cos\theta}{m} y_2, \quad \dot{y} = \frac{\sin\theta}{m} y_2, \quad \dot{\theta} = \frac{1}{J} y_1.$$

The optimal control problem consists on finding an admissible curve satisfying the previous equations given initial and final conditions and minimizing the functional

$$J = \frac{1}{2} \int_0^T \left(u_1^2 + u_2^2\right) dt,$$

for the cost function $C : \mathcal{D} \times U \to \mathbb{R}$ given by

$$C(x, y, \theta, y_1, y_2, u_1, u_2) = \frac{1}{2}(u_1^2 + u_2^2).$$

This optimal control problem is equivalent to the second-order optimization problem determined by $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$, where

$$\mathcal{L}(x, y, \theta, y_1, y_2, \dot{y}_1, \dot{y}_2) = \frac{1}{2}\left(\frac{\dot{y}_1^2}{J^2} + \frac{\dot{y}_2^2}{m^2}\right).$$

Here, $\mathcal{D}^{(2)}$ is a submanifold of the vector bundle $T\mathcal{D}$ over $\mathcal{D}$ defined by

$$\mathcal{D}^{(2)} := \left\{ (x, y, \theta, y_1, y_2, \dot{x}, \dot{y}, \dot{\theta}, \dot{y}_1, \dot{y}_2) \in T\mathcal{D} \;\middle|\; \dot{x} - \frac{\cos\theta}{m} y_2 = 0, \quad \dot{y} - \frac{\sin\theta}{m} y_2 = 0, \quad \dot{\theta} - \frac{1}{J} y_1 = 0 \right\},$$

where the inclusion $i_{\mathcal{D}^{(2)}} : \mathcal{D}^{(2)} \hookrightarrow T\mathcal{D}$, is given by the map

$$i_{\mathcal{D}^{(2)}}(x, y, \theta, y_1, y_2, \dot{y}_1, \dot{y}_2) = \left(x, y, \theta, y_1, y_2, \frac{\cos\theta}{m} y_2, \frac{\sin\theta}{m} y_2, \frac{1}{J} y_1, \dot{y}_1, \dot{y}_2 \right).$$

The equations of motion for the second-order extended Lagrangian

$$\widetilde{\mathcal{L}}(x, y, \theta, y_1, y_2, \dot{x}, \dot{y}, \dot{\theta}, \dot{y}_1, \dot{y}_2, \lambda) = \mathcal{L} + \lambda_1 \left(\dot{x} - \frac{\cos\theta}{m} y_2 \right) + \lambda_2 \left(\dot{y} - \frac{\sin\theta}{m} y_2 \right) + \lambda_3 \left(\dot{\theta} - \frac{1}{J} y_1 \right)$$

are

$$\begin{aligned}
\dot{\lambda}_1 &= 0, \quad \dot{\lambda}_2 = 0, \\
\dot{\lambda}_3 &= -\frac{y_2}{m} \left(\lambda_2 \cos\theta - \lambda_1 \sin\theta \right), \\
\lambda_3 &= -\frac{\ddot{y}_1}{J}, \\
\ddot{y}_2 &= -m(\lambda_1 \cos\theta + \lambda_2 \sin\theta)
\end{aligned}$$

with

$$\dot{x} = \frac{\cos\theta}{m} y_2, \quad \dot{y} = \frac{\sin\theta}{m} y_2, \quad \dot{\theta} = \frac{1}{J} y_1.$$

The first and second equations can be integrated as $\lambda_1 = c_1$ and $\lambda_2 = c_2$ where c_1 and c_2 are constants, then the system can be rewritten as

$$\begin{aligned}
\dot{\lambda}_3 &= -\frac{y_2}{m} \left(c_2 \cos\theta - c_1 \sin\theta \right), \\
\lambda_3 &= -\frac{\ddot{y}_1}{J}, \\
\ddot{y}_2 &= -m(c_1 \cos\theta + c_2 \sin\theta).
\end{aligned}$$

Now, differentiating the second equation with respect the time and replacing into the first one the problem is reduced to solve the system

$$\begin{aligned}
\frac{\dddot{y}_1}{J} &= \frac{y_2}{m} \left(c_2 \cos\theta - c_1 \sin\theta \right), \\
\ddot{y}_2 &= -m(c_1 \cos\theta + c_2 \sin\theta),
\end{aligned}$$

with

$$\dot{x} = \frac{\cos\theta}{m} y_2, \quad \dot{y} = \frac{\sin\theta}{m} y_2, \quad \dot{\theta} = \frac{1}{J} y_1.$$

If we suppose, $\lambda_1 = 0, \lambda_2 = 0$ (that is, $c_1 = c_2 = 0$) then the system can be reduced to

$$\begin{aligned} \dddot{y}_1 &= 0, \\ \ddot{y}_2 &= 0, \end{aligned}$$

Integrating these equations and using the admissibility conditions we obtain for c_i, $i = 3, \ldots, 8$ constants,

$$\begin{aligned} \theta(t) &= \frac{c_3 t^3}{6J} + \frac{c_4 t^2}{2J} + \frac{c_5 t + c_6}{J}, \\ x(t) &= \frac{1}{m}\int_0^t \cos\left(\frac{c_3 s^3 + 3c_4 s^2 + 6c_5 s + 6c_6}{6J}\right)(c_7 s + c_8)\, ds, \\ y(t) &= \frac{1}{m}\int_0^t \sin\left(\frac{c_3 s^3 + 3c_4 s^2 + 6c_5 s + 6c_6}{6J}\right)(c_7 s + c_8)\, ds. \end{aligned}$$

Therefore the controls u_1 and u_2 are

$$u_1(t) = \frac{c_7}{m}, \quad u_2(t) = \frac{c_3 t + c_4}{J}.$$

Remark 4.2.3. The last optimal control problem was studied in [24]. The authors have been used the theory of affine connections to study the optimal control problem of underactuated nonholonomic mechanical systems. The main difference is given by the fact that with our formalism we are working on the distribution $\mathcal{D}$ itself as we have commented in Remark 4.2.2. As in [24] we impose the extra condition $\lambda_1 = \lambda_2 = 0$ and we obtain the same controls minimizing the cost function.

This is only because we want to compare the proposed method with [24]. In fact, we think that Equations 4.3 applied to the vertical coin example are equivalent to Equations (38) in [24]. Moreover, in equations (39) of [24] the authors shown that the solutions of Equations (32), (33) and (34) are included in the set of solutions of (39), but, in principle (39) includes more solutions that (32), (33) and (34). Imposing the condition of the vanishing of two Lagrange multipliers we arrive to their subset of solutions.

Another argument is related with the optimization problem for this kind of optimal control problems when typically we impose initial and final boundary conditions. Usually, initial boundary condition on $\mathcal{D}$ and final boundary condition on $\mathcal{D}$. For the vertical coin example we impose conditions $(x(0), y(0), \theta(0), y_1(0), y_2(0))$ and $(x(T), y(T), \theta(T), y_1(T), y_2(T))$. Heuristically, observe that if we transform these conditions into initial conditions we will need to take the initial condition $(x(0), y(0), \theta(0), y_1(0), y_2(0), \dot{y}_1(0), \dot{y}_2(0), \lambda_1(0), \lambda_2(0), \lambda_3(0))$ and it is not clear that some of the multipliers are zero from the very beginning.

$\diamond$

4.2.2 Optimal control of underactuated nonholonomic mechanical systems

The case of underactuated nonholonomic mechanical systems can be derived in the same way. We assume that the distribution $\mathcal{D} \subseteq TQ$ is

$$\mathcal{D} = \text{ span}\{e_a, e_\alpha\} = \text{ span}\{e_A\}$$

where e_A are sections of $\mathcal{D}$. Taking this into account the control distribution $\mathcal{D}_c \subseteq \mathcal{D}$ is just $\mathcal{D}_c = \text{ span}\{e_a\}$.

Definition 4.2.4. *A solution of an underactuated nonholonomic problem is an admissible curve $\gamma : I \subset \mathbb{R} \to \mathcal{D}$ such that*

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))) = u^a(t) e_a(\tau_{\mathcal{D}}(\gamma(t))).$$

We denote by $\{e^a, e^\alpha\}$ the dual basis of $\{e_a, e_\alpha\}$. The basis of $\mathcal{D}$ induces local coordinates (q^i, y^a, y^α) on $\mathcal{D}$, that is, if $e \in \mathcal{D}$ then $e = y^A e_A = y^a e_a + y^\alpha e_\alpha$. Therefore, an admissible curve has a local representation $\gamma(t) = (q^i(t), y^a(t), y^\alpha(t))$.

The solution of an underactuated nonholonomic problem is characterized by the admissible curves which solve

$$\begin{aligned}
\left\langle \nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))), e^a(\tau_{\mathcal{D}}(\gamma(t)) \right\rangle &= u^a(t)\\
\left\langle \nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))), e^\alpha(\tau_{\mathcal{D}}(\gamma(t)) \right\rangle &= 0.
\end{aligned}$$

The last set of equations are interpreted as constraints, therefore we can denote by $\mathcal{M} \subset \mathcal{D}^{(2)}$ the submanifold of $\mathcal{D}^{(2)}$ determined by these constraints.

Given a cost function $C : \mathcal{D} \times U \to \mathbb{R}$ the optimal control problem consists on finding an admissible curve $\gamma : I \subset \mathbb{R} \to \mathcal{D}$ solving the previous equations, given boundary conditions and extremizing the cost functional

$$\mathcal{J}(\gamma(t), u(t)) = \int_0^T C(\gamma(t), u(t)) dt.$$

The proposed nonholonomic optimal control problem is equivalent to a second-order variational problem with second-order constraints, determined by the lagrangian $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$, given in the selected coordinates by

$$\mathcal{L}(q^i, y^A, \dot{y}^A) = C\left(q^i, y^A, \dot{y}^c + \Gamma^c_{AB} y^A y^B + (\mathcal{G}^{\mathcal{D}})^{cB} \rho^i_B \frac{\partial L}{\partial q^i}\right)$$

subjected to the second-order constraints

$$\begin{aligned}
\Phi^\gamma(q^i, y^A, \dot{y}^A) &= \dot{y}^\gamma + \Gamma^\gamma_{AB} y^A y^B + (\mathcal{G}^{\mathcal{D}})^{\gamma B} \rho^i_B \frac{\partial L}{\partial q^i},\\
0 &= \dot{q}^i - \rho^i_A y^A.
\end{aligned}$$

To derive the equations of motion of this second-order variational problem with second-order constraints we can use standard variational calculus defining the extended Lagrangian with the Lagrange multipliers λ_i and $\overline{\lambda}_\alpha$ by

$$\widetilde{\mathcal{L}} := \mathcal{L} + \lambda_i(\dot{q}^i - \rho^i_A(q) y^A) + \overline{\lambda}_\gamma \Phi^\gamma(q^i, y^A, \dot{y}^\gamma).$$

Therefore the equations of motion are the admissible curves satisfying

$$
\begin{aligned}
0 &= \dot{\lambda}_i + \lambda_j \frac{\partial \rho^j_A}{\partial q^i} y^A - \frac{\partial \mathcal{L}}{\partial q^i} - \overline{\lambda}_\alpha \frac{\partial \Phi^\alpha}{\partial q^i}, \\
0 &= \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^a}\right) + \rho^i_a \lambda_i - \frac{\partial \mathcal{L}}{\partial y^a} - \overline{\lambda}_\gamma \left(\Gamma^\gamma_{aB} + \Gamma^\gamma_{Ba}\right) y^B, \\
0 &= \dot{\overline{\lambda}}_\alpha + \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^\alpha}\right) + \lambda_i \rho^i_\alpha - \frac{\partial \mathcal{L}}{\partial y^\alpha} - \overline{\lambda}_\gamma \left(\Gamma^\gamma_{\alpha B} + \Gamma^\gamma_{B\alpha}\right) y^B, \qquad (4.4) \\
0 &= \Phi^\gamma(q^i, y^A, \dot{y}^A), \\
0 &= \dot{q}^i - \rho^i_A y^A.
\end{aligned}
$$

4.3 Lagrangian submanifolds and nonholonomic optimal control problems

4.3.1 Intrinsic equations of motion for a nonholonomic mechanical control problems

Given the Lagrangian function $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$, following Theorem (1.6.11) when $N = \mathcal{D}$, we can construct the Lagrangian submanifold $\Sigma_{\mathcal{L}} = Im(d\mathcal{L}(T\mathcal{D})) \subset T^*T\mathcal{D}$. Therefore, $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ generates a Lagrangian submanifold $\Sigma_{\mathcal{L}} \subset T^*T\mathcal{D}$ of the symplectic manifold $(T^*T\mathcal{D}, \omega_{T\mathcal{D}})$ where $\omega_{T\mathcal{D}}$ is the canonical symplectic 2-form on $T^*T\mathcal{D}$.

The relationship between these spaces is summarized in the following diagram:

$$
\begin{tikzcd}
\Sigma_{\mathcal{L}} \arrow[r, hook, "i_{\Sigma_L}"] \arrow[d, "(\pi_{T^*T\mathcal{D}})|_{\Sigma_{\mathcal{L}}}"'] & T^*T\mathcal{D} \arrow[d, "\pi_{T^*T\mathcal{D}}"] \\
\mathcal{D}^{(2)} \arrow[r, hook, "i_{D^{(2)}}"] & T\mathcal{D}
\end{tikzcd}
$$

Proposition 4.3.1. *Let $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ be a second order Lagrangian. Consider the inclusion $i_{\mathcal{D}^{(2)}} : \mathcal{D}^{(2)} \to T\mathcal{D}$ and $\omega_{T\mathcal{D}}$ is the canonical symplectic 2-form in $T^*T\mathcal{D}$. Then*

$$\Sigma_{\mathcal{L}} = \{\mu \in T^*T\mathcal{D} | i^*_{\mathcal{D}^{(2)}}\mu = d\mathcal{L}\} \subset T^*T\mathcal{D}$$

is a Lagrangian submanifold of $(T^*T\mathcal{D}, \omega_{T\mathcal{D}})$.

Definition 4.3.2. *Let $\mathcal{D}$ be a non-integrable distribution, $T\mathcal{D}$ its tangent bundle and $\mathcal{D}^{(2)}$ the subbundle of $T\mathcal{D}$ defined on (4.1). A second-order nonholonomic system is a triple $(\mathcal{D}^{(2)}, \Sigma_{\mathcal{L}}, \mathcal{L})$ where $\Sigma_{\mathcal{L}} \subset T^*T\mathcal{D}$ is the Lagrangian submanifold generated by $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$.*

Consider local coordinates $(q^i, y^A, \dot{q}^i, \dot{y}^A)$ on $T\mathcal{D}$. These coordinates induce local coordinates $(q^i, y^A, \dot{q}^i, \dot{y}^A, \mu_i, \mu_A, \gamma_i, \gamma_A)$ on $T^*T\mathcal{D}$. Therefore, locally, the system is characterized by the following set of equations on $T^*T\mathcal{D}$

$$\begin{aligned} \mu_i + \gamma_j \frac{\partial \rho^j_A}{\partial q^i} y^A &= \frac{\partial \mathcal{L}}{\partial q^i}, \\ \mu_A + \gamma_j \rho^j_A &= \frac{\partial \mathcal{L}}{\partial y^A}, \\ \gamma_A &= \frac{\partial \mathcal{L}}{\partial \dot{y}^A}, \\ \dot{q}^i &= \rho^i_A y^A. \end{aligned} \tag{4.5}$$

Remark 4.3.3. Typically local coordinates on $\Sigma_{\mathcal{L}} \subset T^*T\mathcal{D}$ are $(q^i, y^A, \dot{y}^A, \gamma_i)$ where γ_i plays the role of Lagrange multipliers. $\diamond$

We define the map $\Psi : T^*T\mathcal{D} \to T^*\mathcal{D}$ as

$$\langle \Psi(\mu_{v_x}), X(x) \rangle = \langle \mu_{v_x}, X^V(v_x) \rangle,$$

where $\mu \in T^*T\mathcal{D}, v_x \in T_x\mathcal{D}$, $X(x) \in T_x\mathcal{D}$ and $X^V(v_x) \in T_{v_x}T\mathcal{D}$ is its vertical lift to v_x. Locally,

$$\Psi(q^i, y^A, \dot{q}^i, \dot{y}^A, \mu_i, \mu_A, \gamma_i, \gamma_A) = (q^i, y^A, \gamma_i, \gamma_A).$$

Definition 4.3.4. *Define the Legendre transform associated with the second-order nonholonomic system as the map $\mathbb{F}\mathcal{L} : \Sigma_{\mathcal{L}} \to T^*\mathcal{D}$ given by $\mathbb{F}\mathcal{L} = \Psi \circ i_{\Sigma_{\mathcal{L}}}$. In local coordinates, it is given by*

$$\mathbb{F}\mathcal{L}(q^i, y^A, \dot{y}^A, \gamma_i) = \left(q^i, y^A, \gamma_i, \frac{\partial \mathcal{L}}{\partial \dot{y}^A} \right).$$

The following diagram summarizes the situation

$$\begin{tikzcd} \Sigma_{\mathcal{L}} \arrow[r, hook, "i_{\Sigma_{\mathcal{L}}}"] \arrow[rr, bend right, "\mathbb{F}\mathcal{L}"'] & T^*T\mathcal{D} \arrow[r, "\Psi"] & T^*\mathcal{D} \end{tikzcd}$$

Definition 4.3.5. *We say that the second-order nonholonomic system is **regular** if $\mathbb{F}\mathcal{L} : \Sigma_{\mathcal{L}} \to T^*\mathcal{D}$ is a local diffeomorphism and **hyperregular** if $\mathbb{F}\mathcal{L}$ is a global diffeomorphism.*

From the local expression of $\mathbb{F}\mathcal{L}$ we can observe that from a direct application of the implicit function theorem we have:

Proposition 4.3.6. *The second-order nonholonomic system determined by $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ is regular if and only if the matrix $\left(\dfrac{\partial^2 \mathcal{L}}{\partial \dot{y}^A \partial \dot{y}^B} \right)$ is non singular.*

Remark 4.3.7. Observe that if the Lagrangian $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ is determined from an optimal control problem and its expression is given by (4.2) then the regularity of the matrix $\left(\dfrac{\partial^2 \mathcal{L}}{\partial \dot{y}^A \partial \dot{y}^B} \right)$ is equivalent to

$$\det \left(\frac{\partial^2 C}{\partial u^A \partial u^B} \right) \neq 0$$

for the cost function. $\diamond$

4.3.2 Hamiltonian formalism

Assume that the system is regular. Then if we denote by $p_i = \gamma_i$ and $p_A = \dfrac{\partial \mathcal{L}}{\partial \dot{y}^A}$ we can write $\dot{y}^A = \dot{y}^A(q^i, y^A, p_A)$. Define the *Hamiltonian function* $\mathcal{H} : T^*\mathcal{D} \to \mathbb{R}$ by

$$\mathcal{H}(\alpha) = \langle \alpha, \pi_{T^*T\mathcal{D}} |_{\Sigma_\mathcal{L}} \left(\mathbb{F}\mathcal{L}^{-1}(\alpha)\right)\rangle - \mathcal{L}\left(\pi_{T^*T\mathcal{D}} |_{\Sigma_\mathcal{L}} \left(\mathbb{F}\mathcal{L}^{-1}(\alpha)\right)\right)$$

where $\alpha \in T^*\mathcal{D}$ is a one-form on $\mathcal{D}$, and $\pi_{T^*T\mathcal{D}} |_{\Sigma_\mathcal{L}} : \Sigma_\mathcal{L} \to \mathcal{D}^{(2)}$ is the projection locally given by $\pi_{T^*T\mathcal{D}} |_{\Sigma_\mathcal{L}} (q^i, y^A, \dot{y}^A, \gamma_i) = (q^i, y^A, \dot{y}^A)$. Locally the Hamiltonian is given by

$$\mathcal{H}(q^i, y^A, p_i, p_A) = p_A \dot{y}^A + p_i \rho^i_A y^A - \mathcal{L}(q^i, y^A, \dot{y}^A(q^i, y^A, p_A)),$$

where we are using that

$$\mathbb{F}\mathcal{L}^{-1}(q^i, y^A, p_i, p_A) = \left(q^i, y^A, \rho^i_A, \dot{y}^A(q^i, y^A, p_A), \frac{\partial \mathcal{L}}{\partial q^i} - p_j \frac{\partial \rho^j_A}{\partial q^i} y^A, \frac{\partial \mathcal{L}}{\partial y^A} - p_j \rho^j_A, p_i, p_A\right).$$

In the next, we will see that the dynamics of the nonholonomic optimal control problem is determined by the Hamiltonian system given by the triple $(T^*\mathcal{D}, \omega_\mathcal{D}, \mathcal{H})$ where $\omega_\mathcal{D}$ is the standard symplectic $2-$form on $T^*\mathcal{D}$.

The dynamics of the optimal control problem for the second-order nonholonomic system is given by the symplectic hamiltonian dynamics determined by the dynamical equation

$$i_{X_\mathcal{H}} \omega_\mathcal{D} = d\mathcal{H}. \tag{4.6}$$

Therefore, if we look the integral curves of $X_\mathcal{H}$, this is a curve of the type $t \mapsto (\dot{q}^i(t), \dot{y}^A(t), \dot{p}_i(t), \dot{p}_A(t))$; the solutions of the nonholonomic Hamiltonian system is specified by the Hamilton's equations on $T^*\mathcal{D}$

$$\begin{aligned} \dot{q}^i &= \frac{\partial \mathcal{H}}{\partial p_i}, & \dot{y}^A = \frac{\partial \mathcal{H}}{\partial p_A}, \\ \dot{p}_i &= -\frac{\partial \mathcal{H}}{\partial q^i}, & \dot{p}_A = -\frac{\partial \mathcal{H}}{\partial y^A}; \end{aligned}$$

that is,

$$\begin{aligned} \dot{q}^i &= \rho^i_A y^A, \\ \dot{p}_i &= \frac{\partial \mathcal{L}}{\partial q^i}(q^i, y^A, \dot{y}^A(q^i, y^A, p_A)) - p_j \frac{\partial \rho^j_A}{\partial q^i} y^A, \\ \dot{p}_A &= \frac{\partial \mathcal{L}}{\partial y^A}(q^i, y^A, \dot{y}^A(q^i, y^A, p_A)) - p_j \rho^j_A. \end{aligned} \tag{4.7}$$

From equation (4.6) it is clear that the flow is preserving the symplectic $2-$form $\omega_\mathcal{D}$. Moreover, these equations are equivalent to equations given in (4.3) using the identification between the Lagrange multipliers with the variables p_i and the relation for $p_A = \dfrac{\partial \mathcal{L}}{\partial \dot{y}^A}$.

Remark 4.3.8. We want to point out that in our formalism the optimal control dynamics is deduced using a *constrained variational procedure* and equivalently it is possible to apply Hamilton-Pontryagin's principle (see [81] for example), but, in any case, this "variational procedure" implies the preservation of the symplectic 2-form, and this is reflected in the Lagrangian submanifold character. Moreover, in our case, under the regularity condition, we have seen that the Lagrangian submanifold expresses that the system can be written as a Hamiltonian system (which is obviously simplectic).

Additionally, we use the Lagrangian submanifold $\Sigma_{\mathcal{L}}$ as a way to define intrinsically the Hamiltonian side since we define the Legendre transformation using the Lagrange submanifold $\Sigma_{\mathcal{L}}$. However there exists other possibilities, for instance, in [6] the authors had used a way to define the corresponding momenta for a vakonomic system. Both are equivalent, but we thought that the our derivation is more intrinsic and geometric, that is, independent of coordinates and without Lagrange multipliers. $\diamond$

4.3.3 Example: The vertical coin (cont'd)

Recall that the constraint distribution for the vertical coin optimal control problem is given by $\mathcal{D} \subset TSE(2)$

$$\mathcal{D} = \left\{ \frac{1}{J}\frac{\partial}{\partial\theta}, \frac{\cos\theta}{m}\frac{\partial}{\partial x} + \frac{\sin\theta}{m}\frac{\partial}{\partial y} \right\}.$$

The system is regular since

$$\det\left(\frac{\partial^2\mathcal{L}}{\partial\dot{y}^A\partial\dot{y}^B}\right) = \frac{1}{m^2J^2} \neq 0.$$

Denoting by $(x, y, \theta, y_1, y_2, p_x, p_y, p_\theta, p_1, p_2)$ local coordinates on $T^*\mathcal{D}$ the dynamic of the optimal control problem for this nonholonomic system is determined by the Hamiltonian function $\mathcal{H} : T^*\mathcal{D} \to \mathbb{R}$,

$$\mathcal{H}(x, y, \theta, y_1, y_2, p_x, p_y, p_\theta, p_1, p_2) = \frac{J^2}{2}p_1^2 + \frac{m^2}{2}p_2^2 + p_x\frac{\cos\theta}{m}y^1 + \frac{p_\theta}{J}y^1 + p_y\frac{\sin\theta}{m}y^2.$$

The dynamical equations are

$$\begin{aligned}
\dot{y}_1 &= J^2p_1, \quad \dot{p}_x = 0,\\
\dot{y}_2 &= m^2p_2, \quad \dot{p}_y = 0,\\
\dot{p}_\theta &= -p_x\frac{\sin\theta}{m}y^2 + p_y\frac{\cos\theta}{m}y^2,\\
\dot{p}_1 &= = -\frac{p_\theta}{J}, \quad \dot{p}_2 = -p_x\frac{\cos\theta}{m} - p_y\frac{\sin\theta}{m}.
\end{aligned}$$

Integrating the equations $\dot{p}_x = 0$ and $\dot{p}_y = 0$ as $p_x = c_1$ and $p_y = c_2$ where c_1 and c_2 are constants the last system of differential equations is rewritten as

$$\begin{aligned}
\dot{y}_1 &= J^2p_1, \quad \dot{p}_\theta = -c_1\frac{\sin\theta}{m}y_1 + c_2\frac{\cos\theta}{m}y_2,\\
\dot{y}_2 &= m^2p_2, \quad \dot{p}_1 = -\frac{p_\theta}{J}, \quad \dot{p}_2 = -c_1\frac{\cos\theta}{m} - c_2\frac{\sin\theta}{m}.
\end{aligned}$$

If we impose the condition $c_1 = c_2 = 0$ then we obtain $\dot{p}_\theta = 0$. Thus, if we take $p_\theta = c_3$ the system can be rewritten as

$$\begin{aligned} \dot{y}_1 &= J^2 p_1, \quad \dot{p}_1 = -\frac{c_3}{J} \\ \dot{y}_2 &= m^2 p_2, \quad \dot{p}_2 = 0. \end{aligned}$$

Differentiating with respect to the time $\dot{y}_1$ and $\dot{y}_2$ and replacing the another equations we obtain

$$\begin{aligned} \dddot{y}_1 &= 0 \\ \ddot{y}_2 &= 0 \end{aligned}$$

as in the Lagrangian setting.

4.4 Extension to Lie algebroids

Now, instead of working on TQ we can consider an arbitrary Lie algebroid.

Definition 4.4.1. *A nonholonomic system on a Lie algebroid $(E, \rho_E, [\![\cdot,\cdot]\!]_E)$ over a manifold Q with bundle projection $\tau_E : E \to Q$ is determined by the following three data*

1. *a subbundle $\mathcal{D}$ of E,*
2. *a nondegenerate bundle metric $\mathcal{G}$ on E*

$$\mathcal{G} : \mathcal{D} \times_Q \mathcal{D} \to \mathbb{R},$$

3. *a smooth function $V : Q \to \mathbb{R}$.*

Using the bundle metric we can construct two complementary projectors with respect to the orthogonal decomposition $E = \mathcal{D} \oplus \mathcal{D}^\perp$,

$$\begin{aligned} \mathcal{P} &: E \to \mathcal{D}, \\ \mathcal{Q} &: E \to \mathcal{D}^\perp; \end{aligned}$$

and modifying the Lie bracket on $\Gamma(\tau_E)$, we obtain a new Lie bracket over $\Gamma(\tau_\mathcal{D})$ as

$$[\![X, Y]\!]_\mathcal{D} := \mathcal{P}[\![i_\mathcal{D}(X), i_\mathcal{D}(Y)]\!]_E,$$

where $X, Y \in \Gamma(\tau_\mathcal{D})$, $\tau_\mathcal{D} : \mathcal{D} \to Q$ is the canonical projection of $\mathcal{D}$ over Q and $i_D : \mathcal{D} \to E$ is the inclusion of the subbundle $\mathcal{D}$ on E.

Suppose that (x^i) are local coordinates on Q and $\{e_A\}$ is a local basis of the space of sections $\Gamma(\tau_\mathcal{D})$, then

$$[\![e_A, e_B]\!]_\mathcal{D} = \mathcal{C}_{AB}^C e_C, \quad \rho_\mathcal{D}(e_A) = (\rho_\mathcal{D})_A^i \frac{\partial}{\partial x^i},$$

where $\rho_{\mathcal{D}} : \mathcal{D} \to TQ$ is the restriction of ρ_E to $\mathcal{D}$. The functions $\mathcal{C}^C_{AB}, (\rho_{\mathcal{D}})^i_A \in C^\infty(Q)$ are called the *local structure functions* of $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ (see [9], [11] and [75] for example).

Also using the bundle metric we can construct a unique torsion-less connection $\nabla^{\mathcal{G}^{\mathcal{D}}}$ on $\mathcal{D}$ which is metric with respect to $\mathcal{G}$ (see [57] for the standard case on Lie algebroids). The construction mimics the classical construction of the Levi-Civita connection for a Riemannian metric on a differentiable manifold.

The *Levi-Civita connection* $\nabla^{\mathcal{G}^{\mathcal{D}}} : \Gamma(\tau_{\mathcal{D}}) \times \Gamma(\tau_{\mathcal{D}}) \to \Gamma(\tau_{\mathcal{D}})$ associated to the bundle metric $\mathcal{G}^{\mathcal{D}}$ is defined by the formula

$$\begin{array}{rcl} 2\mathcal{G}^{\mathcal{D}}(\nabla^{\mathcal{G}^{\mathcal{D}}}_X Y, Z) & = & \rho_{\mathcal{D}}(X)(\mathcal{G}^{\mathcal{D}}(Y,Z)) + \rho_{\mathcal{D}}(Y)(\mathcal{G}^{\mathcal{D}}(X,Z)) \\ & & -\rho_{\mathcal{D}}(Z)(\mathcal{G}^{\mathcal{D}}(X,Y)) + \mathcal{G}^{\mathcal{D}}(X, [\![Z,Y]\!]_{\mathcal{D}}) \\ & & +\mathcal{G}^{\mathcal{D}}(Y, [\![Z,X]\!]_{\mathcal{D}}) - \mathcal{G}^{\mathcal{D}}(Z, [\![Y,X]\!]_{\mathcal{D}}) \end{array}$$

for $X, Y, Z \in \Gamma(\tau_{\mathcal{D}})$.

Alternatively, $\nabla^{\mathcal{G}^{\mathcal{D}}}$ is determined by the properties

$$\begin{array}{l} [\![X,Y]\!]_{\mathcal{D}} = \nabla^{\mathcal{G}^{\mathcal{D}}}_X Y - \nabla^{\mathcal{G}^{\mathcal{D}}}_Y X \text{ (symmetry)} \\ \rho_{\mathcal{D}}(X)(\mathcal{G}^{\mathcal{D}}(Y,Z)) = \mathcal{G}^{\mathcal{D}}(\nabla^{\mathcal{G}^{\mathcal{D}}}_X Y, Z) + \mathcal{G}^{\mathcal{D}}(Y, \nabla^{\mathcal{G}^{\mathcal{D}}}_X Z) \text{ (metricity)}\ , \end{array}$$

These two properties allow to determine the *Christoffel symbols* associated with the connection $\nabla^{\mathcal{G}^{\mathcal{D}}}$ that satisfy

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{e_B} e_C = \Gamma^A_{BC} e_A.$$

A $\rho_{\mathcal{D}}$*-admissible curve* is a curve $\gamma : I \subseteq \mathbb{R} \longrightarrow \mathcal{D}$ such that

$$\frac{\mathrm{d}(\tau_{\mathcal{D}} \circ \gamma)}{\mathrm{d}t}(t) = \rho_{\mathcal{D}}(\gamma(t))\ .$$

Locally, if we take local coordinates (x^i) on Q and a local basis $\{e_A\}$ of sections of $\mathcal{D}$, then we have the corresponding induced coordinates (x^i, y^A) on $\mathcal{D}$, where $y^A(a)$ is the A-th coordinate of $a \in \mathcal{D}$ in the given basis. Therefore, $\gamma(t) = (x^i(t), y^A(t))$ is $\rho_{\mathcal{D}}-$admissible if, and only if

$$\dot{x}^i = (\rho_{\mathcal{D}})^i_A y^A.$$

Definition 4.4.2. *A solution of the nonholonomic problem is a $\rho_{\mathcal{D}}-$admissible curve $\gamma : I \subset \mathbb{R} \to \mathcal{D}$ such that*

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)} \gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))) = 0.$$

Here, $grad_{\mathcal{G}^{\mathcal{D}}} V$ is a section of $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ characterized by

$$\mathcal{G}^{\mathcal{D}}(grad_{\mathcal{G}^{\mathcal{D}}} V, X) = \rho_{\mathcal{D}}(X)V, \quad \text{for every } X \in \Gamma(\tau_{\mathcal{D}}).$$

Locally, the solution satisfy the *Lagrange-D'Alembert's* equations

$$\begin{array}{rcl} \dot{x}^i & = & (\rho_{\mathcal{D}})^i_A y^A, \\ \dot{y}^C & = & -\Gamma^C_{AB} y^A y^B - (\mathcal{G}^{\mathcal{D}})^{CB} (\rho_{\mathcal{D}})^i_B \dfrac{\partial V}{\partial x^i}\ , \end{array}$$

Example 4.4.3. As a particular example, we include the case of finite dimension Lie algebras $\mathfrak{g}$ (it is clear that $\mathfrak{g}$ is a Lie algebroid over a single point). Now, suppose that $(\ell, \mathfrak{D})$ is a nonholonomic Lagrangian system on $\mathfrak{g}$, where $\ell : \mathfrak{g} \to \mathbb{R}$ is a Lagrangian function defined by

$$\ell(\xi) = \frac{1}{2}\langle \mathbb{I}\xi, \xi\rangle,$$

where $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$ is a symmetric positive definite inertia operator and $\mathfrak{D}$ is a vector subspace of $\mathfrak{g}$. We have the orthogonal decomposition

$$\mathfrak{g} := \mathfrak{D} \oplus \mathfrak{D}^\perp,$$

where $\mathfrak{D}^\perp = \{\eta \in \mathfrak{g} \,|\, \langle \mathbb{I}\eta, \xi\rangle = 0 \;\forall \xi \in \mathfrak{D}\}$ and the associated orthogonal projector $\mathcal{P} : \mathfrak{g} \to \mathfrak{D}$; then the nonholonomic bracket is given by $[\![\ ,\ ]\!] = \mathcal{P}[\cdot, \cdot]$. Take now an adapted basis $\mathfrak{D} =$ span $\{e_A\}$. Then, the *Euler-Poincaré-Suslov equations* for $(\ell, \mathfrak{D})$ are

$$\dot{y}^C = -\Gamma^C_{AB} y^A y^B$$

(see for example [75]).

4.4.1 Optimal control of nonholonomic mechanical systems on Lie algebroids

Assume that the nonholonomic system determined by $(\mathcal{G}^{\mathcal{D}}, V, \mathcal{D})$ also contains some input sections $Y_1, \ldots, Y_k$ with $k \leq \operatorname{rank}\mathcal{D}$. Therefore the control distribution is given by $\mathcal{D}_{(c)} :=$ *span*$\{Y_a\}$, where $Y_a \in \Gamma(\tau_{\mathcal{D}})$. We complete $\{Y_a\}$ to be a basis of $\Gamma(\tau_{\mathcal{D}})$ as $\{Y_a, Y_\alpha\}$ and take its dual basis $\{Y^a, Y^\alpha\}$, that is, $\{Y^a, Y^\alpha\}$ is a basis of $\Gamma(\tau_{\mathcal{D}^*})$, where $\mathcal{D}^*$ is the dual space of the bundle $\mathcal{D}$ with projection $\tau_{\mathcal{D}^*} : \mathcal{D}^* \to Q$.

The *equations of motion for a nonholonomic system with input sections* are as follows (see [11])

$$\nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))) \in \mathcal{D}_{(c)}(\gamma(t)), \quad \forall\, t \in I \subseteq \mathbb{R}, \tag{4.8}$$

where $\gamma : I \subset \mathbb{R} \to \mathcal{D}$ is a $\rho_{\mathcal{D}}$-admissible curve.

In terms of the control inputs, Equation (4.8) can be rewritten as

$$\nabla^{\mathcal{G}^{D}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))) = \sum_{a=1}^{k} u^a(t) Y_a(\tau_D(\gamma(t))) \tag{4.9}$$

for some $u : I \subseteq \mathbb{R} \to \mathbb{R}^k$, playing the role of control parameters.

The solutions of an underactuated nonholonomic problem are characterized by the admissible curves which solve

$$\begin{aligned}
\left\langle \nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))), Y^a(\tau_{\mathcal{D}}(\gamma(t)))\right\rangle &= u^a(t)\\
\left\langle \nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))), Y^\alpha(\tau_{\mathcal{D}}(\gamma(t)))\right\rangle &= 0.
\end{aligned} \tag{4.10}$$

Definition 4.4.4. *The 4-tuple $(\mathcal{D}, \mathcal{G}^{\mathcal{D}}, V, \mathcal{D}_{(c)})$ is called an underactuated nonholonomic mechanical control system on a Lie algebroid.*

Given a cost function

$$\begin{aligned} C \quad : \quad & \mathcal{D} \times U \to \mathbb{R} \\ & (q^i, y^A, u^a) \mapsto C(q^i, y^A, u^a) \end{aligned}$$

the *optimal control problem* consists on finding a $\rho_{\mathcal{D}}-$admissible curve $\gamma : I \to \mathcal{D}$ solution of (4.10) given initial and final boundary conditions and minimizing the cost functional

$$\mathcal{J}(\gamma(t), u(t)) := \int_0^T C(\gamma(t), u(t))dt.$$

Consider the subbundle $\mathcal{D}^{(2)}$ of $T\mathcal{D}$

$$\mathcal{D}^{(2)} := \{v \in T\mathcal{D} \mid v = \dot{\gamma}(0) \text{ where } \gamma : I \to \mathcal{D} \text{ is admissible}\}.$$

As before, $\mathcal{D}^{(2)}$ is locally described by the vanishing of the constraints

$$\dot{q}^i - \rho^i_A y^A = 0 \text{ on } T\mathcal{D}.$$

Local coordinates on $T\mathcal{D}$ are $(x^i, y^a, y^\alpha, \dot{x}^i, \dot{y}^a, \dot{y}^\alpha)$.

The proposed underactuated nonholonomic optimal control problem is equivalent to a second-order variational problem with second-order constraints, determined by the Lagrangian $\mathcal{L} : \mathcal{D}^{(2)} \to \mathbb{R}$ given, in the selected coordinates, by

$$\mathcal{L}(q^i, y^A, \dot{y}^A) = C\left(q^i, y^A, \dot{y}^c + \Gamma^c_{AB} y^A y^B + (\mathcal{G}^{\mathcal{D}})^{cB} \rho^i_B \frac{\partial L}{\partial q^i}\right)$$

and subjected to the second-order constraints

$$\Phi^\gamma(q^i, y^A, \dot{y}^A) = \dot{y}^\gamma + \Gamma^\gamma_{AB} y^A y^B + (\mathcal{G}^{\mathcal{D}})^{\gamma B} \rho^i_B \frac{\partial L}{\partial q^i}$$

As is well know, to derive the equations of motion of this second-order variational problem with second-order constraints we can use standard variational calculus defining the extended Lagrangian

$$\widetilde{\mathcal{L}} = \mathcal{L} + \lambda_i(\dot{q}^i - \rho^i_A(q)y^A) + \overline{\lambda}_\alpha \Phi^\alpha(q^i, y^A, \dot{y}^\gamma),$$

and therefore the equations of motion are the admissible curves satisfying

$$\begin{aligned} 0 &= \dot{\lambda}_i + \lambda_j \frac{\partial \rho^j_A}{\partial q^i} y^A - \frac{\partial \mathcal{L}}{\partial q^i} - \overline{\lambda}_\alpha \frac{\partial \Phi^\alpha}{\partial q^i}, \\ 0 &= \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^a}\right) + \rho^i_a \lambda_i - \frac{\partial \mathcal{L}}{\partial y^a} - \overline{\lambda}_\gamma \left(\Gamma^\gamma_{aB} + \Gamma^\gamma_{Ba}\right) y^B, \\ 0 &= \dot{\overline{\lambda}}_\alpha + \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^\alpha}\right) + \lambda_i \rho^i_\alpha - \frac{\partial \mathcal{L}}{\partial y^\alpha} - \overline{\lambda}_\gamma \left(\Gamma^\gamma_{\alpha B} + \Gamma^\gamma_{B\alpha}\right) y^B, \\ 0 &= \Phi^\alpha(q^i, y^A, \dot{y}^\gamma), \\ \dot{q}^i &= \rho^i_A y^A. \end{aligned}$$

Remark 4.4.5. If we consider the case when the input section $\{Y^a\}$ are just a basis of $\Gamma(\tau_{\mathcal{D}})$, that is, the system is total-actuated, the solution of the total-actuated nonholonomic problem are characterized by the admissible curves which solve

$$\left\langle \nabla^{\mathcal{G}^{\mathcal{D}}}_{\gamma(t)}\gamma(t) + grad_{\mathcal{G}^{\mathcal{D}}} V(\tau_{\mathcal{D}}(\gamma(t))), Y^a(\tau_{\mathcal{D}}(\gamma(t)))\right\rangle = u^a(t).$$

Then the optimal control problem is determined by the Lagrangian

$$\mathcal{L}(q^i, y^A, \dot{y}^A) = C\left(q^i, y^A, \dot{y}^C + \Gamma^C_{AB} y^A y^B + (\mathcal{G}^{\mathcal{D}})^{CB}\rho^i_B \frac{\partial L}{\partial q^i}\right)$$

subjected to the constraint $\dot{q}^i - \rho^i_A(q)y^A = 0$.

To derive the equations of motion of this variational problem with constraints we can use standard variational calculus defining the extended Lagrangian

$$\widetilde{\mathcal{L}} := \mathcal{L} + \lambda_i(\dot{q}^i - \rho^i_A(q)y^A),$$

and therefore the equations of motion are

$$\begin{aligned} 0 &= \dot{\lambda}_i + \lambda_j \frac{\partial \rho^j_A}{\partial q^i} y^A - \frac{\partial \mathcal{L}}{\partial q^i}, \\ 0 &= \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^A}\right) + \rho^i_A \lambda_i - \frac{\partial \mathcal{L}}{\partial y^A}, \\ 0 &= \dot{q}^i - \rho^i_A y^A. \end{aligned}$$

$\diamond$

In the same way that in the previous section we can a Hamiltonian formalism for this type of systems using similar techniques.

Chapter 5

Higher-order variational systems on Lie groupoids

The topic of discrete Lagrangian mechanics concerns the study of certain discrete dynamical systems on manifolds. As the name suggests, these discrete systems exhibit many geometric features which are analogous to those in continuous Lagrangian mechanics: in particular, the dynamics of these systems satisfy variational principles, have symplectic or Poisson flow maps, conserve momentum maps associated to Noether-type symmetries, and admit a theory of reduction. While discrete Lagrangian systems are quite mathematically interesting, in their own right, they also have important applications to structure-preserving numerical simulation of dynamical systems in geometric mechanics and optimal control theory.

In this chapter, we generalize the theory of discrete higher-order Lagrangian mechanics and variational integrators in two main directions. First, we develop variational principles for second-order variational problems on Lie groupoids and we show how to apply this theory to the construction of variational integrators for optimal control problems of mechanical systems. Secondly, we show that Lagrangian submanifolds of a symplectic groupoid (cotangent groupoid) give rise to discrete dynamical second-order systems, and we study the properties of these systems, including their regularity and reversibility, from the perspective of symplectic and Poisson geometry. We also develop a theory of reduction and Noether symmetries, and study the relationship between the dynamics and variational principles for these second-order variational problems. Next, we use this framework along with a generalized notion of generating function due to Tulczyjew to develop a theory of discrete constrained Lagrangian mechanics. This allows for systems with arbitrary constraints, including those which are nonholonomic (in an appropriate discrete, variational sense). We would like to point out that our results are strongly based on the paper of J.C. Marrero, D. Martín de Diego and A. Stern [126] but in higher-order theory.

5.1 Discrete Mechanics

This section briefly reviews some key results of discrete mechanics following Marsden and West [131] and Marrero, Martínez and Martín de Diego [124], [125].

5.1.1 Discrete Lagrangian Mechanics

A discrete path $\{q_k\}_{k=0}^N$, on an n-dimensional differentiable manifold Q, can be described by the following discrete variational principle. Denote by S_d the action sum defined from $L_\mathrm{d} : Q \times Q \to \mathbb{R}$:

$$S_\mathrm{d}(\{q_k\}_{k=0}^N) = \sum_{k=0}^{N-1} L_\mathrm{d}(q_k, q_{k+1}) \approx \int_0^{t_N} L(q(t), \dot{q}(t))\, dt, \tag{5.1}$$

which is an approximation of the action integral as shown above.

Consider discrete variations $q_k \mapsto q_k + \varepsilon\, \delta q_k$, for $k = 0, 1, \ldots, N$, with $\delta q_0 = \delta q_N = 0$ and $\delta q_k \in T_{q_k} Q$ arbitrary. Then, the discrete variational principle $\delta S_\mathrm{d} = 0$ gives the discrete Euler–Lagrange equations:

$$D_2 L_\mathrm{d}(q_{k-1}, q_k) + D_1 L_\mathrm{d}(q_k, q_{k+1}) = 0. \tag{5.2}$$

This determines implicitly the discrete flow $F_{L_\mathrm{d}} : Q \times Q \to Q \times Q$:

$$F_{L_\mathrm{d}} : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1}) \tag{5.3}$$

when the matrix $(D_{12} L_d(q_k, q_{k+1}))$ is regular. Let us define the discrete Lagrangian 1-forms $\Theta^{\pm}_{L_\mathrm{d}} : Q \times Q \to T^*(Q \times Q)$ by

$$\Theta^{+}_{L_\mathrm{d}} : (q_k, q_{k+1}) \mapsto D_2 L_\mathrm{d}(q_k, q_{k+1})\, dq_{k+1}, \tag{5.4a}$$

$$\Theta^{-}_{L_\mathrm{d}} : (q_k, q_{k+1}) \mapsto -D_1 L_\mathrm{d}(q_k, q_{k+1})\, dq_k. \tag{5.4b}$$

Then, the discrete flow F_{L_d} preserves the discrete Lagrangian form

$$\Omega_{L_\mathrm{d}}(q_k, q_{k+1}) = -d\Theta^{+}_{L_\mathrm{d}} = -d\Theta^{-}_{L_\mathrm{d}} = D_1 D_2 L_\mathrm{d}(q_k, q_{k+1})\, dq_k \wedge dq_{k+1}. \tag{5.5}$$

Specifically, we have

$$(F_{L_\mathrm{d}})^* \Omega_{L_\mathrm{d}} = \Omega_{L_\mathrm{d}}.$$

5.1.2 Discrete Hamiltonian Mechanics

Introduce the *right and left discrete Legendre transformations* $\mathbb{F}L^{\pm}_\mathrm{d} : Q \times Q \to T^*Q$ by

$$\mathbb{F}L^{+}_\mathrm{d} : (q_k, q_{k+1}) \mapsto (q_{k+1}, D_2 L_\mathrm{d}(q_k, q_{k+1})), \tag{5.6a}$$

$$\mathbb{F}L^{-}_\mathrm{d} : (q_k, q_{k+1}) \mapsto (q_k, -D_1 L_\mathrm{d}(q_k, q_{k+1})), \tag{5.6b}$$

respectively. Then we find that the Eq. (5.4) and (5.5) are pull-backs by these maps of the Liouville 1-forms and the canonical symplectic 2-form on T^*Q as follows:

$$\Theta^{\pm}_{L_\mathrm{d}} = (\mathbb{F}L^{\pm}_\mathrm{d})^* \lambda_Q, \qquad \Omega^{\pm}_{L_\mathrm{d}} = (\mathbb{F}L^{\pm}_\mathrm{d})^* \omega_Q.$$

Let us define the momenta

$$p^{-}_{k,k+1} = -D_1 L_\mathrm{d}(q_k, q_{k+1}), \qquad p^{+}_{k,k+1} = D_2 L_\mathrm{d}(q_k, q_{k+1}).$$

Then, the discrete Euler–Lagrange equations (5.2) become simply $p^+_{k-1,k} = p^-_{k,k+1}$. So defining

$$p_k = p^+_{k-1,k} = p^-_{k,k+1},$$

one can rewrite the discrete Euler–Lagrange equations (5.2) as follows:

$$\begin{aligned} p_k &= -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \\ p_{k+1} &= D_2 L_{\mathrm{d}}(q_k, q_{k+1}). \end{aligned} \tag{5.7}$$

Furthermore, define the *discrete Hamiltonian map* $\tilde{F}_{L_{\mathrm{d}}} : T^*Q \to T^*Q$ by

$$\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}). \tag{5.8}$$

Then, one may relate this map with the discrete Legendre transforms in Eq. (5.6) as follows:

$$\tilde{F}_{L_{\mathrm{d}}} = \mathbb{F}L^+_{\mathrm{d}} \circ (\mathbb{F}L^-_{\mathrm{d}})^{-1}. \tag{5.9}$$

Furthermore, one can also show that this map is symplectic, i.e.,

$$(\tilde{F}_{L_{\mathrm{d}}})^* \omega_Q = \omega_Q.$$

This corresponds to the Hamiltonian description of the dynamics defined by the discrete Euler–Lagrange equation (5.2) introduced by Marsden and West in [131]. Notice, however, that no discrete analogue of Hamilton's equations is introduced here, although the flow is now on the cotangent bundle T^*Q.

Numerical methods which are constructed in this way are called variational integrators, due to the key role played by the variational principle. This approach to discretizing Lagrangian systems was put forward in seminal papers by Suris [164], Moser and Veselov [140], and others in the early 1990s, and the general theory was developed over the subsequent decade (see Marsden and West [131] for a comprehensive overview).

A. Weinstein [176] observed that these systems could be understood as a special case of a more general theory, describing discrete Lagrangian mechanics on arbitrary Lie groupoids.

5.1.3 Lie Groupoids and Discrete Mechanics

In this Section, we will review some generalities on discrete mechanics on Lie groupoids strongly based in [124] and [125].

Discrete Euler-Lagrange equations

Let G be a Lie groupoid with structural maps

$$\alpha, \beta : G \to M, \quad \epsilon : M \to G, \quad i : G \to G, \quad m : G_2 \to G.$$

Denote by $\tau : AG \to M$ the Lie algebroid of G.

A *discrete Lagrangian* is a function $L_d : G \longrightarrow \mathbb{R}$. Fixed $g \in G$, we define the set of admissible sequences with values in G:

$$\mathcal{C}_g^N = \{(g_1, \dots, g_N) \in G^N \ / \ (g_k, g_{k+1}) \in G_2 \text{ for } k = 1, \dots, N-1 \\ \text{and } g_1 \dots g_N = g\}.$$

An admissible sequence $(g_1, \dots, g_N) \in \mathcal{C}_g^N$ is a solution of *the discrete Euler-Lagrange equations* if

$$0 = \sum_{k=1}^{N-1} \left[\overleftarrow{X}_k(g_k)(L_d) - \overrightarrow{X}_k(g_{k+1})(L_d) \right], \quad \text{for } X_1, \dots, X_{N-1} \in \Gamma(\tau_{AG}).$$

For $N = 2$ we obtain that $(g, h) \in G_2$ is a solution if

$$\overleftarrow{X}(g)(L_d) - \overrightarrow{X}(h)(L_d) = 0$$

for every section X of AG.

Discrete Poincaré-Cartan sections

Given a Lagrangian function $L_d : G \longrightarrow \mathbb{R}$, we will study the geometrical properties of the discrete Euler-Lagrange equations.

Consider the vector bundle

$$\pi^{\tau_{AG}} : P^{\tau_{AG}} G = V\beta \oplus V\alpha \to G$$

where $V\alpha$ (respectively, $V\beta$) is the vertical bundle of the source map $\alpha : G \to M$ (respectively, the target map $\beta : G \to M$). Then, one may introduce a Lie algebroid structure on $\pi^{\tau_{AG}} : P^{\tau_{AG}} G = V\beta \oplus V\alpha \to G$ (see section 3 in [124] and subsection 1.9.3). The anchor map $\rho^{P^{\tau_{AG}} G} : P^{\tau_{AG}} G = V\beta \oplus V\alpha \to TG$ is given by

$$\rho^{P^{\tau_{AG}} G}(X_g, Y_g) = X_g + Y_g, \quad \text{for } (X_g, Y_g) \in V_g\beta \oplus V_g\alpha$$

and the Lie bracket $[\![\cdot, \cdot]\!]^{P^{\tau_{AG}} G}$ on the space $\Gamma(\pi^{\tau_{AG}})$ is characterized by the following relation

$$[\![(\overrightarrow{X}, \overleftarrow{Y}), (\overrightarrow{X'}, \overleftarrow{Y'})]\!]^{P^{\tau_{AG}} G} = (-\overrightarrow{[\![X, X']\!]}, \overleftarrow{[\![Y, Y']\!]}), \tag{5.10}$$

for $X, Y, X', Y' \in \Gamma(\tau_{AG})$ (see [124]).

Now, define the *Poincaré-Cartan 1-sections* $\Theta_{L_d}^-, \Theta_{L_d}^+ \in \Gamma((\pi^{\tau_{AG}})^*)$ as follows

$$\Theta_{L_d}^-(g)(X_g, Y_g) = -X_g(L_d), \qquad \Theta_{L_d}^+(g)(X_g, Y_g) = Y_g(L_d), \tag{5.11}$$

for each $g \in G$ and $(X_g, Y_g) \in V_g\beta \oplus V_g\alpha$.

If d is the differential of the Lie algebroid $\pi^{\tau_{AG}} : P^{\tau_{AG}} G = V\beta \oplus V\alpha \to G$ we have that $dL_d = \Theta_{L_d}^+ - \Theta_{L_d}^-$ and so, using $d^2 = 0$, it follows that $d\Theta_{L_d}^+ = d\Theta_{L_d}^-$. This means that there exists a unique 2-section $\Omega_{L_d} = -d\Theta_{L_d}^+ = -d\Theta_{L_d}^-$, that will be called the *Poincaré-Cartan* 2-section. This 2-section will be important to study the symplecticity of the discrete Euler-Lagrange equations.

Let X be a section of the Lie algebroid $\tau_{AG} : AG \to M$. Then, one may consider the sections $X^{(1,0)}$ and $X^{(0,1)}$ of the vector bundle $\pi^{\tau_{AG}} : P^{\tau_{AG}}G \simeq V\beta \oplus V\alpha \to G$ given by

$$X^{(1,0)}(g) = (\overrightarrow{X}(g), 0_g), \quad X^{(0,1)}(g) = (0_g, \overleftarrow{X}(g)), \quad \text{for } g \in G.$$

Moreover, if $g \in G$, $\{X_\gamma\}$ (respectively, $\{Y_\mu\}$) is a local basis of $\Gamma(\tau_{AG})$ in an open subset U (respectively, V) of M such that $\alpha(g) \in U$ (respectively, $\beta(g) \in V$) then $\{X_\gamma^{(1,0)}, Y_\mu^{(0,1)}\}$ is a local basis of $\Gamma(\pi^{\tau_{AG}})$ in $\alpha^{-1}(U) \cap \beta^{-1}(V)$ and

$$\Omega_{L_d}(X_\gamma^{(1,0)}, Y_\mu^{(1,0)}) = \Omega_{L_d}(X_\gamma^{(0,1)}, Y_\mu^{(0,1)}) = 0, \tag{5.12}$$

and

$$\Omega_{L_d}(X_\gamma^{(1,0)}, Y_\mu^{(0,1)}) = \overleftarrow{Y_\mu}(\overrightarrow{X_\gamma}(L_d)) = \overrightarrow{X_\gamma}(\overleftarrow{Y_\mu}(L_d)). \tag{5.13}$$

(for more details, see [124]).

Discrete Lagrangian evolution operator

We say that a differentiable mapping $\Psi : G \longrightarrow G$ is a *discrete flow* or a *discrete Lagrangian evolution operator for L_d* if it verifies the following properties:

- graph$(\Psi) \subseteq G_2$, that is, $(g, \Psi(g)) \in G_2$, $\forall g \in G$.
- $(g, \Psi(g))$ is a solution of the discrete Euler-Lagrange equations, for all $g \in G$, that is,

$$\overleftarrow{X}(g)(L_d) - \overrightarrow{X}(\Psi(g))(L_d) = 0 \tag{5.14}$$

  for every section X of AG and every $g \in G$.

Discrete Legendre transformations

Given a discrete Lagrangian $L_d : G \longrightarrow \mathbb{R}$ we define two *discrete Legendre transformations* $\mathbb{F}^-L_d : G \longrightarrow A^*G$ and $\mathbb{F}^+L_d : G \longrightarrow A^*G$ as follows (see [124])

$$(\mathbb{F}^-L_d)(h)(v_{\epsilon(\alpha(h))}) = -v_{\epsilon(\alpha(h))}(L_d \circ r_h \circ i), \text{ for } v_{\epsilon(\alpha(h))} \in A_{\alpha(h)}G, \tag{5.15}$$

$$(\mathbb{F}^+L_d)(g)(v_{\epsilon(\beta(g))}) = v_{\epsilon(\beta(g))}(L_d \circ l_g), \text{ for } v_{\epsilon(\beta(g))} \in A_{\beta(g)}G. \tag{5.16}$$

Remark 5.1.1. Note that $(\mathbb{F}^+L_d)(g) \in A^*_{\beta(g)}G$ and $(\mathbb{F}^-L_d)(h) \in A^*_{\alpha(h)}G$. Furthermore, if $\{X_\gamma\}$ (respectively, $\{Y_\mu\}$) is a local basis of $\Gamma(\tau)$ in an open subset U such that $\alpha(h) \in U$ (respectively, $\beta(g) \in V$) and $\{X^\gamma\}$ (respectively, $\{Y^\mu\}$) is the dual basis of $\Gamma(\tau_{A^*G})$, it follows that

$$\mathbb{F}^-L_d(h) = \overrightarrow{X}_\gamma(h)(L_d)X^\gamma(\alpha(h)), \quad \mathbb{F}^+L_d(g) = \overleftarrow{Y}_\mu(g)(L_d)Y^\mu(\beta(g)).$$

$\diamond$

Discrete regular Lagrangians

A Lagrangian $L_d : G \to \mathbb{R}$ on a Lie groupoid G is said to be *regular* if the Poincaré-Cartan 2-section Ω_{L_d} is symplectic on the Lie algebroid $\pi^{\tau_{AG}} : P^{\tau}_{AG}G \equiv V\beta \oplus_G V\alpha \to G$, that is, Ω_{L_d} is nondegenerate (see [124]).

Using (5.13), we deduce that the Lagrangian L_d is regular if and only if for every $g \in G$ and every local basis $\{X_\gamma\}$ (respectively, $\{Y_\mu\}$) of $\Gamma(\tau_{AG})$ on an open subset U (respectively, V) of M such that $\alpha(g) \in U$ (respectively, $\beta(g) \in V$) we have that the matrix $\overrightarrow{X_\gamma}(\overleftarrow{Y_\mu}(L_d))$ is regular on $\alpha^{-1}(U) \cap \beta^{-1}(V)$.

In [124], the authors have proved that the following conditions are equivalent:

- $L_d : G \to \mathbb{R}$ is a regular discrete Lagrangian function.
- The Legendre transformation $\mathbb{F}^{-}L_d$ is a local diffeomorphism.
- The Legendre transformation $\mathbb{F}^{+}L_d$ is a local diffeomorphism.

Moreover, if $L_d : G \to \mathbb{R}$ is regular and $(g_0, h_0) \in G_2$ is a solution of the discrete Euler-Lagrange equations for L_d then there exist two open subsets U_0 and V_0 of G, with $g_0 \in U_0$ and $h_0 \in V_0$, and there exists a (local) discrete Lagrangian evolution operator $\Psi_{L_d} : U_0 \to V_0$ such that:

- $\Psi_{L_d}(g_0) = h_0$,
- Ψ_{L_d} is a diffeomorphism and
- Ψ_{L_d} is unique, that is, if U_0' is an open subset of G, with $g_0 \in U_0'$ and $\Psi'_{L_d} : U_0' \to G$ is a (local) discrete Lagrangian evolution operator then $\Psi'_{L_d}|U_0 \cap U_0' = \Psi_{L_d}|U_0 \cap U_0'$.

5.2 Second-order variational problems on Lie groupoids

In this section, we will discuss discrete second-order Lagrangian mechanics using techniques of variational calculus on Lie groupoids (see [85] and [124] for first order variational calculus on Lie groupoids) and we will illustrate the results obtained in this section with some examples and applications.

5.2.1 Second-order variational principle on Lie groupoids

Let G be a Lie groupoid with structural maps

$$\alpha, \beta : G \to Q, \quad \epsilon : Q \to G, \quad i : G \to G, \quad m : G_2 \to G.$$

Denote by $\tau_{AG} : AG \to Q$ the Lie algebroid of G.

Definition 5.2.1. *A discrete second-order Lagrangian $L : G_2 \to \mathbb{R}$ is a differentiable function defined on the set of composable elements describing the dynamics of the mechanical system.*

In the following we will use the notation $G^4 := G \times G \times G \times G$. Fixed $g \in G$, we define the set of admissible sequences with values in G

$$C_g^4 = \{(g_1, g_2, g_3, g_4) \in G^4 \mid (g_k, g_{k+1}) \in G_2 \text{ for } k = 1, 2, 3$$
$$\text{with } g_1 \text{ and } g_4 \text{ fixed and } g_1 g_2 g_3 g_4 = g\}.$$

Given a tangent vector at the point $\bar{g} = (g_1, g_2, g_3, g_4)$ to the manifold C_g^4, we may write it as the tangent vector at $t = 0$ of a curve $c(t)$ in C_g^4, $t \in (-\epsilon, \epsilon) \subset \mathbb{R} \to c(t)$ which passes through $\bar{g}$ at $t = 0$. This type of curves has the form

$$c(t) = (g_1, g_2 h_2(t), h_2^{-1}(t) g_3, g_4),$$

where $h_2(t) \in \alpha^{-1}(\beta(g_2))$, for all t, and $h_2(0) = \epsilon(\beta(g_2))$. The curve c is called a *variation* of $\bar{g}$. Therefore we may identify the tangent space to C_g^4 at $\bar{g}$ with

$$T_{\bar{g}} C_g^4 \equiv \{v_2 \mid v_2 \in A_{x_2} G \text{ where } x_2 = \beta(g_2)\}.$$

The curve v_2 is called *infinitesimal variation* of $\bar{g}$ and is the tangent vector to the α-vertical curve h_2 at $t = 0$.

Now, we define the *discrete action sum* associated to the discrete second-order Lagrangian $L : G_2 \to \mathbb{R}$ as

$$\begin{array}{rcl} S_L : C_g^4 & \to & \mathbb{R} \\ \bar{g} & \mapsto & \displaystyle\sum_{k=1}^{3} L(g_k, g_{k+1}). \end{array} \tag{5.17}$$

We now procedure, to derive the discrete equations of motion applying Hamilton's principle of critical action. To do this, we need to consider the variations of the discrete action sum.

Definition 5.2.2. ***Discrete Hamilton's principle on Lie groupoids***

Given $g \in G$ an admissible sequence $\bar{g} \in C_g^4$ is a solution of the Lagrangian system determined by $L : G_2 \to \mathbb{R}$ if and only if $\bar{g}$ is a critical point of S_L.

In order to characterize the critical points, we calculate,

$$\begin{array}{rcl} \left.\dfrac{d}{dt}\right|_{t=0} S_L(c(t)) & = & \left.\dfrac{d}{dt}\right|_{t=0} \{L(g_1, g_2 h_2(t)) + L(g_2 h_2(t), h_2^{-1}(t) g_3) \\ & + & L(h_2^{-1}(t) g_3, g_4)\}. \end{array}$$

Then, the condition

$$\left.\frac{d}{dt}\right|_{t=0} S_L(c(t)) = 0 \tag{5.18}$$

is equivalent to

$$\begin{array}{rcl} 0 & = & d^\circ(L \circ \ell_{g_2})(\epsilon(\beta(g_2)))(v_2) + d^\circ(L \circ r_{g_3} \circ i)(\epsilon(\beta(g_2)))(v_2) + d^\circ(L \circ \ell_{g_2})(\epsilon(\beta(g_1)))(v_2) \\ & + & d^\circ(L \circ r_{g_3} \circ i)(\epsilon(\beta(g_3)))(v_2) \end{array} \tag{5.19}$$

where d° is the standard differential on G, that is, the differential of the Lie algebroid $\tau_{TG}: TG \to G$ and ℓ_g and r_g were defined on 1.9.3.

Then, $\bar{g}$ is a solution of the Lagrangian system determined by the discrete second-order Lagrangian $L: G_2 \to \mathbb{R}$ if and only if

$$
\begin{aligned}
0 &= d^\circ(L\circ \ell_{g_2})(\epsilon(\beta(g_2)))(v_2) + d^\circ(L\circ r_{g_3}\circ i)(\epsilon(\beta(g_2)))(v_2) + d^\circ(L\circ \ell_{g_2})(\epsilon(\beta(g_1)))(v_2)\\
&+ d^\circ(L\circ r_{g_3}\circ i)(\epsilon(\beta(g_3)))(v_2).
\end{aligned}
$$

Or alternatively, $\bar{g}$ is a solution of the Lagrangian system determined by $L: G_2 \to \mathbb{R}$ if and only if $\bar{g}$ satisfies

$$\ell_{g_2}^*\left(D_1L(g_2,g_3)+D_2L(g_1,g_2)\right) + (r_{g_3}\circ i)^*\left(D_1L(g_3,g_4)+D_2L(g_2,g_3)\right) = 0. \tag{5.20}$$

These equations will be called *discrete second-order Euler-Lagrange equations* on the Lie groupoid G.

Example 5.2.3. Let $Q\times Q \rightrightarrows Q$ be the pair groupoid. An admissible path is the 4-tuple $((q_1,q_2),(q_2,q_3),(q_3,q_4),(q_4,q_5)) \in C^4_{(q,\tilde{q})}$. The inclusion of $3Q$ into $(Q\times Q)_2$ is given by the map $(q,\tilde{q},\bar{q}) \hookrightarrow ((q,\tilde{q}),(\tilde{q},\bar{q}))$. Applying Hamilton's principle for $L:(Q\times Q)_2 \simeq 3Q \to \mathbb{R}$ we have that

$$\frac{d}{dt}\Big|_{t=0} S_L = \frac{d}{dt}\Big|_{t=0}\sum_{k=1}^{3} L(q_k,q_{k+1},q_{k+2}).$$

Then the path $((q_1,q_2),(q_2,q_3),(q_3,q_4),(q_4,q_5)) \in C^4_{(q,\tilde{q})}$ is a critical point of S_L if and only if it satisfies the difference equation

$$D_3L(q_1,q_2,q_3)+D_2L(q_2,q_3,q_4)+D_1L(q_3,q_4,q_5)=0. \tag{5.21}$$

These equations are just the *discrete second-order Euler-Lagrange equations* (See for example [14]).

Now, If we consider the case when the Lie groupoid G is a Lie group, then the discrete equations for the Lagrangian L are just

$$\ell^*_{g_k}D_1L_d(g_k,g_{k+1}) + \ell^*_{g_k}D_2L_d(g_{k-1},g_k) = r^*_{g_{k+1}}D_2L_d(g_k,g_{k+1}) + r^*_{g_{k+1}}D_1L_d(g_{k+1},g_{k+2}). \tag{5.22}$$

These equations are the *discrete second-order Euler-Poincaré equations.* See for example [53] and [32].

5.2.2 Application to optimal control of mechanical systems subject to external forces

In the general situation, the dynamics is specified fixed a lagrangian $L: AG \to \mathbb{R}$; where AG is the Lie algebroid associated with a Lie groupoid G over Q. The external forces are modeled, in this case, by curves $u_F: \mathbb{R} \to A^*G$.

It is possible to adapt the derivation of the Lagrange-d'Alembert principle to the case of total actuated mechanical systems defined on Lie algebroids (see [57] and [136]). Let $q_0, q_T \in Q$, then consider an admissible curve on AG, $\xi : I \subset \mathbb{R} \to AG$; which satisfies the principle:

$$0 = \delta \int_0^T L(\xi(t))dt + \int_0^T \langle u_F, \eta \rangle dt,$$

where $\eta \in \Gamma(\tau_{AG})$ and $u_F(t) \in A^*G$ defines the control force that we are assuming that are arbitrary (fully-actuated case). The infinitesimal variations are $\delta\xi = \eta^C$, for all time-dependent sections $\eta \in \Gamma(\tau_{AG})$, with $\eta(0) = 0$ and $\eta(T) = 0$; where η^C is a time-dependent vector field on AG, the *complete lift;* locally defined by

$$\eta^C = \rho^i_\alpha \eta^\alpha \frac{\partial}{\partial x^i} + (\dot{\eta} + \mathcal{C}^\alpha_{\beta\gamma}\eta^\beta y^\gamma)\frac{\partial}{\partial y^\alpha};$$

where we choose coordinates (q^i) on Q, and fixed a basis of sections $\{e_\alpha\}$ of $\tau_{AG} : AG \to Q$ we have induced coordinates (x^i, y^α) on AG (see [57], [87], [132] and [134]).

From this variational principle we can derive the controlled Euler-Lagrange equations

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) - \rho^i_\alpha \frac{\partial L}{\partial x^i} + \mathcal{C}^\gamma_{\alpha\beta} y^\beta \frac{\partial L}{\partial y^\gamma} &= (u_F)_\alpha, \\ \frac{dx^i}{dt} &= \rho^i_\alpha y^\alpha. \end{aligned}$$

The control force u_F is chosen in such a way it minimizes the cost functional

$$\int_0^T C(x^i, y^\alpha, (u_F)_\alpha)dt,$$

where $C : AG \oplus A^*G \to \mathbb{R}$ is the cost functional.

We define the second-order lagrangian (see section 3.6) $\tilde{L} : A^{(2)}G \to \mathbb{R}$ as

$$\tilde{L}(x^i, y^\alpha, \dot{y}^\alpha) = \widetilde{C}\left(x^i, y^\alpha, \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) - \rho^i_\alpha \frac{\partial L}{\partial x^i} + \mathcal{C}^\gamma_{\alpha\beta} y^\beta \frac{\partial L}{\partial y^\gamma}\right).$$

Here $A^{(2)}G$ denotes the set of admissible elements of the Lie algebroid AG.

Also, in a more intrinsic way, we can define the Lagrangian $\widetilde{L} : A^{(2)}G \to \mathbb{R}$ as

$$\widetilde{L} = C \circ (\tau^{A^{(2)}G}_{AG} \oplus \mathcal{EL}(L)) : A^{(2)}G \to \mathbb{R},$$

where $\mathcal{EL}(L) : A^{(2)}G \to A^*G$ is the Euler-Lagrange operator which locally reads as

$$\mathcal{EL}(L) = \left(\frac{d}{dt}\frac{\partial L}{\partial y^\alpha} - \rho^i_\alpha \frac{\partial L}{\partial x^i} + \mathcal{C}^\gamma_{\alpha\beta} y^\beta \frac{\partial L}{\partial y^\gamma}\right) e^\alpha.$$

Here $\{e^\alpha\}$ is the dual basis of $\{e_\alpha\}$, the basis of sections of AG and $\tau^{A^{(2)}G}_{AG} : A^{(2)}G \to AG$ is the canonical projection between $A^{(2)}G$ and AG given by the map $A^{(2)}G \ni (x^i, y^\alpha, v^\alpha) \mapsto (x^i, y^\alpha) \in AG$.

The optimal control problem consists on finding an admissible trajectory of the state variables and controls input given initial and final boundary conditions, solving the controlled Euler-Lagrange equations and minimizing the cost function.

Now we will state the discrete problem. Let $L_d : G \to \mathbb{R}$ be an approximation of the continuous Lagrangian,

$$L_d(g_k) = \int_{kh}^{(k+1)h} L(\xi(t))dt$$

where $h > 0$ is the time step with $T = Nh$. The discrete Euler-Lagrange equations with controls are given by

$$\ell_{g_k}^* dL_d(g_k) - (r_{g_{k+1}} \circ i)^* dL_d(g_{k+1}) = u_k \quad \in A^*_{\beta(g_k)}G, \tag{5.23}$$

for all k, where g_0 and g_N are fixed.

The discrete optimal control problem is determined prescribing the discrete cost functional $C_d : G_\beta \times_{\tau_{A^*G}} A^*G \to \mathbb{R}$,

$$\mathcal{J}_d(g_0, g_1, \dots, g_N) := \sum_{k=0}^{N-1} C_d(g_k, u_k), \tag{5.24}$$

for $(g_0, g_1, \dots, g_N) \in G^{N+1}$, $(g_k, g_{k+1}) \in G_2$, $k = 0, \dots, N-1$; g_0, g_1, g_{N-1}, g_N and $g = g_0 g_1 \dots g_N \in G$ are fixed points in G and satisfy the equations (5.23). Here $G_\beta \times_{\tau_{A^*G}} A^*G := \{(g, u) \in G \times A^*G \mid \beta(g) = \tau_{A^*G}(u)\}$.

We define the discrete second order lagrangian $\tilde{L}_d : G_2 \to \mathbb{R}$ as

$$\tilde{L}_d(g_k, g_{k+1}) := C\left(g_k, \ell_{g_k}^* dL_d(g_k) - (r_{g_{k+1}} \circ i)^* dL_d(g_{k+1})\right). \tag{5.25}$$

Thus, the discrete optimal control problem consists on find a path $(g_0, g_1, \dots, g_N) \in G^{N+1}$ such that minimize the discrete action sum $\mathcal{J}_d$ for the discrete second-order Lagrangian $\tilde{L}_d : G_2 \to \mathbb{R}$ where g_0, g_1, g_{N-1}, g_N and $g = g_0 g_1 \dots g_N \in G$ are fixed points in G.

By discrete Hamilton's principle (5.2.2) the path which minimize $\mathcal{J}_d$ subject fixed points $g_0, g_1, g_{N-1}, g_N \in G$ satisfy the discrete second order Euler-Lagrange equations for $\tilde{L}_d : G_2 \to \mathbb{R}$ given by

$$\begin{aligned} 0 &= \ell_{g_k}^* \left(D_1 \tilde{L}_d(g_k, g_{k+1}) + D_2 \tilde{L}_d(g_{k-1}, g_k)\right) \\ &+ (r_{g_{k+1}} \circ i)^* \left(D_1 \tilde{L}_d(g_{k+1}, g_{k+2}) + D_2 \tilde{L}_d(g_k, g_{k+1})\right). \end{aligned} \tag{5.26}$$

Optimal control of a rigid body on SO(3)

We consider the optimal control problem of a rigid body on the Lie groupoid $SO(3)$ over the 3×3 identity matrix Id. The Lie groupoid structure is given by

$$\alpha(R) = Id, \quad \beta(R) = Id, \quad \epsilon(Id) = Id, \quad i(R) = R^{-1} \text{ and } m(RG) = RG$$

for $R, G \in SO(3)$. The Lie algebroid associated with the Lie groupoid $SO(3)$ is the Lie algebra $\mathfrak{so}(3)$.

The equations of motion of the controlled rigid body are

$$\begin{aligned} \dot{\Omega}_1 &= P_1\Omega_2\Omega_3 + u_1, \\ \dot{\Omega}_2 &= P_2\Omega_1\Omega_3 + u_2, \\ \dot{\Omega}_3 &= P_3\Omega_1\Omega_2 + u_3, \end{aligned} \tag{5.27}$$

where $(\Omega_1, \Omega_2, \Omega_3) = \Omega \in \mathbb{R}^3$ and $(\dot{\Omega}_1, \dot{\Omega}_2, \dot{\Omega}_3) = \dot{\Omega} \in \mathbb{R}^3$, u_i are the control inputs or torques and $P_i \in \mathbb{R}$ $i = 1, 2, 3$ are given by $P_1 = \frac{I_1}{I_2 - I_3}$, $P_2 = \frac{I_2}{I_3 - I_1}$, $P_3 = \frac{I_3}{I_1 - I_2}$, where I_1, I_2, I_3 are the moments of inertia of the body. In the following we will use the typical identification of the Lie algebra of $SO(3), \mathfrak{so}(3)$ with $\mathbb{R}^3$ by the hat map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ (see [82] for example), and with some abuse of notation, we will directly identify $\mathbb{R}^3$ with $\mathfrak{so}(3)$ by omitting the hat notation.

Our fixed boundary conditions are $(R(0), \Omega(0))$ and $(R(T), \Omega(T))$, where $R(t) \in SO(3)$ is the attitude of the rigid body subject to the conditions $\dot{R} = R\Omega$ and $\delta R = R\eta$, with η an arbitrary curve on $\mathfrak{so}(3)$. Besides the equations, the cost functional is

$$\mathcal{C} = \frac{1}{2}\int_0^T \left(u_1^2 + u_2^2 + u_3^3\right)\, dt,$$

From eqs. (5.27) we can work out u in terms of Ω and $\dot{\Omega}$. Consequently, we can define the function $\ell : \mathfrak{so}(3) \times \mathfrak{so}(3) \to \mathbb{R}$ in the following way

$$\ell(\Omega, \dot{\Omega}) = \frac{1}{2} u(\Omega, \dot{\Omega}) \cdot u(\Omega, \dot{\Omega})$$

where $u = (u_1, u_2, u_3)$. Therefore, the Lagrangian function has the following form:

$$\begin{aligned} \ell(\Omega, \dot{\Omega}) &= \frac{1}{2}\left(\dot{\Omega}_1 - P_1\Omega_2\Omega_3\right)^2 + \frac{1}{2}\left(\dot{\Omega}_2 - P_2\Omega_1\Omega_3\right)^2 + \\ &+ \frac{1}{2}\left(\dot{\Omega}_3 - P_3\Omega_1\Omega_2\right)^2. \end{aligned} \tag{5.28}$$

With this redefinition, the cost functional becomes

$$\mathcal{C} = \int_0^T \ell(\Omega, \dot{\Omega})\, dt.$$

We discretize the Lagrangian ℓ as $\tilde{L} : \mathfrak{so}(3) \times \mathfrak{so}(3) \to \mathbb{R}$ by

$$\begin{aligned} \frac{\xi_k + \xi_{k+1}}{2} &\simeq \Omega(kh) \\ \xi_{k+1} &\simeq \xi_k + h\dot{\Omega}(kh) \end{aligned}$$

that is,

$$\tilde{L}(\xi_k, \xi_{k+1}) = \ell\left(\frac{\xi_k + \xi_{k+1}}{2}, \frac{\xi_{k+1} - \xi_k}{h}\right)$$

where $\xi_k, \xi_{k+1} \in \mathfrak{so}(3)$ and $h > 0$ is a fixed real number.

Now, we want to derive the discrete associated optimal control problem, then, we need to minimize the cost function associated with $L_d : SO(3)\times SO(3)\to\mathbb{R}$ where

$$L_d(w_k,w_{k+1}) = h\tilde{L}(\text{cay}^{-1}(w_k),\text{cay}^{-1}(w_{k+1})), \tag{5.29}$$

$(w_k,w_{k+1})\in SO(3)\times SO(3)$ and $\text{cay}:\mathfrak{so}(3)\to SO(3)$ denote the Cayley map (see Appendix C);

$$\begin{aligned} h\xi_k &= \text{cay}^{-1}(w_k)\in\mathfrak{so}(3) && (5.30)\\ h\xi_{k+1} &= \text{cay}^{-1}(w_{k+1})\in\mathfrak{so}(3) && (5.31)\end{aligned}$$

Therefore, we have the following discrete Lagrangian

$$L_d(w_k,w_{k+1}) = h\ell\left(\frac{\text{cay}^{-1}(w_k)+\text{cay}^{-1}(w_{k+1})}{2h},\frac{\text{cay}^{-1}(w_{k+1})-\text{cay}^{-1}(w_k)}{h^2}\right). \tag{5.32}$$

The geometric integrator is given by discrete Hamilton's principle (5.2.2) for $L_d : SO(3)\times SO(3)\to\mathbb{R}$ minimizing the cost function

$$\mathcal{C}_d = \sum_{k=0}^{N-1} L_d(w_k,w_{k+1}). \tag{5.33}$$

Instead of the discrete sum 5.33 we will take

$$\mathcal{C}_d = \sum_{k=0}^{N-1}\ell\left(\frac{\xi_k+\xi_{k+1}}{2},\frac{\xi_{k+1}-\xi_k}{h}\right). \tag{5.34}$$

where we take variations of $\xi_k = \text{cay}^{-1}(w_k)\in\mathfrak{so}(3)$. These variations are (see for example [99])

$$\begin{aligned}\delta\xi_k &= \frac{1}{h}\left(\text{Ad}_{w_k}\eta_{k+1}-\eta_k+\frac{h}{2}\text{ad}_{\xi_k}\eta_k-\frac{h}{2}\text{ad}_{\xi_k}(\text{Ad}_{w_k}\eta_{k+1})+\frac{h^2}{4}\xi_k\eta_k\xi_k\right.\\ &- \left.\frac{h^2}{4}\xi_k(\text{Ad}_{w_k}\eta_{k+1})\xi_k\right). \end{aligned}\tag{5.35}$$

Therefore,

$$\begin{aligned}\delta\mathcal{C}_d &= \frac{1}{2h}D_1\ell\left(\frac{\xi_k+\xi_{k+1}}{2h},\frac{\xi_{k+1}-\xi_k}{h^2}\right)\delta\xi_k-\frac{1}{h^2}D_2\ell\left(\frac{\xi_k+\xi_{k+1}}{2h},\frac{\xi_{k+1}-\xi_k}{h^2}\right)\delta\xi_k\\ &+ \frac{1}{2h}D_1\ell\left(\frac{\xi_k+\xi_{k-1}}{2h},\frac{\xi_k-\xi_{k-1}}{h^2}\right)\delta\xi_k+\frac{1}{h^2}D_2\ell\left(\frac{\xi_k+\xi_{k-1}}{2h},\frac{\xi_k-\xi_{k-1}}{h^2}\right)\delta\xi_k.\end{aligned}$$

Now, if we denote by

$$\begin{aligned}W_1(\xi_{k-1},\xi_k) &= D_1\ell\left(\frac{\xi_k+\xi_{k-1}}{2h},\frac{\xi_k-\xi_{k-1}}{h^2}\right), W_2(\xi_{k-1},\xi_k) = D_2\ell\left(\frac{\xi_k+\xi_{k-1}}{2h},\frac{\xi_k-\xi_{k-1}}{h^2}\right),\\ \widetilde{W}_1(\xi_k,\xi_{k+1}) &= D_2\ell\left(\frac{\xi_k+\xi_{k+1}}{2h},\frac{\xi_{k+1}-\xi_k}{h^2}\right), \widetilde{W}_2(\xi_k,\xi_{k+1}) = D_2\ell\left(\frac{\xi_k+\xi_{k+1}}{2h},\frac{\xi_{k+1}-\xi_k}{h^2}\right),\end{aligned}$$

after straightforward computations, we finally arrive to the algorithm

$$
\begin{aligned}
&\frac{1}{2}\Bigg(\mathrm{Ad}^*_{w_{k-1}}\widetilde{W}_1(\xi_{k-1},\xi_k)-\widetilde{W}_1(\xi_k,\xi_{k+1})+\frac{h}{2}\mathrm{ad}^*_{\xi_k}\widetilde{W}_1(\xi_k,\xi_{k+1})+\frac{h^2}{4}\xi_k^*\widetilde{W}_1(\xi_k,\xi_{k+1})\xi_k^*\\
&-\frac{h}{2}\mathrm{Ad}^*_{w_{k-1}}\mathrm{ad}^*_{\xi_{k-1}}\widetilde{W}_1(\xi_{k-1},\xi_k)-\frac{h^2}{4}\mathrm{Ad}^*_{w_{k-1}}\xi_{k-1}^*\widetilde{W}_1(\xi_{k-1},\xi_k)\xi_{k-1}^*\Bigg)-\\
&\frac{1}{h}\Bigg(-\frac{h}{2}\mathrm{Ad}^*_{w_{k-1}}\mathrm{ad}^*_{\xi_{k-1}}\widetilde{W}_2(\xi_{k-1},\xi_k)-\widetilde{W}_2(\xi_k,\xi_{k+1})-\frac{h^2}{4}\mathrm{Ad}^*_{w_{k-1}}(\xi_{k-1}^*\widetilde{W}_2(\xi_{k-1},\xi_k)\xi_{k-1}^*)+\\
&\frac{h}{2}\xi_k^*\widetilde{W}_2(\xi_k,\xi_{k+1})\xi_k^*+\frac{h}{2}\mathrm{ad}^*_{\xi_k}\widetilde{W}_2(\xi_k,\xi_{k+1})+\mathrm{Ad}_{w_{k-1}}\widetilde{W}_2(\xi_{k-1},\xi_k)\Bigg)+\\
&\frac{1}{2}\Bigg(\mathrm{Ad}^*_{w_{k-1}}W_1(\xi_{k-2},\xi_{k-1})-W_1(\xi_{k-1},\xi_k)+\frac{h}{2}\mathrm{ad}^*_{\xi_k}W_1(\xi_{k-1},\xi_k)+\frac{h^2}{4}\xi_k^*W_1(\xi_{k-1},\xi_k)\xi_k^*-\\
&\frac{h}{2}\mathrm{Ad}^*_{w_{k-1}}\mathrm{ad}^*_{\xi_{k-1}}W_1(\xi_{k-2},\xi_{k-1})-\frac{h^2}{4}\mathrm{Ad}^*_{w_{k-1}}\xi_{k-1}^*W_1(\xi_{k-2},\xi_{k-1})\xi_{k-1}^*\Bigg)\\
&\frac{1}{h}\Bigg(-\frac{h}{2}\mathrm{Ad}^*_{w_{k-1}}\mathrm{ad}^*_{\xi_{k-1}}W_2(\xi_{k-2},\xi_{k-1})-W_2(\xi_{k-1},\xi_k)-\\
&\frac{h^2}{4}\mathrm{Ad}^*_{w_{k-1}}(\xi_{k-1}^*W_2(\xi_{k-2},\xi_{k-1})\xi_{k-1}^*)+\mathrm{Ad}_{w_{k-1}}W_2(\xi_{k-2},\xi_{k-1})\\
&\frac{h^2}{4}\xi_k^*\widetilde{W}_2(\xi_k,\xi_{k+1})\xi_k^*+\frac{h}{2}\mathrm{ad}^*_{\xi_k}W_2(\xi_{k-1},\xi_k)\Bigg)=0 \text{ for } k=2,...,N,
\end{aligned}
\tag{5.36}
$$

$$
R_{k+1}=R_kw_k \qquad\qquad k=0,...,N-1, \tag{5.37}
$$

subjet to boundary conditions (R_0,ξ_0) and (R_N,ξ_{N-1}).

Here Ad^*_g and ad^*_ξ are the adjoint of the usual $\mathrm{Ad}_g:\mathfrak{g}\to\mathfrak{g}$, $\mathrm{ad}_\xi:\mathfrak{g}\to\mathfrak{g}$ operations, $\xi^*\omega\xi^*\in\mathfrak{g}^*$ is defined such that $\langle\xi^*\omega\xi^*,\eta\rangle=\langle\omega,\xi\eta\xi\rangle$ for $\omega\in\mathfrak{g}^*$, $\xi,\eta\in\mathfrak{g}$ and $\langle\cdot,\cdot\rangle$ is the natural pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$.

• *Boundary conditions*: From our discretization choice $R_{k+1}=R_kw_k$, is clear that fixing ξ_k implies constraints in the neighboring points, in this case R_{k+1} and R_k. If we allow ξ_N, that means constraints at the points R_N and R_{N+1}. Since we only consider time points up to $t=Nh$, having a constraint in the beyond-terminal configuration point R_{N+1} makes no sense. Hence, to ensure that the effect of the constraint on Ω is correctely accounted for, the set of unknown algebra points (*velocities*) must be reduced to $\xi_{0:N-1}$. Moreover, we can set $\xi_0=\Omega(0)$, which reduces again, since $\Omega(0)$ is fixed, the unknown velocities to $\xi_{1:N-1}$. We will discuss more about that in Chapter 6 giving an alternative way to study higher-order problems.

The boundary condition $R(T)$ is enforced by the relation $\mathrm{cay}^{-1}(R_N^{-1}R(T))=0$. Recalling that $\mathrm{cay}(0)=e$, this last expression just means that $R_N=R(T)$. Moreover, it is possible to translate it in terms of ξ_k such that there is no need to optimize over any of the configurations R_k. In that sense, (5.36) together with

$$
\mathrm{cay}^{-1}\left(w_{N-1}^{-1}...w_0^{-1}R_0^{-1}R(T)\right)=0,
$$

form a set of $3(N-1)$ equations (since dim $(\mathfrak{so}(3)) = 3$) for the $3(N-1)$ unknowns $\xi_{1:N-1}$. Consequently, the optimal control problem has become a nonlinear root finding problem. From the set of velocities $\xi_{0:N-1}$ and boundary conditions $(R(0), R(T))$, we are able to reconstruct the configuration trajectory by means of the reconstruction equation $R_{k+1} = R_k w_k$.

5.2.3 Second-order mechanical systems on Lie groupoids subject to constraints

Let $L : G_2 \to \mathbb{R}$ be a discrete second-order Lagrangian describing the dynamics of a mechanical system. We suppose that the dynamics is restricted. This restriction is given by the vanishing of m smooth constraints functions $\Phi^\alpha : G_2 \to \mathbb{R}$, $\alpha = 1, \ldots, m$. Then one can consider the augmented Lagrangian $\widehat{L} : G_2 \times \mathbb{R}^m \to \mathbb{R}$

$$\widehat{L}(g_k, g_{k+1}) = L(g_k, g_{k+1}) + \lambda_\alpha \Phi^\alpha(g_k, g_{k+1})$$

where λ_α takes the roll of Lagrange multipliers (see subsection 5.3.4 for an intrinsic approach).

An easy adaptation of variational principle (5.26) can be done to obtain the discrete second-order Euler-Lagrange equations for systems subject to second-order constraints only changing L by $\widehat{L}$. The resulting equations are:

$$\begin{aligned}
0 &= \Phi^\alpha(g_k, g_{k+1}), \text{ for all } \alpha = 1, \ldots, m, \quad k = 1, \ldots N-1; \\
0 &= \ell^*_{g_k} \left(D_1 L_d(g_k, g_{k+1}) + (\lambda_k)_a D_1 \Phi^a(g_k, g_{k+1}) + D_2 L_d(g_{k-1}, g_k) + (\lambda_{k-1})_a D_2 \Phi^a(g_{k-1}, g_k)\right) \\
&\quad + (r_{g_{k+1}} \circ i)^* \left(D_1 L_d(g_{k+1}, g_{k+2}) + (\lambda_{k+1})_a D_1 \Phi^a(g_{k+1}, g_{k+2})\right. \\
&\quad \left. + D_2 L_d(g_k, g_{k+1}) + (\lambda_k)_a D_2 \Phi^a(g_k, g_{k+1})\right), \text{ for } k = 2, \ldots, N-2.
\end{aligned}$$

In the following we will use our second-order variational calculus with second-order constraints on Lie groupoids to design variational integrators to solve optimal control problems for underactuated mechanical system.

Example: optimal control of a heavy top with two internal rotors

Now, we apply the previous theory to the optimal control of the upright spinning of the heavy top (see [46] and reference therein) seen as a second-order problem with second-order constraints.

First, we describe the heavy top with two rotors. Consider the top with two rotors so that each rotor's rotation axis is parallel to the first and the second principal axes of the top. Let I_1, I_2, I_3 be the moments of inertia of the top in the body fixed frame. Let J_1, J_2 be the moments of inertia of the rotors around their rotation axes and J_{i1}, J_{i2}, J_{i3} be the moments of inertia of the i-th rotor with $i = 1, 2$ around the first, the second and the third principal axes, respectively. Also we define the quantities $\bar{I}_1 = I_1 + J_{11} + J_{21}$, $\bar{I}_2 = I_2 + J_{12} + J_{22}$, $\bar{I}_3 = I_3 + J_{13} + J_{23}$, $\lambda_1 = \bar{I}_1 + J_1$ and $\lambda_2 = \bar{I}_2 + J_2$.

Let M be the total mass of the system, g the magnitude of the gravitational acceleration and h the distance from the origin to the center of mass of the system.

The system is modeled on the transformation Lie algebroid $E = \mathbb{S}^2 \times \mathfrak{so}(3) \times T(\mathbb{S}^1 \times \mathbb{S}^1)$ over $\mathbb{S}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ where the anchor map $\rho : \mathbb{S}^2 \times \mathfrak{so}(3) \times T(\mathbb{S}^1 \times \mathbb{S}^1) \to T(\mathbb{S}^2 \times \mathbb{S}^1 \times \mathbb{S}^1)$ is given by

$$\rho(\Gamma, \Omega, \theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = (\Gamma, \theta_1, \theta_2, \Gamma \times \Omega, \dot{\theta}_1, \dot{\theta}_2).$$

Here $\Omega = (\Omega_1, \Omega_2, \Omega_3) \in \mathfrak{so}(3) \simeq \mathbb{R}^3$ is the angular velocity of the top in the body fixed frame, $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3)$ represents the unit vector with the direction opposite to the gravity as seen from the body and $\theta = (\theta_1, \theta_2)$ is the rotation angle of rotors around their axes. $\Gamma \times \Omega \in T\mathbb{S}^2$ where $\times$ denotes the cross product.

If we denote by E_i, $i = 1, 2, 3$ the standard basis of matrices of $\mathfrak{so}(3)$,

$$E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

then the basis of sections of E is given by the elements $X^{E_i}(\Gamma, \theta_1, \theta_2) = (\Gamma, E_i, \theta_1, \theta_2, 0, 0)$; $X^{\theta_1}(\Gamma, \theta_1, \theta_2) = (\Gamma, 0, \theta_1, \theta_2, 1, 0)$, $X^{\theta_1}(\Gamma, \theta_1, \theta_2) = (\Gamma, 0, \theta_1, \theta_2, 0, 1)$ with $i = 1, 2, 3$. Finally, the Lie bracket of sections of E is determined by $[\![X^{E_1}, X^{E_2}]\!] = X^{[E_1,E_2]} = X^{E_3}$, $[\![X^{E_1}, X^{E_3}]\!] = X^{[E_1,E_3]} = X^{E_2}$, $[\![X^{E_2}, X^{E_3}]\!] = X^{[E_2,E_3]} = X^{E_1}$, $[\![X^{\theta_r}, X^{\theta_j}]\!] = 0$, $r, j = 1, 2$ and $[\![X^{\theta_j}, X^{E_i}]\!] = 0$, for $j = 1, 2$ and $i = 1, 2, 3$.

The reduced Lagrangian $\ell : \mathbb{S}^2 \times \mathfrak{so}(3) \times T(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$ is given by

$$\ell(\Gamma, \Omega, \dot{\theta}) = \frac{1}{2} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix}^T \begin{pmatrix} \lambda_1 & 0 & 0 & J_1 & 0 \\ 0 & \lambda_2 & 0 & 0 & J_2 \\ 0 & 0 & \bar{I}_3 & 0 & 0 \\ J_1 & 0 & 0 & J_1 & 0 \\ 0 & J_2 & 0 & 0 & J_2 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} - Mgh\Gamma_3.$$

The Euler-Lagrange equations are given by

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \ell}{\partial \Omega}\right) &= \frac{\partial \ell}{\partial \Omega} \times \Omega + Mgh\Gamma \times e_3, \\ \frac{d}{dt}\left(\frac{\partial \ell}{\partial \dot{\theta}_i}\right) &= 0, \quad i = 1, 2; \end{aligned}$$

together with the admissibility condition $\dot{\Gamma} = \Gamma \times \Omega$.

Now we add controls in our picture. We suppose that the rotors can be controlled, then the controlled Euler-Lagrange equations can be rewritten as

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \ell}{\partial \Omega}\right) &= \frac{\partial \ell}{\partial \Omega} \times \Omega + Mgh\Gamma \times e_3, \\ \frac{d}{dt}\left(\frac{\partial \ell}{\partial \dot{\theta}_i}\right) &= u_i, \quad i = 1, 2; \\ \dot{\Gamma} &= \Gamma \times \Omega. \end{aligned}$$

That is,

$$
\begin{aligned}
\lambda_1\dot{\Omega}_1 + J_1\ddot{\theta}_1 - \lambda_2\Omega_2\Omega_3 + \dot{\Omega}_3\bar{I}_3\Omega_2 &= MghY,\\
\lambda_2\dot{\Omega}_2 + J_2\ddot{\theta}_2 + \lambda_1\Omega_1\Omega_3 - J_1\dot{\theta}_1\Omega_3 &= -MghX,\\
\bar{I}_3\dot{\Omega}_3 - \lambda_1\Omega_1\Omega_2 - J_1\dot{\theta}_1\Omega_2 + \lambda_2\Omega_2\Omega_1 + J_2\dot{\theta}_2\Omega_1 &= 0,\\
J_1(\dot{\Omega}_1 + \ddot{\theta}_1) &= u_1,\\
J_2(\dot{\Omega}_2 + \ddot{\theta}_2) &= u_2,\\
\dot{X} &= Y\Omega_3 - Z\Omega_2,\\
\dot{Y} &= Z\Omega_1 - X\Omega_3,\\
\dot{Z} &= X\Omega_2 - Y\Omega_1
\end{aligned}
$$

where $\Gamma = (X, Y, Z) \in \mathbb{S}^2$.

The optimal control problem consists on finding an admissible curve $\gamma(t) = (\Gamma(t), \Omega(t), \theta(t), u_i)$ of the state variables and control inputs, given boundary conditions solving the previous equations and minimizing

$$\mathcal{J} = \frac{1}{2}\int_0^T \left(u_1^2 + u_2^2\right) dt.$$

This optimal control problem is equivalent to solve the following second-order problem with second-order constraints:

$$\text{extremize} \int_0^T \mathcal{L}\left(\Omega, \theta, \dot{\Omega}, \dot{\theta}, \ddot{\theta}\right) dt,$$

subject to the second-order constraints $\Phi^\alpha : T\mathbb{S}^2 \times 2\mathfrak{so}(3) \times T^{(2)}(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$, $\alpha = 1, 2$;

$$
\begin{aligned}
\Phi^1 &= \lambda_1\dot{\Omega}_1 + J_1\ddot{\theta}_1 - \lambda_2\Omega_2\Omega_3 + \dot{\Omega}_3\bar{I}_3\Omega_2 - MghY,\\
\Phi^2 &= \lambda_2\dot{\Omega}_2 + J_2\ddot{\theta}_2 + \lambda_1\Omega_1\Omega_3 - J_1\dot{\theta}_1\Omega_3 + MghX,\\
\Phi^3 &= \bar{I}_3\dot{\Omega}_3 - \lambda_1\Omega_1\Omega_2 - J_1\dot{\theta}_1\Omega_2 + \lambda_2\Omega_2\Omega_1 + J_2\dot{\theta}_2\Omega_1\\
\Phi^4 &= \dot{\Gamma} - \Gamma \times \Omega
\end{aligned}
$$

where $\mathcal{L} : 2\mathfrak{so}(3) \times T^{(2)}(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$, is defined by

$$
\begin{aligned}
\mathcal{L}\left(\Omega, \theta, \dot{\Omega}, \dot{\theta}, \ddot{\theta}\right) &= \mathcal{C}\left(\Omega, \theta, \dot{\Omega}, \dot{\theta}, \frac{d}{dt}\left(\frac{\partial \ell}{\partial \dot{\theta}}\right)\right)\\
&= \frac{1}{2}(u_1^2 + u_2^2) = \frac{J_1^2}{2}(\dot{\Omega}_1 + \ddot{\theta}_1)^2 + \frac{J_2^2}{2}(\dot{\Omega}_2 + \ddot{\theta}_2)^2.
\end{aligned}
$$

We will use the Cayley transformation on $SO(3)$ to describe the discrete optimal control problem for the heavy top with internal rotors. We redefine the Lagrangian $\mathcal{L}$ and the constraints Φ^α as $\widetilde{L} : 2\mathfrak{so}(3) \times T^{(2)}(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$ and $\widetilde{\Phi}^\alpha : T\mathbb{S}^2 \times 2\mathfrak{so}(3) \times T^{(2)}(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$ by

$$\mathcal{L}\left(\frac{\xi_k + \xi_{k+1}}{2}, \theta, \frac{\xi_{k+1} - \xi_k}{h}, \dot{\theta}, \ddot{\theta}\right) =: \widetilde{L}(\xi_k, \theta, \xi_{k+1}, \dot{\theta}, \ddot{\theta})$$

and

$$\Phi^{\alpha}\left(\Gamma, \frac{\xi_k+\xi_{k+1}}{2}, \theta, \dot{\Gamma}, \frac{\xi_{k+1}-\xi_k}{h}, \dot{\theta}, \ddot{\theta}\right) =: \widetilde{\Phi}^{\alpha}(\Gamma, \xi_k, \theta, \dot{\Gamma}, \xi_{k+1}, \dot{\theta}, \ddot{\theta}),$$

where $\xi_k, \xi_{k+1} \in \mathfrak{so}(3)$ and $h > 0$ is a fixed real number.

To derive the associated discrete optimal control problem we need to consider the discrete second-order Lagrangian $L_d : 2\mathfrak{so}(3) \times 3(\mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$ where

$$\begin{aligned}&L_d(\xi_k, \theta_k^1, \theta_k^2, \xi_{k+1}, \theta_{k+1}^1, \theta_{k+1}^2, \theta_{k+2}^1, \theta_{k+2}^2) = \\ &h\widetilde{L}\left(\frac{\xi_k+\xi_{k+1}}{2}, \frac{\theta_k^i+\theta_{k+1}^i+\theta_{k+2}^i}{3}, \frac{\xi_{k+1}-\xi_k}{h}, \frac{\theta_{k+2}^i-\theta_k^i}{2h}, \frac{\theta_{k+2}^i-2\theta_{k+1}^i+\theta_k^i}{h^2}\right)\end{aligned}$$

and the discrete constraints

$$\begin{aligned}&\Phi_d^{\alpha}(\Gamma_k, \xi_k, \theta_k^1, \theta_k^2, \xi_{k+1}, \Gamma_{k+1}, \theta_{k+1}^1, \theta_{k+1}^2, \theta_{k+2}^1, \theta_{k+2}^2) = \\ &h\widetilde{\Phi}^{\alpha}\left(\Gamma_k, \frac{\xi_k+\xi_{k+1}}{2}, \frac{\theta_k^i+\theta_{k+1}^i+\theta_{k+2}^i}{3}, \frac{\Gamma_{k+1}-\Gamma_k}{h}, \frac{\xi_{k+1}-\xi_k}{h}, \frac{\theta_{k+2}^i-\theta_k^i}{2h}, \frac{\theta_{k+2}^i-2\theta_{k+1}^i+\theta_k^i}{h^2}\right)\end{aligned}$$

where $i = 1, 2$ and $h\xi_k = cay^{-1}(\omega_k) \in \mathfrak{so}(3)$ with $\omega_k \in SO(3)$. Here

$$\xi_k = \begin{pmatrix} 0 & -(\xi_3)_k & (\xi_2)_k \\ (\xi_3)_k & 0 & -(\xi_1)_k \\ -(\xi_2)_k & (\xi_1)_k & 0 \end{pmatrix} \in \mathfrak{so}(3),$$

$$\begin{aligned}\Gamma(kh) &\simeq \frac{\Gamma_k+\Gamma_{k+1}}{2}, \quad \dot{\Gamma}(kh) \simeq \frac{\Gamma_{k+1}-\Gamma_k}{h}, \quad \theta(kh) \simeq \frac{\theta_{k+1}+\theta_k}{2}, \\ \dot{\theta}(kh) &\simeq \frac{\theta_{k+2}-\theta_k}{2h}, \quad \ddot{\theta}(kh) \simeq \frac{\theta_{k+2}-2\theta_{k+1}+\theta_k}{h^2}.\end{aligned}$$

The geometric integrator is given by extremize

$$C_d = \sum_{k=0}^{N-1} L_d(\xi_k, \theta_k^i, \xi_{k+1}, \theta_{k+1}^i, \theta_{k+2}^i) + \lambda_\alpha \Phi_d^{\alpha}(\Gamma_k, \xi_k, \theta_k^i, \Gamma_{k+1}, \xi_{k+1}, \theta_{k+1}^i, \theta_{k+2}^i)$$

where λ_α are Lagrange multipliers,

$$\begin{aligned}L_d &= \frac{J_1^2}{2}\left(\frac{(\xi_{k+1})_1-(\xi_k)_1}{h} + \frac{\theta_{k+2}^1-2\theta_{k+1}^1+\theta_k^1}{h^2}\right)^2 + \frac{J_2^2}{2}\left(\frac{(\xi_{k+1})_2-(\xi_k)_2}{h} + \frac{\theta_{k+2}^2-2\theta_{k+1}^2+\theta_k^2}{h^2}\right)^2, \\ \Phi_d^1 &= \lambda_1\left(\frac{(\xi_{k+1})_1-(\xi_k)_1}{h}\right) + J_1\left(\frac{\theta_{k+2}^1-2\theta_{k+1}^1+\theta_k^1}{h^2}\right) - \lambda_2\left(\frac{((\xi_k)_2+(\xi_{k+1})_2)((\xi_k)_3+(\xi_{k+1})_3)}{4}\right) \\ &\quad +\bar{I}_3\left(\frac{((\xi_{k+1})_3-(\xi_k)_3)((\xi_{k+1})_2+(\xi_k)_2)}{2h}\right) - MghY_k,\end{aligned}$$

$$\begin{aligned}
\Phi_d^2 &= \lambda_2\left(\frac{(\xi_{k+1})_2-(\xi_k)_2}{h}\right)+J_2\left(\frac{\theta^2_{k+2}-2\theta^2_{k+1}+\theta^2_k}{h^2}\right)+\lambda_1\left(\frac{((\xi_k)_1+(\xi_{k+1})_1)((\xi_k)_3+(\xi_{k+1})_3)}{4}\right)\\
&\quad -J_1\left(\frac{(\theta^1_{k+2}-\theta^1_k)((\xi_{k+1})_3+(\xi_k)_3)}{2h}\right)-MghY_k,\\
\Phi_d^3 &= \bar{I}_3\left(\frac{(\xi_{k+1})_3-(\xi_k)_3}{h}\right)-\lambda_1\left(\frac{((\xi_{k+1})_1+(\xi_k)_1)((\xi_{k+1})_2+(\xi_k)_2)}{4}\right)\\
&\quad +\lambda_2\left(\frac{((\xi_{k+1})_2+(\xi_k)_2)((\xi_{k+1})_1+(\xi_k)_1)}{4}\right)+J_2\left(\frac{(\theta^2_{k+2}-\theta^2_k)((\xi_k)_1+(\xi_{k+1})_1)}{4h}\right)\\
&\quad -J_1\left(\frac{(\theta^1_{k+2}-\theta^1_k)((\xi_k)_2+(\xi_{k+1})_2)}{4h}\right),\\
\Phi_d^4 &= \Gamma_{k+1}-\Gamma_k\left(\frac{cay^{-1}(\omega_k)+cay^{-1}(\omega_{k+1})}{2h}\right).
\end{aligned}$$

Remark 5.2.4. Simulating this optimal control problem remains work for future study. ◇

Example: Optimal Control of a Homogeneous Ball on a Rotating Plate

We consider the following well-known problem (see [17, 97, 114]), namely the model of a homogeneous ball on a rotating plate. A (homogeneous) ball of radius $r > 0$, mass m and inertia mk^2 about any axis rolls without slipping on a horizontal table which rotates with angular velocity Ω about a vertical axis x_3 through one of its points. Apart from the constant gravitational force, no other external forces are assumed to act on the sphere. Let (x, y) be denote the position of the point of contact of the sphere with the table. The configuration space of the sphere is $Q = \mathbb{R}^2 \times SO(3)$ where may be parametrized Q by (x, y, g), $g \in SO(3)$, all measured with respect to the inertial frame. Let $\omega = (\omega_1, \omega_2, \omega_3)$ be the angular velocity vector of the sphere measured also with respect to the inertial frame. The potential energy is constant, so we may put $V = 0$.

The nonholonomic constraints are given by the non-slipping condition by

$$\dot{x}+\frac{r}{2}Tr(\dot{g}g^TE_2)=-\Omega y,\quad \dot{y}-\frac{r}{2}Tr(\dot{g}g^TE_1)=\Omega x,$$

where $\{E_1, E_2, E_3\}$ is the standard basis of $\mathfrak{so}(3)$.

The matrix $\dot{g}g^T$ is skew-symmetric therefore we may write

$$\dot{g}g^T=\begin{pmatrix}0&-\omega_3&\omega_2\\ \omega_3&0&-\omega_1\\ -\omega_2&\omega_1&0\end{pmatrix}$$

where $(\omega_1, \omega_2, \omega_3)$ represents the angular velocity vector of the sphere measured with respect to the inertial frame. Then, we may rewrite the constraints in the usual form:

$$\dot{x}+r\omega_2=-\Omega y,\quad \dot{y}-r\omega_1=\Omega x.$$

In addition, since we do not consider external forces the Lagrangian of the system corresponds with the kinetic energy

$$K(x,y,g,\dot{x},\dot{y},\dot{g}) = \frac{1}{2}(m\dot{x}^2 + m\dot{y}^2 + mk^2(\omega_1^2 + \omega_2^2 + \omega_3^2)).$$

Observe that the Lagrangian is metric on Q which is bi-invariant on $SO(3)$ as the ball is homogeneous.

Now, it is clear that $Q = \mathbb{R}^2 \times SO(3)$ is the total space of a trivial principal $SO(3)$-bundle over $\mathbb{R}^2$ with respect the right $SO(3)-$action given by $(x,y,R) \mapsto (x,y,RS)$ for all $S \in SO(3)$ and $(x,y,R) \in \mathbb{R}^2 \times SO(3)$. The action is in the right side since the symmetries are material symmetries.

The bundle projection $\phi : Q \to M = \mathbb{R}^2$ is just the canonical projection on the first factor. Therefore, we may consider the corresponding Atiyah algebroid $A = TQ/SO(3)$ over $M = \mathbb{R}^2$. We will identify the tangent bundle to $SO(3)$ with $\mathfrak{so}(3) \times SO(3)$ by using right translation. Note that throughout the previous exposition we have employed the left trivialization. However, we would like to point out that the right trivialization just implies minor changes in the derivation of the equations of motion (see [81]).

Under this identification between $T(SO(3))$ and $\mathfrak{so}(3) \times SO(3)$, the tangent action of $SO(3)$ on $T(SO(3)) \cong \mathfrak{so}(3) \times SO(3)$ is the trivial action

$$(\mathfrak{so}(3) \times SO(3)) \times SO(3) \to \mathfrak{so}(3) \times SO(3), \quad ((\omega, g), h) \mapsto (\omega, gh). \tag{5.38}$$

Thus, the Atiyah algebroid $TQ/SO(3)$ is isomorphic to the real vector bundle $T\mathbb{R}^2 \times \mathfrak{so}(3) \to \mathbb{R}^2$, and the vector bundle projection is $\tau_{\mathbb{R}^2} \circ pr_1$, where $pr_1 : T\mathbb{R}^2 \times \mathfrak{so}(3) \to T\mathbb{R}^2$ and $\tau_{\mathbb{R}^2} : T\mathbb{R}^2 \to \mathbb{R}^2$ are the canonical projections. The anchor map $\rho : A \simeq T\mathbb{R}^2 \times \mathfrak{so}(3) \to \mathbb{R}^2$ is just the projection onto the first factor.

A section of $A = TQ/SO(3) \cong T\mathbb{R}^2 \times \mathfrak{so}(3) \to \mathbb{R}^2$ is a pair (X, f), where X is a vector field on $\mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathfrak{so}(3)$ is a smooth map. Therefore, a global basis of sections of $T\mathbb{R}^2 \times \mathfrak{so}(3) \to \mathbb{R}^2$ is

$$e_1 = \left(\frac{\partial}{\partial x}, 0\right), e_2 = \left(\frac{\partial}{\partial y}, 0\right), e_3 = (0, E_1), e_4 = (0, E_2), e_5 = (0, E_3).$$

There exists a one-to-one correspondence between the space $\Gamma(A = TQ/SO(3))$ and the G-invariant vector fields on Q.

If $[\![\cdot,\cdot]\!]$ is the Lie bracket on the space $\Gamma(A = TQ/SO(3))$, then the only non-zero fundamental Lie brackets are

$$[\![e_4, e_3]\!] = e_5, \quad [\![e_5, e_4]\!] = e_3, \quad [\![e_3, e_5]\!] = e_4.$$

Moreover, it follows that the Lagrangian function $L = K$ and the constraints are $SO(3)$-invariant. Consequently, L induces a Lagrangian function ℓ on $A = TQ/SO(3) \simeq T\mathbb{R}^2 \times \mathfrak{so}(3)$.

We have a constrained system on $A = TQ/SO(3) \simeq T\mathbb{R}^2 \times \mathfrak{so}(3)$ and note that in this case the constraints are nonholonomic and affine in the velocities. This kind of systems was analyzed by J. Cortés *et al* [55] (in particular, this example was carefully studied). The constraints define an affine subbundle of the vector bundle $A \simeq T\mathbb{R}^2 \times \mathfrak{so}(3) \to \mathbb{R}^2$ which is modeled over the vector subbundle $\mathcal{D}$ generated by the sections

$$\mathcal{D} = \operatorname{span}\{e_5; re_1 + e_4; re_2 - e_3\}.$$

Moreover, the angular momentum of the ball about the axis x_3 is a conserved quantity since the Lagrangian is invariant under rotations about the axis x_3 and the infinitesimal generator for these rotations lies in the distribution $\mathcal{D}$. The conservation law is written as $\omega_z = c$, where c is a constant or as $\dot{\omega}_z = 0$. Then by the conservation of the angular momentum the second-order constraints appear.

After some computations the equations of motion for this constrained system are precisely

$$\left\{\begin{array}{rcl} \dot{x} - r\omega_2 & = & -\Omega y, \\ \dot{y} + r\omega_1 & = & \Omega x, \\ \dot{\omega}_3 & = & 0 \end{array}\right. \tag{5.39}$$

together with

$$\ddot{x} + \frac{k^2\Omega}{r^2+k^2}\dot{y} = 0, \quad \ddot{y} - \frac{k^2\Omega}{r^2+k^2}\dot{x} = 0.$$

Now, we pass to an optimization problem. Assume full controls over the motion of the center of the ball (the shape variables). The controlled system can be written as,

$$\ddot{x} + \frac{k^2\Omega}{r^2+k^2}\dot{y} = u_1, \quad \ddot{y} - \frac{k^2\Omega}{r^2+k^2}\dot{x} = u_2,$$

subject to

$$\left\{\begin{array}{rcl} \omega_2 - \frac{1}{r}\dot{x} & = & \frac{\Omega y}{r}, \\ \omega_1 + \frac{1}{r}\dot{y} & = & \frac{\Omega x}{r}, \\ \dot{\omega}_3 & = & 0. \end{array}\right. \tag{5.40}$$

Next, we consider the optimal control problem for this system: Let C be the cost function given by

$$C = \frac{1}{2}\left(u_1^2 + u_2^2\right).$$

Given $q(0), q(T) \in \mathbb{R}^2$, $\dot{q}(0) \in T_{q(0)}\mathbb{R}^2, \dot{q}(T) \in T_{q(T)}\mathbb{R}^2$, $q = (x, y) \in \mathbb{R}^2$, $\omega(0), \omega(T) \in \mathfrak{so}(3)$, we look for an optimal control curve $(q(t), \omega(t), u(t))$ on the reduced space that steers the system from $(q(0), \omega(0))$ to $(q(T), \omega(T))$ minimizing

$$\int_0^T \frac{1}{2}\left(u_1^2 + u_2^2\right) dt,$$

subject to the constraints given by equations (5.40). Note that $R(0), R(T) \in SO(3)$, the initial and final configurations of the problem, are also fixed. Its dynamics is given by the continuous reconstruction equation $\dot{R}(t) = R(t)\omega(t)$.

As in the previous example, we define the second order Lagrangian $\widetilde{L} : T^{(2)}\mathbb{R}^2 \times 2\mathfrak{so}(3) \to \mathbb{R}$ given by

$$\widetilde{L}(x, y, \dot{x}, \dot{y}, \ddot{x}, \ddot{y}, \omega_1, \omega_2, \omega_3, \dot{\omega}_1, \dot{\omega}_2, \dot{\omega}_3) = \frac{1}{2}\left(\ddot{x} + \frac{k^2\Omega}{r^2+k^2}\dot{y}\right)^2 + \frac{1}{2}\left(\ddot{y} - \frac{k^2\Omega}{r^2+k^2}\dot{x}\right)^2 \tag{5.41}$$

subject to second-order constraints $\Phi^\alpha : T^{(2)}\mathbb{R}^2 \times 2\mathfrak{so}(3) \to \mathbb{R}$, $\alpha = 1, 2, 3$,

$$\Phi^1 = \omega_1 + \frac{1}{r}\dot{y} - \frac{\Omega x}{r}, \tag{5.42a}$$

$$\Phi^2 = \omega_2 - \frac{1}{r}\dot{x} - \frac{\Omega y}{r}, \tag{5.42b}$$

$$\Phi^3 = \dot{\omega}_3. \tag{5.42c}$$

The optimal control problem is prescribed by solving the following system of 4-order differential equations (ODEs).

$$\begin{aligned}
0 &= \lambda_1\frac{\Omega}{r} + \frac{\dot{\lambda}_2}{r} + x^{(iv)} + \frac{2k^2\Omega\dddot{y}}{r^2+k^2} - \frac{k^4\Omega^2\ddot{x}}{(r^2+k^2)^2}\\
0 &= \lambda_2\frac{\Omega}{r} + \frac{\dot{\lambda}_1}{r} + y^{(iv)} - \frac{2k^2\Omega\dddot{x}}{r^2+k^2} - \frac{k^4\Omega^2\ddot{y}}{(r^2+y^2)^2},\\
0 &= \dot{\lambda}_1 + \lambda_2\omega_3 - \lambda_3\omega_2,\\
0 &= \dot{\lambda}_2 - \lambda_1\omega_3 + \lambda_3\omega_1,\\
0 &= \dot{\lambda}_3 + \lambda_1\omega_2 - \lambda_2\omega_1,\\
0 &= \omega_1 + \frac{1}{r}\dot{y} - \frac{\Omega x}{r},\\
0 &= \omega_2 - \frac{1}{r}\dot{x} - \frac{\Omega y}{r},\\
0 &= \dot{\omega}_3.
\end{aligned}$$

In addition, the configurations $R \in SO(3)$ are given by the continuous reconstruction equation $\dot{R} = R\omega$.

Remark 5.2.5. In the particular case when the angular velocity Ω depends on the time (see [17, 97]), the equations of motion are rewritten as

$$\begin{aligned}
0 &= \lambda_1\frac{\Omega(t)}{r} + \frac{\dot{\lambda}_2}{r} + x^{(iv)} + \frac{k^2\Omega''(t)\dot{y}}{r^2+k^2} + \frac{2k^2\Omega'(t)\ddot{y}}{r^2+k^2} + \frac{2k^2\Omega(t)\dddot{y}}{r^2+k^2},\\
&+ \frac{k^2\Omega'(t)\dddot{y}}{r^2+k^2} - \frac{k^4\Omega^2(t)\ddot{x}}{(r^2+k^2)^2} - \frac{2k^4\Omega'(t)\Omega(t)\dot{x}}{(r^2+k^2)^2}\\
0 &= \lambda_2\frac{\Omega(t)}{r} + \frac{\dot{\lambda}_1}{r} + y^{(iv)} - \frac{k^2\Omega''(t)\dot{x}}{r^2+k^2} - \frac{3k^2\Omega'(t)\ddot{x}}{r^2+k^2} - \frac{2k^2\Omega(t)\dddot{x}}{r^2+k^2},\\
&- \frac{k^4\Omega^2(t)\ddot{y}}{(r^2+y^2)^2} - \frac{2k^4\Omega(t)\Omega'(t)\dot{y}}{(r^2+k^2)^2},\\
0 &= \dot{\lambda}_1 + \lambda_2\omega_3 - \lambda_3\omega_2,\\
0 &= \dot{\lambda}_2 - \lambda_1\omega_3 + \lambda_3\omega_1,\\
0 &= \dot{\lambda}_3 + \lambda_1\omega_2 - \lambda_2\omega_1,\\
0 &= \omega_1 + \frac{1}{r}\dot{y} - \frac{\Omega(t)x}{r},\\
0 &= \omega_2 - \frac{1}{r}\dot{x} - \frac{\Omega(t)y}{r},\\
0 &= \dot{\omega}_3.
\end{aligned}$$

$\diamond$

• **Discrete setting:** As in the previous example, we discretize this problem by choosing a discrete Lagrangian $\widetilde{L}_d$ and discrete constraints Φ_d^α. Employing equivalent arguments than in the previous example, we set $\widetilde{L}_d : 3(\mathbb{R}^2)\times 2\mathfrak{so}(3)\to\mathbb{R}$ and $\Phi_d^\alpha : 3(\mathbb{R}^2)\times 2\mathfrak{so}(3)\to\mathbb{R}$, $\alpha=1,2,3$, as

$$
\begin{aligned}
&\widetilde{L}_d(q_k,q_{k+1},q_{k+2},\omega_k,\omega_{k+1})+\lambda_\alpha^k\Phi_d^\alpha(q_k,q_{k+1},q_{k+2},\omega_k,\omega_{k+1})=\\
&h\widetilde{L}\left(\frac{q_k+q_{k+1}+q_{k+2}}{3},\frac{q_{k+2}-q_k}{2h},\frac{q_{k+2}-2q_{k+1}+q_k}{h^2},\frac{\omega_k+\omega_{k+1}}{2},\frac{\omega_{k+1}-\omega_k}{h}\right)\\
&+\lambda_\alpha^k\Phi^\alpha\left(\frac{q_k+q_{k+1}+q_{k+2}}{3},\frac{q_{k+2}-q_k}{2h},\frac{q_{k+2}-2q_{k+1}+q_k}{h^2},\frac{\omega_k+\omega_{k+1}}{2},\frac{\omega_{k+1}-\omega_k}{h}\right),
\end{aligned}
$$

We employ the same unknowns-equations counting process than in the previous example to find out that the number of unknowns matches the number of equations. Therefore, our discrete variational problem (which comes from the original optimal control problem) has become again a nonlinear root finding problem. For computational reasons is useful to consider the retraction map τ as the Cayley map for $SO(3)$ instead of a truncation of the exponential map (see Appendix C for further details).

Now we show some simulations to test our method for $T=4$, $r=1$, $\Omega=0.3$ and $\omega_3=m=k=1$,

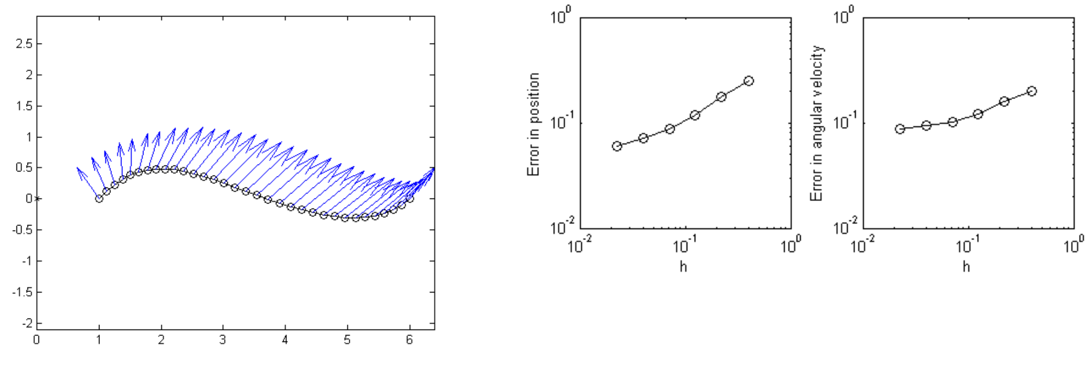


Figure 5.1: Left: Simulation of the method with $q_0=(1,0)$ $v_0=(1,1)$, $q_N=(6,0)$, $v_N=(1,1)$, $N=33$. Blue arrows shown the (scaled) angular velocity. Right: Error (root-mean-square error) in position and angular velocity for different values of h between our method and with and a Runge-Kutta 4.

The following table show the root mean square error in positions and angular velocities between our method and a Runge-Kutta 4:

h	Error in position	Error in angular velocity
0.4	0.2471	0.1995
0.22	0.1746	0.1576
0.1250	0.1173	0.1204
0.0714	0.0866	0.1020
0.04	0.0705	0.0932
0.0225	0.0606	0.0875

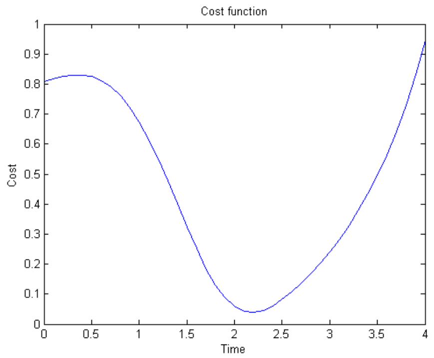


Figure 5.2: cost function $\frac{1}{2}(u_1^2 + u_2^2)$

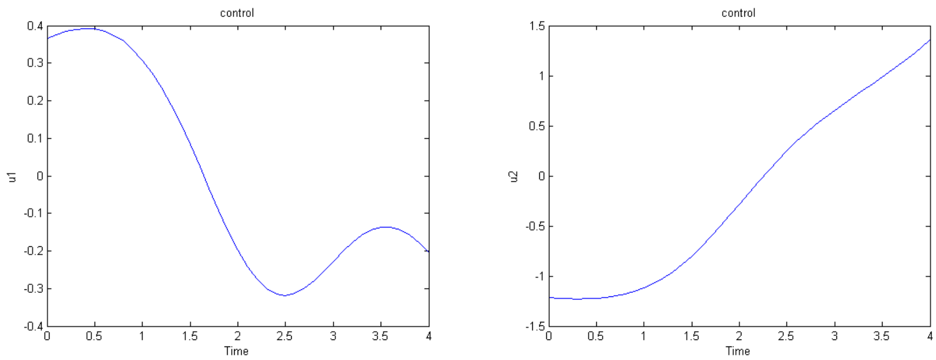


Figure 5.3: controls u_1 and u_2

5.3 Lagrangian submanifolds generating discrete dynamics

In this section we will see how a Lagrangian submanifold $\Sigma_L \subset T^*(\mathcal{P}^\alpha G)$ of the cotangent groupoid $T^*(\mathcal{P}^\alpha G) \rightrightarrows A^*(\mathcal{P}^\alpha G)$ will give the second-order discrete dynamics associated with a discrete second-order Lagrangian $L : G_2 \to \mathbb{R}$.

In general, the dynamics will be defined implicitly rather than a discrete explicit flow map (see [126]). This dynamic can be interpreted as a discrete second-order Lagrangian dynamic on the Lagrangian submanifold $\Sigma_L \subset T^*(\mathcal{P}^\alpha G)$ or as a Hamiltonian second-order dynamic on $A^*(\mathcal{P}^\alpha G)$.

The motivation of this section is to show an alternative and geometric approach to obtain the dynamic of discrete second-order variational problems on Lie groupoids instead of use standard discrete variational calculus. We will clarify the case of discrete second-order systems subject to constraints in this framework since, as is well know, the use of Lagrange multipliers for discrete constrained systems is not the better way to show the geometric properties of the flow map. In this sense, with our framework, we will study the regularity and reversibility of this kind of discrete systems, from the perspective of symplectic and Poisson geometry and we will also study the theory of reduction under Noether symmetries.

5.3.1 Generating Lagrangian submanifolds and dynamics on Lie groupoids

Let $G \rightrightarrows Q$ be a Lie groupoid with source and target map $\alpha, \beta : G \to Q$ respectively, and we consider the prolongation Lie groupoid $\mathcal{P}^\alpha G \rightrightarrows G$ over the source map of G where we denote $\alpha^\alpha, \beta^\alpha : \mathcal{P}^\alpha G \to G$ the source and target maps of this prolongation Lie groupoid. Let $\tau_{A^*(\mathcal{P}^\alpha G)} : A^*(\mathcal{P}^\alpha G) \to G$ be the dual of the vector bundle associated with the Lie algebroid $\tau_{A(\mathcal{P}^\alpha G)} : A(\mathcal{P}^\alpha G) \to G$. Then the Lie groupoid $T^*(\mathcal{P}^\alpha G) \rightrightarrows A^*(\mathcal{P}^\alpha G)$ is a symplectic groupoid (see example 3 in section 1.9.1).

By Tulczyjew's theorem (1.6.11), $L : G_2 \to \mathbb{R}$ generates a Lagrangian submanifold $\Sigma_L \subset T^*(\mathcal{P}^\alpha G)$ of the symplectic Lie groupoid $(T^*(\mathcal{P}^\alpha G), \omega_{\mathcal{P}^\alpha G})$ where $\omega_{\mathcal{P}^\alpha G}$ denotes the canonical symplectic 2-form on $T^*(\mathcal{P}^\alpha G)$. That is, denoting by $i_{G_2} : G_2 \to \mathcal{P}^\alpha G$ the inclusion defined by $i_{G_2}(g_1, g_2) = (g_1, g_1, g_2)$,

$$\Sigma_L = \{\mu \in T^*(\mathcal{P}^\alpha G) / i^*_{G_2}\mu = dL\} \subset T^*(\mathcal{P}^\alpha G)$$

is a Lagrangian submanifold of $(T^*(\mathcal{P}^\alpha G), \omega_{\mathcal{P}^\alpha G})$.

The relationship among these spaces is summarized in the following diagram

$$\begin{tikzcd}
 & & \Sigma_L \arrow[r, hook] & T^*(\mathcal{P}^\alpha G) \arrow[r, shift left, "\tilde{\alpha}"] \arrow[r, shift right, "\tilde{\beta}"'] \arrow[d] & A^*(\mathcal{P}^\alpha G) \arrow[d] \\
\mathbb{R} & G_2 \arrow[l, "L"'] \arrow[rr, hook, "i_{G_2}"] & & \mathcal{P}^\alpha G \arrow[r, shift left, "\alpha^\alpha"] \arrow[r, shift right, "\beta^\alpha"'] & G
\end{tikzcd}$$

From now on, we will denote $\tilde{\alpha}$ and $\tilde{\beta}$ the source and target maps of the Lie groupoid $T^*(\mathcal{P}^\alpha G) \rightrightarrows A^*(\mathcal{P}^\alpha G)$ respectively. Given an element $\mu \in T^*_{(g,h,r)}(\mathcal{P}^\alpha G)$ with $(g, h, r) \in \mathcal{P}^\alpha G$ the source and target are defined, such that for all sections $Z \in \Gamma(\tau_{A(\mathcal{P}^\alpha G)})$

$$\begin{aligned}
\langle \tilde{\alpha}(\mu), Z(\alpha(g)) \rangle &= \langle \mu, \overrightarrow{Z}(g, h, r) \rangle \\
\langle \tilde{\beta}(\mu), Z(\beta(g)) \rangle &= \langle \mu, \overleftarrow{Z}(g, h, r) \rangle,
\end{aligned}$$

where $\overleftarrow{Z}$ and $\overrightarrow{Z}$ are the corresponding left and right invariant vector fields associated with the section Z of $A(\mathcal{P}^\alpha G)$ according to (1.25) and (1.26) in section 1.9.3.

Denote by

$$\gamma_{(g_k,g_{k+1})} := (\mu_{g_k}, \tilde{\mu}_{g_k}, \bar{\mu}_{g_{k+1}}) \in T^*(\mathcal{P}^\alpha G),$$

with $(g_k, g_{k+1}) \in G_2$.

Definition 5.3.1. *A sequence $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$ satisfy the second-order dynamics on Σ_L if $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in \Sigma_L$ and*

$$\tilde{\alpha}(\gamma_{(g_k,g_{k+1})}) = \tilde{\beta}(\gamma_{(g_{k-1},g_k)}) \text{ for } k = 2, \ldots, N-1.$$

That is, $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)}$ are a composable sequence on $T^*(\mathcal{P}^\alpha G)$.

Applying the definition of $\tilde{\alpha}$ and $\tilde{\beta}$ we have that for any section $Z \in \Gamma(\tau_{A(\mathcal{P}^\alpha G)})$, the last equation is equivalent to

$$\begin{aligned} \langle \overleftarrow{Z}(g_k, g_k, g_{k+1}); \gamma_{(g_k,g_{k+1})} \rangle &= \langle \overrightarrow{Z}(g_{k+1}, g_{k+1}, g_{k+2}); \gamma_{(g_{k+1},g_{k+2})} \rangle & (5.43)\\ \mu_{g_k} + \tilde{\mu}_{g_k} &= D_1 L(g_k, g_{k+1}) & (5.44)\\ \bar{\mu}_{g_{k+1}} &= D_2 L(g_k, g_{k+1}) \text{ for } k = 1, \ldots, N-1; & (5.45) \end{aligned}$$

and using (1.25) and (1.26),

$$\begin{aligned} \langle \overleftarrow{X}(g_k), \tilde{\mu}_{g_k} \rangle + \langle Y(g_{k+1}); \bar{\mu}_{g_{k+1}} \rangle &= \langle \overrightarrow{X}(g_{k+1}); \tilde{\mu}_{g_{k+1}} \rangle - \langle Y(g_{k+1}); \mu_{g_{k+1}} \rangle & (5.46)\\ \mu_{g_k} + \tilde{\mu}_{g_k} &= D_1 L(g_k, g_{k+1}) & (5.47)\\ \bar{\mu}_{g_{k+1}} &= D_2 L(g_k, g_{k+1}) \text{ for } k = 1, \ldots, N-1. & (5.48) \end{aligned}$$

Therefore, we can state the following result:

Theorem 5.3.2. *Let $G \rightrightarrows Q$ be a Lie groupoid and $L : G_2 \to \mathbb{R}$ be a discrete second-order Lagrangian. If we denote by Σ_L the Lagrangian submanifold of $T^*(\mathcal{P}^\alpha G)$ generated by L, then a sequence $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)}$ satisfies the discrete second-order dynamics on Σ_L if and only if $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)}$ satisfy*

$$\begin{aligned} \langle \overleftarrow{X}(g_k), \tilde{\mu}_{g_k} \rangle + \langle Y(g_k); \bar{\mu}_{g_{k+1}} \rangle &= \langle \overrightarrow{X}(g_{k+1}); \tilde{\mu}_{g_{k+1}} \rangle - \langle Y(g_{k+1}); \mu_{g_{k+1}} \rangle\\ \mu_{g_k} + \tilde{\mu}_{g_k} &= D_1 L(g_k, g_{k+1})\\ \bar{\mu}_{g_{k+1}} &= D_2 L(g_k, g_{k+1}) \text{ for } k = 1, \ldots, N-1 \end{aligned}$$

for any section $Z \in \Gamma(\tau_{A(\mathcal{P}^\alpha G)})$, $Z = (X, Y), X \in \Gamma(\tau_{AG}), Y \in \mathfrak{X}(G)$ such that $T\beta(X) = T\alpha(Y)$.

Remark 5.3.3. We have seen how the dynamics is given implicitly by a relation in $T^*(\mathcal{P}^\alpha G)$ rather that given by a discrete flow map. Therefore, $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$ satisfy the discrete second-order dynamics on Σ_L if and only if for each pair of successive elements satisfies the relation

$$(\gamma_{(g_{k-1},g_k)}; \gamma_{(g_k,g_{k+1})}) \in (T^*(\mathcal{P}^\alpha G))_2 \cap (\Sigma_L \times \Sigma_L).$$

$\diamond$

Now, given a basis of sections of AG one can obtain the basis of sections $\{Z_1, Z_2\}$ of $A(\mathcal{P}^\alpha G)$ with

$$Z_1 = (-X, \overrightarrow{X}) \text{ and } Z_2 = (0, \overleftarrow{X})$$

where $X \in \Gamma(\tau_{AG}), \overrightarrow{X} \in \overrightarrow{\mathfrak{X}}(G)$ and $\overleftarrow{X} \in \overleftarrow{\mathfrak{X}}(G)$. We use the notation $\overrightarrow{\mathfrak{X}}(G)$ (respect. $\overleftarrow{\mathfrak{X}}(G)$) for the set of right-invariant (respect. left-invariant) vector fields on G.

The condition $T\beta(X) = T\alpha(Y)$ is obviously satisfied by Z_1 and Z_2. In fact, observe that for Z_2 we have that $T\beta(0) = 0 = T\alpha(\overleftarrow{X})$ since $\overleftarrow{X}$ is α-vertical. For $Z_1 = (-X, \overrightarrow{X})$, using the identification between sections of AG and vector fields on G; the condition $T\beta(-X) = T\alpha(\overrightarrow{X})$ is equivalent to see that $\beta(h) = \alpha((r_g \circ i)(h))$ for $h \in G$; that is, $\beta(h) = \alpha(r_g(h^{-1}))$ and using that $\alpha(h^{-1}) = \alpha(h^{-1}g)$; we obtain $\beta(h) = \alpha(h^{-1})$; that is h and h^{-1} are composables; $h = \beta^{-1}(\alpha(h^{-1}))$.

In what follows we will derive the discrete second-order equations for $L : G_2 \to \mathbb{R}$ in terms of L. We will need to use the next result which is a direct consequence of (1.25) and (1.26),

Lemma 5.3.4. *Let as consider a section $Z \in \Gamma(\tau_{A(\mathcal{P}^\alpha G)})$, $Z = (Z_1, Z_2)$ where $Z_1 = (-X, \overrightarrow{X})$ and $Z_2 = (0, \overleftarrow{X})$, $X \in \Gamma(\tau_{AG}), \overrightarrow{X} \in \mathfrak{X}(G)$ and $\overleftarrow{X} \in \mathfrak{X}(G)$. Then the associated left and right invariant vector fields are*

$$\begin{aligned} \overrightarrow{Z_1}(g,h,r) &= (-\overrightarrow{X}(g), -\overrightarrow{X}(h), 0_r) \qquad & \overrightarrow{Z_2}(g,h,r) &= (-\overleftarrow{X}(g), 0_h, 0_r) \\ \overleftarrow{Z_1}(g,h,r) &= (0_g, -\overleftarrow{X}(h), \overrightarrow{X}(r)) & \overleftarrow{Z_2}(g,h,r) &= (0_g, 0_h, \overleftarrow{X}(r)). \end{aligned}$$

Using the lemma for $\overrightarrow{Z_2}$ and $\overleftarrow{Z_2}$, and replacing into (5.44) we obtain that

$$\langle \overleftarrow{Z_2}(g_k, g_k, g_{k+1}); (\gamma_{(g_k,g_{k+1})}\rangle = \langle \overrightarrow{Z_2}(g_{k+1}, g_{k+1}, g_{k+2}); (\gamma_{(g_{k+1},g_{k+2})})\rangle$$

if and only if

$$\begin{aligned} \langle \overleftarrow{X}(g_{k+1}); \bar{\mu}_{g_{k+1}}\rangle &= -\langle \overleftarrow{X}(g_{k+1}); \mu_{g_{k+1}}\rangle & (5.49)\\ \mu_{g_{k+1}} + \tilde{\mu}_{g_{k+1}} &= D_1L(g_{k+1}, g_{k+2}) & (5.50)\\ \bar{\mu}_{g_{k+1}} &= D_2L(g_k, g_{k+1}) \qquad k = 1, \dots, N-1. & (5.51) \end{aligned}$$

That is,

$$\langle \overleftarrow{X}(g_{k+1}); D_2L(g_k, g_{k+1})\rangle = -\langle \overleftarrow{X}(g_{k+1}); D_1L(g_{k+1}, g_{k+2}) - \tilde{\mu}_{g_{k+1}}\rangle. \qquad (5.52)$$

As before, using the definition of $\overleftarrow{Z_1}$ and $\overrightarrow{Z_1}$ we can observe that

$$\langle \overleftarrow{Z}_1(g_k, g_k, g_{k+1}); (\gamma_{(g_k,g_{k+1})})\rangle = \langle \overrightarrow{Z}_1(g_{k+1}, g_{k+1}, g_{k+2}); (\gamma_{(g_{k+1},g_{k+2})})\rangle$$

if and only if

$$-\langle \overleftarrow{X}(g_k); \tilde{\mu}_{g_k}\rangle + \langle \overrightarrow{X}(g_{k+1}), \bar{\mu}_{g_{k+1}}\rangle = -\langle \overrightarrow{X}(g_{k+1}); \mu_{g_{k+1}}\rangle - \langle \overrightarrow{X}(g_{k+1}); \tilde{\mu}_{g_{k+1}}\rangle; \qquad (5.53)$$

and using (5.50),(5.51) and (5.52) we can rewrite (5.53) as,

$$\langle \overrightarrow{X}(g_{k+1}), D_1L(g_{k+1}, g_{k+2}) + D_2L(g_k, g_{k+1})\rangle - \langle \overleftarrow{X}(g_k); D_1L(g_k, g_{k+1}) + D_2L(g_{k-1}, g_k)\rangle = 0.$$

After some computations, we deduce that

$$\ell_{g_k}^* \left(D_1L(g_k, g_{k+1}) + D_2L(g_{k-1}, g_k)\right) + (r_{g_{k+1}} \circ i)^* \left(D_1L(g_{k+1}, g_{k+2}) + D_2L(g_k, g_{k+1})\right) = 0,$$

for $k = 2, \ldots, N-2$.

We can summarize these developments in the following theorem

Theorem 5.3.5. *Let $L : G_2 \to \mathbb{R}$ be a discrete second order Lagrangian. For every section Z of $\Gamma(\tau_{A(\mathcal{P}^\alpha G)})$ as in Proposition (5.3.4) the discrete second order Euler-Lagrange equation are*

$$\ell_{g_k}^* \left(D_1L(g_k, g_{k+1}) + D_2L(g_{k-1}, g_k)\right) + (r_{g_{k+1}} \circ i)^* \left(D_1L(g_{k+1}, g_{k+2}) + D_2L(g_k, g_{k+1})\right) = 0,$$

for $k = 2, \ldots, N-2$.

These equations are just (5.20), the same equations that we have obtained from a variational point of view.

Example 5.3.6. Let G be a Lie group. G is a Lie groupoid over $\{\mathfrak{e}\}$, the identity element of G. Let $L : G_2 \to \mathbb{R}$. In this case $\mathcal{P}^\alpha G = 3G$ and $\mathcal{P}^\alpha G$ is a Lie groupoid over G. Then we can construct the Lagrangian submanifold of $T^*(3G)$ as

$$\Sigma_L := \{\mu \in T^*(2G)/i_{3G}^*\mu = dL\} \subseteq T^*(3G). \tag{5.54}$$

Σ_L is a Lagrangian submanifold of $T^*(3G)$ and $T^*(3G)$ is a Lie groupoid over $\mathfrak{g}^* \times T^*G$. The following diagram illustrates the situation,

$$\begin{tikzcd}
 & \Sigma_L \arrow[r, hook] & T^*(3G) \arrow[r, shift left, "\tilde{\alpha}"] \arrow[r, shift right, "\tilde{\beta}"'] \arrow[d] & \mathfrak{g}^* \times T^*G \arrow[d] \\
\mathbb{R} & 2G \arrow[l, "L"'] \arrow[r, hook, "i_{2G}"] & 3G \arrow[r, shift left, "\alpha^\alpha"] \arrow[r, shift right, "\beta^\alpha"'] & G
\end{tikzcd}$$

Applying the results given before; $\forall Z \in \Gamma(\mathfrak{g} \times TG)$ we obtain the discrete second-order dynamics for the discrete second-order Lagrangian $L : G_2 \to \mathbb{R}$,

$$\langle \overleftarrow{X}(g_k), \tilde{\mu}_{g_k} \rangle = \langle \overrightarrow{X}(g_{k+1}); \tilde{\mu}_{g_{k+1}} \rangle \tag{5.55}$$
$$\mu_{g_k} + \tilde{\mu}_{g_k} = D_1L(g_k, g_{k+1}) \tag{5.56}$$
$$\bar{\mu}_{g_{k+1}} = D_2L(g_k, g_{k+1}) \text{ for } k = 1, \ldots, N-1. \tag{5.57}$$

Now, we observe that $\overleftarrow{X}(g_k) = T\ell_{g_k}(X)$, $\overrightarrow{X}(g_k) = -T(r_{g_k} \circ i)(X) = Tr_{g_k}(X)$, $Ti(X) = -X$, because $T(\epsilon \circ \beta)(X) = 0$. Therefore, replacing and using the equations (5.55), (5.56), (5.57) we have that

$$\ell_{g_k}^* D_1L_d(g_k, g_{k+1}) + \ell_{g_k}^* D_2L_d(g_{k-1}, g_k) = r_{g_{k+1}}^* D_2L_d(g_k, g_{k+1}) + r_{g_{k+1}}^* D_1L_d(g_{k+1}, g_{k+2}) \tag{5.58}$$

Example 5.3.7. Consider the banal groupoid $M \times M \rightrightarrows M$, where the source and target maps are given by the projections onto the fist and second factor, respectively. The set of admissible elements is given by

$$(M \times M)_2 = \{((m_0, m_1), (\bar{m}_1, m_2)) \in (M \times M) \times (M \times M) \mid m_1 = \bar{m}_1\} \simeq 3M.$$

Moreover, in this case, the prolongation Lie groupoid is

$$\mathcal{P}^\alpha(M \times M) = \{((m_0, m_1), (m_2, m_3), (m_4, m_5)) \in 3(M \times M) \mid m_0 = m_2 \text{ and } m_3 = m_4\} \simeq 4M,$$

where we have the inclusion of $3M$ into $4M$ given by

$$\begin{array}{rcl} i_{3M} : 3M & \hookrightarrow & 4M \\ (m_0, m_1, m_2) & \mapsto & (m_0, m_1, m_1, m_2). \end{array}$$

Let $L : (M \times M)_2 \to \mathbb{R}$ be a discrete second-order Lagrangian and we construct the Lagrangian submanifold Σ_L of the symplectic groupoid $(T^*(4M), \omega_{4M})$, (where ω_{4M} denotes the canonical 2-form on $T^*(4M)$),

$$\Sigma_L = \{\mu \in T^*(4M) \mid i_{3M}^* \mu = dL\}$$

where $\mu = \mu_0 dm_0 + \mu_1 dm_1 + \bar{\mu}_1 d\bar{m}_1 + \mu_2 dm_2$. The following diagram illustrates the situation:

$$\begin{tikzcd} & \Sigma_L \arrow[r, hook] & T^*(4M) \arrow[r, shift left, "\tilde{\alpha}"] \arrow[r, shift right, "\tilde{\beta}"'] \arrow[d] & T^*(M \times M) \arrow[d] \\ \mathbb{R} & 3M \arrow[l, "L"'] \arrow[r, hook, "i_{3M}"] & 4M \arrow[r, shift left, "\alpha^\alpha"] \arrow[r, shift right, "\beta^\alpha"'] & (M \times M) \end{tikzcd}$$

Therefore, $\mu \in \Sigma_L$ if it satisfies

$$\begin{array}{rcl} \mu_0 & = & \dfrac{\partial L}{\partial m_0}(m_0, m_1, m_2), \\ \mu_1 + \bar{\mu}_1 & = & \dfrac{\partial L}{\partial m_1}(m_0, m_1, m_2), \\ \mu_2 & = & \dfrac{\partial L}{\partial m_2}(m_0, m_1, m_2). \end{array}$$

Using the source and target map given by

$$\begin{array}{rcl} \tilde{\alpha} & : & T^*(4M) \to T^*(M \times M) \quad (\mu_0, \mu_1, \bar{\mu}_1, \mu_2) \to (-\mu_0, -\mu_1,) \\ \tilde{\beta} & : & T^*(4M) \to T^*(M \times M) \quad (\mu_0, \mu_1, \bar{\mu}_1, \mu_2) \to (\bar{\mu}_1, \mu_2); \end{array}$$

we have that the second-order discrete dynamics is satisfies if and only if the following equations holds:

$$D_2 L(m_{k-1}, m_k, m_{k+1}) + D_1 L(m_k, m_{k+1}, m_{k+2}) + D_3 L(m_{k-2}, m_{k-1}, m_k) = 0,$$

for $k = 2, \ldots, N-2$.

5.3.2 Regularity conditions and Poisson structure

We have seen how the dynamics is implicitly defined by a relation on $T^*(\mathcal{P}^\alpha G)$ rather than an explicitly defined map. Also we have seen that $\gamma_{(g_1,g_2)},\ldots,\gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$ satisfies the discrete second-order dynamics if and only if each pair of successive elements in $T^*(\mathcal{P}^\alpha G)$ those satisfy

$$(\gamma_{(g_k,g_{k+1})},\gamma_{(g_{k+1},g_{k+2})}) \in (T^*(\mathcal{P}^\alpha G))_2 \cap (\Sigma_L \times \Sigma_L), \quad k=1,\ldots,N-2.$$

Weinstein [176] first raised the question of how regularity results for the pair groupoid $Q\times Q$ might be generalized to arbitrary Lie groupoids $G \rightrightarrows Q$, and this question was answered by Marrero et al. [[124], Theorem 4.13]. Here, we extend this answer to discrete second order systems following [126]. Then we ask: Under which conditions the last relation is the graph of an explicit flow

$$\gamma_{(g_{k-1},g_k)} \mapsto \gamma_{(g_k,g_{k+1})}$$

(at least locally) and what properties does this application have?

First, consider the source map of the cotangent groupoid $T^*(\mathcal{P}^\alpha G)$ restrict to the Lagrangian submanifold, that is, $\tilde{\alpha}\mid_{\Sigma_L}\colon \Sigma_L \to A^*(\mathcal{P}^\alpha G)$. If this map is a local diffeomorphism, then the Lagrangian flow is locally given by $\Gamma_L := (\tilde{\alpha}\mid_{\Sigma_L})^{-1}\circ\tilde{\beta}\mid_{\Sigma_L}\colon \Sigma_L \to \Sigma_L$. Moreover, if $\tilde{\beta}\mid_{\Sigma_L}\colon \Sigma_L \to A^*(\mathcal{P}^\alpha G)$ is also a local diffeomorphism the flow is reversible and its local inverse is given by $(\tilde{\beta}\mid_{\Sigma_L})^{-1}\circ\tilde{\alpha}\mid_{\Sigma_L}\colon \Sigma_L \to \Sigma_L$.

Proposition 5.3.8. *Given the symplectic groupoid $(T^*(\mathcal{P}^\alpha G);\omega_{\mathcal{P}^\alpha G})$ and $\Sigma_L \subset T^*(\mathcal{P}^\alpha G)$ is the Lagrangian submanifold generated by the second-order discrete Lagrangian $L: G_2 \to \mathbb{R}$, then $\tilde{\alpha}\mid_{\Sigma_L}\colon \Sigma_L \to \Sigma_L$ is a local diffeomorphism if and only if $\tilde{\beta}\mid_{\Sigma_L}\colon \Sigma_L \to \Sigma_L$ is a local diffeomorphism.*

Proof. Apply Th. 2.15 in [126] when $G = \mathcal{P}^\alpha G$. □

The applications $\tilde{\alpha}\mid_{\Sigma_L}$ and $\tilde{\beta}\mid_{\Sigma_L}$ plays the roll of $\mathbb{F}L^+$ and $\mathbb{F}L^-$ in discrete mechanics. We define the *Hamiltonian flow map* as the map given by $\tilde{\Gamma}_L := \tilde{\beta}\mid_{\Sigma_L}\circ(\tilde{\alpha}\mid_{\Sigma_L})^{-1} : A^*(\mathcal{P}^\alpha G) \to A^*(\mathcal{P}^\alpha G)$.

Proposition 5.3.9. *Let $T^*(\mathcal{P}^\alpha G) \rightrightarrows A^*(\mathcal{P}^\alpha G)$ be a symplectic Lie groupoid and $L: G_2 \to \mathbb{R}$ be a discrete second-order Lagrangian. If $\tilde{\alpha}\mid_{\Sigma_L}$ or $\tilde{\beta}\mid_{\Sigma_L}$ are local diffeomorphisms, the discrete Hamiltonian evolution operator $\widetilde{\Gamma}_L : A^*(\mathcal{P}^\alpha G) \to A^*(\mathcal{P}^\alpha G)$ preserves the Poisson structure on $A^*(\mathcal{P}^\alpha G)$.*

Proof. We will seen that $\widetilde{\Gamma}_L$ is a local Poisson automorphism. By theorem (5.3.8) if $\tilde{\alpha}\mid_{\Sigma_L}$ (alternatively $\tilde{\beta}\mid_{\Sigma_L}$) is a local diffeomorphism then the Lagrangian flow map $(\tilde{\alpha}\mid_{\Sigma_L})^{-1}\circ\tilde{\beta}\mid_{\Sigma_L}$ is a local automorphism on Σ_L and reversing the order of composition the Hamiltonian flow is a local automorphism on $A^*(\mathcal{P}^\alpha G)$. To see that the Hamiltonian evolution operator is a local Poisson automorphism consider the Poisson map

$$\begin{aligned}(\tilde{\alpha},\tilde{\beta}) : T^*(\mathcal{P}^\alpha G) &\longrightarrow \overline{A^*(\mathcal{P}^\alpha G)} \times A^*(\mathcal{P}^\alpha G)\\ \mu &\mapsto (\tilde{\alpha}(\mu),\tilde{\beta}(\mu));\end{aligned}$$

where $\overline{A^*(\mathcal{P}^\alpha G)}$ denotes $A^*(\mathcal{P}^\alpha G)$ endowed with the linear Poisson structure changed of sign. The image of Σ_L is just the graph of $\tilde{\beta}\mid_{\Sigma_L} \circ(\tilde{\alpha}\mid_{\Sigma_L})^{-1}$ in $\overline{A^*(\mathcal{P}^\alpha G)} \times A^*(\mathcal{P}^\alpha G)$. As Σ_L is a Lagrangian submanifold, its image under the Poisson map application $(\tilde{\alpha}, \tilde{\beta})$ is coisotropic. Then $\tilde{\beta}\mid_{\Sigma_L} \circ(\tilde{\alpha}\mid_{\Sigma_L})^{-1}$ is a local Poisson automorphism on $A^*(\mathcal{P}^\alpha G)$ since its graph is coisotropic in $\overline{A^*(\mathcal{P}^\alpha G)} \times A^*(\mathcal{P}^\alpha G)$. □

Definition 5.3.10. *Let $G \rightrightarrows Q$ be a Lie groupoid. A discrete second-order Lagrangian system is a pair (G, L) where $L : G_2 \to \mathbb{R}$ is a discrete second-order Lagrangian.*

Definition 5.3.11. *Let (G, L) be a discrete second-order Lagrangian system. The discrete Legendre transformations $\mathbb{F}L^{\pm} : \Sigma_L \to A^*(\mathcal{P}^\alpha G)$ are defined as*

$$\mathbb{F}L^+ = \tilde{\beta}\mid_{\Sigma_L} \quad \textit{and} \; \mathbb{F}L^- = \tilde{\alpha}\mid_{\Sigma_L} .$$

Remark 5.3.12. From (5.3.8) $\mathbb{F}L^+$ is a local diffeomorphism if and only if $\mathbb{F}L^-$ is a local diffeomorphism. ◇

Definition 5.3.13. *Let (G, L) be a discrete second-order Lagrangian system; this said to be regular if $\mathbb{F}L^{\pm}$ are local diffeomorphisms; and hyperregular if $\mathbb{F}L^{\pm}$ are global diffeomorphisms.*

Given a sequence $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$; this satisfies (by (5.3.1)) the discrete second-order dynamics on Σ_L if and only if $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in \Sigma_L$ and

$$\mathbb{F}L^+(\gamma_{(g_k,g_{k+1})}) = \mathbb{F}L^-(\gamma_{(g_{k+1},g_{k+2})}), \quad k = 1, \ldots, N-2.$$

5.3.3 Morphism, reduction and Noether symmetries

In following paragraphs we study the reduction of discrete second-order Lagrangian systems. To do this, we need to define the notion of morphism of discrete second-order Lagrangian systems in the dual of the cotangent Lie groupoid.

Let us consider two Lie groupoids, $G \rightrightarrows Q$ with source map denoted by α and $G' \rightrightarrows Q'$ with source map denoted by α'.

Definition 5.3.14. *Given two discrete second-order Lagrangian systems (G, L) and $(G', \tilde{L})$, a smooth map $\chi : \mathcal{P}^\alpha G \to \mathcal{P}^{\alpha'} G'$ is a morphism of discrete second-order Lagrangian systems if it is a morphism of Lie groupoids and, satisfies that $G_2 = \chi^{-1}(G'_2)$ and $L = \tilde{L} \circ \chi\mid_{G'_2}$.*

Definition 5.3.15. *Let $\chi : \mathcal{P}^\alpha G \to \mathcal{P}^{\alpha'} G'$ be a morphism of Lie groupoids. Two covectors $\mu \in T^*_{(g,h,r)}(\mathcal{P}^\alpha G)$ and $\mu' \in T^*_{\chi(g,h,r)}(\mathcal{P}^{\alpha'} G')$, where $(g, h, r) \in \mathcal{P}^\alpha G$ and $\chi(g, h, r) \in \mathcal{P}^{\alpha'} G'$ are said to be χ^*-related if*

$$\langle \mu, \xi \rangle = \langle \mu', T\chi(\xi) \rangle \qquad \forall \xi \in T_{(g,h,r)}(\mathcal{P}^\alpha G).$$

Also, if $z \in A^*_{(g,h,r)}(\mathcal{P}^\alpha G)$ and $z' \in A^*_{\chi_0(g,h,r)}(\mathcal{P}^{\alpha'} G')$, are $A^*\chi$−related if

$$\langle z, \widetilde{\xi} \rangle = \langle z', A\chi(\widetilde{\xi}) \rangle \qquad \forall \widetilde{\xi} \in A_{(g,h,r)}(\mathcal{P}^\alpha G).$$

Here $\chi_0 : \mathcal{P}^\alpha G \to \mathcal{P}^{\alpha'} G'$ denotes the smooth map on the base induced by the morphism χ and $A\chi : A(\mathcal{P}^\alpha G) \to A(\mathcal{P}^{\alpha'} G')$ is the associated Lie algebroid morphism (see section 1.9.2).

In the following we will give the theorem of reduction of second-order discrete Lagrangian systems on Lie groupoids.

Theorem 5.3.16. *Consider two second-order discrete Lagrangian systems (G, L) and $(G', \tilde{L})$ and let $\chi : \mathcal{P}^{\alpha}G \to \mathcal{P}^{\alpha'}G'$ be a morphism of discrete second-order Lagrangian systems. Suppose that $\mu \in T^*(\mathcal{P}^{\alpha}G)$ and $\mu' \in T^*(\mathcal{P}^{\alpha'}G')$ are χ^*-related, then the following properties hold*

- *If $\mu' \in \Sigma_{\tilde{L}}$ then $\mu \in \Sigma_L$*
- *The sources $\tilde{\alpha}(\mu) \in A^*(\mathcal{P}^{\alpha}G)$ and $\tilde{\alpha}'(\mu') \in A^*(\mathcal{P}^{\alpha'}G')$ are $A^*\chi$-related;*
- *The targets $\tilde{\beta}(\mu) \in A^*(\mathcal{P}^{\alpha}G)$ and $\tilde{\beta}'(\mu') \in A^*(\mathcal{P}^{\alpha'}G')$ are $A^*\chi$-related.*

Proof. To prove the first statement, let us consider $\mu \in T^*(\mathcal{P}^{\alpha}G)$ and $v \in T_{(g,h,r)}(\mathcal{P}^{\alpha}G)$, then since μ and μ' are χ^*-related we have

$$\langle \mu, v\rangle = \langle \mu', T\chi(v)\rangle = \langle dL', T\chi\mid_{G_2}(v)\rangle = \langle \chi^*(dL'), v\rangle = \langle d(L' \circ \chi\mid_{G_2}), v\rangle = \langle dL, v\rangle$$

and therefore $\mu \in \Sigma_L$.

To prove the point (ii) let us consider $v \in A_{\alpha(g,h,r)}(\mathcal{P}^{\alpha}G)$. Then $\langle \tilde{\alpha}(\mu), v\rangle = \langle \mu, \overrightarrow{v}\rangle = \langle \mu', T\chi(\overrightarrow{v})\rangle$, because are χ^*-related. Also, using the properties given in 1.9.2 we have that $\langle \mu', T\chi(\overrightarrow{v})\rangle = \langle \mu', \overrightarrow{A\chi(v)}\rangle = \langle \tilde{\alpha}'(\mu'), A\chi(v)\rangle$. Therefore, $\langle \tilde{\alpha}'(\mu), v\rangle = \langle \tilde{\alpha}(\mu'), A\chi(v)\rangle$, and then $\tilde{\alpha}(\mu)$ and $\tilde{\alpha}'(\mu')$ are $A^*\chi-$related.

Finally, to prove the point (iii),

$$\langle \tilde{\beta}(\mu), v\rangle = \langle \mu, \overleftarrow{v}\rangle = \langle \mu', T\chi(\overleftarrow{v})\rangle = \langle \mu', \overleftarrow{A\chi(v)}\rangle = \langle \tilde{\beta}'(\mu'), A\chi(v)\rangle.$$

Thus, $\tilde{\beta}(\mu)$ and $\tilde{\beta}'(\mu')$ are $A^*\chi-$related. $\square$

The following diagram shown the relations involved in the reduction theorem,

$$\begin{tikzcd}
& \Sigma_{\tilde{L}} \arrow[rr, hook] & & T^*(\mathcal{P}^{\alpha'}G') \arrow[rr, shift left, "\tilde{\alpha}'"] \arrow[rr, shift right, "\tilde{\beta}'"'] \arrow[dl, "T^*\Phi"'] \arrow[dd, "\pi_{\mathcal{P}^{\alpha'}G'}"'] & & A^*(\mathcal{P}^{\alpha'}G') \arrow[dl, squiggly, "A^*\Phi"'] \arrow[dd, "\tau_{A^*(\mathcal{P}^{\alpha'}G')}"] \\
\Sigma_L \arrow[rr, hook] & & T^*(\mathcal{P}^{\alpha}G) \arrow[rr, shift left, "\tilde{\alpha}"] \arrow[rr, shift right, "\tilde{\beta}"'] \arrow[dd, "\pi_{\mathcal{P}^{\alpha}G}"'] & & A^*(\mathcal{P}^{\alpha}G) \arrow[dd, "\tau_{A^*(\mathcal{P}^{\alpha}G)}"] & \\
& G'_2 \arrow[rr, hook] \arrow[dd, "\tilde{L}"'] & & \mathcal{P}^{\alpha'}G' \arrow[rr, shift left, "\alpha'"] \arrow[rr, shift right, "\beta'"'] & & G' \\
G_2 \arrow[rr, hook] \arrow[dd, "L"'] & & \mathcal{P}^{\alpha}G \arrow[ur, "\Phi"] \arrow[rr, shift left, "\alpha^{\alpha}"] \arrow[rr, shift right, "\beta^{\alpha}"'] & & G \arrow[ur, "\Phi_0"] & \\
& \mathbb{R} \arrow[dl, equal] & & & & \\
\mathbb{R} & & & & &
\end{tikzcd}$$

Corollary 5.3.17. *Let $\chi : \mathcal{P}^\alpha G \to \mathcal{P}^{\alpha'} G'$ be a morphism of discrete second-order Lagrangian systems. If $\gamma'_{(g_1,g_2)}, \ldots, \gamma'_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^{\alpha'} G')$ satisfy the discrete second order dynamics for $(G', \tilde{L})$, then any sequence χ^*−related $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$ satisfy the discrete second order dynamics for (G, L).*

Proof. By theorem 5.3.16, if $\mu'_k \in \Sigma_{\tilde{L}}$ then $\mu_k \in \Sigma_L$ for $k = 1, \ldots, n$. Moreover, for all $v \in A_{\beta(g_k,g_k,g_{k+1})}(\mathcal{P}^\alpha G) = A_{\alpha(g_{k+1},g_{k+1},g_{k+2})}(\mathcal{P}^\alpha G)$, using that μ and μ' are χ^*−related, we have that

$$\begin{aligned}
\langle \tilde{\beta}(\gamma_{(g_k,g_{k+1})}), \xi \rangle &= \langle \tilde{\beta}'(\gamma'_{(g_k,g_{k+1})}); A\chi(\xi) \rangle \\
&= \langle \tilde{\alpha}'(\gamma'_{(g_{k+1},g_{k+2})}); A\chi(\xi) \rangle \\
&= \langle \tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})}); \xi \rangle;
\end{aligned}$$

then, $\tilde{\beta}(\gamma_{(g_k,g_{k+1})}) = \tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})})$ for $k = 1, \ldots, N-2$. Therefore, $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)}$ satisfy the discrete second-order dynamics for (G, L). □

In what follows we will give the notion of Noether symmetry and constant of motion for discrete second order Lagrangian systems.

Definition 5.3.18. *A section $Z \in \Gamma(A(\mathcal{P}^\alpha G))$ is said to be a Noether symmetry of the discrete second-order Lagrangian system (G, L) if there exists a function $f \in C^\infty(G)$ such that*

$$\langle \tilde{\alpha}(\mu), Z(\tilde{\alpha}(g,h,r)) \rangle + f(\tilde{\alpha}(g,h,r)) = \langle \tilde{\beta}(\mu), Z(\tilde{\beta}(g,h,r)) \rangle + f(\tilde{\beta}(g,h,r))$$

for all $\mu \in \Sigma_L \subset T^*(\mathcal{P}^\alpha G)$, where $(g,h,r) = \pi_{\mathcal{P}^\alpha G}(\mu)$, where $\pi_{\mathcal{P}^\alpha G} : T^*(\mathcal{P}^\alpha G) \to \mathcal{P}^\alpha G$ is the cotangent bundle projection.

In the following theorem we will prove that for all Noether symmetry of the discrete second-order Lagrangian system there is a corresponding constant of motion which is preserved by the discrete second order dynamics.

Theorem 5.3.19. *If $Z \in \Gamma(A(\mathcal{P}^\alpha G))$ is a Noether symmetry of a discrete second-order Lagrangian system (G, L) then, the function $F_X : \Sigma_L \to \mathbb{R}$ given by*

$$F_X(\mu) = \langle \tilde{\alpha}(\mu), Z(\tilde{\alpha}(g,h,r)) \rangle + f(\tilde{\alpha}(g,h,r)) = \langle \tilde{\beta}(\mu), Z(\tilde{\beta}(g,h,r)) \rangle + f(\tilde{\beta}(g,h,r)),$$

where $(g,h,r) = \pi_{\mathcal{P}^\alpha G}(\mu)$; is a constant of motion. That is, if $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)} \in T^*(\mathcal{P}^\alpha G)$ satisfy the discrete second-order dynamics then, $F_X(\gamma_{(g_k,g_{k+1})}) = F_X(\gamma_{(g_{k+1},g_{k+2})})$ for $k = 1, \ldots, N-2$.

Proof. Since, $\gamma_{(g_1,g_2)}, \ldots, \gamma_{(g_{N-1},g_N)}$ satisfy the discrete second-order dynamics, $\tilde{\beta}(\gamma_{(g_k,g_{k+1})}) = \tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})})$ where $\gamma_{(g_k,g_{k+1})} \in \Sigma_L$ for $k = 1, \ldots, N-1$. Then, we have that

$$
\begin{aligned}
F_X(\gamma_{(g_k,g_{k+1})}) &= \langle \tilde{\beta}(\gamma_{(g_k,g_{k+1})}), Z(\tilde{\beta}(\gamma_{(g_k,g_{k+1})}))\rangle + f(\tilde{\beta}(\gamma_{(g_k,g_{k+1})})) \\
&= \langle \tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})}), Z(\tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})}))\rangle + f(\tilde{\alpha}(\gamma_{(g_{k+1},g_{k+2})})) \\
&= F_X(\gamma_{(g_{k+1},g_{k+2})}), \quad k = 1, \dots, N-2.
\end{aligned}
$$

□

5.3.4 Discrete second-order constrained mechanics

In this subsection we apply the previous techniques to the case of second-order constrained systems.

Definition 5.3.20. *A discrete second-order constrained Lagrangian system consists of a triple (G, N, L) where $G \rightrightarrows Q$ is a Lie groupoid, $N \subset G_2$ is a submanifold and $L : N \to \mathbb{R}$ is a discrete second-order Lagrangian.*

From Tulczyjew theorem (1.6.11) it is clear that

$$\Sigma_{L,N} = \{\mu \in T^*(\mathcal{P}^\alpha G) \mid i_N^*\mu = dL_N\} \subset T^*(\mathcal{P}^\alpha G)$$

is a Lagrangian submanifold. The following diagram shows the situation:

$$
\begin{tikzcd}
 & \Sigma_{L,N} \arrow[r, hook] & T^*(\mathcal{P}^\alpha G) \arrow[r, shift left, "\tilde{\alpha}"] \arrow[r, shift right, "\tilde{\beta}"'] \arrow[d] & A^*(\mathcal{P}^\alpha G) \arrow[d] \\
\mathbb{R} & N \arrow[l, "L"'] \arrow[r, hook, "i_N"] & \mathcal{P}^\alpha G \arrow[r, shift left, "\alpha^\alpha"] \arrow[r, shift right, "\beta^\alpha"'] & G
\end{tikzcd}
$$

Note that $\Sigma_{L,N}$ is also a bundle over N taking the projection $\pi_{\mathcal{P}^\alpha G} |_{\Sigma_{L,N}} : \Sigma_{L,N} \to N$ to be the restriction of the cotangent bundle projection.

Consider the *conormal bundle* $\nu^*N = \{\mu \in T^*(\mathcal{P}^\alpha G) | i_N^*\mu = 0\}$. From the definition of conormal bundle its follows that $\Sigma_{L,N}$ is isomorphic to the conormal bundle where the isomorphism is given to specify a distinguished section $\sigma : N \to \Sigma_{L,N}$ and then

$$\Sigma_{L,N} = \{\sigma(g_1, g_2) + \Lambda \mid \Lambda \in \nu^*N, \text{ and } (g_1, g_2) = \pi_{\mathcal{P}^\alpha G}(\Lambda)\} \simeq \nu^*N.$$

Now we take an arbitrary extension $\widehat{L}$ of L to a neighborhood of N in G_2, so that $L = \widehat{L} |_N$. Therefore, an alternative description of $\Sigma_{L,N}$ is

$$\Sigma_{L,N} = \{\mu \in T^*(\mathcal{P}^\alpha G) | \mu - d\widehat{L} \in \nu^*N\}.$$

Suppose that the constraint submanifold $N \subset G_2$ is given by

$$N = \{(g_1, g_2) \in G_2 \mid \Phi^a(g_1, g_2) = 0, \text{ with } a \in A\},$$

where $\{\Phi^a\}_{a\in A}$ is a family of real functions defined in a neighborhood of N and A is an index set. It follows that $\{d\Phi^a |_N\}$ is a basis of sections of the conormal bundle ν^*N. Therefore, a section Λ of the conormal bundle can be written as $\Lambda = \lambda_a d\Phi^a \mid_N$, where the real functions $\lambda_a, a \in A$ are the Lagrange multipliers. Since $\Phi^a \mid_N = 0$ we can deduce that $\Lambda = d(\lambda_a \Phi^a \mid_N)$.

As before, suppose that $L : N \to \mathbb{R}$ is the restriction to N of $\widehat{L} : G_2 \to \mathbb{R}$. Then, an element $\mu \in \Sigma_{L,N}$ with $(g_1, g_2) = \pi_{\mathcal{P}^\alpha G}(\mu)$, can be written as

$$\mu = d\widehat{L}(g_1, g_2) + \lambda_\alpha d\Phi^\alpha(g_1, g_2) = d(\widehat{L} + \lambda_a \Phi^a)(g_1, g_2) \in \Sigma_{L,N}.$$

In this sense, $\Sigma_{L,N}$ can be seen as the space consisting of the elements $(g_1, g_2) \in N$ together the Lagrange multipliers λ_a constraining (g_1, g_2) to N.

Therefore applying Proposition (5.3.4) and (5.50), (5.51) and (5.52) the sequence of composable elements and Lagrange multipliers determined by the elements $(g_1, g_2), (g_2, g_3), \ldots, (g_{N-1}, g_N),\ \lambda_1, \lambda_2, \ldots, \lambda_{N-1}$ is a solution of the discrete second-order Lagrangian dynamic when $g_1, g_2, \ldots, g_N \in N$ and

$$\begin{aligned} 0 \;=\; & \Big\langle \overrightarrow{X}(g_{k+1}), D_1\widehat{L}_d(g_{k+1}, g_{k+2}) + (\lambda_{k+1})_a D_1\Phi^a(g_{k+1}, g_{k+2}) + D_2\widehat{L}_d(g_k, g_{k+1}) \\ & +(\lambda_k)_a D_2\Phi^a(g_k, g_{k+1})\Big\rangle - \Big\langle \overleftarrow{X}(g_k), D_1\widehat{L}_d(g_k, g_{k+1}) + (\lambda_k)_a D_1\Phi^a(g_k, g_{k+1}) \\ & +D_2\widehat{L}_d(g_{k-1}, g_k) + (\lambda_{k-1})_a D_2\Phi^a(g_{k-1}, g_k)\Big\rangle \end{aligned}$$

That is, the sequence $(g_1, g_2, \lambda_1), (g_2, g_3, \lambda_2), \ldots, (g_{N-1}, g_N, \lambda_{N-1})$ satisfies

$$\begin{aligned} 0 &= \Phi^a(g_k, g_{k+1}),\ \text{for all } a \in A, \quad k = 1, \ldots N-1; \\ 0 &= \ell^*_{g_k}\Big(D_1\widehat{L}_d(g_k, g_{k+1}) + (\lambda_k)_a D_1\Phi^a(g_k, g_{k+1}) \\ &\quad +D_2\widehat{L}_d(g_{k-1}, g_k) + (\lambda_{k-1})_a D_2\Phi^a(g_{k-1}, g_k)\Big),\ \text{for } k = 2, \ldots, N-2; \\ 0 &= (r_{g_{k+1}} \circ i)^*\Big(D_1\widehat{L}_d(g_{k+1}, g_{k+2}) + (\lambda_{k+1})_a D_1\Phi^a(g_{k+1}, g_{k+2}) \\ &\quad +D_2\widehat{L}_d(g_k, g_{k+1}) + (\lambda_k)_a D_2\Phi^a(g_k, g_{k+1})\Big),\ \text{for } k = 2, \ldots, N-2. \end{aligned}$$

Remark 5.3.21. Observe that, when the Lie groupoid is a Lie group, we obtain the second-order Euler-Poincaré equations for systems with constraints (see also [53])

$$\begin{aligned} 0 &= \Phi^a(g_k, g_{k+1}), \quad 0 = \Phi^a(g_{k-1}, g_k), \quad 0 = \Phi^a(g_{k+1}, g_{k+2}) \text{ for all } a \in A; \\ 0 &= \ell^*_{g_k}\Big(D_1\widehat{L}_d(g_k, g_{k+1}) + (\lambda_k)_a D_1\Phi^a(g_k, g_{k+1}) \\ &\quad +D_2\widehat{L}_d(g_{k-1}, g_k) + (\lambda_{k-1})_a D_2\Phi^a(g_{k-1}, g_k)\Big); \\ 0 &= (r_{g_{k+1}})^*\Big(D_1\widehat{L}_d(g_{k+1}, g_{k+2}) + (\lambda_{k+1})_a D_1\Phi^a(g_{k+1}, g_{k+2}) \\ &\quad +D_2\widehat{L}_d(g_k, g_{k+1}) + (\lambda_k)_a D_2\Phi^a(g_k, g_{k+1})\Big),\ \text{for } k = 2, \ldots, N-2. \end{aligned}$$

Moreover, when the Lie groupoid is the Banal groupoid we have second-order Euler-Lagrange equations for systems with constraints given by

$$\begin{aligned}
0 &= \Phi^a(q_k, q_{k+1}, q_{k+2}), 0 = \Phi^a(q_{k-1}, q_k, q_{k+1}), 0 = \Phi^a(q_{k-2}, q_{k-1}, q_k) \text{ for all } a \in A;\\
0 &= D_1\widehat{L}_d(q_k, q_{k+1}, q_{k+2}) + (\lambda_k)_a D_1\Phi^a(q_k, q_{k+1}, q_{k+2}) + D_2\widehat{L}_d(q_{k-1}, q_k, q_{k+1})\\
&+ (\lambda_{k-1})_a D_2\Phi^a(q_{k-1}, q_k, q_{k+1}) + D_3\widehat{L}_d(q_{k-2}, q_{k-1}, q_k) + (\lambda_{k-2})_a D_3\Phi^a(q_{k-2}, q_{k-1}, q_k),
\end{aligned}$$

for $k = 2, \ldots, N-2$.

These equations are just the discrete second-order Euler-Lagrange equations for systems with second-order constraints (See for example [50], [51] and [52]).

$\diamond$

Chapter 6

Discrete mechanics and optimal control

In previous approaches (see for example [14] and [52]), the theory of discrete variational mechanics for higher-order systems was derived using a discrete lagrangian $L_d : Q^{k+1} \to \mathbb{R}$ where Q^{k+1} is the cartesian product of $k+1$- copies of the configuration manifold Q. In some sense, this is a very natural discretization since we are using $k+1$-points to approximate the positions and the higher-order velocities (such as the standard velocities,more general accelerations, jerks...) which represents the higher-order tangent bundle $T^{(k)}Q$.

We will see in this chapter, that other possibility more general is to take a Lagrangian function $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ since really the discrete variational calculus is not based on the discretization of the Lagrangian itself, but on the discretization of the associated action. We will see that the appropriate approximation of the action

$$\int_0^h L(q, \dot{q}, \dots, q^{(k)})\, dt \tag{6.1}$$

is given by a Lagrangian of the form $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$. Moreover, we will derive a particular choice of discrete Lagrangian which gives an exact correspondence between discrete and continuous systems, the exact discrete Lagrangian. For instance, if we take the Lagrangian $L(q, \dot{q}, \ddot{q}) = \frac{1}{2}\ddot{q}^2$, the corresponding exact discrete Lagrangian $L_d^e : TQ \times TQ \to \mathbb{R}$ is

$$\begin{aligned} L_d^e(q_0, v_0, q_h, v_h) &= \int_0^h L(q(t), \dot{q}(t), \ddot{q}(t))\, dt \\ &= \frac{6}{h^3}(q_0 - q_h)^2 + \frac{6}{h^2}(q_0 - q_h)(v_0 + v_h) + \frac{2}{h}(v_0^2 + v_0 v_h + v_h^2) \end{aligned}$$

where $q(t)$ is the unique solution of the Euler-Lagrange equations for L verifying $q(0) = q_0$, $\dot{q}(0) = v_0$, $q(h) = q_h$, $\dot{q}(h) = v_h$ for h enough small.

Observe from the previous example that now this theory of variational integrators for higher-order system is even simpler, since it fits directly into the standard discrete mechanics theory for a discrete lagrangian of the form $L_d : M \times M \to \mathbb{R}$ where $M = T^{(k-1)}Q$.

6.1 Existence and uniqueness of solutions for higher order-systems

Given k nearby pairs $(q_0, v_0), (q_1, v_1), \ldots, (q_{k-1}, v_{k-1}) \in TQ$ does there exists a unique evolution curve solution of an explicit $2k$ order ordinary differential equation $q(t)$ such that $q(0), \dot{q}(0), \ldots, q^{(k-1)}(0), q(h), \ldots, q^{(k-1)}(h)$ for h enough small?

Standard ODE theory provides existence and uniqueness of the corresponding initial value problem for an explicit higher-order ordinary differential equation, but also it is possible to show that given an explicit $2k$ order differential equation one can choose different boundary conditions to guarantee the existence and uniqueness of solutions for the initial value problem. For instance, in [4] the author stated that there exists an unique solution of an explicit $2k$ ODE verifying the boundary values $q_0, \ldots, q_{k-1}, v_0, \ldots, v_{k-1}$ assuming C^{2k}-differentiability. Nevertheless, we will follow a different way, using the variational origin of the equations of motion. In our case, we have a Lagrangian function $L : T^{(k)}Q \to \mathbb{R}$, the action functional (6.1) and the corresponding higher-order Euler-Lagrange equations (6.2) assuming only C^k-differentiability.

Here we show that a regular higher-order Lagrangian system has a unique solution for given nearby endpoint conditions using a direct variational proof of existence and uniqueness of the local boundary values problem using a regularization procedure. It results in the replacement of the variational problem with an equivalent one which is regular at $h = 0$. The argument follows closely the proof by Patrick [150] for first-order Lagrangians; the formulas, of course, reduce to those in [150] for order 1, but we introduce an additional modification using orthonormal polynomials.

6.1.1 Non-regularity of Hamilton's principle

Let Q be a finite-dimensional manifold and $L\colon T^{(k)}Q \to \mathbb{R}$ be a smooth Lagrangian of order $k \geq 1$. We take L to be a C^r function, $r \geq 2$. Since the result will be local, we assume from now on that Q is an open subset of $\mathbb{R}^n$. Let $(q^{[0]}, q^{[1]}, \ldots, q^{[k]})$ be coordinates on $T^{(k)}Q \equiv Q \times (\mathbb{R}^n)^k$. We suppose that L is regular in the sense that the hessian matrix

$$\left(\frac{\partial^2 L}{\partial q^{[k]2}}\right)$$

is a regular matrix.

In this setting, we can formulate Hamilton's principle as follows.

Variational Principle 1. *Find a C^k curve $q\colon [0,h] \to Q$ among those whose first $k-1$ derivatives are fixed at the endpoints, such that it is a critical point of the action*

$$S_h = \int_0^h L\left(q(t), \dot{q}(t), \ldots, q^{(k)}(t)\right)\, dt. \qquad \triangle$$

It is known that the critical points are the solutions of the k^{th}-order Euler-Lagrange equations

$$\sum_{j=0}^{k}(-1)^j \frac{d^j}{dt^j}\frac{\partial L}{\partial q^{[j]}}(q(t),\dot{q}(t),\ldots,q^{(k)}(t)) = 0. \tag{6.2}$$

We want to determine whether there exists a unique solution curve for this variational principle, given endpoint conditions that are close enough. The main obstacle for a straightforward affirmative answer is that the local boundary value problem as stated above is non-regular at $h = 0$. That is, the constraint function $g : C^k \to (\mathbb{R}^n)^k \times (\mathbb{R}^n)^n$

$$g : q(\cdot) \mapsto \left(q(0), \dot{q}(0), \ldots, q^{(k-1)}(0); q(h), \dot{q}(h), \ldots, q^{(k-1)}(h)\right),$$

maps into the diagonal of $T^{(k-1)}Q \times T^{(k-1)}Q$ for $h = 0$ and is not therefore a submersion. For $h \neq 0$, the constraint function is a submersion.

The approach consists in replacing this problem by an equivalent one that is regular at $h = 0$, and show that locally there is a unique solution to the regularized problem.

6.1.2 Regularization

First we replace the space of curves on Q in the variational problem by the space of curves on $T^{(k)}Q$, and include additional constraints. Denote an arbitrary curve by

$$\left(q(t) = q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t)\right) \in T^{(k)}Q \equiv Q \times (\mathbb{R}^n)^k,$$

$t \in [0, h]$. Here we have used superscripts in square brackets to make a distinction with the actual derivatives of $q(t)$.

Variational Principle 2. *Find a curve $(q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t))$ on $T^{(k)}Q$, where $q^{[l]} \in C^{k-l}([0,h],\mathbb{R}^n)$, $l = 0, \ldots, k$, such that it is a critical point of*

$$S_h = \int_0^h L\left(q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t)\right) dt$$

subject to the constraints

$$q^{[j+1]}(t) = \frac{dq^{[j]}}{dt}(t), \quad q^{[j]}(0) = q_1^{[j]}, \quad q^{[j]}(h) = q_2^{[j]}, \quad j = 0, \ldots, k-1,$$

where $(q_i^{[0]}, q_i^{[1]}, \ldots, q_i^{[k-1]})$, $i = 1, 2$, are given points in $T^{(k-1)}Q$. △

Now reparameterize the curve by defining

$$Q^{[j]}(u) = q^{[j]}(hu), \quad j = 0, \ldots, k, \quad u \in [0, 1].$$

For $h > 0$, the curve $(Q^{[0]}(u), \ldots, Q^{[k]}(u))$ satisfies an equivalent variational problem as follows. Since h is a constant for each instance of the problem, we can use

$$\frac{1}{h}\int_0^h L\left(q^{[0]}(t), q^{[1]}(t), \ldots, q^{[k]}(t)\right) dt = \int_0^1 L\left(Q^{[0]}(u), \ldots, Q^{[k]}(u)\right) du$$

as an objective function. The first set of constraints becomes

$$0 = \frac{dq^{[j]}}{dt}(t) - q^{[j+1]}(t) = \left(\frac{1}{h} \frac{dQ^{[j]}}{du}(u) - Q^{[j+1]}(u) \right)_{u=t/h}$$

where $j = 0, \ldots, k-1$.

The reparametrized variational principle is the following.

Variational Principle 3. *Find a curve* $(Q^{[0]}(u), \ldots, Q^{[k]}(u))$ *on* $T^{(k)}Q$, $Q^{[l]} \in C^{k-l}([0,1], \mathbb{R}^n)$, $l = 0, \ldots, k$, *that is a critical point of*

$$S = \int_0^1 L(Q^{[0]}(u), \ldots, Q^{[k]}(u))\, du,$$

subject to the constraints

$$\frac{dQ^{[j]}}{du}(u) = hQ^{[j+1]}(u), \tag{6.3}$$

$$Q^{[j]}(0) = q_1^{[j]}, \tag{6.4}$$

$$Q^{[j]}(1) = q_2^{[j]}, \tag{6.5}$$

where $j = 0, \ldots, k-1$, *and* $(q_i^{[0]}, q_i^{[1]}, \ldots, q_i^{[k-1]})$, $i = 1, 2$, *are given points in* $T^{(k-1)}Q$. $\triangle$

The objective S does not depend on h. Moreover, denoting by

$$\widetilde{C} = \{(Q^{[0]}(u), \ldots, Q^{[k]}(u)) \in T^{(k)}Q, Q^{[l]} \in C^{k-l}([0,1], \mathbb{R}^n), l = 0, \ldots, k\},$$

the constraints $g_{3,h} : \widetilde{C} \to \{0\}^k \times (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k$, given by

$$g_{3,h}(Q^{[0]}, \ldots, Q^{[k]}) =$$
$$\left(\frac{dQ^{[0]}(u)}{du} - hQ^{[1]}(u), \ldots, \frac{dQ^{[k-1]}(u)}{du} - hQ^{[k]}(u), Q^{[0]}(0), \ldots, Q^{[k-1]}(0), Q^{[0]}(1), \ldots, Q^{[k-1]}(1) \right)$$

is smooth through $h = 0$. That is, the application $h \mapsto g_{3,h} \in L(\widetilde{C}, \{0\}^k \times (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k)$ is smooth through $h = 0$. Here $L(\widetilde{C}, \{0\}^k \times (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k)$ denotes the set of linear transformations form $\widetilde{C}$ to $\{0\}^k \times (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k)$.

Remark 6.1.1. For $h = 0$, the constraints (6.3) imply that $Q^{[0]}(u), \ldots, Q^{[k-1]}(u)$ remain constant, which restricts the possible values of the endpoint conditions in order to have a compatible set of constraints. More precisely, $q_1^{[j]} = q_2^{[j]}$ for $j = 0, \ldots, k-1$; otherwise there would be no curves satisfying the constraints. This kind of restriction also appears in the original variational principle 1. Moreover, the problem becomes the unconstrained problem of finding a curve $Q^{[k]}(u) \in C^0([0,1], \mathbb{R}^n)$ such that it is a critical point of

$$\int_0^1 L(q^{[0]}, \ldots, q^{[k-1]}, Q^{[k]}(u)) du.$$

This means
$$\frac{\partial L}{\partial q^{[k]}}(q^{[0]}, q^{[1]}, \ldots, q^{[k-1]}, Q^{[k]}(u)) = 0. \tag{6.6}$$

Differentiating with respect to u,

$$\frac{\partial^2 L}{\partial {q^{[k]}}^2}(q^{[0]}, q^{[1]}, \ldots, q^{[k-1]}, Q^{[k]}(u))\frac{dQ^{[k]}(u)}{du} = 0.$$

Since the Lagrangian is regular, we obtain that $\frac{dQ^{[k]}(u)}{du} = 0$, that it, $Q^{[k]}$ is constant.

Therefore $Q^{[k]}(u)$ belongs without moving at critical points of $L(q^{[0]}, \ldots, q^{[k-1]}, \cdot) : \mathbb{R}^n \to \mathbb{R}$, which are isolated points in $\mathbb{R}^n$. ◇

In preparation for the next step for regularization, let us solve the constraints (6.3) to get

$$Q^{[j]}(u) = Q^{[j]}(0) + h\int_0^u Q^{[j+1]}(s)\, ds, \quad j = 0, \ldots, k-1.$$

This means that the functions $Q^{[j]}(u)$, $j = 0, \ldots, k-1$, can be expressed in terms of $Q^{[j]}(0)$, $\ldots$, $Q^{[k-1]}(0)$, the function $Q^{[k]}(u)$ and h. For example, for $k = 2$ we have

$$\begin{aligned}
Q^{[1]}(u) &= Q^{[1]}(0) + h\int_0^u Q^{[2]}(s)\, ds \\
Q^{[0]}(u) &= Q^{[0]}(0) + h\int_0^u Q^{[1]}(s)\, ds \\
&= Q^{[0]}(0) + huQ^{[1]}(0) + h^2\int_0^u\int_0^s Q^{[2]}(\tau)\, d\tau\, ds \\
&= Q^{[0]}(0) + huQ^{[1]}(0) + h^2\int_0^u (u-\tau)Q^{[2]}(\tau)\, d\tau
\end{aligned}$$

For a general k, and for $j = 0, \ldots, k-1$, an iterated change of order of integration yields

$$Q^{[j]}(u) = Q^{[j]}(0) + \sum_{i=1}^{k-j-1}\frac{h^i u^i}{i!}Q^{[j+i]}(0) + h^{k-j}\int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!}Q^{[k]}(s)\, ds. \tag{6.7}$$

If the upper bound of summation is less than the lower bound, by a standard convention the sum is 0.

Note that the final endpoint data $(q_2^{[0]}, \ldots, q_2^{[k-1]})$ can now be written as

$$q_2^{[j]} = Q^{[j]}(1) = q_1^{[j]} + \sum_{i=1}^{k-j-1}\frac{h^i}{i!}q_1^{[j+i]} + h^{k-j}\int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!}Q^{[k]}(s)\, ds, \tag{6.8}$$

so we define

$$z^{[j]} = \int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!}Q^{[k]}(s)\, ds = \frac{1}{h^{k-j}}\left(q_2^{[j]} - \sum_{i=0}^{k-j-1}\frac{h^i}{i!}q_1^{[j+i]}\right). \tag{6.9}$$

We will discuss the case $h = 0$ in Remark 6.1.2.

Now replace the curves and endpoint data by just $Q^{[k]}(u)$, $(q_1^{[0]}, \dots, q_1^{[k-1]})$, and $(z^{[0]}, \dots, z^{[k-1]})$, to get a new variational principle.

Variational Principle 4. *Given h, $(q_1^{[0]}, \dots, q_1^{[k-1]})$ and $(z^{[0]}, \dots, z^{[k-1]})$, find a continuous curve $Q^{[k]} \colon [0,1] \to \mathbb{R}^n$ that is a critical point of*

$$\mathcal{S} = \int_0^1 L(Q^{[0]}(u), \dots, Q^{[k]}(u))\, du,$$

where $Q^{[0]}(u)$, …, $Q^{[k-1]}(u)$ are defined as in (6.7) by

$$Q^{[j]}(u) = q_1^{[j]} + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} q_1^{[j+i]} + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds, \quad j = 0, \dots, k-1$$

subject to the constraints

$$\int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds = z^{[j]}, \quad j = 0, \dots, k-1. \qquad \triangle$$

Observe that the constraints $z^{[j]}$ with $j = 0, \dots, k-1$, do not depend of h and they are linear combinations of the curve $Q^{[k]}$.

This variational principle is regular through $h = 0$. Indeed, at $h = 0$, the problem becomes finding $Q^{[k]}$ such that it is a critical point of

$$\mathcal{S} = \int_0^1 L(q_1^{[0]}, \dots, q_1^{[k-1]}, Q^{[k]}(u))\, du,$$

subject to the constraints

$$\int_0^1 \frac{(1-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s)\, ds = z^{[j]}, \quad j = 0, \dots, k-1.$$

Using Lagrange multipliers $\lambda_0, \dots, \lambda_{k-1} \in \mathbb{R}^n$, this is solved by asking that for all $\delta Q^{[k]}(u)$,

$$\int_0^1 \frac{\partial L}{\partial q^{[k]}}(q_1^{[0]}, \dots, q_1^{[k-1]}, Q^{[k]}(u)) \delta Q^{[k]}(u)\, du = \sum_{j=0}^{k-1} \int_0^1 \lambda_j \cdot \frac{(1-u)^{k-j-1}}{(k-j-1)!} \delta Q^{[k]}(u)\, du.$$

This means

$$\frac{\partial L}{\partial q^{[k]}}(q_1^{[0]}, \dots, q_1^{[k-1]}, Q^{[k]}(u)) = \sum_{j=0}^{k-1} \lambda_j \frac{(1-u)^{k-j-1}}{(k-j-1)!},$$

which can be solved for $Q^{[k]}(u)$ if L is regular.

Remark 6.1.2. The data $q_1^{[0]}, \dots, q_1^{[k-1]}, z^{[0]}, \dots, z^{[k-1]}$ can be transformed into the endpoint conditions for the variational principle 3 in a straightforward way, for any h, using (6.8) and (6.9). The converse (6.9) is possible only for $h \neq 0$, in principle. However, if $h = 0$

let $(Q^{[0]}(u),\dots,Q^{[k]}(u))$ a solution for the variational principle 3 with boundary conditions $(q_1^{[0]},\dots q_1^{[k-1]})$ and $(q_2^{[0]},\dots,q_2^{[k-1]})$. Define $z^{[j]}$ by the constraint in 4. Since $Q^{[k]}$ is constant and $\frac{(1-s)^{k-j-1}}{(k-j-1)!}>0$ in $(0,1)$, from different values of $Q^{[k]}$ corresponds different values of $z^{[j]}$. Then $Q^{[k]}$ is a solution of 4 with boundary conditions $q_1^{[0]}$, …, $q_1^{[k-1]}$, $z^{[0]}$, …, $z^{[k-1]}$.

◇

Finally, we will introduce a modification that will enable us to carry out the computations in the next section easily. Consider the inner product on $C^0([0,1],\mathbb{R})$ given by

$$\langle f,g\rangle=\int_0^1 f(s)g(s)\,ds.$$

If $f\in C^0([0,1],\mathbb{R})$ and $V=(V_1,\dots,V_n)\in C^0([0,1],\mathbb{R}^n)$ we define the bilinear operation

$$\langle f,V\rangle\!\rangle=\int_0^1 f(s)V(s)\,ds=(\langle f,V_0\rangle,\dots,\langle f,V_n\rangle)\in\mathbb{R}^n.$$

The integrals appearing in the constraints in the variational principle 4 are $\langle a_j^{[k]},Q^{[k]}\rangle\!\rangle$, where $a_j^{[k]}$ are the polynomials

$$a_j^{[k]}(s)=\frac{(1-s)^{k-j-1}}{(k-j-1)!},\quad j=0,\dots,k-1.$$

These form a basis of the space of polynomials of degree at most $k-1$. Let us consider a basis $b_j^{[k]}(s)$, $j=0,\dots,k-1$, of the same space of polynomials consisting of orthonormal polynomials on $[0,1]$, and let $(\gamma_j^{[k],i})$, where $i,j=0,\dots,k-1$, be the invertible real matrix such that $a_j^{[k]}(s)=\gamma_j^{[k],i}b_i^{[k]}(s)$. For example, for $k=2$,

$$a_0^{[2]}(s)=1-s,\quad a_1^{[2]}(s)=1,$$

and we can take for instance the orthonormal basis

$$b_0^{[2]}(s)=\sqrt{3}(1-2s),\quad b_1^{[2]}(s)=1;$$

therefore,

$$\gamma_0^{[2],0}=\tfrac{1}{2\sqrt{3}},\quad \gamma_0^{[2],1}=\tfrac{1}{2},\quad \gamma_1^{[2],0}=0,\quad \gamma_1^{[2],1}=1.$$

Using this matrix, the constraints can be rewritten as

$$z^{[j]}=\langle a_j^{[k]},Q^{[k]}\rangle\!\rangle=\gamma_j^{[k],i}\langle b_i^{[k]}(s),Q^{[k]}\rangle\!\rangle,$$

for $j=0,\dots,k-1$. This allows us to reformulate the variational principle in an equivalent way by replacing the data $(z^{[0]},\dots,z^{[k-1]})$ and constraints $\langle a_j^{[k]},Q^{[k]}\rangle\!\rangle=z^{[j]}$ by new data $(w^{[0]},\dots,w^{[k-1]})$ and constraints $\langle b_j^{[k]},Q^{[k]}\rangle\!\rangle=w^{[j]}$, $j=0,\dots,k-1$. The old and new data are related by

$$\sum_{i=0}^{k-1}\gamma_j^{[k],i}w^{[i]}=z^{[j]}. \tag{6.10}$$

Remark 6.1.3. Given q_1 and q_2 we can construct $z^{[j]}$ from (6.9) $j = 0, \ldots, k-1$. Also we know the relation

$$\sum_{j=0}^{k-1} \tilde{\gamma}_m^{[k],j} \gamma_j^{[k],i} = \delta_m^i,$$

where $(\tilde{\gamma}_m^{[k],j})$ is the inverse matrix of $(\gamma_j^{[k],i})$. Therefore multiplying this last expression in (6.10) we can find a way to write the new data $w^{[j]}$ given $z^{[j]}$ as

$$w^{[m]} = \sum_{j=0}^{k-1} \tilde{\gamma}_m^{[k],j} z^{[j]}.$$

$\diamond$

Variational Principle 5. *Given h, $(q_1^{[0]}, \ldots, q_1^{[k-1]})$ and $(w^{[0]}, \ldots, w^{[k-1]})$, find a continuous curve $Q^{[k]} \colon [0,1] \to \mathbb{R}^n$ that is a critical point of*

$$S_h = \int_0^1 L(Q^{[0]}(u), \ldots, Q^{[k]}(u)) \, du,$$

where $Q^{[0]}(u)$, …, $Q^{[k-1]}(u)$ are defined by

$$Q^{[j]}(u) = q_1^{[j]} + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} q_1^{[j+i]} + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}(s) \, ds, \tag{6.11}$$

subject to the constraints

$$\int_0^1 b_j^{[k]}(s) Q^{[k]}(s) \, ds = w^{[j]}, \quad j = 0, \ldots, k-1.$$

$\triangle$

Solution of the regularized problem

Let S_h be given as in the variational principle 5, regarded as a real-valued map defined on the space $C^0([0,1], \mathbb{R}^n)$ of curves $Q^{[k]}(u)$. We can also think of it as defined on any of the Banach spaces $C^l([0,1], \mathbb{R}^n)$, where $0 \leq l < \infty$. We are going to use the following lemma [2].

Lemma 6.1.4 (Omega Lemma [2])**.** *Let E, F be Banach spaces, U open in E, and M a compact topological space. Let $g \colon U \to F$ be a C^r map, $r > 0$. The map*

$$\Omega_g \colon C^0(M, U) \to C^0(M, F) \quad \textit{defined by} \quad \Omega_g(f) = g \circ f$$

is also C^r, and $D\Omega_g(f) \cdot h = [(Dg) \circ f] \cdot h$.

The objective S_h is the composition of the maps

$$C^0([0,1], \mathbb{R}^n) \xrightarrow{i} C^0([0,1], T^{(k)}Q) \xrightarrow{\Omega_L} C^0([0,1], \mathbb{R}) \xrightarrow{\int} \mathbb{R}$$

where i is defined by $Q^{[k]}(u) \mapsto (Q^{[0]}(u), \ldots, Q^{[k]}(u))$ (observe that i depends on the initial values). Here $Q^{[0]}(u), \ldots,$ $Q^{[k-1]}(u)$ stand for the right-hand sides of (6.11). Both i and $\int$ are bounded affine and therefore C^∞ where $||Q^{[k]}||_{C^0} = sup_{u\in[0,1]}||Q^{[k]}(u)||$. By the Omega Lemma, Ω_L is C^r because L is C^r and therefore so is S_h.

If we regard S_h as defined on $C^l([0,1], \mathbb{R}^n)$, $0 \leq l < \infty$, then we should append the inclusion $C^l([0,1], \mathbb{R}^n) \hookrightarrow C^0([0,1], \mathbb{R}^n)$ to the left side of the diagram above. This inclusion is C^∞ because it is linear and bounded ($\|Q^{[k]}\|_{C^0} \leq \|Q^{[k]}\|_{C^l}$ for all $Q^{[k]}$). Then S_h is C^r for any l, $0 \leq l < \infty$.

Let us now compute $\mathbf{d}S_h$. The functions $Q^{[0]}(u), \ldots, Q^{[k-1]}(u)$ are defined by (6.11). Take a deformation $Q^{[k]}_\epsilon(u) = Q^{[k]}(u) + \epsilon\delta Q^{[k]}(u)$ of $Q^{[k]}(u)$. For $j = 0, \ldots, k-1$, define the corresponding lower order curves as in (6.11) by

$$Q^{[j]}_\epsilon(u) = q^{[j]}_1 + \sum_{i=1}^{k-j-1} \frac{h^i u^i}{i!} q^{[j+i]}_1 + h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} Q^{[k]}_\epsilon(s)\, ds, \tag{6.12}$$

so $Q^{[j]}_0(u) = Q^{[j]}(u)$ and

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} Q^{[j]}_\epsilon(u) = h^{k-j} \int_0^u \frac{(u-s)^{k-j-1}}{(k-j-1)!} \delta Q^{[k]}(s)\, ds.$$

Denoting $a^{[k]}_j(u,s) = (u-s)^{k-j-1}/(k-j-1)!$ and $Q(u) = (Q^{[0]}(u), \ldots,\ Q^{[k]}(u))$ for short, we have

$$\begin{aligned}
&\mathbf{d}S_h[Q^{[k]}(u)] \cdot \delta Q^{[k]}(u) = \\
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_0^1 L(Q^{[0]}_\epsilon(u), \ldots, Q^{[k]}_\epsilon(u))\, du \\
&= \int_0^1 \left(\sum_{j=0}^{k-1} \frac{\partial L}{\partial q^{[j]}}(Q(u)) h^{k-j} \int_0^u a^{[k]}_j(u,s) \delta Q^{[k]}(s)\, ds + \frac{\partial L}{\partial q^{[k]}}(Q(u)) \delta Q^{[k]}(u) \right) du \\
&= \sum_{j=0}^{k-1} \int_0^1 \int_s^1 \frac{\partial L}{\partial q^{[j]}}(Q(u)) h^{k-j} a^{[k]}_j(u,s) \delta Q^{[k]}(s)\, du\, ds + \int_0^1 \frac{\partial L}{\partial q^{[k]}}(Q(u)) \delta Q^{[k]}(u)\, du \\
&= \sum_{j=0}^{k-1} \int_0^1 \int_u^1 \frac{\partial L}{\partial q^{[j]}}(Q(s)) h^{k-j} a^{[k]}_j(s,u) \delta Q^{[k]}(u)\, ds\, du + \int_0^1 \frac{\partial L}{\partial q^{[k]}}(Q(u)) \delta Q^{[k]}(u)\, du \\
&= \int_0^1 \left(\sum_{j=0}^{k-1} \int_u^1 \frac{\partial L}{\partial q^{[j]}}(Q(s)) h^{k-j} a^{[k]}_j(s,u)\, ds + \frac{\partial L}{\partial q^{[k]}}(Q(u)) \right) \delta Q^{[k]}(u)\, du.
\end{aligned}$$

For each $u \in [0,1]$, the first factor in the integrand of the last expression is in $(\mathbb{R}^n)^*$. If $\sharp : (\mathbb{R}^n)^* \to \mathbb{R}^n$ denotes the index raising operator associated to the Euclidean inner product, define

$$\nabla S_h[Q^{[k]}(u)] := \left(\sum_{j=0}^{k-1} \int_u^1 \frac{\partial L}{\partial q^{[j]}}(Q(s)) h^{k-j} a_j^{[k]}(s,u)\, ds + \frac{\partial L}{\partial q^{[k]}}(Q(u)) \right)^{\sharp}. \tag{6.13}$$

Since $\partial L/\partial q^{[0]}$, ..., $\partial L/\partial q^{[k]}$ are C^{r-1}, then $\nabla S_h[Q^{[k]}(u)]$ is $C^{\min(l,r-1)}([0,1],\mathbb{R}^n)$. Then we have a vector field

$$\nabla S_h : C^0([0,1],\mathbb{R}^n) \to C^{\min(l,r-1)}([0,1],\mathbb{R}^n) \subset C^0([0,1],\mathbb{R}^n)$$

which we call the gradient of S_h. Also, an argument similar to the one for S_h shows that ∇S_h is a C^{r-1} map from C^l to $C^{\min(l,r-1)}$.

If we consider the inner product on $C^{\min(l,r-1)}([0,1],\mathbb{R}^n)$ given by

$$\langle\!\langle V, W \rangle\!\rangle = \int_0^1 V(u) \cdot W(u)\, du,$$

then $\mathbf{d}S_h[Q^{[k]}(u)] \cdot \delta Q^{[k]}(u)$ is the inner product of $\delta Q^{[k]}(u)$ and $\nabla S_h[Q^{[k]}(u)]$; that is,

$$\mathbf{d}S_h[Q^{[k]}(u)] \cdot \delta Q^{[k]}(u) = \langle\!\langle \nabla S_h[Q^{[k]}(u)], \delta Q^{[k]}(u) \rangle\!\rangle.$$

The constraints $g_j[Q^{[k]}(s)] := \langle b_j^{[k]}, Q^{[k]} \rangle\!\rangle = w^{[j]}$, $j = 0, \ldots, k-1$, in the variational principle 5 are bounded, linear and therefore C^∞, and the corresponding derivatives are the same functions g_j. Define

$$g = (g_0, \ldots, g_{k-1})\colon C^l([0,1],\mathbb{R}^n) \to (\mathbb{R}^n)^k$$

so

$$E = \operatorname{Ker} g \subset C^l([0,1],\mathbb{R}^n).$$

is the tangent space to the constraint set. It is actually parallel to it since the constraints are linear. It is not difficult to show using the definitions that the space

$$E^\perp = \{ c^j b_j^{[k]} \mid c^0, \ldots, c^{k-1} \in \mathbb{R}^n \}$$

of $\mathbb{R}^n$-valued polynomials of degree at most $k-1$ is indeed the $\langle\!\langle , \rangle\!\rangle$-orthogonal complement of E, which is then a split subspace (see Appendix B). The orthogonal projection $P\colon C^l([0,1],\mathbb{R}^n) = E \oplus E^\perp \to E$ is given by

$$P(\delta Q^{[k]}(u)) = \delta Q^{[k]}(u) - \sum_{j=0}^{k-1} \langle b_j^{[k]}, \delta Q^{[k]} \rangle\!\rangle\, b_j^{[k]}.$$

Now S_h has a critical point on the constraint set (for any value of the constraints) if and only if the projection $P\nabla S_h$ of ∇S_h to the tangent space E of the constraint set is 0. That is, in order to find solutions to the variational principle 5, we solve

$$P\nabla S_h(Q^{[k]}) = P\nabla S_h(Q_E^{[k]} \oplus Q_{E^\perp}^{[k]}) = 0 \tag{6.14}$$

for $Q_E^{[k]}$, near

$$Q^{[k]} = 0, \quad w^{[0]} = \cdots = w^{[k-1]} = 0,$$
$$q_1^{[0]} = \bar{q}^{[0]}, \ldots, q_1^{[k-1]} = \bar{q}^{[k-1]}, \quad h = 0.$$

Remark 6.1.5. Observe that $Q_{E^\perp}^{[k]}$ is determined by the initial data $w^{[j]}$. That is, using the decomposition $Q^{[k]} = Q_E^{[k]} \oplus Q_{E^\perp}^{[k]}$ we have $w^{[j]} = g_j(Q_{E^\perp}^{[k]}) = \langle b_j^{[k]}, Q_{E^\perp}^{[k]} \rangle\!\rangle$ since g_j is linear and $g_j(Q_E^{[k]}) = 0$. Therefore, noting that $\langle b_j^{[k]}, Q_{E^\perp}^{[k]} \rangle\!\rangle = \langle b_j^{[k]}, c^{j'} b_{j'}^{[k]} \rangle\!\rangle = c^{j'}$ one obtains $Q_{E^\perp}^{[k]} = w^{[j]} b_j^{[k]}$. $\diamond$

Equation (6.14) can be solved using the implicit function theorem by requiring that the appropriate partial derivative of $P\nabla S_h(Q^{[k]})$ at this point is a linear isomorphism. In order to compute this derivative, take a deformation of $Q^{[k]} = 0$ of the form $Q_\epsilon^{[k]} = \epsilon \delta Q_E^{[k]}$, where $\delta Q_E^{[k]} \in E$. Recalling (6.12) and noting that $h = 0$, we have

$$\begin{aligned}
\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} & P \frac{\partial L}{\partial q^{[k]}}(Q_\epsilon^{[0]}(u), \ldots, Q_\epsilon^{[k]}(u)) = \\
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} P \frac{\partial L}{\partial q^{[k]}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, Q_\epsilon^{[k]}(u)) \\
&= P \frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0) \delta Q_E^{[k]}(u) \\
&= \frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0) \delta Q_E^{[k]}(u) \\
&\qquad - \sum_{j=0}^{k-1} \left\langle\!\!\left\langle b_j^{[k]}, \frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0) \delta Q_E^{[k]} \right\rangle\!\!\right\rangle b_j^{[k]} \\
&= \frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0) \delta Q_E^{[k]}(u).
\end{aligned}$$

Here the inner products vanish because $\frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0)$ is a constant matrix (that is, it does not depend of u) and $\langle b^{[j]}, \delta Q_E^{[k]} \rangle\!\rangle = 0$ for $j = 0, \ldots, k-1$.

Then the derivative is precisely $\frac{\partial^2 L}{\partial q^{[k]2}}(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}, 0)$, seen as a linear map from E into itself, and if L is regular then it is an isomorphism.

Now we will specialize to the cases $l = 0$ and $l = r-1$. By the implicit function theorem, there are neighborhoods $W_1 \subseteq (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k \times \mathbb{R}$ (with variables $(q_1^{[0]}, \ldots, q_1^{[k-1]}; w^{[0]}, \ldots, w^{[k-1]}; h))$ containing $(\bar{q}^{[0]}, \ldots, \bar{q}^{[k-1]}; 0, \ldots, 0; 0)$ and $W_2^l \subseteq C^l([0,1], \mathbb{R}^n)$ containing the constant curve $Q^{[k]}(u) = 0$, and a C^{r-1} map $\psi\colon W_1 \to W_2^l$ such that for each $(q_1^{[0]}, \ldots, q_1^{[k-1]}; w^{[0]}, \ldots, w^{[k-1]}; h) \in W_1$, the curve

$$Q^{[k]} = \psi(q_1^{[0]}, \ldots, q_1^{[k-1]}; w^{[0]}, \ldots, w^{[k-1]}; h) \in C^l([0,1], \mathbb{R}^n)$$

is the unique critical point in W_2^l of the variational problem 5.

By taking $l = r-1$, ψ has values in $W_2^{r-1} \subseteq C^{r-1}([0,1],\mathbb{R}^n)$. Taking $l=0$, ψ has values in $W_2^0 \subseteq C^0([0,1],\mathbb{R}^n)$. However, since $C^{r-1} \subset C^0$, this ψ also provides the unique solution among the C^0 curves in a C^0-open neighborhood of the curve $u \mapsto 0$, say $\{Q^{[k]}(u) \mid \|Q^{[k]}\|_0 < \epsilon\}$.

Let us now reverse the regularization in order to obtain a unique C^k solution of the variational principle 1. For $(q_1,q_2) = ((q_1^{[0]},\dots,q_1^{[k-1]}),(q_2^{[0]},\dots,q_2^{[k-1]})) \in (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k$ the corresponding values of $z^{[0]},\dots,z^{[k-1]}$ are given by (6.9) and the values of $w^{[0]},\dots,w^{[k-1]}$ can be computed from (6.10) using the inverse matrix of $\left(\gamma_j^{[k],i}\right)$. This defines a function $(w^{[0]},\dots,w^{[k-1]}) = \varpi(q_1,q_2,h)$. We write $\bar{q} = (\bar{q}^{[0]},\dots,\bar{q}^{[k-1]}) \in (\mathbb{R}^n)^k$ and let $h>0$ be such that $(\bar{q};0;h) \in W_1$. Define

$$\widetilde{W}_1 = \{(q_1,q_2) \in (\mathbb{R}^n)^k \times (\mathbb{R}^n)^k \mid (q_1;\varpi(q_1,q_2,h);h) \in W_1\}$$

and for each $(q_1,q_2) = \left((q_1^{[0]},\dots,q_1^{[k-1]}),(q_2^{[0]},\dots,q_2^{[k-1]})\right) \in W_1$ define the curve $Q^{[0]}_{(q_1,q_2)}(u)$ according to (6.7) as

$$Q^{[0]}_{(q_1,q_2)}(u) = q_1^{[0]} + \sum_{i=1}^{k-1} \frac{h^i u^i}{i!} q_1^{[i]} + h^k \int_0^u \frac{(u-s)^{k-1}}{(k-1)!} \psi\left(q_1;\varpi(q_1,q_2,h);h\right)(s)\,ds.$$

Since ψ takes values in the C^l curves, $Q^{[0]}_{(q_1,q_2)}(u)$ is C^{k+l} by the reasoning leading to equation (6.7).

Now reparameterize with $t = hu$ to get a C^{k+l} curve

$$q^{[0]}_{(q_1,q_2)}(t) = q_1^{[0]} + \sum_{i=1}^{k-1} \frac{t^i}{i!} q_1^{[i]} + (t/u)^k \int_0^{t/h} \frac{(t/h-s)^{k-1}}{(k-1)!} \psi\left(q_1;\varpi(q_1,q_2,h);h\right)(s)\,ds.$$

on Q, defined for $t \in [0,h]$. This curve is the unique solution of the variational principle 1 with endpoint conditions q_1 and q_2.

This solution is unique among the curves corresponding to $Q^{[k]}$ continuous with $\|Q^{[k]}\|_0 < \epsilon$. These are the C^k curves $q(t)$ on Q with $\|q^{(k)}\|_0 < \epsilon/h^k$, which are the C^k curves in some C^k neighborhood of the constant curve $u \mapsto \bar{q}^{[0]}$.

6.2 The exact discrete Lagrangian and discrete equations for second-order systems

In this section we will restrict ourselves to second-order Lagrangians and the extension to higher-order theories is straightforward.

Let Q be the configuration manifold and let $L : T^{(2)}Q \to \mathbb{R}$ be a regular lagrangian

Definition 6.2.1. *The exact discrete lagrangian $L_d^e : TQ \times TQ \to \mathbb{R}$, is defined by*

$$L_d^e(q_0,\dot{q}_0,q_1,\dot{q}_1) = \int_0^h L(q(t),\dot{q}(t),\ddot{q}(t))dt,$$

where $q(t) : I \subset \mathbb{R} \to Q$ *is the unique solution of the Euler-Lagrange equations for the second-order lagrangian* L*:*

$$\frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{q}}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{\partial L}{\partial q} = 0 \tag{6.15}$$

for $h > 0$*, satisfying the boundary conditions* $q(0) = q_0, q(h) = q_1, \dot{q}(0) = \dot{q}_0$ *and* $\dot{q}(h) = \dot{q}_1$*.*

Observe that the exact discrete lagrangian $L_d^e : TQ \times TQ \to \mathbb{R}$ is defined on the cartesian product of two copies of TQ. Our idea is to take approximations of $L_d : TQ \times TQ \to \mathbb{R}$ for $L_d^e : TQ \times TQ \to \mathbb{R}$ to construct variational integrators in the same way that in discrete mechanics (see subsection 6.3). In other words, for given $h > 0$ we define $L_d(q_0, v_0, q_1, v_1)$ as an approximation of the action integral along the exact solution curve segment $q(t)$ with boundary conditions $q(0) = q_0$, $\dot{q}(0) = v_0$, $q(h) = q_1$, and $\dot{q}(h) = v_1$. For example, we can use the formula

$$L_d(q_0, v_0, q_1, v_1) = hL\left(\kappa(q_0, v_0, q_1, v_1), \chi(q_0, v_0, q_1, v_1), \zeta(q_0, v_0, q_1, v_1)\right),$$

where κ, χ and ζ are functions of $(q_0, v_0, q_1, v_1) \in TQ \times TQ$ which approximate the configuration $q(t) \in Q$, the velocity $\dot{q}(t) \in T_qQ$ and the acceleration $\ddot{q}(t) \in T^{(2)}Q$, respectively, in terms of the initial and final positions and velocities. We can also, for instance, consider suitable linear combinations of discrete Lagrangians of this type, for instance, weighted averages of the type

$$L_d(q_0, v_0, q_1, v_1) = \frac{1}{2}L\left(q_0, v_0, \frac{v_1 - v_0}{h}\right) + \frac{1}{2}L\left(q_1, v_1, \frac{v_1 - v_0}{h}\right),$$

or other combinations.

For completeness, we will derive the discrete equations for the Lagrangian $L_d : TQ \times TQ \to \mathbb{R}$, but these results are a direct translation of Marsden and West [131] to our case.

Construct the grid $\{t_k = kh \mid k = 0, \ldots, N\}$, $Nh = T$ and define the discrete path space $\mathcal{P}_d(TQ) := \{(q_d, v_d) : \{t_k\}_{k=0}^N \to TQ\}$. We will identify a discrete trajectory $(q_d, v_d) \in \mathcal{P}_d(TQ)$ with its image $(q_d, v_d) = \{(q_k, v_k)\}_{k=0}^N$ where $(q_k, v_k) := (q_d(t_k), v_d(t_k))$. The discrete action $\mathcal{A}_d : \mathcal{P}_d(TQ) \to \mathbb{R}$ along this sequence is calculated by summing the discrete Lagrangian on each adjacent pair and defined by

$$\mathcal{A}_d(q_d, v_d) := \sum_{k=0}^{N-1} L_d(q_k, v_k, q_{k+1}, v_{k+1}).$$

We would like to point out that the discrete path space is isomorphic to the smooth product manifold which consists on $N + 1$ copies of TQ, the discrete action inherits the smoothness of the discrete Lagrangian and the tangent space $T_{(q_d, v_d)}\mathcal{P}_d(TQ)$ at (q_d, v_d) is the set of maps $a_{(q_d, v_d)} : \{t_k\}_{k=0}^N \to TTQ$ such that $\tau_{TQ} \circ a_{(q_d, v_d)} = (q_d, v_d)$.

Hamilton's principle seeks discrete curves $\{(q_k, v_k)\}_{k=0}^N$ that satisfy

$$\delta \sum_{k=0}^{N-1} L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0$$

for all variations $\{(\delta q_k, \delta v_k)\}_{k=0}^N$ vanishing at the endpoints. This is equivalent to the *discrete Euler-Lagrange equations*

$$D_3 L_d(q_{k-1}, v_{k-1}, q_k, v_k) + D_1 L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \tag{6.16a}$$
$$D_4 L_d(q_{k-1}, v_{k-1}, q_k, v_k) + D_2 L_d(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \tag{6.16b}$$

for $1 \leq k \leq N-1$.

Given a solution $\{q_k^*, v_k^*\}_{k\in\mathbb{Z}}$ of equations (6.16) and assuming that the $2n \times 2n$ matrix

$$\begin{pmatrix} D_{13} L_d(q_k, v_k, q_{k+1}, v_{k+1}) & D_{14} L_d(q_k, v_k, q_{k+1}, v_{k+1}) \\ D_{23} L_d(q_k, v_k, q_{k+1}, v_{k+1}) & D_{24} L_d(q_k, v_k, q_{k+1}, v_{k+1}) \end{pmatrix}$$

is nonsingular, it is possible to define the (local) *discrete Lagrangian map* $F_{L_d} : \mathcal{U}_k \subset TQ \times TQ \to TQ \times TQ$ mapping $(q_{k-1}, v_{k-1}, q_k, v_k)$ to $(q_k, v_k, q_{k+1}, v_{k+1})$ from (6.16) where $\mathcal{U}_k$ is a neighborhood of the point $(q_{k-1}^*, v_{k-1}^*, q_k^*, v_k^*)$.

6.2.1 Discrete Legendre transforms

We define the *discrete Legendre transformations* $\mathbb{F}^+ L_d, \mathbb{F}^- L_d : TQ \times TQ \to T^*TQ$ which maps the space $TQ \times TQ$ into T^*TQ. These are given by

$$\begin{aligned} \mathbb{F}^+ L_d(q_0, v_0, q_1, v_1) &= (q_0, v_0, -D_1 L_d(q_0, v_0, q_1, v_1), -D_2 L_d(q_0, v_0, q_1, v_1)), \\ \mathbb{F}^- L_d(q_0, v_0, q_1, v_1) &= (q_1, v_1, D_3 L_d(q_0, v_0, q_1, v_1), D_4 L_d(q_0, v_0, q_1, v_1)). \end{aligned}$$

If both discrete fibre derivatives are locally isomorphisms for nearby (q_0, v_0) and (q_1, v_1), then we say that L_d is *regular*. If Q is a vector space and both discrete fibre derivatives are global isomorphisms we say that L_d is *hyperregular*.

Using the discrete Legendre transforms the discrete second-order Euler-Lagrange equations (6.16) can be rewritten as

$$\mathbb{F}^- L_d(q_k, v_k, q_{k+1}, v_{k+1}) = \mathbb{F}^+ L_d(q_{k-1}, v_{k-1}, q_k, v_k).$$

It will be useful to note that

$$\begin{aligned} \mathbb{F}^- L_d \circ F_{L_d}(q_0, v_0, q_1, v_1) &= \mathbb{F}^- L_d(q_1, v_1, q_2, v_2) \\ &= (q_1, v_1, -D_1 L_d(q_1, v_1, q_2, v_2), -D_2 L_d(q_1, v_1, q_2, v_2)) \\ &= (q_1, v_1, D_3 L_d(q_0, v_0, q_1, v_1), D_4 L_d(q_0, v_0, q_1, v_1)) \\ &= \mathbb{F}^+ L_d(q_0, v_0, q_1, v_1), \end{aligned}$$

that is

$$\mathbb{F}^+ L_d = \mathbb{F}^- L_d \circ F_{L_d}. \tag{6.17}$$

6.2.2 Example: Cubic splines

Let $Q = \mathbb{R}^n$ and $L : T^{(2)}Q \equiv (\mathbb{R}^n)^3 \to \mathbb{R}$ the second-order Lagrangian given by $L(q, \dot{q}, \ddot{q}) = \frac{1}{2}\ddot{q}^2$.

It is well known that the solutions to the corresponding Euler-Lagrange equations $q^{(4)} = 0$ are the so-called cubic splines $q(t) = at^3 + bt^2 + ct + d$, for $a, b, c, d \in \mathbb{R}^n$. We define $L_d : (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^n) \to \mathbb{R}$ as follows. Write

$$q(0) = q(h) - h\dot{q}(h) + \frac{h^2}{2}\ddot{q}(h) + \mathcal{O}(h^3), \tag{6.18a}$$

$$q(h) = q(0) + h\dot{q}(0) + \frac{h^2}{2}\ddot{q}(0) + \mathcal{O}(h^3). \tag{6.18b}$$

Given sufficiently close $(q_0, v_0), (q_1, v_1) \in TQ$ we can use equations (6.18) to obtain approximations of the acceleration of the exact solution joining these boundary conditions at time h, which we call

$$a_0 = \frac{2}{h^2}(q_1 - q_0 - hv_0) \text{ and } a_1 = \frac{2}{h^2}(q_0 - q_1 + hv_1).$$

Then we define

$$L_d(q_0, v_0, q_1, v_1) = \frac{h}{2}\left(L(q_0, v_0, a_0) + L(q_1, v_1, a_1)\right) = \frac{(hv_1 + q_0 - q_1)^2}{h^3} + \frac{(-hv_0 - q_0 + q_1)^2}{h^3}.$$

Solving the discrete Euler-Lagrange equations for this discrete Lagrangian, the evolution of the discrete trajectory is

$$q_{k+1} = q_{k-1} + 2hv_k \tag{6.19a}$$

$$v_{k+1} = v_{k-1} + 4\left(v_k - \frac{q_k - q_{k-1}}{h}\right). \tag{6.19b}$$

In the following section we will continue this example and we will show some simulations.

6.3 Relation between discrete and continuous variational systems

Let $L : T^{(2)}Q \to \mathbb{R}$ be a regular lagrangian and consider the *exact discrete lagrangian* $L_d^e : TQ \times TQ \to \mathbb{R}$, given by

$$L_d^e(q_0, \dot{q}_0, q_1, \dot{q}_1) = \int_0^h L(q(t), \dot{q}(t), \ddot{q}(t))dt,$$

where $q(t) : I \subset \mathbb{R} \to Q$ is the unique solution of the Euler-Lagrange equations for the second-order lagrangian L

$$\frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{q}}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{\partial L}{\partial q} = 0 \tag{6.20}$$

for $h > 0$, satisfying the boundary conditions $q(0) = q_0$, $q(h) = q_1$, $\dot{q}(0) = \dot{q}_0$ and $\dot{q}(h) = \dot{q}_1$.

The Legendre transformation associated to L is defined to be the map $\mathbb{F}L : T^{(3)}Q \to T^*TQ$ given by (see [50])

$$\mathbb{F}L(q, \dot{q}, \ddot{q}, q^{(3)}) = \left(q, \dot{q}, \frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{q}}\right), \frac{\partial L}{\partial \ddot{q}}\right). \tag{6.21}$$

We will see that there is a special relationship between the Legendre transform of a regular Lagrangian and the discrete Legendre transforms of the corresponding exact discrete lagrangian L_d^e.

Theorem 6.3.1. *Let $L : T^{(2)}Q \to \mathbb{R}$ be a regular Lagrangian and $L_d^e : TQ \times TQ \to \mathbb{R}$, the corresponding exact discrete Lagrangian. Then L and L_d^e have Legendre transformations related by*

$$\begin{aligned} \mathbb{F}^-L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= \mathbb{F}L(q(0), \dot{q}(0), \ddot{q}(0), q^{(3)}(0)) \\ \mathbb{F}^+L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= \mathbb{F}L(q(h), \dot{q}(h), \ddot{q}(h), q^{(3)}(0)). \end{aligned}$$

Proof. We begin by computing the derivatives of L_d^e :

$$\begin{aligned} \frac{\partial L_d^e}{\partial q_0} &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q_0} + \frac{\partial L}{\partial \ddot{q}}\frac{\partial \ddot{q}}{\partial q_0}\right) dt \\ &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q_0} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial q_0}\right) dt + \frac{\partial L}{\partial \ddot{q}}\frac{\partial \dot{q}}{\partial q_0}\Big|_0^h \\ &= \int_0^h \frac{\partial L}{\partial q}\frac{\partial q}{\partial q_0} + \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial \dot{q}}{\partial q_0} dt \end{aligned}$$

using that

$$\frac{\partial \dot{q}}{\partial q_0}(0) = 0 \text{ and } \frac{\partial \dot{q}}{\partial q_0}(h) = 0$$

and integration by parts.

Therefore,

$$\frac{\partial L_d^e}{\partial q_0} = \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial q_0}\Big|_0^h + \int_0^h \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{q}}\right)\right)\frac{\partial q}{\partial q_0} dt.$$

Since $q(t)$ is a solution of the Euler-Lagrange equations for $L : T^{(2)}Q \to \mathbb{R}$, the last term in the previous equality is zero. Therefore

$$\frac{\partial L_d^e}{\partial q_0} = \left(\frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}\right)\frac{\partial q}{\partial q_0}\Big|_0^h = -\frac{\partial L}{\partial \dot{q}}(q(0), \dot{q}(0), \ddot{q}(0)) + \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{q}}\right)(q(0), \dot{q}(0), \ddot{q}(0), q^{(3)}(0)) \tag{6.22}$$

because

$$\frac{\partial q}{\partial q_0}(0) = \mathrm{Id} \text{ and } \frac{\partial q}{\partial q_0}(h) = 0.$$

Therefore, we have that

$$\frac{\partial L_d^e}{\partial q_0} = \left(-\frac{\partial L}{\partial \dot q} + \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot q}\right)\right)(q(0), \dot q(0), \ddot q(0), q^{(3)}(0)).$$

On the other hand, using twice integration by parts,

$$\begin{aligned}
\frac{\partial L_d^e}{\partial \dot q_0} &= \int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot q_0} + \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial \dot q_0} + \frac{\partial L}{\partial \ddot q}\frac{\partial \ddot q}{\partial \dot q_0}\right) dt = \\
&\int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot q_0} + \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial \dot q_0} - \frac{d}{dt}\frac{\partial L}{\partial \ddot q}\frac{\partial \dot q}{\partial \dot q_0}\right) dt + \frac{\partial L}{\partial \ddot q}\frac{\partial \dot q}{\partial \dot q_0}\Big|_0^h = \\
&\int_0^h \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial \dot q_0} + \left(\frac{\partial L}{\partial \dot q} - \frac{d}{dt}\frac{\partial L}{\partial \ddot q}\right)\frac{\partial \dot q}{\partial \dot q_0}\right) dt + \frac{\partial L}{\partial \ddot q}\frac{\partial \dot q}{\partial \dot q_0}\Big|_0^h = \\
&\int_0^h \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot q} + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot q}\right)\right)\frac{\partial q}{\partial \dot q_0} dt + \frac{\partial L}{\partial \ddot q}\frac{\partial \dot q}{\partial \dot q_0}\Big|_0^h + \left(\frac{\partial L}{\partial \dot q} - \frac{d}{dt}\frac{\partial L}{\partial \ddot q}\right)\frac{\partial q}{\partial \dot q_0}\Big|_0^h.
\end{aligned}$$

As before, since $q(t)$ is a solution of the Euler-Lagrange equations the first term is zero, and using that

$$\frac{\partial \dot q}{\partial \dot q_0}(0) = Id, \quad \frac{\partial \dot q}{\partial \dot q_0}(h) = 0, \quad \frac{\partial q}{\partial \dot q_0}(0) = 0, \text{ and } \frac{\partial q}{\partial \dot q_0}(h) = 0$$

we have

$$\frac{\partial L_d^e}{\partial \dot q_0} = -\frac{\partial L}{\partial \ddot q}(q(0), \dot q(0), \ddot q(0)).$$

Therefore,

$$\mathbb{F}^- L_d^e(q(0), \dot q(0), q(h), \dot q(h)) = \mathbb{F}L(q(0), \dot q(0), \ddot q(0), q^{(3)}(0)).$$

With similar arguments, we can also prove that

$$\frac{\partial L_d^e}{\partial q_1} = \left(\frac{\partial L}{\partial \dot q} - \frac{d}{dt}\frac{\partial L}{\partial \ddot q}\right)(q(h), \dot q(h), \ddot q(h), q^{(3)}(h))$$

and

$$\frac{\partial L_d^e}{\partial \dot q_1} = \frac{\partial L}{\partial \ddot q}(q(h), \dot q(h), \ddot q(h))$$

and in consequence,

$$\mathbb{F}^+ L_d^e(q(0), \dot q(0), q(h), \dot q(h)) = \mathbb{F}L(q(h), \dot q(h), \ddot q(h), q^{(3)}(h)).$$

□

In what follows we will study the relation between the regularity of the continuous Lagrangian, given by the hessian matrix

$$\mathcal{W} = \left(\frac{\partial^2 L}{\partial \ddot q\, \partial \ddot q}\right)$$

and the the regularity condition corresponding to the exact discrete lagrangian $L_d^e : TQ \times TQ \to \mathbb{R}$

$$\mathcal{W}_d = \begin{pmatrix} D_{13}L_d^e & D_{14}L_d^e \\ D_{23}L_d^e & D_{24}L_d^e \end{pmatrix}.$$

Theorem 6.3.2. *The exact discrete Lagrangian $L_d^e \colon TQ \times TQ \to \mathbb{R}$ corresponding to a regular Lagrangian $L \colon T^{(2)}Q \to \mathbb{R}$ is also regular.*

Proof. Taking into account the Taylor expansions

$$\begin{aligned} q(h) &= q(0) + h\dot{q}(0) + \frac{h^2}{2}\ddot{q}(0) + \frac{h^3}{6}q^{(3)}(0) + \mathcal{O}\left(h^4\right) \\ \dot{q}(h) &= \dot{q}(0) + h\ddot{q}(0) + \frac{h^2}{2}q^{(3)}(0) + \mathcal{O}\left(h^3\right) \end{aligned}$$

we have that

$$\begin{aligned} \frac{\partial \ddot{q}}{\partial q_1}(0) &= \frac{6}{h^2} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial q_1}(0) = -\frac{12}{h^3} + \mathcal{O}(h), \\ \frac{\partial \ddot{q}}{\partial \dot{q}_1}(0) &= -\frac{2}{h} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial \dot{q}_1}(0) = \frac{6}{h^2} + \mathcal{O}(h). \end{aligned}$$

Analogously,

$$\begin{aligned} \frac{\partial \ddot{q}}{\partial q_0}(h) &= \frac{6}{h^2} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial q_0}(h) = \frac{12}{h^3} + \mathcal{O}(h), \\ \frac{\partial \ddot{q}}{\partial \dot{q}_0}(h) &= \frac{2}{h} + \mathcal{O}\left(h^2\right), \quad \frac{\partial q^{(3)}}{\partial \dot{q}_0}(h) = \frac{6}{h^2} + \mathcal{O}(h). \end{aligned}$$

Using Theorem 6.3.1 and the previous expansions is easy to show that

$$h^3 D_{13}L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) = h^3 \frac{\partial^2 L_d^e}{\partial q_0 \partial q_1}(q(0), \dot{q}(0), q(h), \dot{q}(h)) = -12\mathcal{W} + \mathcal{O}(h).$$

Now, using a similar procedure, we derive that

$$\begin{aligned} h^2 D_{14}L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= h^2 \frac{\partial^2 L_d^e}{\partial q_0 \partial \dot{q}_1}(q(0), \dot{q}(0), q(h), \dot{q}(h)) = 6\mathcal{W} + \mathcal{O}(h) \\ h^2 D_{23}L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= h^2 \frac{\partial^2 L_d^e}{\partial \dot{q}_0 \partial q_1}(q(0), \dot{q}(0), q(h), \dot{q}(h)) = 6\mathcal{W} + \mathcal{O}(h) \\ h D_{24}L_d^e(q(0), \dot{q}(0), q(h), \dot{q}(h)) &= h \frac{\partial^2 L_d^e}{\partial \dot{q}_0 \partial \dot{q}_1}(q(0), \dot{q}(0), q(h), \dot{q}(h)) = -2\mathcal{W} + \mathcal{O}(h). \end{aligned}$$

Therefore, we conclude that

$$\det \mathcal{W}_d = -\frac{12\mathcal{W}}{h^4} + \mathcal{O}\left(\frac{1}{h^3}\right).$$

That is, if L is regular then L_d^e is regular. □

In what follows we denote $(TQ \times TQ)_2$ the subset of $(TQ \times TQ) \times (TQ \times TQ)$ given by

$$(TQ \times TQ)_2 := \{(q_0, \dot{q}_0, q_1, \dot{q}_1, q_1, \dot{q}_1, q_2, \dot{q}_2) \mid (q_i, \dot{q}_i) \in TQ \text{ with } i = 1, 2, 3\}.$$

If $L : T^{(2)}Q \to \mathbb{R}$ is a regular Lagrangian then the Euler-Lagrange equations for L gives rise to a system of explicit 4-order differential equations

$$q^{(4)} = \Psi(q, \dot{q}, \ddot{q}, q^{(3)}).$$

Therefore, for h given, it is possible to derive the following map (see [4])

$$\Psi_L^h : T^{(3)}Q \to T^{(3)}Q$$

which maps $(q(0), \dot{q}(0), \ddot{q}(0), q^{(3)}(0)) \in T^{(3)}Q$ into $(q(h), \dot{q}(h), \ddot{q}(h), q^{(3)}(h)) \in T^{(3)}Q$. Therefore, from Theorem (6.3.1) we deduce the commutativity of diagram 6.1.

$$\begin{tikzcd}
& (q(0), \dot{q}(0), q(h), \dot{q}(h)) \arrow[dl, mapsto, "\mathbb{F}^- L_d^e"'] \arrow[dr, mapsto, "\mathbb{F}^+ L_d^e"] & \\
(q(0), \dot{q}(0), -D_1 L_d^e, -D_2 L_d^e) & & (q(h), \dot{q}(h), D_3 L_d^e, D_4 L_d^e) \\
(q(0), \dot{q}(0), \ddot{q}(0), q^{(3)}(0)) \arrow[u, mapsto, "\mathbb{F}L"] \arrow[rr, mapsto, "\Psi_L^h"'] & & (q(h), \dot{q}(h), \ddot{q}(h), q^{(3)}(h)) \arrow[u, mapsto, "\mathbb{F}L"']
\end{tikzcd}$$

Figure 6.1: Correspondence between the discrete Legendre transforms and the continuous Hamiltonian flow.

Definition 6.3.3. *The discrete Hamiltonian flow is defined by* $\widetilde{F}_{L_d} : T^*TQ \to T^*TQ$ *as*

$$\widetilde{F}_{L_d} = \mathbb{F}^- L_d \circ F_{L_d} \circ (\mathbb{F}^- L_d)^{-1}. \tag{6.23}$$

Alternatively, it can also be defined as $\widetilde{F}_{L_d} = \mathbb{F}^+ L_d \circ F_{L_d} \circ (\mathbb{F}^+ L_d)^{-1}$.

Theorem 6.3.4. *The diagram in Figure 6.2 is commutative.*

Proof. The central triangle is (6.17). The parallelogram on the left-hand side is commutative by (6.23), so the triangle on the left is commutative. The triangle on the right is the same as the triangle on the left, with shifted indices. Then parallelogram on the right-hand side is commutative, which gives the equivalence stated in the definition of the discrete Hamiltonian flow. □

$$\begin{tikzcd}
 & (q_0,\dot q_0,q_1,\dot q_1) \arrow[rr, mapsto, "F_{L_d}"] \arrow[dl, mapsto, "\mathbb{F}^- L_d"'] \arrow[dr, mapsto, "\mathbb{F}^+ L_d"] & & (q_1,\dot q_1,q_2,\dot q_2) \arrow[dl, mapsto, "\mathbb{F}^- L_d"'] \arrow[dr, mapsto, "\mathbb{F}^+ L_d"] & \\
(q_0,\dot q_0,-D_1L_d,-D_2L_d) \arrow[rr, mapsto, "\tilde F_{L_d}"'] & & (q_1,\dot q_1,D_3L_d,D_4L_d) \arrow[rr, mapsto, "\tilde F_{L_d}"'] & & (q_2,\dot q_2,-D_1L_d,-D_2L_d)
\end{tikzcd}$$

Figure 6.2: Correspondence between the discrete Lagrangian and the discrete Hamiltonian maps.

Corollary 6.3.5. *The following definitions of the discrete Hamiltonian map are equivalent*

$$\begin{aligned}
\widetilde{F}_{L_d} &= \mathbb{F}^+ L_d \circ F_{L_d} \circ (\mathbb{F}^+ L_d)^{-1}, \\
\widetilde{F}_{L_d} &= \mathbb{F}^- L_d \circ F_{L_d} \circ (\mathbb{F}^- L_d)^{-1}, \\
\widetilde{F}_{L_d} &= \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1},
\end{aligned}$$

and have the coordinate expression $\widetilde{F}_{L_d} : (q_0, \dot q_0, p_0, \tilde p_0) \mapsto (q_1, \dot q_1, p_1, \tilde p_1)$*, where we use the notation*

$$\begin{aligned}
p_0 &= -D_1 L_d(q_0, \dot q_0, q_1, \dot q_1), \\
\tilde p_0 &= -D_2 L_d(q_0, \dot q_0, q_1, \dot q_1), \\
p_1 &= D_3 L_d(q_0, \dot q_0, q_1, \dot q_1), \\
\tilde p_1 &= D_4 L_d(q_0, \dot q_0, q_1, \dot q_1).
\end{aligned}$$

Combining Theorem (6.3.1) with the diagram in Figure 6.2 gives the commutative diagram shown in Figure 6.3 for the exact discrete Lagrangian.

Here, F_H^h denotes the flow of the Hamiltonian vector field X_H associated with the Hamiltonian $H : T^*TQ \to \mathbb{R}$ given by $H = E_L \circ (\mathbb{F}L)^{-1}$ where $E_L : T^{(3)}Q \to \mathbb{R}$ denotes the energy function associated to L (see [112]).

This proves the following theorem

Theorem 6.3.6. *Let* $L : T^{(2)}Q \to \mathbb{R}$ *be a regular Lagrangian, its corresponding exact discrete Lagrangian* $L_d^e : TQ \times TQ \to \mathbb{R}$ *and consider the pushforward of both the continuous Lagrangian and discrete system to* T^*TQ*, yielding a Hamiltonian system with Hamiltonian* H *and discrete Hamiltonian map* $\tilde F_{L_d^e}$*, respectively. Then, for a sufficiently small time-step* $h \in \mathbb{R}$*,* $F_H^h = \tilde F_{L_d^e}$*.*

6.3.1 Example: Cubic spline (cont'd.)

Recall that in this example $Q = \mathbb{R}^n$ and $L = \frac{1}{2}\ddot q^2$. Since the exact solutions for the second-order Euler-Lagrange equation for L can be found explicitly, it is easy to show that the

$$
\begin{tikzcd}
& (q_0,\dot q_0,q_1,\dot q_1) \arrow[rr, mapsto, "F_{L_d^e}"] \arrow[dl, mapsto, "\mathbb{F}^-L_d^e"'] \arrow[dr, mapsto, "\mathbb{F}^+L_d^e"] & & (q_1,\dot q_1,q_2,\dot q_2) \arrow[dl, mapsto, "\mathbb{F}^-L_d^e"'] \arrow[dr, mapsto, "\mathbb{F}^+L_d^e"] & \\
(q_0,\dot q_0,p_0,\tilde p_0) \arrow[rr, mapsto, "\tilde F_{L_d^e}=F_H^h"'] & & (q_1,\dot q_1,p_1,\tilde p_1) \arrow[rr, mapsto, "\tilde F_{L_d^e}=F_H^h"'] & & (q_2,\dot q_2,p_2,\tilde p_2) \\
(q(0),\dot q(0),\ddot q(0),q^{(3)}(0)) \arrow[u, mapsto, "\mathbb{F}L"] \arrow[rr, mapsto, "\Psi_L^h"'] & & (q(h),\dot q(h),\ddot q(h),q^{(3)}(h)) \arrow[u, mapsto, "\mathbb{F}L"] \arrow[rr, mapsto, "\Psi_L^h"'] & & (q(2h),\dot q(2h),\ddot q(2h),q^{(3)}(2h)) \arrow[u, mapsto, "\mathbb{F}L"]
\end{tikzcd}
$$

Figure 6.3: Correspondence between the exact discrete Lagrangian and the continuous Hamiltonian flow, where $q(0)=q_0$, $\dot q(0)=\dot q_0$, $q(h)=q_1$, $\dot q(h)=\dot q_1$, $q(2h)=q_2$ and $\dot q(2h)=\dot q_2$.

discrete exact Lagrangian is

$$L_d^e(q_0,v_0,q_1,v_1)=\frac{6}{h^3}(q_0-q_1)^2+\frac{6}{h^2}(q_0-q_1)(v_0+v_1)+\frac{2}{h}(v_0^2+v_0v_1+v_1^2).$$

From the corresponding discrete Euler-Lagrange equation, the evolution is

$$\begin{aligned} q_{k+1}&=5q_{k-1}-4q_k+2h(v_{k-1}+2v_k)\\ v_{k+1}&=v_{k-1}+\frac{2}{h}(q_{k-1}-2q_k+q_{k+1}). \end{aligned}$$

It is easy to check that both this exact method and method (6.19) preserve the quantity

$$\varphi(q_k,v_k,q_{k+1},v_{k+1})=\frac{q_{k+1}-q_k}{h}-\frac{v_k+v_{k+1}}{2}.$$

6.3.2 Variational error analysis

Now, we rewrite the result of Patrick, Marsden and West [150], [131] to the particular case of a Lagrangian $L_d: TQ\times TQ\to\mathbb{R}$.

Definition 6.3.7. *Let $L_d: TQ\times TQ\to\mathbb{R}$ be a discrete Lagrangian. We say that L_d has an order r discretization if there exists an open subset $U_1\subset T^{(2)}Q$ with compact closure and constants $C_1>0$, $h_1>0$ so that*

$$||L_d(q(0),\dot q(0),q(h),\dot q(h),h)-L_d^e(q(0),\dot q(0),q(h),\dot q(h),h)||\le C_1h^{r+1} \tag{6.24}$$

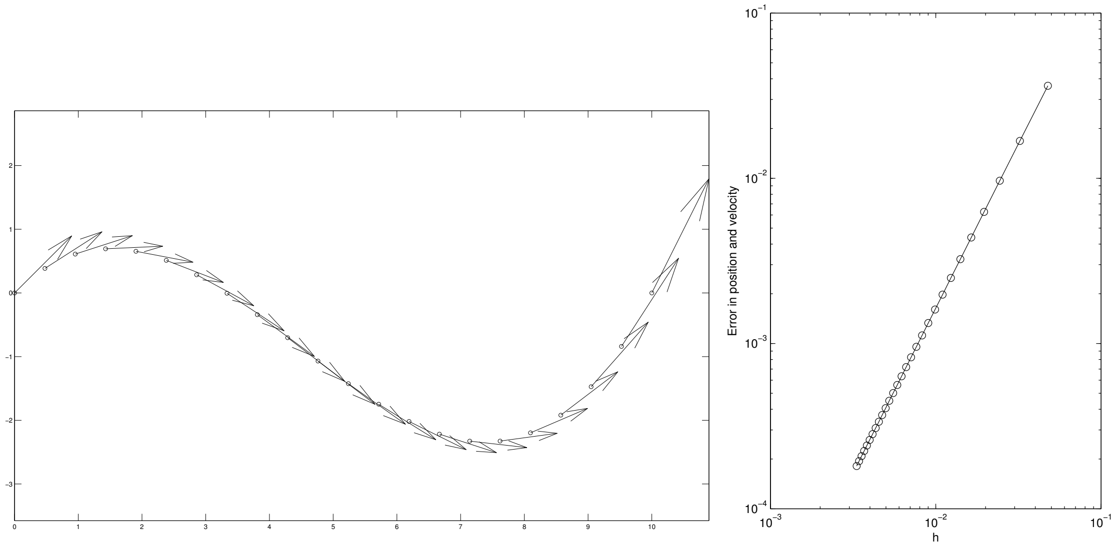


Figure 6.4: Left: simulation of the method (6.19) with $q_0 = (0,0)$ $v_0 = (10,10)$, $q_N = (10,0)$, $v_N = (10,20)$, $N = 21$ (velocities are scaled). Right: Error in position and velocity for different values of h.

for all solutions $q(t)$ of the second-order Euler-Lagrange equations with initial conditions $(q_0, \dot{q}_0, \ddot{q}_0) \in U_1$ and for all $h \leq h_1$.

Theorem 6.3.8 (See Theorem 4.8 of [151])**.** *If $\widetilde{F}_{L_d}$ is the the evolution map of an order r discretization $L_d : TQ \times TQ \to \mathbb{R}$ of the exact discrete Lagrangian $L_d^e : TQ \times TQ \to \mathbb{R}$, then*

$$\widetilde{F}_{L_d} = \widetilde{F}_{L_d^e} + \mathcal{O}(h^{r+1}).$$

In other words, $\widetilde{F}_{L_d}$ gives an integrator of order r for $\widetilde{F}_{L_d^e} = F_H^h$.

Note that given a discrete Lagrangian $L_d : TQ \times TQ \to \mathbb{R}$ its order can be calculated by expanding the expressions for $L_d(q(0), \dot{q}(0), q(h), \dot{q}(h), h)$ in a Taylor series in h and comparing this to the same expansions for the exact Lagrangian. If the series agree up to r terms, then the discrete Lagrangian is of order r.

6.4 Discrete mechanical systems with constraints

In the case of systems with constraints, the principle seeks to find a discrete curve $\{(q_k, v_k)\}_{k=0}^N$ which is a critical point of the discrete action subject to some discrete constraint functions (see [15]).

Let us consider the m-independent smooth functions $\Phi^\alpha : T^{(2)}Q \to \mathbb{R}$, $\alpha = 1, \ldots, m$ (that is $\{d\Phi^\alpha\}$ is of maximum rank at each point) and define the $n - m$ dimensional constraint submanifold $\mathcal{M}$ of $T^{(2)}Q$ by the vanishing of these functions, that is,

$$\mathcal{M} = \{(q, \dot{q}, \ddot{q}) \in T^{(2)}Q \mid \Phi^\alpha(q, \dot{q}, \ddot{q}) = 0, \text{ where } 1 \leq \alpha \leq m\}.$$

To establish the discrete setting we discretize the continuous Lagrangian and the constraint submanifold to a discrete lagrangian function $L_d : TQ \times TQ \to \mathbb{R}$ and a submanifold $\mathcal{M}_d$ of $TQ \times TQ$ determined by the vanishing of the m discrete constraints functions $\Phi_d^\alpha : TQ \times TQ \to \mathbb{R}$,

$$\mathcal{M}_d = \{(q_k, v_k, q_{k+1}, v_{k+1}) \in TQ \times TQ \mid \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \text{ where } 1 \leq \alpha \leq m\}.$$

We compute the critical points of the discrete action subjected to the constraint equations; that is,

$$\begin{cases} \min \mathcal{A}_d(q_d, v_d) \text{ with } (q_0, v_0) \text{ and } (q_N, v_N) \text{ fixed} \\ \text{subject to } \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}) = 0, \quad 1 \leq \alpha \leq m \text{ and } \ 0 \leq k \leq N-1\ . \end{cases} \tag{6.25}$$

Now, we define the *augmented Lagrangian* $\widetilde{L}_d : TQ \times TQ \times \mathbb{R}^m \to \mathbb{R}$ by

$$\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}, \lambda^k) = L_d(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda_\alpha^k \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}),$$

where $\lambda_\alpha^k \in \mathbb{R}^m$ with $\alpha = 1, \ldots, m$ are the Lagrange multipliers.

This new Lagrangian gives rise to the following unconstrained discrete variational problem

$$\begin{cases} \min \overline{\mathcal{A}}_d\left((q_d, v_d), \lambda^0, \lambda^1, \ldots, \lambda^{N-1}\right) \text{ with } (q_0, v_0) \text{ and } (q_N, v_N) \text{ fixed in } TQ\ , \\ \lambda^k \in \mathbb{R}^m \text{ and } k = 0, \ldots, N-1. \end{cases} \tag{6.26}$$

where

$$\overline{\mathcal{A}}_d\left((q_d, v_d), \lambda^0, \lambda^1, \ldots, \lambda^{N-1}\right) = \sum_{k=0}^{N-1} \left[L_d(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda_\alpha^k \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}) \right]\ ,$$

where λ^k is a m-vector with components λ_α^k with $1 \leq \alpha \leq m$.

From the classical lagrangian multiplier theorem, we have that the regular extremals of Problem (6.25) are the same than in Problem (6.26). Therefore, applying standard discrete variational calculus we deduce that the solutions of problem (6.25) verify the following set of difference equations

$$\begin{aligned} 0 &= D_1 L_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_3 L_d(q_{k-1}, v_{k-1}, q_k, v_k) \\ &+ \lambda_\alpha^k D_1 \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda_\alpha^{k-1} D_3 \Phi_d^\alpha(q_{k-1}, v_{k-1}, q_k, v_k), \\ 0 &= D_2 L_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_4 L_d(q_{k-1}, v_{k-1}, q_k, v_k) \\ &+ \lambda_\alpha^k D_2 \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda_\alpha^{k-1} D_4 \Phi_d^\alpha(q_{k-1}, v_{k-1}, q_k, v_k), \\ 0 &= \Phi_d^\alpha(q_k, v_k, q_{k+1}, v_{k+1}), \quad 0 = \Phi_d^\alpha(q_{k-1}, v_{k-1}, q_k, v_k), \quad 1 \leq \alpha \leq m \text{ and } 1 \leq k \leq N-1. \end{aligned} \tag{6.27}$$

If the $(2n+2m)\times(2n+2m)$ matrix

$$\begin{pmatrix} D_{13}L_d+\lambda_\alpha D_{13}\Phi_d^\alpha & D_{14}L_d+\lambda_\alpha D_{14}\Phi_d^\alpha & D_1\Phi_d^\alpha \\ D_{23}L_d+\lambda_\alpha D_{23}\Phi_d^\alpha & D_{24}L_d+\lambda_\alpha D_{24}\Phi_d^\alpha & D_2\Phi_d^\alpha \\ (D_3\Phi_d^\alpha)^T & (D_a\Phi_d^\alpha)^T & 0 \end{pmatrix}$$

is regular along $\mathcal{M}_d\times\mathbb{R}^m$, by the implicit function theorem; if the element $(q^*_{k-1}, v^*_{k-1}, q^*_k, v^*_k, q^*_{k+1}, v^*_{k+1}, \lambda^{k-1}_*, \lambda^k_*)$ satisfies equation (6.27), there exists a neighborhood $\widetilde{\mathcal{U}}_k\subset\mathcal{M}_d\times\mathbb{R}^m$ of the point $(q^*_{k-1}, v^*_{k-1}, q^*_k, v^*_k, \lambda^{k-1}_*)$ and an unique (local) application

$$\begin{array}{rccc} F_{\widetilde{L}_d}: & \widetilde{\mathcal{U}}_k & \longrightarrow & \mathcal{M}_d\times\mathbb{R}^m \\ & (q_{k-1}, v_{k-1}, q_k, v_k, \lambda^{k-1}_\alpha) & \longmapsto & (q_k, v_k, q_{k+1}, v_{k+1}, \lambda^k_\alpha)\ , \end{array}$$

where $(q_{k-1}, v_{k-1}, q_k, v_k, q_{k+1}, v_{k+1}, \lambda^{k-1}, \lambda^k)$ satisfies equations (6.27). Thus,

$$F_{\widetilde{L}_d}(q_{k-1}, v_{k-1}, q_k, v_k, \lambda^{k-1}) = (q_k, v_k, q_{k+1}, v_{k+1}, \lambda^k)$$

is the *discrete Lagrangian map* for equations (6.27).

6.5 Optimal control of mechanical systems

In this section we will study how to apply our variational integrator to optimal control problems. First, we will study optimal control problems for total actuated mechanical systems and we will show how our methods can be applied to the optimal control of a robotic leg. Secondly we will apply our techniques to underactuated mechanical control systems showing, as an implementation of our integrator, the control of a cart with an inverted pendulum on it.

6.5.1 Optimal control of fully actuated systems

Let $L: TQ\to\mathbb{R}$ be a regular Lagrangian and we take local coordinates (q^A) on Q where $1\leq A\leq n$. For this Lagrangian the *controlled Euler-Lagrange equations* are (see section 3.6)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^A}\right)-\frac{\partial L}{\partial q^A}=u_A, \tag{6.28}$$

where $u\in U\subset\mathbb{R}^n$ is an open subset of $\mathbb{R}^n$, the set of control parameters.

The optimal control problem consists on finding a trajectory of the sates variables and controls input $(q^{(A)}(t), u^A(t))$ satisfying (6.28) given initial and final conditions $(q^A(t_0), \dot{q}^A(t_0)), (q^A(t_f), \dot{q}^A(t_f))$ respectively, minimizing the cost function

$$\mathcal{A}=\int_{t_0}^{t_f} C(q^A, \dot{q}^A, u_A)dt,$$

where $C: TQ\times U\to\mathbb{R}$.

From (6.28) we can rewrite the cost function as a second-order Lagrangian $\widetilde{L} : T^{(2)}Q \to \mathbb{R}$ given by

$$\widetilde{L}(q^A, \dot{q}^A, \ddot{q}^A) = C\left(q^A, \dot{q}^A, \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^A} - \frac{\partial L}{\partial q^A}\right)$$

replacing the the controls by the Euler-Lagrange equations in the cost function (see [17] for example).

We define the discrete Lagrangian $\widetilde{L}_d : TQ \times TQ \to \mathbb{R}$,

$$\begin{aligned}
\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) &= \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k+1}}{2}, \frac{v_k + v_{k+1}}{2}, \frac{2}{h^2}(q_{k+1} - q_k - hv_k)\right) \\
&+ \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k+1}}{2}, \frac{v_k + v_{k+1}}{2}, \frac{2}{h^2}(q_k - q_{k+1} + hv_{k+1})\right).
\end{aligned}$$

Other natural possibilities are

$$\begin{aligned}
\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) &= hL\left(\frac{q_k + q_{k+1}}{2}, \frac{q_{k+1} - q_k}{h}, \frac{v_{k+1} - v_k}{h}\right) \quad \text{or} \\
\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) &= \frac{1}{2}L\left(q_k, v_k, \frac{v_{k+1} - v_k}{h}\right) + \frac{1}{2}L\left(q_{k+1}, v_{k+1}, \frac{v_{k+1} - v_k}{h}\right)
\end{aligned}$$

Applying the results given in Section 6.2, we know that the extremals of the optimal control problem are obtained solving the discrete Euler-Lagrange equations

$$\begin{aligned}
D_1\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_3\widetilde{L}_d(q_{k-1}, v_{k-1}, q_k, v_k) &= 0, \\
D_2\widetilde{L}_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_4\widetilde{L}_d(q_{k-1}, v_{k-1}, q_k, v_k) &= 0.
\end{aligned}$$

If the matrix

$$\begin{pmatrix} D_{13}\widetilde{L}_d & D_{14}\widetilde{L}_d \\ D_{23}\widetilde{L}_d & D_{24}\widetilde{L}_d \end{pmatrix}$$

is regular, then one can define the discrete Lagrangian map, to solve the optimal control problem.

Example 6.5.1. Two-link manipulator

We consider the optimal control of a two-link manipulator which is a classical example studied in robotics (see for example [142] and [146]). The two-link manipulator consists of two coupled (planar) rigid bodies with mass m_i, length l_i and inertia J_i, with $i = 1, 2$, respectively. Let θ_1 and θ_2 be the angles in which the first link is measured conterclockwise from the positive horizontal axis rotating about the origin and the angle of the second link rotating about the endpoint of the first link, respectively. If we assume one end of the first link to be fixed in an inertial reference frame, the configuration of the system is locally specified by the coordinate $q = (\theta_1, \theta_2) \in \mathbb{S}^1 \times \mathbb{S}^1$. The Lagrangian is, as is usual, given by the kinetic energy of the system minus the potential energy, that is,

$$\begin{aligned}
L(q, \dot{q}) &= \frac{1}{8}(m_1 + 4m_2)l_1^2\dot{\theta}_1^2 + \frac{1}{8}m_2 l_2^2(\dot{\theta}_1 + \dot{\theta}_2)^2 + \frac{1}{2}m_2 l_1 l_2 \cos(\theta_2)\dot{\theta}_1(\dot{\theta}_1 + \dot{\theta}_2) + \frac{1}{2}J_1\dot{\theta}_1^2 \\
&+ \frac{1}{2}J_2(\dot{\theta}_1 + \dot{\theta}_2)^2 + g\left(\frac{1}{2}m_1 l_1 \sin\theta_1 + m_2 l_1 \sin\theta_1 + \frac{1}{2}m_2 l_2(\theta_1 + \theta_2)\right)
\end{aligned}$$

where g is the gravitational constant acceleration.

The control torques u_1 and u_2 are applied at the base of the first link and the joint between the two links. The equations of motion of the controlled system are

$$\begin{aligned}
u_1 &= -\sin\theta_2 l_1 l_2 m_2 \dot{\theta}_2\dot{\theta}_1 - \frac{1}{2}\sin\theta_2\dot{\theta}_2^2 l_1 l_2 m_2 + \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g \\
&+ \left(m_2 g\cos\theta_1 + \frac{1}{2}g\cos\theta_1 m_1\right)l_1 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot{\theta}_2 \\
&+ \left(\cos\theta_2 l_1 l_2 m_2 + \left(\frac{m_1}{4}+m_2\right)l_1^2 + \frac{m_2 l_2^2}{4} + J_1 + J_2\right)\ddot{\theta}_1, \\
u_2 &= \frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot{\theta}_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot{\theta}_1 \\
&+ \frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot{\theta}_2.
\end{aligned}$$

We look for trajectories $(\theta_1(t), \theta_2(t), u(t))$ of the state variables and control inputs with initial and final conditions, $(\theta_1(0), \theta_2(0), \dot{\theta}_1(0), \dot{\theta}_2(0)) = (-\pi/2+0.2, 0, 0, 0)$ and $(\theta_1(T), \theta_2(T), \dot{\theta}_1(T), \dot{\theta}_2(T)) = (-\pi/2, 0, 0, 0)$ respectively, and minimizing the cost functional

$$\mathcal{A} = \frac{1}{2}\int_0^T (u_1^2 + u_2^2)dt.$$

We construct the discrete Lagrangian $\widetilde{L}_d : T(\mathbb{S}^1\times\mathbb{S}^1)\times T(\mathbb{S}^1\times\mathbb{S}^1)\to\mathbb{R}$, discretizing the Lagrangian $\widetilde{L} : T^{(2)}Q \to \mathbb{R}$ given by

$$\begin{aligned}
\widetilde{L}(\theta_1,\theta_2,\dot{\theta}_1,\dot{\theta}_2,\ddot{\theta}_1,\ddot{\theta}_2) &= \frac{1}{2}\left[\frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot{\theta}_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot{\theta}_1\right. \\
&+ \left.\frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot{\theta}_2\right]^2 \\
&+ \frac{1}{2}\left[\frac{1}{2}\sin\theta_2 l_1 l_2 m_2\dot{\theta}_1^2 + \left(\frac{1}{4}m_2 l_2^2 + J_2 + \frac{1}{2}\cos\theta_2 l_1 l_2 m_2\right)\ddot{\theta}_1\right. \\
&+ \left.\frac{1}{2}m_2 l_2\cos(\theta_1+\theta_2)g + \left(\frac{1}{4}m_2 l_2^2 + J_2\right)\ddot{\theta}_2\right]^2
\end{aligned}$$

taking the same discretization than in Equation (6.18) to approximate the acceleration and taking midpoint averages to approximate the position and velocity.

Now we show some simulations to test our method in Figure 6.5 for $T=1$, $m_1 = 1.5$, $m_2 = 1$, $l_1 = l_2 = 1$, $J_1 = \frac{m_1 l_1^2}{12}$, $J_2 = \frac{m_2 l_2^2}{12}$ and $g = 10$.

The following table show the root mean square error in positions and velocities:

h	Error in position	Error in velocity
0.1000	0.0128	0.0655
0.0556	0.0042	0.0238
0.0312	0.0014	0.0080

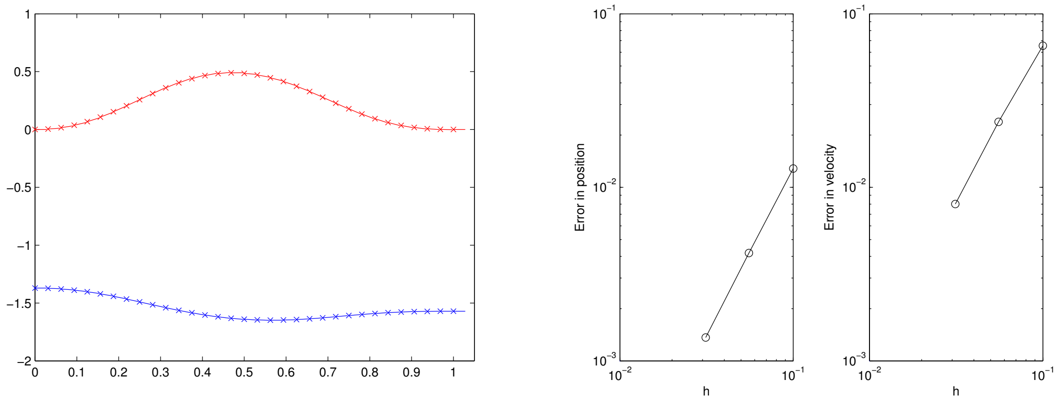


Figure 6.5: Left: Simulation of the method with $q_0 = (-\pi/2 + .2, 0)$ $v_0 = (0, 0)$, $q_N = (-\pi/2, 0)$, $v_N = (0, 0)$, $N = 30$. Blue and red lines show the positions of the first and second links respectively, and the slope between the segments corresponds to the (scaled) velocity. Right: Error (root-mean-square error) in position and velocity for different values of h.

6.5.2 Optimal control of underactuated systems

Let Q be a configuration manifold of a mechanical system with local coordinates (q^A) with $1 \leq A \leq n$. Given a Lagrangian $L : TQ \to \mathbb{R}$, the corresponding control system is called *underactuated* if the number of independent control inputs is less than the dimension of Q. We denote these control forces by $\bar{X}^a = \bar{X}^a_A dq^A$, $1 \leq a \leq r < n$. We complete this basis of the control forces to a (local) basis $\{\bar{X}^a, \bar{X}^\alpha\}$ of one-forms on Q, where $\alpha = r+1, \ldots, n$. The controlled Euler-Lagrange equations are given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = u_a \bar{X}^a_A.$$

Taking the dual basis $\{X_a, X_\alpha\}$ of vector fields on Q, these equations can be rewritten as

$$\begin{aligned} \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A}\right) X^A_a(q) &= u_a, \\ \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A}\right) X^A_\alpha(q) &= 0\,, \end{aligned} \tag{6.29}$$

where $a = 1, \ldots, r$, and $\alpha = r+1, \ldots, n$. Also, from this basis of vector fields we can induce local coordinates $(q^A, \dot{q}^A) = (q^A, (\dot{q}^a, \dot{q}^\alpha))$ on TQ as $\dot{q}^A = \dot{q}^a X^A_a(q) + \dot{q}^\alpha X^A_\alpha(q)$.

We will study the optimal control problem which consist on finding a trajectory $(q^A(t), u^a(t))$ of the state variables and control inputs satisfying equations (6.29) from given initial and final conditions, $(q^A(t_0), \dot{q}^A(t_0))$, $(q^A(t_f), \dot{q}^A(t_f),)$ respectively, and minimizing the cost functional

$$\mathcal{A} = \int_{t_0}^{t_f} C(q^A, \dot{q}^A, u_a)\, dt,$$

where $C : TQ \times U \to \mathbb{R}$.

This optimal control problem is equivalent to the following second-order variational problem with second-order constraints:

$$\text{extemize } \widetilde{\mathcal{A}} = \int_{t_0}^{t_f} \widetilde{L}(q^A(t), \dot{q}^A(t), \ddot{q}^A(t))\, dt$$

subject to the second-order constraints

$$\Phi^\alpha(q^A, \dot{q}^A, \ddot{q}^A) = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A} \right) - \frac{\partial L}{\partial q^A} \right) X^A_\alpha(q) \ ,$$

and the boundary conditions, where $\widetilde{L} : T^{(2)}Q \to \mathbb{R}$ is defined as

$$\widetilde{L}(q^a, q^\alpha, \dot{q}^a, \dot{q}^\alpha, \ddot{q}^a, \ddot{q}^\alpha) = C\left(q^a, q^\alpha, \dot{q}^a, \dot{q}^\alpha, F_a(q^A, \dot{q}^A, \ddot{q}^A)\right) .$$

Here $F_a : T^{(2)}Q \to \mathbb{R}$ is the function defined as

$$F_a(q^A, \dot{q}^A, \ddot{q}^A) = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A} \right) - \frac{\partial L}{\partial q^A} \right) X^A_a(q)$$

We define the discrete lagrangian $L_d : TQ \times TQ \to \mathbb{R}$ and the discrete constraints $\Phi^\alpha_d : TQ \times TQ \to \mathbb{R}$ by

$$\begin{aligned}
L_d(q_{k-1}, v_{k-1}, q_k, v_k) &= \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k-1}}{2}, \frac{v_k + v_{k-1}}{2}, \frac{2}{h^2}(q_k - q_{k-1} - hv_{k-1}) \right) \\
&+ \frac{h}{2}\widetilde{L}\left(\frac{q_k + q_{k-1}}{2}, \frac{v_k + v_{k-1}}{2}, \frac{2}{h^2}(q_{k-1} - q_k + hv_k) \right), \\
\Phi^\alpha_d(q_{k-1}, v_{k-1}, q_k, v_k) &= \frac{h}{2}\Phi^\alpha\left(\frac{q_k + q_{k-1}}{2}, \frac{v_k + v_{k-1}}{2}, \frac{2}{h^2}(q_k - q_{k-1} - hv_{k-1}) \right) \\
&+ \frac{h}{2}\Phi^\alpha\left(\frac{q_k + q_{k-1}}{2}, \frac{v_k + v_{k-1}}{2}, \frac{2}{h^2}(q_{k-1} - q_k + hv_k) \right).
\end{aligned}$$

Then then discrete algorithm is given by solving the following system of algebraic difference equations

$$\begin{aligned}
0 &= D_1L_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_3L_d(q_{k-1}, v_{k-1}, q_k, v_k) \\
&\quad + \lambda^k_\alpha D_1\Phi^\alpha_d(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda^{k-1}_\alpha D_3\Phi^\alpha_d(q_{k-1}, v_{k-1}, q_k, v_k), \\
0 &= D_2L_d(q_k, v_k, q_{k+1}, v_{k+1}) + D_4L_d(q_{k-1}, v_{k-1}, q_k, v_k) \\
&\quad + \lambda^k_\alpha D_2\Phi^\alpha_d(q_k, v_k, q_{k+1}, v_{k+1}) + \lambda^{k-1}_\alpha D_4\Phi^\alpha_d(q_{k-1}, v_{k-1}, q_k, v_k), \\
0 &= \Phi^\alpha_d(q_k, v_k, q_{k+1}, v_{k+1}), \quad 0 = \Phi^\alpha_d(q_{k-1}, v_{k-1}, q_k, v_k), \quad 1 \leq \alpha \leq m \text{ and } 1 \leq k \leq N-1.
\end{aligned} \tag{6.30}$$

If the matrix

$$\begin{pmatrix} D_{13}L_d + \lambda_\alpha D_{13}\Phi^\alpha_d & D_{14}L_d + \lambda_\alpha D_{14}\Phi^\alpha_d & D_1\Phi^\alpha_d \\ D_{23}L_d + \lambda_\alpha D_{23}\Phi^\alpha_d & D_{24}L_d + \lambda_\alpha D_{24}\Phi^\alpha_d & D_2\Phi^\alpha_d \\ (D_3\Phi^\alpha_d)^T & (D_4\Phi^\alpha_d)^T & 0 \end{pmatrix}$$

is regular, then we can define the discrete Lagrange map

$$\begin{array}{rccc} F_d : & \mathcal{M}_d \times \mathbb{R}^m & \longrightarrow & \mathcal{M}_d \times \mathbb{R}^m \\ & (q_k, v_k, q_{k+1}, v_{k+1}, \lambda_{\alpha,k}) & \longmapsto & (q_{k+1}, v_{k+1}, q_{k+2}, v_{k+2}, \lambda_{\alpha,k+1}) \end{array}$$

where $\mathcal{M}_d$ denotes the submanifold of $TQ \times TQ$ determined by the constraint equations $\Phi^\alpha_d = 0$.

Example 6.5.2. Cart-Pole system

A Cart-Pole System consists on a cart and an inverted pendulum on it. The coordinate x denotes the position of the cart on the x-axis and θ denotes the angle of the pendulum with the upright vertical. The configuration space is $Q = \mathbb{R} \times \mathbb{S}^1$.

The inertia matrix of this mechanical system is given by

$$\mathbb{I} = \begin{pmatrix} M+m & ml\cos\theta \\ ml\cos\theta & ml^2 \end{pmatrix}$$

where M is the mass of the cart and m, l are the mass and length of the pendulum, respectively. The potential energy of the cart-pole system is $V(\theta) = mgl\cos(\theta)$. Then the Lagrangian of this system is given by (kinetic energy minus potential energy)

$$L(q, \dot q) = L(x, \theta, \dot x, \dot\theta) = \frac{1}{2}M\dot x^2 + \frac{1}{2}m(\dot x^2 + 2\dot x l\dot\theta\cos\theta + l^2\dot\theta^2) - mgl\cos\theta - mg\widetilde{h}\ ,$$

where $\widetilde{h}$ is the car height.

We apply a control force u to our picture. The control input is parallel to the track remaining the joint angle θ unactuated. Therefore, the equations of motion of the controlled system are

$$\begin{array}{rcl} (M+m)\ddot x - ml\dot\theta^2\sin\theta + ml\ddot\theta\cos\theta & = & u \\ \ddot x\cos\theta + l\ddot\theta - g\sin\theta & = & 0 \end{array}$$

Now we look for trajectories $(x(t), \theta(t)), u(t))$ on the state variables and the control 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

$$\mathcal{A} = \frac{1}{2}\int_0^T u^2 dt.$$

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

$$\widetilde{L}(x,\theta,\dot{x},\dot{\theta},\ddot{x},\ddot{\theta}) = \frac{1}{2}\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x}\right)^2 = \frac{1}{2}\left[(M+m)\ddot{x} - ml\dot{\theta}^2\sin\theta + ml\ddot{\theta}\cos\theta\right]^2 .$$

As before, we construct the discrete Lagrangian and discrete constraint $\widetilde{L}_d : TQ \times TQ \to \mathbb{R}$, $\Phi_d : TQ \times TQ \to \mathbb{R}$ in the variables $(x_0, \dot{x}_0, \theta_0, \dot{\theta}_0, x_1, \dot{x}_1, \theta_1, \dot{\theta}_1)$ given by

$$\begin{aligned}
\widetilde{L}_d &= \frac{h}{4}\left(\frac{2(M+m)(x_1 - h\dot{x}_0 - x_0)}{h^2} - \frac{ml}{4}(\dot{\theta}_1 + \dot{\theta}_0)^2 \sin\left(\frac{\theta_1+\theta_0}{2}\right)\right. \\
&+ \left.\frac{2ml(\theta_1 - \theta_0 - h\dot{\theta}_0)\cos(\frac{1}{2}(\theta_1+\theta_2))}{h^2}\right)^2 + \frac{h}{4}\left(\frac{2(M+m)(h\dot{x}_1 + x_0 - x_1)}{h^2}\right. \\
&- \left.\frac{ml}{4}(\dot{\theta}_1 + \dot{\theta}_0)^2 \sin\left(\frac{\theta_1+\theta_2}{2}\right) + \frac{2ml(\theta_0 - \theta_1 + h\dot{\theta}_1)\cos(\frac{1}{2}(\theta_1+\theta_0))}{h^2}\right)^2 ,
\end{aligned}$$

$$\begin{aligned}
\Phi_d &= \frac{h}{2}\left(\frac{2(x_1 - x_0 - h\dot{x}_0)\cos(\frac{1}{2}(\theta_1+\theta_2))}{h^2} + \frac{2l(\theta_1 - \theta_0 - h\dot{\theta}_0)}{h^2} + g\sin(\frac{1}{2}(\theta_1+\theta_0))\right. \\
&+ \left.\frac{2(x_0 - x_1 + h\dot{x}_1)\cos(\frac{1}{2}(\theta_1+\theta_0))}{h^2} + \frac{2l(\theta_0 - \theta_1 + h\dot{\theta}_1)}{h^2} + g\sin(\frac{1}{2}(\theta_1+\theta_0))\right).
\end{aligned}$$

Now, we can use the equations given in (6.30) to simulate the behavior of the system as we have seen in the previous example.

Appendix A: Higher-order Euler-Arnold equations on $T^*(T^{(k-1)}G)$

This appendix deals with the construction of the Liouville 1-form and the canonical symplectic 2-form on $G\times(k-1)\mathfrak{g}\times k\mathfrak{g}^*$ to obtain higher-order *Euler-Arnold* equations. This construction was very used in Chapter 3.

Euler-Arnold equations

Let G be a finite dimensional Lie group. The left multiplication $L : G \to G$ allows us to trivialize the tangent bundle TG and the cotangent bundle T^*G as follows

$$\begin{array}{rcll} \Lambda : TG & \to & G\times\mathfrak{g}, & (g,\dot g) \longmapsto (g, g^{-1}\dot g) = (g, T_gL_{g^{-1}}\dot g) = (g,\xi)\ , \\ \Lambda^* : T^*G & \to & G\times\mathfrak{g}^*, & (g,\alpha_g) \longmapsto (g, T_e^*L_g(\alpha_g)) = (g,\alpha)\ , \end{array}$$

where $\mathfrak{g} = T_eG$ is the Lie algebra of G and e is the neutral element of G (see, for instance, [13]). In the same way, we have the following identifications: $TTG \equiv G\times 3\mathfrak{g}$, $T^*TG = G\times\mathfrak{g}\times 2\mathfrak{g}^*$. $TT^*G = G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ and $T^*T^*G = G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ (the same is valid for the right-translation, but in the sequel we only work with the left-translation, for sake of simplicity).

Using this left trivialization it is possible to write the classical Hamiltonian equations for a Hamiltonian function $H : T^*G \to \mathbb{R}$ from a different and interesting perspective. For instance, it is easy to show that (see [13]) the canonical structures of the cotangent bundle: the Liouville 1-form θ_G and the canonical symplectic 2-form ω_G, are now rewritten using this left-trivialization as follows:

$$\begin{aligned} (\theta_G)_{(g,\alpha)}(\xi_1,\nu_1) &= \langle\alpha,\xi_1\rangle\ , &(6.31)\\ (\omega_G)_{(g,\alpha)}\left((\xi_1,\nu_1),(\xi_2,\nu_2)\right) &= -\langle\nu_1,\xi_2\rangle + \langle\nu_2,\xi_1\rangle + \langle\alpha,[\xi_1,\xi_2]\rangle\ , &(6.32)\end{aligned}$$

with $(g,\alpha)\in G\times\mathfrak{g}^*$, where $\xi_i\in\mathfrak{g}$ and $\nu_i\in\mathfrak{g}^*$, $i=1,2$ and we have used the previous identifications. Observe that we are identifying the elements of $T_{\alpha_g}T^*G$ with the pairs $(\xi,\nu)\in\mathfrak{g}\times\mathfrak{g}^*$.

Therefore given the Hamiltonian $H : T^*G \equiv G\times\mathfrak{g}^* \longrightarrow \mathbb{R}$, we compute

$$dH_{(g,\alpha)}(\xi_2,\nu_2) = \langle L_g^*\left(\frac{\delta H}{\delta g}(g,\alpha)\right),\xi_2\rangle + \langle\nu_2,\frac{\delta H}{\delta\alpha}(g,\alpha)\rangle\ , \tag{6.33}$$

since $\frac{\delta H}{\delta\alpha}(g,\alpha) \in \mathfrak{g}^{**} = \mathfrak{g}$.

We now derive the Hamilton's equations which are satisfied by the integral curves of the Hamiltonian vector field X_H on T^*G. After a left-trivialization, $X_H(g,\alpha) = (\xi_1, \nu_1)$ where $\xi_1 \in \mathfrak{g}$ and $\nu_1 \in \mathfrak{g}^*$ are elements to be determined using the Hamilton's equations

$$i_{X_H}\omega_G = dH \ .$$

Therefore, from expressions (6.32) and (6.33) we deduce that

$$\begin{aligned} \xi_1 &= \frac{\delta H}{\delta\alpha}(g,\alpha) \ , \\ \nu_1 &= -L_g^*\left(\frac{\delta H}{\delta g}(g,\alpha)\right) + ad^*_{\xi_1}\alpha \ . \end{aligned}$$

In other words, taking $\dot g = g\xi_1$ we obtain the *Euler-Arnold equations*:

$$\begin{aligned} \dot g &= T_eL_g(\frac{\delta H}{\delta\alpha}(g,\alpha)) \equiv g\frac{\delta H}{\delta\alpha}(g,\alpha) \ , \\ \dot\alpha &= -L_g^*\left(\frac{\delta H}{\delta g}(g,\alpha)\right) + ad^*_{\frac{\delta H}{\delta\alpha}(g,\alpha)}\alpha \ . \end{aligned}$$

If the Hamiltonian is left-invariant, that is, $h : \mathfrak{g}^* \to \mathbb{R}$ where $h(\alpha) = H(e,\alpha)$ then we deduce that

$$\begin{aligned} \dot g &= g\frac{\delta h}{\delta\alpha} \ , \\ \dot\alpha &= ad^*_{\delta h/\delta\alpha}\alpha \ . \end{aligned}$$

The last equation is known as the *Lie-Poisson equation* for a Hamiltonian $h : \mathfrak{g}^* \to \mathbb{R}$.

Higher-order Euler-Arnold equations on $T^*(T^{(k-1)}G)$

Combining the results of the previous subsection and subsection 1.7 we have that

$$T^*(T^{(k-1)}G) \equiv T^*(G \times (k-1)\mathfrak{g}) \equiv T^*G \times (k-1)T^*\mathfrak{g} \equiv G \times (k-1)\mathfrak{g} \times k\mathfrak{g}^* \ .$$

Denote by $\boldsymbol{\xi} \in (k-1)\mathfrak{g}$ and $\boldsymbol{\alpha} \in k\mathfrak{g}^*$ with components $\boldsymbol{\xi} = (\xi^{(0)}, \dots, \xi^{(k-2)})$ and $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{k-1})$.

As we work in a vector space, the Liouville $1-$form $\theta_{G\times(k-1)\mathfrak{g}} \in \Lambda^1(G \times \mathfrak{g}^* \times (k-1)(\mathfrak{g}\times\mathfrak{g}^*))$ is expressed as

$$\theta_{G\times(k-1)\mathfrak{g}} = \theta_G + \theta_{(k-1)\mathfrak{g}}.$$

We want to know $\theta_{G\times(k-1)\mathfrak{g}}$; this 1-form at the point $(g, \boldsymbol{\xi}, \boldsymbol{\alpha})$ is applied to elements $(\boldsymbol{\xi}_1, \boldsymbol{\nu}^1) \in T_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}(G \times \mathfrak{g}^* \times (k-1)(\mathfrak{g}\times\mathfrak{g}^*))$, where $\boldsymbol{\xi}_a \in k\mathfrak{g}$ and $\boldsymbol{\nu}^a \in k\mathfrak{g}^*$, $a = 1,2$ with components $\boldsymbol{\xi}_a = (\xi_a^{(i)})_{0\le i\le k-1}$ and $\boldsymbol{\nu}^a = (\nu^a_{(i)})_{0\le i\le k-1}$ where each component $\xi_a^{(i)} \in \mathfrak{g}$ and $\nu^a_{(i)} \in \mathfrak{g}^*$. Observe that α_0 comes from the identification $T^*G = G \times \mathfrak{g}^*$.

To compute θ_G we need to find the tangent application to $\tau \circ Pr_{(1,2)}\Lambda^*_{(2)}$ where $\mathrm{Pr}_{(1,2)} : G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \to G \times \mathfrak{g}^*$ is a canonical projection onto the first and second factors, and $\tau :$

$T^*G \simeq G\times\mathfrak{g}^* \to G$ is the fibration which defines T^*G and where $\Lambda^*_{(2)} : T^*TG \to G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ is the cotangent left-trivialization. We consider the application

$$\varphi_t^{(\xi_1^0,\nu_0^1)} : G \times \mathfrak{g}^* \to G \times \mathfrak{g}^*.$$

This is applied to an element $(g, \alpha_0) \in G \times \mathfrak{g}^*$ and gives an element $(g\exp(t\xi_1^0), \alpha_0 + t\nu_0^1) \in G \times \mathfrak{g}^*$. $\varphi_t^{(\xi_1^0,\nu_0^1)}$ is the flow of the vector field $X^{(\xi_1^0,\nu_0^1)}(g, \alpha_0) = (g\xi_1^0, \nu_0^1)$.

Therefore, the tangent application for $\tau \circ Pr_{(1,2)}\Lambda^*_{(2)}$ is

$$\begin{aligned}
T_{(g,\alpha_0)}(\tau \circ Pr_{(1,2)}\Lambda^*_{(2)})(g\xi_1^0, \nu_0^1) &= \left.\frac{d}{dt}\right|_{t=0} \tau \circ Pr_{(1,2)}\Lambda^*_{(2)}(\varphi_t^{(\xi_1^0,\nu_0^1)}(g, \alpha_0)) \\
&= \left.\frac{d}{dt}\right|_{t=0} g\exp(t\xi_1^0) = g\xi_1^0,
\end{aligned}$$

Now, we can compute θ_G

$$\begin{aligned}
\langle\theta_{(g,\alpha_0)}, (g\xi_1^0, \nu_0^1)\rangle &= \langle\theta_{(Pr_{(1,2)}\Lambda^*_{(2)})}(g, \alpha_0), T_{(g,\alpha_0)}(Pr_{(1,2)}\Lambda^*_{(2)})(g\xi_1^0, \nu_0^1)\rangle \\
&= \langle\alpha_0, \xi_1^0\rangle = \alpha_0(\xi_1^0).
\end{aligned}$$

In the same way, one can compute $\theta_{(k-1)\mathfrak{g}}$. This is given by $\sum_{i=1}^{k-1} \alpha_i(\xi_1^{(i)})$. Then,

$$(\theta_{G\times(k-1)\mathfrak{g}})_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}(\boldsymbol{\xi}_1, \boldsymbol{\nu}^1) = \langle\boldsymbol{\alpha}, \boldsymbol{\xi}_1\rangle.$$

In the next, we will find the expression of the $2-$form $\omega_{G\times(k-1)\mathfrak{g}}$. To do this, we will use the followings formula

$$-d\theta_{G\times(k-1)\mathfrak{g}} = -d(\theta_G + \theta_{(k-1)\mathfrak{g}}) = -d(\theta_G) - d(\theta_{(k-1)\mathfrak{g}}).$$

To compute $-d\theta_G$, we use the formula

$$\begin{aligned}
-d\theta_G(X^{(\xi_1^0,\nu_0^1)}, X^{(\xi_0^1,\nu_0^1)}) &= -i_{X^{(\xi_1^0,\nu_0^1)}} d\left(i_{X^{(\xi_1^0,\nu_0^1)}}\theta_G\right) \\
&+ i_{X^{(\xi_1^0,\nu_0^1)}} d\left(i_{X^{(\xi_1^0,\nu_0^1)}}\theta_G\right) \\
&+ i_{[X^{(\xi_1^0,\nu_0^1)},X^{(\xi_1^0,\nu_0^1)}]}\theta_G.
\end{aligned}$$

In what follows, we compute each term of the equality las equality,

$$
\begin{aligned}
& i_{X^{(\xi_1^0,\nu_0^1)}} d\left(i_{X^{(\xi_2^0,\nu_0^2)}}\theta_G\right)(g,\alpha_0) = \\
& L_{X^{(\xi_1^0,\nu_0^1)}}\left(i_{X^{(\xi_2^0,\nu_0^2)}}\theta_G\right)(g,\alpha_0) = \\
& \left.\frac{d}{dt}\right|_{t=0}\left(\varphi_t^{(\xi_1^0,\nu_0^1)}\right)^*\left(i_{X^{(\xi_2^0,\nu_0^2)}}\theta_G\right)(g,\alpha_0) = \\
& \left.\frac{d}{dt}\right|_{t=0}\langle\theta_G(g\exp(t\xi_1^0),\alpha_0+t\nu_0^1), X^{(\xi_2^0,\nu_0^2)}(g\exp(t\xi_1^0),\alpha_0+t\nu_0^1)\rangle = \\
& \left.\frac{d}{dt}\right|_{t=0}\langle\theta_G(g\exp(t\xi_1^0),\alpha_0+t\nu_0^1),(g\exp\nu_0^1\xi_2^0,\nu_0^2)\rangle = \\
& \left.\frac{d}{dt}\right|_{t=0}(\alpha_0+t\nu_0^1)(\nu_0^2) = \nu_0^1(\xi_2^0).
\end{aligned}
$$

The second term is computed in a similar way, and is given by $\nu_0^2(\xi_1^0)$. To compute the third term, we observe that

$$
\begin{aligned}
& [X^{(\xi_1^0,\nu_0^1)}, X^{(\xi_2^0,\nu_0^2)}](g,\alpha_0) = \left.\frac{d}{dt}\right|_{t=0}\left(\varphi_{-\sqrt{t}}^{(\xi_2^0,\nu_0^2)}\circ\varphi_{-\sqrt{t}}^{(\xi_1^0,\nu_0^1)}\circ\varphi_{\sqrt{t}}^{(\xi_2^0,\nu_0^2)}\circ\varphi_{\sqrt{t}}^{(\xi_1^0,\nu_0^1)}\right)(g,\alpha_0) \\
& = \left.\frac{d}{dt}\right|_{t=0}(g\exp(\sqrt{t}\xi_1^0)\exp(\sqrt{t}\xi_2^0)\exp-\sqrt{t}\xi_1^0\exp-\sqrt{t}\xi_2^0,\alpha_0) = (T_eL_g[\xi_1^0,\xi_2^0],0) = \\
& (g[\xi_1^0,\xi_2^0],0).
\end{aligned}
$$

Then,

$$
\begin{aligned}
\theta_G\left([X^{(\xi_1^0,\nu_0^1)}, X^{(\xi_2^0,\nu_0^2)}]\right)(g,\alpha_0) = \; & = \theta_G(g,\alpha_0)(g[\xi_1^0,\xi_2^0],0) \\
& = \alpha_0([\xi_1^0,\xi_2^0]).
\end{aligned}
$$

Therefore,

$$
-d\theta_G\left(X^{(\xi_1^0,\nu_0^1)}, X^{(\xi_2^0,\nu_0^2)}\right) = \omega_{(g,\alpha_0)}((g\xi_1^0,\nu_0^1),(g\xi_2^0,\nu_0^2)) = -\nu_0^1(\xi_2^0)+\nu_0^2(\xi_1^0)+\alpha_0([\xi_1^0,\xi_2^0]).
$$

Applying, as before, the same formula we have

$$
\omega_{(k-1)\mathfrak{g}} = \sum_{i=1}^{k-1}\langle\nu_{(i)}^1,\xi_2^{(i)}\rangle + \langle\nu_{(i)}^2,\xi_1^{(i)}\rangle.
$$

Since $\omega_{G\times(k-1)\mathfrak{g}} = \omega_G + \omega_{(k-1)\mathfrak{g}}$, we have the identities,

$$
\begin{aligned}
(\theta_{G\times(k-1)\mathfrak{g}})_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}(\boldsymbol{\xi}_1,\boldsymbol{\nu}^1) &= \langle\boldsymbol{\alpha},\boldsymbol{\xi}_1\rangle\,, \\
(\omega_{G\times(k-1)\mathfrak{g}})_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}\left((\boldsymbol{\xi}_1,\boldsymbol{\nu}^1),(\boldsymbol{\xi}_2,\boldsymbol{\nu}^2)\right) &= -\langle\boldsymbol{\nu}^1,\boldsymbol{\xi}_2\rangle + \langle\boldsymbol{\nu}^2,\boldsymbol{\xi}_1\rangle + \langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle \\
&= -\sum_{i=0}^{k-1}\left[\langle\nu_{(i)}^1,\xi_2^{(i)}\rangle + \langle\nu_{(i)}^2,\xi_1^{(i)}\rangle\right] + \langle\alpha_0,[\xi_1^{(0)},\xi_2^{(0)}]\rangle.
\end{aligned}
$$

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

$$\begin{aligned} dH_{(g,\boldsymbol{\xi},\boldsymbol{\alpha})}(\boldsymbol{\xi}_2, \boldsymbol{\nu}^2) &= \langle L_g^*\left(\frac{\delta H}{\delta g}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\right), \xi_2^{(0)}\rangle + \sum_{i=0}^{k-2}\langle \frac{\delta H}{\delta \xi^{(i)}}(g,\boldsymbol{\xi},\boldsymbol{\alpha}), \xi_2^{(i+1)}\rangle \\ &\quad + \langle \boldsymbol{\nu}^2, \frac{\delta H}{\delta \boldsymbol{\alpha}}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\rangle. \end{aligned}$$

As in the last subsection, we can derive the Hamilton's equations which are satisfied by the integral curves of the Hamiltonian vector field X_H defined by $X_H(g,\boldsymbol{\xi},\boldsymbol{\alpha}) = (\boldsymbol{\xi}_1, \boldsymbol{\nu}^1)$. Therefore, we deduce that

$$\begin{aligned} \boldsymbol{\xi}_1 &= \frac{\delta H}{\delta \boldsymbol{\alpha}}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\ , \\ \nu^1_{(0)} &= -L_g^*\left(\frac{\delta H}{\delta g}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\right) + ad^*_{\xi_1^{(0)}}\alpha_0\ , \\ \nu^1_{(i+1)} &= -\frac{\delta H}{\delta \xi^{(i)}}(g,\boldsymbol{\xi},\boldsymbol{\alpha}), \qquad 0 \le i \le k-2\ . \end{aligned}$$

In other words, taking $\dot{g} = g\xi^{(0)}$ we obtain the *higher-order Euler-Arnold equations*:

$$\begin{aligned} \dot{g} &= g\frac{\delta H}{\delta \alpha_0}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\ , \\ \frac{d\xi^{(i)}}{dt} &= \frac{\delta H}{\delta \alpha_i}(g,\boldsymbol{\xi},\boldsymbol{\alpha}), \qquad 1 \le i \le k-1\ , \\ \frac{d\alpha_0}{dt} &= -L_g^*\left(\frac{\delta H}{\delta g}(g,\boldsymbol{\xi},\boldsymbol{\alpha})\right) + ad^*_{\delta H/\delta \alpha_0}\alpha_0\ , \\ \frac{d\alpha_{i+1}}{dt} &= -\frac{\delta H}{\delta \xi^{(i)}}(g,\boldsymbol{\xi},\boldsymbol{\alpha}), \qquad 0 \le i \le k-2\ . \end{aligned}$$

Appendix B: Technical results for Chapter 6

This Appendix deals with some technical results which have used in subsection 6.1.2.

Let F be the space of $\mathbb{R}^n$-valued polynomials of degree at most $k-1$,

$$F = \mathrm{span}_{\mathbb{R}^n}(b_0^{[j]}, \ldots, b_{k-1}^{[j]}) = \{c^j b_j^{[k]} | c^0, \ldots, c^{k-1} \in \mathbb{R}^n\}$$

where b_j, with $j = 0, \ldots, k-1$ is a basis of this space of polynomials consisting of orthonormal polynomials on $[0,1]$.

Let E be the kernel of g, where $g = (g_0, \ldots, g_{k-1}) : C^l([0,1], \mathbb{R}^n) \to (\mathbb{R}^n)^k$, this minds that E is the tangent space of the constraint set determined by $g_j[\cdot] = \langle b_j^{[k]}, \cdot \rangle\!\rangle$.

Lemma 6.5.3. *$F = E^\perp$ where the orthogonal complement of E is taking with respect to the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$.*

Proof. First we show that $F \subset E^\perp$. Taking $e \in E$,

$$\begin{aligned}
\langle\!\langle c^j b_j^{[k]}, e \rangle\!\rangle &= \int_0^1 (c^j b_j^{[k]}(u)) \cdot e(u) du = \sum_{i=1}^n \int_0^1 c_i^j b_j^{[k]}(u) e_i(u) du \\
&= c^j \cdot \left(\int_0^1 b_j^{[k]}(u) e_1(u), \ldots, \int_0^1 b_j^{[k]}(u) e_n(u) \right) \\
&= c^j \cdot \langle b_j^{[k]}, e \rangle\!\rangle = c^j \cdot g_j[e] = 0,
\end{aligned}$$

since $e \in E = \mathrm{Ker} g$.

Also, if $e' \in E^\perp$ then $E^\perp \subset F$ since $\langle\!\langle e', e \rangle\!\rangle = 0$ for all $e \in E$; therefore $F = E^\perp$. $\square$

Lemma 6.5.4. *There exist an orthogonal decomposition of the space $C^l([0,1], \mathbb{R}^n)$ between E and F. That is,*

$$C^l([0,1], \mathbb{R}^n) = E \oplus F.$$

Proof. Let $Q^{[k]}$ be an element of E, then $g(Q^{[k]}) = 0$, that is $g_j[Q^{[k]}(s)] = \langle b_j^{[k]}, Q^{[k]} \rangle\!\rangle = 0$ for all j. Let $c^j b_j^{[k]}$ be an element of F, then

$$\langle\!\langle c^j b_j^{[k]}, Q^{[k]} \rangle\!\rangle = c^j \cdot \langle b_j^{[k]}, Q^{[k]} \rangle\!\rangle = 0.$$

Therefore all element of E is orthogonal to all element of F and conversely.

Now, if $Q^{[k]} \in E \cap F$ since $Q^{[k]} \in F$ it follows that $Q^{[k]} = c^j b_j$. Therefore $\forall j'$, $0 = \langle b_{j'}, c^j b_j \rangle = c^j$, that is, $c^j = 0$ for all j and then $Q^{[k]} = 0$. Thus, $E \cap F = 0$.

Finally, let $Q^{[k]} \in C^l([0,1], \mathbb{R}^n)$. Taking into account that $\sum_{j=0}^{k-1} \langle b_j, Q^{[k]} \rangle b_j \in F$ we write

$$Q^{[k]} = \left(Q^{[k]} - \sum_{j=0}^{k-1} \langle b_j, Q^{[k]} \rangle b_j \right) + \sum_{j=0}^{k-1} \langle b_j, Q^{[k]} \rangle b_j.$$

Observe that $Q^{[k]} - \sum_{j=0}^{k-1} \langle b_j, Q^{[k]} \rangle b_j \in E$ since

$$\langle b_{j'}, Q^{[k]} - \sum_{j=0}^{k-1} \langle b_j, Q^{[k]} \rangle b_j \rangle = \langle b_{j'}, Q^{[k]} \rangle - \sum_{j=0}^{k-1} \delta^j_{j'} \langle b_j, Q^{[k]} \rangle = 0.$$

Therefore $C^l([0,1], \mathbb{R}^n) = E + F$ and then $C^l([0,1], \mathbb{R}^n) = E \oplus F$. $\square$

Appendix C: Retraction maps

In this appendix we will review the basics notions about retraction maps and the Cayley transformation (as an example of retraction map) which we use along this thesis (see [53], [54] and [66] for example).

A *retraction map* $\tau : \mathfrak{g} \to G$ is an analytic local diffeomorphism which maps a neighborhood of $0 \in \mathfrak{g}$ onto a neighborhood of the neutral element $e \in G$, such that $\tau(0) = e$ and $\tau(\xi)\tau(-\xi) = e$, for $\xi \in \mathfrak{g}$. There are many choices for the map τ such as the Cayley map, the exponential map, etc. The retraction map is used to express small discrete changes in the group configuration through unique Lie algebra elements, say $\xi_k = \tau^{-1}(g_k^{-1}g_{k+1})/h$. That is, if ξ_k were regarded as an average velocity between g_k and g_{k+1}, then τ is an approximation to the integral flow of the dynamics. The difference $g_k^{-1}\, g_{k+1} \in G$, which is an element of a nonlinear space, can now be represented by the vector ξ_k. (See [33, 99] for further details.)

Of great importance is the *right trivialized tangent* of the retraction map.

Definition 6.5.5. *Given a retraction map $\tau\colon \mathfrak{g} \to G$, its right trivialized tangent $d\tau_\xi\colon \mathfrak{g} \to \mathfrak{g}$ is defined as the ξ-dependent linear map obtained by composition of the linear maps*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\{\xi\}\times\mathrm{id}"] \arrow[rrr, bend right=15, "\mathrm{d}\tau_\xi"'] & \{\xi\}\times\mathfrak{g} \arrow[r, "T_\xi\tau"] & T_{\tau(\xi)}G \arrow[r, "T_{\tau(\xi)}r_{\tau(\xi)^{-1}}"] & T_eG \equiv \mathfrak{g} \end{tikzcd}$$

where r denotes right translation in the group. Since τ is a local diffeomorphism, all the arrows are linear isomorphisms. We denote the inverse of $d\tau_\xi$ as $d\tau_\xi^{-1}$. Therefore, we can write

$$d\tau_\xi = T_{\tau(\xi)}r_{\tau(\xi)^{-1}} \circ T_\xi\tau \tag{6.34}$$

$$d\tau_\xi^{-1} = (T_\xi\tau)^{-1} \circ T_e r_{\tau(\xi)} = T_{\tau(\xi)}(\tau^{-1}) \circ T_e r_{\tau(\xi)} \tag{6.35}$$

Remark 6.5.6. Omitting the identifications $\mathfrak{g} \equiv \{\xi\} \times \mathfrak{g}$, $\xi \in \mathfrak{g}$, can lead to mismatches when using the definitions above explicitly; for example, if we rewrite equation (6.37) below using (6.35), then the left-hand side would be in $\{\xi\} \times \mathfrak{g}$ while the right-hand side would be in $\{-\xi\} \times \mathfrak{g}$. This should cause no problems if the identifications are made explicit when needed. In any case, (6.37) makes sense as an identity in $\mathfrak{g}$. ◇

Lemma 6.5.7. *(See [130]) Let $g \in G$, $\lambda \in \mathfrak{g}$ and δf denote the variation of a function f with respect to its parameters. Assuming λ is constant, the following identity holds*

$$\delta(Ad_g\, \lambda) = -Ad_g\, [\lambda\, ,\, g^{-1}\delta g],$$

where $[\cdot\,,\,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *denotes the Lie bracket operation or equivalently* $[\xi\,,\,\eta] \equiv ad_{\xi}\eta$*, for given* η*,* $\xi \in \mathfrak{g}$*.*

Lemma 6.5.8. *(See [66]) For each* $\lambda \in \mathfrak{g}$*, the derivative of the map* $\psi_\lambda \colon \mathfrak{g} \to \mathfrak{g}$ *defined by* $\psi_\lambda(\xi) = Ad_{\tau(\xi)}\lambda$ *is given by*

$$D\psi_\lambda(\xi) \cdot \eta = -[Ad_{\tau(\xi)}\lambda\,,\, \mathrm{d}\tau_\xi(\eta)],$$

$\eta \in \mathfrak{g}$.

The lemma above holds not only for retraction maps but also for any smooth map $\tau \colon \mathfrak{g} \to G$.

The following lemma relates the right trivialized tangents at ξ and $-\xi$, as well as their inverses.

Lemma 6.5.9. *(See [66]) For a retraction map* $\tau \colon \mathfrak{g} \to G$ *and any* $\xi, \eta \in \mathfrak{g}$*, the following identities hold:*

$$\mathrm{d}\tau_\xi\, \eta = Ad_{\tau(\xi)}\, \mathrm{d}\tau_{-\xi}\, \eta, \tag{6.36}$$

$$\mathrm{d}\tau_\xi^{-1}\, \eta = \mathrm{d}\tau_{-\xi}^{-1} \left(Ad_{\tau(-\xi)}\, \eta\right). \tag{6.37}$$

Some retraction map choices

a) The exponential map $\exp : \mathfrak{g} \to G$, defined by $\exp(\xi) = \gamma(1)$, where $\gamma : \mathbb{R} \to G$ is the integral curve through the identity of the vector field associated with $\xi \in \mathfrak{g}$ (hence, with $\dot\gamma(0) = \xi$). The most natural example of retraction map is the exponential map. We recall that, for a finite-dimensional Lie group, exp is locally a diffeomorphism and gives rise a natural chart [128]. Then, there exists a neighborhood U of the neutral element $e \in G$ such that $\exp^{-1} : U \to \exp^{-1}(U)$ is a local $C^\infty-$diffeomorphism. A chart at $g \in G$ is given by $\Psi_g := \exp^{-1} \circ \ell_{g^{-1}}$.

The right trivialized derivative and its inverse are

$$\mathrm{dexp}_x\, y = \sum_{j=0}^{\infty} \frac{1}{(j+1)!}\, \mathrm{ad}_x^j\, y,$$

$$\mathrm{dexp}_x^{-1}\, y = \sum_{j=0}^{\infty} \frac{B_j}{j!}\, \mathrm{ad}_x^j\, y,$$

where B_j are the Bernoulli numbers (see [79]). Typically, these expressions are truncated in order to achieve a desired order of accuracy.

b) In general, it is not easy to work with the exponential map. In consequence it will be useful to use a different retraction map. More concretely, the Cayley map (see [33, 79] for further details) will provide us a proper framework in the examples shown along the thesis.

The Cayley map $\mathrm{cay} : \mathfrak{g} \to G$ is defined by $\mathrm{cay}(\xi) = (e - \frac{\xi}{2})^{-1}(e + \frac{\xi}{2})$ and is valid for a general class of *quadratic groups*. The quadratic Lie groups are those defined as

$$G = \left\{ Y \in GL(n, \mathbb{R}) \mid Y^T P Y = P \right\},$$

where $P \in GL(n, \mathbb{R})$ is a given matrix (here, $GL(n, \mathbb{R})$ denotes the general linear group of degree n). $O(n)$ or $SO(n)$ are examples of quadratic Lie groups. The corresponding Lie algebra is

$$\mathfrak{g} = \left\{ \Omega \in \mathfrak{g}l(n, \mathbb{R}) \mid P\Omega + \Omega^T P = 0 \right\}.$$

The right trivialized derivative and inverse of the Cayley map are defined by

$$\begin{aligned} \mathrm{dcay}_x\, y &= (e - \frac{x}{2})^{-1}\, y\, (e + \frac{x}{2})^{-1}, \\ \mathrm{dcay}_x^{-1}\, y &= (e - \frac{x}{2})\, y\, (e + \frac{x}{2}). \end{aligned}$$

Applications to matrix groups: $SO(3)$

We specify the exact form of the Cayley transform for the group $SO(3)$. While we have given more than one general choice for τ, for computational efficiency we recommend the Cayley map since it is simple. In addition, it is suitable for iterative integration and optimization problems since its derivatives do not have any singularities that might otherwise cause difficulties for gradient-based methods. The group of rigid body rotations is represented by 3×3 matrices with orthonormal column vectors corresponding to the axes of a right-handed frame attached to the body. Recall the map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ (see [81] for example). A Lie algebra basis for $SO(3)$ can be constructed as $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$, $\hat{e}_i \in \mathfrak{so}(3)$, where $\{e_1, e_2, e_3\}$ is the standard basis for $\mathbb{R}^3$. Elements $\xi \in \mathfrak{so}(3)$ can be identified with the vector $\omega \in \mathbb{R}^3$ through $\xi = \omega^\alpha\, \hat{e}_\alpha$, or $\xi = \hat{\omega}$. Under such identification the Lie bracket coincides with the standard cross product, i.e., $\mathrm{ad}_{\hat{\omega}}\, \hat{\rho} = \omega \times \rho$, for $\omega, \rho \in \mathbb{R}^3$. Using this identification we have

$$\mathrm{cay}(\hat{\omega}) = I_3 + \frac{4}{4 + \parallel \omega \parallel^2} \left(\hat{\omega} + \frac{\hat{\omega}^2}{2} \right),$$

where I_3 is the 3×3 identity matrix. The linear maps $\mathrm{d}\tau_\xi$ and $\mathrm{d}\tau_\xi^{-1}$ are expressed as the 3×3 matrices

$$\mathrm{dcay}_\omega = \frac{2}{4 + \parallel \omega \parallel^2} (2I_3 + \hat{\omega}), \quad \mathrm{dcay}_\omega^{-1} = I_3 - \frac{\hat{\omega}}{2} + \frac{\omega\, \omega^T}{4}.$$

Conclusions and future research

The closing chapter of this memory is devoted to summarize the contributions of the work. An outlook of the future research is also provided.

Conclusions

Chapter 2 has been devoted to the geometric study of the relationship between higher-order Hamiltonian dynamics and higher-order Lagrangian dynamics. After introducing the notion of Legendre transformation on a Lagrangian submanifold in Definition 2.5.5 we find that both are equivalent. Furthermore, we have given an alternative characterization of the dynamics in the Lagrangian submanifols $\Sigma_L \subset T^*TT^{(k-1)}Q$ in terms of the solution of the higher-order Euler-Lagrange equations in Theorem 2.5.6.

In Chapter 3 we have studied higher-order mechanics from the point of view of the Skinner and Rusk formalism to obtain higher-order Euler-Lagrange equations, higher-order Euler-Poincaré equations and higher-order Lagrange-Poincaré equations. This geometric formalism has permitted to define a presymplectic form and an unique Hamiltonian function; and, in consequence, a global and unique formulation of the dynamics. Also, we have developed an intrinsic formulation of the higher-order variational problem equations subject to constraint depending on higher-order derivatives. The extension of these theories to the natural setting of Lie algebroids has been also developed. In particular these results are given by Equations 3.7, 3.17, 3.27, 3.28, 3.34, 3.35, 3.77, 3.79; Theorems 3.2.1, 3.2.3, 3.3.1, 3.3.3, 3.4.1; Proposition 3.5.5 and Remark 3.2.4.

Also, in Section 3.6 we have studied the geometric description of an underactuated mechanical control systems using the Skinner and Rusk formalism and the results commented before. We have stated the geometric formulation of an optimal control problem of underactuated mechanical system in the natural framework of Lie algebroids. 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 such as the symplecticity and the preservation of the hamiltonian. In particular, these results are given in Equations 3.90 and 3.91, Theorem 3.6.2, Equations 3.93, 3.94 and 3.95, Example 3.6.6, Example 3.6.3 and Example 3.6.10.

In Chapter 4 we study optimal control problems of mechanical systems subject to nonholonomic constraints. In our framework we have implicitly a reduction process, that is; after the geometric procedure applied through this chapter to describe the dynamical equations for the optimal control problem we can reduce the degrees of freedom of the lagrangian which describes the control problem using the nonholonomic constraints. Under some regularity conditions we check that our initial Lagrangian formalism is equivalent to a Hamiltonian formalism for the optimal control problem. We have seen how this framework can be easily extended when instead of working on TQ we consider an arbitrary Lie algebroid. The main results in this Chapter are shown in Equations 4.3, 4.5 and 4.7; Definition 4.3.4; Remarks 4.2.3 and 4.3.7; and Propositions 4.3.1 and 4.3.6.

The aim of Chapter 5 has been devoted to generalize the theory of discrete higher-order Lagrangian mechanics and variational integrators in two directions. First, we have developed variational principles for second-order variational problems on Lie groupoids for mechanical systems with and without second-order constraints, and we have shown how to apply this theory to the construction of variational integrators for optimal control problems of mechanical systems. Secondly, we have shown that Lagrangian submanifolds of a special symplectic groupoids give rise to dynamical second-order equations. Also we study the properties of these systems, including their regularity and reversibility, from the perspective of symplectic and Poisson geometry. We also have developed a theory of reduction and Noether symmetries, and we have studied the relationship between the dynamics and variational principles for these second- order variational problems. Finally we have designed numerical methods for mechanical optimal control problems of total actuated and under-actuated systems. These results are given in Equations 5.20, 5.26, 5.36 and 5.37; Lemma 5.3.4; Remark 5.3.3; Theorems 5.3.2, 5.3.5, 5.3.16, 5.3.19; Propositions 5.3.8 and 5.3.9; and Corollary 5.3.17.

Chapter 6 accounts for new developments regarding geometric variational integrators for higher-order mechanical systems and optimal control applications. We have seen that one possibility to constructing variational integrators for higher-order mechanical systems is to take a Lagrangian function $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ instead of $k+1$-copies of Q, as is usual, since really the discrete variational calculus is not based on the discretization of the Lagrangian itself, but on the discretization of the associated action. First we show that a regular higher-order Lagrangian system has a unique solution for given nearby endpoint conditions using a direct variational proof of existence and uniqueness of the local boundary value problem using a regularization procedure which it results by the replacement of the variational problem with an equivalent one which is regular at $h = 0$. Secondly we define the exact discrete Lagrangian for higher-order Lagrangians and we have shown how the Legendre transformations for this exact discrete lagrangian are related with the second-order Legendre transform and we have shown that if the exact discrete Lagrangian is regular then the continuous Lagrangian is also regular. We have shown how our framework can be seen in the sense of discrete mechanics developed by Marsden and West [131] and we have constructed variational integrators for optimal control problems in this framework. The results are given in Definition 6.2.1, Theorems 6.3.1, 6.3.2, 6.3.4 and 6.3.6, and Corollary 6.3.5.

The results presented in this thesis have been published in

- *Unified formalism for higher-order variational problems and its applications in optimal control,* joint work with P. Prieto-Martínez. International Journal of Geometric Methods in Modern Physics, (2014) DOI: 10.1142/S0219887814500340.

- *Discrete higher-order variational problems with constraints,* joint work with D. Martín de Diego and M. Zuccalli. Journal of Mathematical Physics. Vol 54, 093507 (2013), doi: 10.1063/1.4820817.

- *On Variational Integrators for Optimal Control of Mechanical Control Systems*, joint work with D. Martín de Diego and Marcela Zuccalli. *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* Volume 106, Issue 1, pp 161-171 (2012).

- *Optimal Control of Underactuated Mechanical Systems: A Geometric Approach,* joint work with D. Martín de Diego and Marcela Zuccalli. Journal of Mathematical Physics, Vol 51, 083519 (2010).

- *On the construction of variational integrators for optimal control problems of nonholonomic mechanical systems,* joint work with D. Martín de Diego and Marcela Zuccalli (To appear in Actas del XII Congreso Monteiro, 2014).

- *Optimal control of underactuated mechanical systems with symmetries,* joint work with D. Martín de Diego. Dynamical Systems and Differential Equations, Discrete and Continuous Dynamical Systems, November (2013). Proceedings of the 9th AIMS international conference, Orlando, Florida, USA. 149 – 158.

- *Quasivelocities and optimal contol of underactuated mechanical systems.* joint work with D. Martín de Diego. Proceedings of the American Institute of Physics, Geometry and Physics, 1260, 133-140, (2011).

- *On the Geometry of Higher-Order Problems on Lie Groups.* joint work with David Martín de Diego. (Submitted to SIAM Journal of Control and Optimization). July 2011. Available at http://arxiv.org/abs/1104.3221.

- *Optimal Control and higher-order discrete mechanics for systems with symmetries,* joint work with F. Jiménez and D. Martín de Diego. (Submitted to Journal of Computational Dynamics) June 2013. Available at http://arxiv.org/abs/1209.6315.

Also we are working in the final versions of the following works which involve some results given in this thesis unpublished yet:

- *Lagrangian submanifolds generating second-order dynamics on Lie algebroids,* with David Martín de Diego. (Preprint available for distribution 22 p.)

- *Higher-order problems on Lie groupoids: Optimal control applications,* with D. Martín de Diego (Preprint available for distribution 28 p.)

- *Geometry of optimal control problems of nonholonomic mechanical systems,* with A. Bloch, R. Gupta and D. Martín de Diego. (Preprint available for distribution 18 p.)

- *Discrete mechanics and optimal control: An analysis for higher-order mechanical systems* with S. Ferraro and D. Martín de Diego. (Preprint available for distribution 19 p.)

In addition, the results contained in this thesis have been presented in the following international meetings:

- *Geometry of optimal control problems of nonholonmic mechanical systems* (poster) *deLeónfest.* Madrid, Spain (2013).
- *Optimal control of nonholonomic mechanical systems* (poster) *8th International Young Researchers Workshop on Geometry, Mechanics and Control.* Barcelona, Spain (2013).
- *Optimal control of nonholonomic mechanical systems* (talk) *VII Summer School on Geometry, Mechanics and Control.* La Cristalera, Madrid, Spain (2013).
- *Optimal control of nonholonomic mechanical systems* (talk) *XII Congreso Dr. Antonio Monteiro.* Bahia Blanca, Argentina (2013).
- *On the plate ball optimal control problem* (talk) *IV Congreso de Matemática Aplicada, Computacional e Industrial.* Ciudad Autónoma de Buenos Aires, Argentina (2013).
- *Lagrangian submanifolds generating second-order Lagrangian mechanics on Lie algebroids* (invited talk) *XV Encuentro de Invierno de Geometría, Mecánica y Control.* Zaragoza, Spain (2013).
- *On the geometry of higher-order mechanical systems on Lie groups* (invited talk) *Seminar on Geometry of the Universitat Politècnica de Catalunya.* Facultad de Matemática y Estadistica, UPC, Barcelona, Spain (2013).
- *On the variational discretization of optimal control problems* (invited talk) *XXI Fall workshop on Geometry and Physics.* Burgos, Spain (2012).
- *Optimal control of underactuated mechanical systems with symmetries* (invited talk) *Focus program on geometry, mechanics and dynamics: the Legacy of Jerry Marsden*, Fields Institute, Toronto, Canada (2012).
- *On the geometry of mechanical control systems on Lie groups* (talk) *9th. AIMS Conference on Dynamical systems, differential equations and applications.* Orlando, Florida, USA (2012).
- *On the geometry of discrete higher-order Lagrangian problems* (talk and poster) *6th. Summer school on geometry, mechanics and control.* Miraflores de la Sierra, Madrid, Spain (2012).
- *Groupoids and Mechanics on Lie groupoids* (talk) *Geometry seminar*, Maths department, Universidad Nacional de La Plata, Argentina (2011).
- *Sobre la geometría de las ecuaciones de Euler-Poincaré de orden superior* (talk) *Seminar on Geometry and Physics*, Universidad Complutense de Madrid, Spain (2011).
- *Second order Lagrangian mechanics on Lie algebroids* (talk and poster) *European Mathematical Society and Spanish Royal Academy of Science, joint weekend.* Bilbao, Spain (2011).

- *On variational problems on Lie groups* (talk) *Primer Encuentro de Jovenes Investigadores en Matemticas Universidad de La Laguna (PEJIM 2011).* La Laguna, Tenerife, Spain (2011).
- *Quasivelocities and optimal control of mechanical systems* (talk) *Seminar of Geometry,* Departamento de Matemática Fundamental, Universidad de La Laguna, Spain (2011).
- *Discrete variational problems on Lie groupoids* (talk) *Congress of Young researchers of the Spanish Royal Mathematical Society.* Soria, Castilla y León, Spain (2011).
- *Higher Order Mechanics on Lie Algebroids* (poster) *XX International Workshop on Geometry and Physics.* ICMAT, Madrid, Spain (2011).
- *On the Geometry of higher order problems on Lie groups* (poster) *Poisson Geometry and applications*, Figueira da Foz, Portugal (2011).
- *An introduction to higher-order mechanics on Lie algebroids* (talk) *Meeting of Geometry, Mechanics and Control,* ICMAT, Madrid, Spain (2011).
- *Discrete second order mechanics on Lie groupoids* (poster) 5^{th} *International Summer School on GMC*, La Cristalera, Madrid, Spain (2011).
- *A Variational and Geometric Approach for the Second Order Euler-Poincaré Equations* (talk) *XIII Winter Meeting on Geometry, Mechanics and Control and Thematic day on Fields*, Zaragoza, Spain (2011).
- *Optimal Control of Underactuated Mechanical Systems on Lie Groups* and *Higher Order Discrete Vakonomic Mechanics for Optimal Control of Underactuated Systems* (posters) *Second Iberoamerican Meeting on Geometry, Mechanics and Control, in Honor of Hernán Cendra*, Centro Atómico de San Carlos de Bariloche, Argentina (2011).
- *Optimal Control of Underactuated Mechanical Systems on Lie Groups* (talk) *5th Young Researchers Workshop on Geometry, Mechanics and Control*, Universidad de La Laguna, Tenerife, Spain (2010).
- *Variational Integrators for Optimal Control of Mechanical Systems* (poster) *5th Young Researchers Workshop on Geometry, Mechanics and Control*, Universidad de La Laguna, Tenerife, Spain (2010).

Future work

Relation between the classical higher-order Legendre transformation and the higher-order Legendre transformation on Σ_L

In Chapter 2 we have introduced the notion of Legendre transformation on a Lagrangian submanifold Σ_L generated by a higher-order Lagrangian $L : T^{(k)}Q \to \mathbb{R}$ (see definition 2.5.5). An open question is the intrinsic relation between the Legendre transformation that we have defined in 2.5.5 $\mathbb{F}L : \Sigma_L \to T^*(T^{(k-1)}Q)$ and the classical Legendre transformation $Leg_L : T^{(2k-1)}Q \to T^*(T^{(k-1)}Q)$. More explicitly, given a regular Lagrangian $L : T^{(k)}Q \to \mathbb{R}$, is there an intrinsic map, $\varphi_L : T^{(2k-1)}Q \to \Sigma_L$ such that $\mathbb{F}L \circ \varphi_L = Leg_L$?. A description in local coordinates is easy.

Lagrangian submanifolds generating second-order dynamics on Lie algebroids

Now we will give some ideas about how to study the dynamics of second-order mechanical systems on Lie algebroids as an alternatively characterization of the results given in Chapter 2.

Let E be a Lie algebroid over M with bundle projection denoted by $\tau_E : E \to M$. We consider local coordinates (x^i) where $i = 1, \ldots, m$ and let $\{e_A\}$ be a basis of sections of τ_E. Also we will denote by $\{e^A\}$ its dual basis being a basis of $\tau_{E^*} : E^* \to M$. The basis $\{e_A\}$ and $\{e^A\}$ induces local coordinates (x^i, y^A) and (x^i, p_A) on E and E^* respectively.

Following [111] one can construct the Tulczyjew's triple on Lie algebroids defining the two vector bundles isomorphisms, namely $\mathcal{A}_E : \mathcal{T}^{\tau_{E^*}}E \to (\mathcal{T}^{\tau_E}E)^*$, $\quad \flat_{E^*} : \mathcal{T}^{\tau_{E^*}}E \to (\mathcal{T}^{\tau_{E^*}}E)^*$. Then applying this construction to the case when the Lie algebroid is the E-tangent bundle to a Lie algebroid, that is, when E is $\mathcal{T}^{\tau_E}E$, one can obtain the Tulczyjew's triple that we need to use to obtain the dynamics for second-order systems. Locally, the isomorphism $\mathcal{A}_{\mathcal{T}^{\tau_E}E}$ is given by

$$\mathcal{A}_{\mathcal{T}^{\tau_E}E}(x^i, y^A, p_A, \bar{p}_A; q^A, \bar{q}^A; l_A, \bar{l}_A) = (x^i, y^A, q^A, \bar{q}^A; l_A + \mathcal{C}_{AB}^C p_C q^B, \bar{l}_A, p_A, \bar{p}_A) \tag{6.38}$$

where $\mathcal{C}_{AB}^C$ are the constant structure of the Lie algebroid E and $(x^i, y^A, p_A, \bar{p}_A; q^A, \bar{q}^A; l_A, \bar{l}_A)$ are local coordinates on the Lie algebroid $\mathcal{T}^{\tau_{E^*}^{(1)}}(\mathcal{T}^{\tau_E}E)$ where $\tau_{E^*}^{(1)} : (\mathcal{T}^{\tau_E}E)^* \to E^*$(see [111]).

Consider a Lagrangian function $L : E^{(2)} \to \mathbb{R}$ where $E^{(2)}$ is the set of admissible elements and also consider the following submanifold of $(\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E))^*$,

$$\Sigma_{L,E^{(2)}} = \{\mu \in (\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E))^* \mid i_\Delta^* \mu = d^N L\}$$

where $\Delta = (\rho_1)^{-1}(TN)$, $i_\Delta : (\rho_1)^{-1}(TN) \subset E \to \mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E)$, is the canonical inclusion, $\rho_1 : \mathcal{T}^{\tau_E}E \to TE$ is the anchor map of the Lie algebroid $\mathcal{T}^{\tau_E}E$ over E and $N = E^{(2)}$. Here d^N denotes the differential operator of the Lie algebroid $\Delta := (\rho_1)^{-1}(TN)$ over $E^{(2)}$ (see 3.5.7).

If we take induced coordinates (x^i, y^A, v^A, z^A) on $\mathcal{T}^{\tau_E}E$ the set of admissible elements is characterized by the condition $y^A = v^A$. Then, the induced coordinates on $E^{(2)}$ are (x^i, y^A, z^A). Locally, the submanifold $\Sigma_{L,E^{(2)}}$ is characterized by the equations

$$\mu_A = \rho_A^i \frac{\partial L}{\partial x^i}, \tag{6.39}$$

$$\bar{\mu}_A + \check{\mu}_A = \frac{\partial L}{\partial y^A}, \tag{6.40}$$

$$\widetilde{\mu}_A = \frac{\partial L}{\partial z^A}, \tag{6.41}$$

where $\{\mu_A, \bar{\mu}_A, \check{\mu}_A, \widetilde{\mu}_A\}$ are sections of $(\mathcal{T}^{\tau_E^{(1)}}(\mathcal{T}^{\tau_E}E))^*$.

Theorem 6.5.10. *$S_L = \mathcal{A}^{-1}_{\mathcal{T}^{\tau_E}E}(\Sigma_{L,E^{(2)}}) = \{x \in \mathcal{T}^{\tau^{(1)}_{E^*}}(\mathcal{T}^{\tau_E}E) \mid \mathcal{A}_{\mathcal{T}^{\tau_E}E}(x) \in \Sigma_{L,E^{(2)}}\} \subset \mathcal{T}^{\tau^{(1)}_{E^*}}(\mathcal{T}^{\tau_E}E)$ is a Lagrangian submanifold of $\mathcal{T}^{\tau^{(1)}_{E^*}}(\mathcal{T}^{\tau_E}E)$.*

Using the expression of the Tulczyjew's isomorphism one can show that equations (6.39), (6.40) and (6.41) are equivalent to

$$l_A + \mathcal{C}^C_{AB}p_C q^B = \rho^i_A \frac{\partial L}{\partial x^i}, \tag{6.42}$$

$$\bar{l}_A + p_A = \frac{\partial L}{\partial y^A}, \tag{6.43}$$

$$\bar{p}_A = \frac{\partial L}{\partial z^A}, \tag{6.44}$$

and after some computations one can show that the *second-order Euler-Lagrange equations on Lie algebroids* for $L : E^{(2)} \to \mathbb{R}$ are given by,

$$\begin{aligned} \rho^i_A y^A &= \frac{dx^i}{dt} \\ 0 &= \frac{d^2}{dt^2}\frac{\partial L}{\partial z^A} + \mathcal{C}^C_{AB}y^B\frac{d}{dt}\left(\frac{\partial L}{\partial z^A}\right) - \frac{d}{dt}\frac{\partial L}{\partial y^A} - \mathcal{C}^C_{AB}y^B\left(\frac{\partial L}{\partial y^A}\right) + \rho^i_A\frac{\partial L}{\partial x^i}. \end{aligned}$$

Higher-order geometric Hamilton Jacobi theory and optimal control applications

In [47] and [48] we have worked in the higher-order geometric formulation of the Hamilton-Jacobi theory. With the idea to do further developments in this geometric Hamilton-Jacobi theory we pretend to extend this theory in the Skinner and Rusk formalism to the case of higher-order mechanical systems with higher-order constraints and study the applications to optimal control problems of underactuated mechanical systems. Also we want to apply the results given in [48] to optimal control problems of fully actuated mechanical systems.

Taking into account the developments of Chapter 4 we would like to analyze the possibility of study nonholonomic optimal control problems adapting the results of [39], [47] and [147] to this class of optimal control problems. Finally, we want to construct symplectic and variational integrators to study optimal control problems in the framework of the higher-order geometric Hamilton-Jacobi theory using and extending the ideas given in [148].

Construction of variational integrators for nonholonomic optimal control problems

In Chapter 4 we have seen as an optimal control problem of a nonholonomic system can be seen as a Hamiltonian system on $T^*\mathcal{D}$. In this sense, one can use standard methods for symplectic integration as symplectic Runge-Kutta methods, collocation methods, Störmer-Verlet, symplectic Euler methods, etc; developed and studied in [106], [107], [108], [159], [160] for example, to simulate nonholonomic control problems from the Hamiltonian point of view.

Also, we would like to build variational integrators as an alternative way to construct integration schemes for this kind of optimal control problems following the results given in

Section 4.2. In this sense, recall that in the continuous case we have considered a Lagranian function $L : \mathcal{D}^{(2)} \to \mathbb{R}$. Since the space $\mathcal{D}^{(2)}$ is a subset of $T\mathcal{D}$ we can discretize the tangent bundle $T\mathcal{D}$ by the cartesian product $\mathcal{D} \times \mathcal{D}$. Therefore, our discrete variational approach for optimal control problems of nonholonomic mechanical systems will be determined by the construction of a discrete Lagrangian $L_d : \mathcal{D}_d^{(2)} \to \mathbb{R}$ where $\mathcal{D}_d^{(2)}$ is the subset of $\mathcal{D} \times \mathcal{D}$ locally determined by imposing the discretization of the constraint $\dot{q}^i = \rho^i_A(q) y^A$, for instance we can consider

$$\mathcal{D}_d^{(2)} = \left\{ (q_0^i, y_0^A, q_1^i, y_1^A) \in \mathcal{D} \times \mathcal{D} \left| \frac{q_1^i - q_0^i}{h} = \rho^i_A \left(\frac{q_0 + q_0}{2} \right) \left(\frac{y_0^A + y_1^A}{2} \right) \right. \right\} .$$

Now the system is adequate for the application of discrete variational methods for constrained systems.

Conclusiones

El capítulo final de esta memoria consiste en enumerar las contribuciones más relevantes de este trabajo.

En el Capítulo 2 nos hemos centrado en estudiar la relación entre los sistemas dinámicos Hamiltonianos de orden superior y los sistemas dinámicos Lagrangianos de orden superior. Después de introducir la noción de transformación de Legendre en una subvariedad Lagrangiana en la definición 2.5.5, hemos probado tal equivalencia. Más aún, hemos dado una caracterización alternativa de la dinámica de orden superior en la subvariedad Lagrangiana $\Sigma_L \subset T^*TT^{(k-1)}Q$ en términos de las soluciones de las ecuaciones de Euler-Lagrange de orden superior en el teorema 2.5.6.

En el Capítulo 3 hemos estudiado la mecánica de orden superior en el formalismo de Skinner y Rusk para obtener las ecuaciones de Euler-Lagrange de orden superior, Euler-Poincaré de orden superior y Lagrange-Poincaré de orden superior. Este formalismo geométrico nos ha permitido definir una forma presimpléctica y una función Hamiltoniana, y en consecuencia, una formulación global y única de la dinámica. También, hemos desarrollado una formulación intrínseca para las ecuaciones provenientes de problemas variacionales de orden superior sujetos a ligaduras de orden superior. La extensión de estas teorías al marco de algebroids de Lie ha sido también desarrollada. En particular, estos resultados fueron dados en: Ecuaciones 3.7, 3.17, 3.27, 3.28, 3.34, 3.35, 3.77, 3.79; Teoremas 3.2.1, 3.2.3, 3.3.1, 3.3.3, 3.4.1; Proposición 3.5.5 y Remark 3.2.4.

También en la Sección 3.6 hemos estudiado la descripción geométrica para tratar un problema de control óptimo para sistemas mecánicos infractuados utilizando la formulación de Skinner y Rusk desarrollada en las secciones previas de este capítulo y los resultados comentados anteriormente. Hemos establecido la formulación geométrica de un problema de control óptimo para un sistema mecánico infractuado en el marco de algebroides de Lie. Este formalismo geométrico dió lugar a una versión intrínseca de las ecuaciones diferenciales que dan lugar a las trayectorias que resuelven el problema de control óptimo y permiten detectar propiedades geométricas de preservación tales como la preservación de la forma y sección simpléctica proveniente de los teoremas de simplecticidad dados en este capítulo y la preservación de la función Hamiltoniana. En particular los resultados obtenidos aquí están contenidos en: Ecuaciones 3.90 y 3.91, Teorema 3.6.2, Ecuaciones 3.93, 3.94 y 3.95, Ejemplo 3.6.6, Ejemplo 3.6.3 y Ejemplo 3.6.10.

En el Capítulo 4 hemos estudiado problemas de control óptimo para sistemas mecánicos con ligaduras noholonomas. En nuestro formalismo tenemos definido implícitamente un proceso de reducción, esto es, luego de un proceso geométrico aplicado a lo largo del capítulo para describir las ecuaciones que describen las trayectorias optimales de un problema de control optimo para un sistema mecánico sujeto a ligaduras noholonomas, podemos reducir los grados de libertad del sistema Lagrangiano, en cuestión, para reducir las variables del problema y luego reconstruir la solución. Bajo ciertas condiciones de regularidad, encontramos que nuestro formalismo Lagrangiano inicial es equivalente a uno Hamiltoniano para resolver el problema de control óptimo. Además hemos extendido este formalismo a uno más general en el contexto de un algebroide de Lie arbitrario. Los principales resultados de este capítulo están dados en: Ecuaciones 4.3, 4.5 y 4.7; Definición 4.3.4; Remarks 4.2.3 y 4.3.7; y Proposiciones 4.3.1 y 4.3.6.

El principal objetivo del Capítulo 5 ha sido la generalización de la teoría de sistemas mecánicos discretos de orden superior e integradores variacionales en dos direcciones principales. Primero, hemos establecido y desarrollado principios variacionales para sistemas mecánicos de orden dos en grupoides de Lie y hemos visto cómo aplicar esta teoría a la construcción de integradores variacionales para problemas de control óptimo de sistemas mecánicos totalmente actuados e infractuados. En segundo lugar, hemos probado que una subvariedad Lagrangiana de un particular grupoide simpléctico (el grupoide cotangente) da lugar a las ecuaciones que describen la dinámica discreta de orden dos. También hemos estudiado las propiedades de estos sistemas discretos, incluyendo la regularidad de ellos y reversivilidad desde el punto de vista de la geometría simpléctica y de Poisson. Hemos desarrollado una teoría de reducción mediante simetrias de Noether. Alguno de los resultados obtenidos en este capítulo pueden verse en: Ecuaciones 5.20, 5.26, 5.36 y 5.37; Lema 5.3.4; Remark 5.3.3; Teoremas 5.3.2, 5.3.5, 5.3.16, 5.3.19; Proposiciones 5.3.8 y 5.3.9; y Corolario 5.3.17.

El Capítulo 6 de esta memoria da lugar a la construcción de integradores variacionales para sistemas mecánicos de orden superior y sus aplicaciones en la teoría de control óptimo. En este capítulo hemos visto que hay otra posibilidad más general de construir integradores variacionales para sistemas mecánicos de orden superior construyendo un Lagrangiano discreto $L_d : T^{(k-1)}Q \times T^{(k-1)}Q \to \mathbb{R}$ en lugar de definirlo en $k+1$ copias de Q, como es habitual, dado que realmente el cálculo variacional discreto no esta basado en la discretización del Lagrangiano sino en la discretización de la acción asociada al principio variacional. Primero, hemos probado que un sistema Lagrangiano de orden superior para un Lagrangiano regular tiene una única solución, para condiciones iniciales dadas suficientemente cercanas, usando una prueba puramente variacional, sin la necesidad de utilizar la teoría estandar de ecuaciones diferenciales ordinarias, para probar la existencia y unicidad de soluciones para el problema de condiciones de borde. Realizamos un proceso de regularización que resulta de reemplazar el principio variacional en cuestión por uno que es regular en $h = 0$. En segundo lugar, definimos el Lagrangiano discreto exacto para sistemas Lagrangianos de orden superior y relacionamos las transformaciones de Legendre de éste con las transformación de Legendre estandar para sistemas lagrangianos de orden superior. Además, probamos que el Lagrangiano discreto exacto es regular si y sólo si el Lagrangiano continuo es regular. Más aún, nosotros hemos visto que nuestro formalismo puede ser visto en el marco de la teoría de Mecánica Discreta desarrollada por Marsden y West [131] y hemos contruido integradores variacionales para problemas de control óptimo en el formalismo comentado anteriormente. Los principales resultados de este capítulo estan dados en Definición 6.2.1, Teoremas 6.3.1, 6.3.2, 6.3.4 y 6.3.6; y Corolario 6.3.5.

Los resultados presentados en esta tesis han sido publicados en las siguientes revistas de investigación:

- L. Colombo, P. Prieto Martínez, *Unified formalism for higher-order variational problems and its applications in optimal control.* International Journal of Geometric Methods in Modern Physics, Vol 11, n4 (2014).
- L. Colombo, D. Martín de Diego, M. Zuccalli, *Discrete higher-order variational problems with constraints.* Journal of Mathematical Physics. Vol 54, 093507 (2013).
- L. Colombo, D. Martín de Diego, M. Zuccalli, *On Variational Integrators for Optimal Control of Mechanical Control Systems.* Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matemáticas, Volume 106, Issue 1, pp 161-171 (2012).
- L. Colombo, D. Martín de Diego, M. Zuccalli, *Optimal Control of Underactuated Mechanical Systems: A Geometric Approach.* Journal of Mathematical Physics, Vol 51, 083519 (2010).
- L. Colombo, D. Martín de Diego, M. Zuccalli, *On the construction of variational integrators for optimal control problems of nonholonomic mechanical systems.* Aceptado en Actas del XII Congreso Monteiro, 2014.
- L. Colombo, D. Martín de Diego, *Optimal control of underactuated mechanical systems with symmetries.* Dynamical Systems and Differential Equations, Discrete and Continuous Dynamical Systems, November (2013). Proceedings of the 9th AIMS international conference, Orlando, Florida, USA. 149 – 158.
- L. Colombo, D. Martín de Diego, *Quasivelocities and optimal contol of underactuated mechanical systems.* Proceedings of the American Institute of Physics, Geometry and Physics, 1260, 133-140, (2011).
- L. Colombo, D. Martín de Diego, *On the Geometry of Higher-Order Problems on Lie Groups.* (Sometido a SIAM journal of Control and Optimization). Julio 2011. Artículo disponible en `http://arxiv.org/abs/1104.3221`.
- L. Colombo, F. Jiménez, D. Martín de Diego, *Optimal Control and higher-order discrete mechanics for systems with symmetries.* (Sometido a Journal of Computational Dynamics) Junio 2013. Artículo disponible en `http://arxiv.org/abs/1209.6315`.

También, estamos trabajando en las versiones finales de los siguientes trabajos, que han resultado del estudio realizado en esta tesis y aún no se han enviado para su publicación:

- L. Colombo, D. Martín de Diego, *Lagrangian submanifolds generating second-order dynamics on Lie algebroids.* (Preprint disponible para distribución privada 22 p.)
- L. Colombo, D. Martín de Diego, *Higher-order problems on Lie groupoids: Optimal control applications,* (Preprint disponible para distribución privada 28 p.)

- A. Bloch, L. Colombo, R. Gupta, D. Martín de Diego *Geometry of optimal control problems of nonholonomic mechanical systems.* (Preprint disponible para distribución privada 18 p.)
- L. Colombo, S. Ferraro, D. Martín de Diego, *Discrete mechanics and optimal control: An analysis for higher-order mechanical systems.* (Preprint disponible para distribución privada 19 p.)

Los resultados que aparecen en esta memoria han sido presentado en los siguientes workshops y congresos nacionales e internacionales:

- *Geometry of optimal control problems of nonholonmic mechanical systems* (poster) *deLeónfest.* Madrid, España (2013).
- *Optimal control of nonholonomic mechanical systems* (poster) *8th International Young Researchers Workshop on Geometry, Mechanics and Control.* Barcelona, España (2013).
- *Optimal control of nonholonomic mechanical systems* (ponencia) *VII Summer School on Geometry, Mechanics and Control.* La Cristalera, Madrid, España (2013).
- *Optimal control of nonholonomic mechanical systems* (ponencia) *XII Congreso Dr. Antonio Monteiro.* Bahia Blanca, Argentina (2013).
- *On the plate ball optimal control problem* (ponencia) *IV Congreso de Matemática Aplicada, Computacional e Industrial.* Ciudad Autónoma de Buenos Aires, Argentina (2013).
- *Lagrangian submanifolds generating second-order Lagrangian mechanics on Lie algebroids* (charla invitada) *XV Encuentro de Invierno de Geometría, Mecánica y Control.* Zaragoza, España (2013).
- *On the geometry of higher-order mechanical systems on Lie groups* (charla invitada) *Seminar on Geometry of the Universitat Politècnica de Catalunya.* Facultad de Matemática y Estadistica, UPC, Barcelona, España (2013).
- *On the variational discretization of optimal control problems* (charla invitada) *XXI Fall workshop on Geometry and Physics.* Burgos, España (2012).
- *Optimal control of underactuated mechanical systems with symmetries* (charla invitada) *Focus program on geometry, mechanics and dynamics: the Legacy of Jerry Marsden*, Fields Institute, Toronto, Canada (2012).
- *On the geometry of mechanical control systems on Lie groups* (ponencia) *9th. AIMS Conference on Dynamical systems, differential equations and applications.* Orlando, Florida, USA (2012).
- *On the geometry of discrete higher-order Lagrangian problems* (ponencia y poster) *6th. Summer school on geometry, mechanics and control.* Miraflores de la Sierra, Madrid, España (2012).

- *Groupoids and Mechanics on Lie groupoids* (ponencia) *Geometry seminar*, Maths department, Universidad Nacional de La Plata, Argentina (2011).
- *Sobre la geometría de las ecuaciones de Euler-Poincaré de orden superior* (ponencia) *Seminar on Geometry and Physics*, Universidad Complutense de Madrid, España (2011).
- *Second order Lagrangian mechanics on Lie algebroids* (ponencia y poster) *European Mathematical Society and Spanish Royal Academy of Science, joint weekend.* Bilbao, España (2011).
- *On variational problems on Lie groups* (ponencia) *Primer Encuentro de Jovenes Investigadores en Matemticas Universidad de La Laguna (PEJIM 2011).* La Laguna, Tenerife, España (2011).
- *Quasivelocities and optimal control of mechanical systems* (ponencia) *Seminar of Geometry,* Departamento de Matemática Fundamental, Universidad de La Laguna, España (2011).
- *Discrete variational problems on Lie groupoids* (ponencia) *Congress of Young researchers of the Spanish Royal Mathematical Society.* Soria, Castilla y León, España (2011).
- *Higher Order Mechanics on Lie Algebroids* (poster) *XX International Workshop on Geometry and Physics.* ICMAT, Madrid, España (2011).
- *On the Geometry of higher order problems on Lie groups* (poster) *Poisson Geometry and applications*, Figueira da Foz, Portugal (2011).
- *An introduction to higher-order mechanics on Lie algebroids* (ponencia) *Meeting of Geometry, Mechanics and Control,* ICMAT, Madrid, España (2011).
- *Discrete second order mechanics on Lie groupoids* (poster) 5^{th} *International Summer School on GMC*, La Cristalera, Madrid, España (2011).
- *A Variational and Geometric Approach for the Second Order Euler-Poincaré Equations* (ponencia) *XIII Winter Meeting on Geometry, Mechanics and Control and Thematic day on Fields*, Zaragoza, España (2011).
- *Optimal Control of Underactuated Mechanical Systems on Lie Groups* and *Higher Order Discrete Vakonomic Mechanics for Optimal Control of Underactuated Systems* (posters) *Second Iberoamerican Meeting on Geometry, Mechanics and Control, in Honor of Hernán Cendra*, Centro Atómico de San Carlos de Bariloche, Argentina (2011).
- *Optimal Control of Underactuated Mechanical Systems on Lie Groups* (ponencia) *5th Young Researchers Workshop on Geometry, Mechanics and Control*, Universidad de La Laguna, Tenerife, España (2010).
- *Variational Integrators for Optimal Control of Mechanical Systems* (poster) *5th Young Researchers Workshop on Geometry, Mechanics and Control*, Universidad de La Laguna, Tenerife, España (2010).

Agradecimientos

Parafraseando a Paulo Cohelo, "siempre hay que saber cuándo una etapa llega a su fín. Cerrando ciclos, cerrando puertas, terminando capítulos; no importa el nombre que le demos, lo que importa es dejar en el pasado los momentos de la vida que ya se han acabado." Aunque ninguna cuestion queda nunca concluida hasta que se concluye bien, hoy siento que mil rutas se apartan del fin elegido, pero sin embargo hay una que llega a él. Muchas emociones se juntan: alegría, euforia, tristeza, entusiasmo...pero la mas fuerte es el agradecimiento, ya que esta etapa en mi vida no hubiera sido tan enriquecedora si no os hubiera conocido y compartido momentos que jamás olvidaré junto a vosotros.

Hoy me viene el recuerdo de cuando a principios del año 2008 vi en el hall de entrada del departamento de matemáticas de la Universidad Nacional de La Plata los posters que anunciaban el primer Congreso Iberoamericano de Geometría, Mecánica y Control realizado en Santiago de Compostela y la segunda Escuela de Verano de Geometría, Mecánica y Control en La Palma ambos organizados por la red de Geometría, Mecánica y Control (GMC network). Gracias a la financiación de la red GMC y i-MATH Consolider para asistir al congreso iberoamericano, comenzó mi camino por el mundo jugando en un equipo de primera llamado ICMAT, que me trajo hasta aquí luego de un difícil y tortuoso fichaje, y, que hoy concluye, quizás no con una goleada como el partido Argentina-España $(4-1)$ en septiembre de 2010 pero si con una dulce victoria dada por un gol como aquel que todos conocen de la mano de Maradona a los ingleses en el mundial de México del 86 al haber conseguido un plaza de profesor e investigador post-doctoral en una universidad de prestigio y excelencia como lo es la Universidad de Michigan, y tener, en mi humilde opinion, una buena tesis doctoral de la cual estoy muy orgulloso.

Esta tesis no hubiera podido llevarse a cabo sin la financiacion de los proyecto *MICINN-MTM-2010-21186-C02-01, MTM 2011-15725-E, ICMAT Severo Ochoa Project SEV-2011-0087, IRSES-project "Geomech-246981"* y la financiación del Programa JAE del Consejo Superior de Investigaciones Científicas gracias a la otorgación de la beca doctoral JAE-Predoc.

Y hablando de tesis... en estos años concluí que una tesis es como una novia. Uno se enamora de ella, discute y se pelea, la odia a veces, aunque en el fondo la ame; llora por ella y hay veces que la quiere dejar por ser insoportable e inentendible, aunque la mayoria de las veces se esmera por sacar lo mejor de ella, entenderla e ir por el mismo camino evolucionando juntos. Ese camino de evolución y entendimiento no hubiera sido posible sin la ayuda de mi amigo y director de tesis David Martín. Sin él nada de esto hubiera sido posible. Su invaluable (e inexplicable con palabras) ayuda y comprensión ha logrado sacar lo mejor de mí en los momentos más difíciles dandome confianza para seguir adelante. Gracias por apostar en mi, ser mi guía, maestro, amigo (y todo lo que esas cinco letras significan) y haber luchado conmigo desde el principio para que pueda hacer la tesis en el ICMAT (cuando no era “la mayor infraestructura europea en matemáticas” ni excelencia Severo Ochoa).

Mi familia académica esta formada por varias grandes personalidades, a los que quiero particularmente agradecer por haberme sufrido, ayudado y generado momentos en mi vida que jamáz olvidaré. Esta familia tiene un abuelo y como en la mayoria de las familias, los abuelos suelen tener una gran sabidura por el transcurso de sus caminos y ejercen de patriarca de ella. La sabiduría del patriarca nos ayuda, guía e incentiva, al menos en mi caso. Manuel de León ejerce para mi esa figura de patriarca. Él es una persona que me inspira, incentiva, hace que sienta que confia y apuesta en mí. Su influencia en mi camino me ha ayudado mucho en este tiempo y quiero agradecercelo con estas palabras.

¡ Y que hubiera sido de mi estos años sin los buenos momentos que pasé con mis hermanos académicos! Gracias Cédric, Fernando, Miguel y Marta por tantas alegrias compartidas, por sus sonrisas y palabras de aliento cuando mas lo necesité y aceptarme con mis malas y buenas costumbres,... y ¡aguantar mis ronquidos y olor a tabaco las veces que hemos compartido habitación! A los mayores, gracias por abrirme el camino, protegerme y guiarme, y a los más pequeños, gracias por enseñarme tanto con su paciencia e ilusión. En este párrafo también quiero agracederle a María, quien al ser la mayor del grupo jóven de Geometría, Mecánica y Control del ICMAT nos ha enseñado que con sacrificio y perseverancia todos podemos triunfar en la vida. Gracias María por compartir esos momentos dificiles conmigo y ayudarme a mirar las cosas de otra manera cuando lo necesité para seguir adelante con esta tesis.

Edith Padrón y Juan Carlos Marrero han sido una influencia, ejemplo y compañia muy importante en el transcurso de estos años. Su comprensión y entendimiento me ha ayudado en medio de risas, cañas, excursiones, comidas y cenas. ¡Muchas gracias por todo lo bueno que me han brindado durante este tiempo!

GMC Network, para mi, es una gran familia y sus congresos y eventos significan la oportunidad de encontrarme durante todo el año con los membros de esa gran familia, donde ademas de temas académicos uno puede disfrutar de ellos como personas y pasar muy buenos momentos. Allí he encontrado grandes amigos y colegas. Gracias Miguel R.O, Narciso R.R, Miguel M.L, Xavier, Miguel T, Mónica A, Eduardo, Silvia, Javier de Luca, Diana, Modesto, Pepín, Irina, Ignazio, David I, Elisa, Lucia, Cristina y Romero por haber compartido conmigo estos años y ser buenos amigos y colegas. A Marco Zambón por aceptar ser mi tutor.

Quisiera agradecer a mis colaboradores en esta tesis por haber compartido conmigo este proceso y brindarme sus conocimientos para aprender de ellos y ayudarme a explotar mis cualidades al máximo. Muchas gracias Anthony Bloch, Pere Daniel Prieto Martínez y Marcela Zuccalli. En especial quiero agradecerle a mi amigo y colaborador Sebastián Ferraro, ya que gracias a su ayuda esta tesis tiene el toque que necesitaba para sentirme orgulloso de ella. Muchas gracias por estar simpre disponible para ayudarme con mis dudas matemáticas y de la vida. Siempre recordaré los buenos momentos compartidos juntos en estos años y las charlas acerca de mi nuevo comienzo de vida como un hombre casado.

También tuve el placer de tener varios contactos con gente del área de Geometría, Mecánica y Control, que han fructificado mis conocimientos como Hernán Cendra, Marin Kobilarov, Luis Garcia Naranjo, Francois Gay-Balmaz, Cesare Tronci, Esther Bonnet, Jair

Koiller, Paula Balseiro, Alejandro Cabrera y Sergio Grillo.

A los dos grupos de investigación que me han acogido durante mis estancias de investigación. He pasado muy buenos momentos y aprendido mucho durante los 4 meses en el Imperial College y las tantas visitas a la Universidad Nacional del Sur en Argentina. Especialmente quiero agradecer a los IP de los repectivos grupos: Darryl Holm y Hernán Cendra que me han brindado las comodidades y su tiempo para que tenga el mejor ambiente de trabajo. Tambien a Henry Jacobs, con quien colaboré en Londres y la gente de ambos grupos con los que compartí todo ese tiempo: David Meier, Joris Vankerchaver, Chris Brunnett, Jaap Eldering y Santiago Capriotti. En particular quiero agradecerle a mi amiga Viviana Diaz, por su hospitalidad, humor, comprensión, y viajes y hermosos momentos compartidos.

Quisiera agradecer también a mis compañeros del ICMAT... muchos ya se han ido, pero por eso no dejo de recordar los buenos momentos compartidos juntos en el ICMAT. Gracias por brindarme su compañia y amistad entre almuerzos, meriendas, cenas y salidas: Rafa Granero, Alberto, Luis, Ana Z, Jezabel, Jose, David F, Alfonso, Javi, Eric, Angel, Diego C, Rafa O, Luis (informático), Daniel L y Daniel P. Y a mis compañeros de despacho que han brindado un ambiente más que magnífico estos años soportando la música que sale de mis cascos en pleno ambiente de concentración y el ruido de mi mate: Tania, Jorge, Ana, Emilio y Fernando, ¡muchas gracias!

La vida de un investigador es agobiante cuando, además de lidear con nuestos teoremas y asuntos académicos, tenemos que dedicarnos a la burocracia del sistema de investigación español. Sin la ayuda de los Eduardos, Rosa, Gloria, Sandra, Vanesa, Arucas, Ágata, Bernardo, Esther y Amalia, todo hubiese resultado más difícil. Gracias al equipo de Gestión del ICMAT que tanta ayuda brinda y su disponibilidad para los investigadores. También a Raquel, Juan y Pedro por recibirme tan alegremente todas las mañanas al llegar al ICMAT.

Quisiera agradecerle al Departamento de Matemáticas de la Universidad Autónoma de Madrid por aceptarme como alumno de doctorado de su departamento y brindarme la posibilidad de haber dado clases allí. Quisiera agradecerle tambien a Jesús García Azorero que ha sido el profesor de la asignatura en la que trabajé por ayudarme en todo ese proceso de docencia y enseñarme de sus experiencias con los alumnos. A mis alumnos, por enseñarme a tener paciencia, mejorar como docente, y demostrarme que aunque no lo vea, la docencia es una parte muy importante en mí, que extrañaba en mi rutina de Argentina, y saca un compromiso hacia las personas que había olvidado que tengo incorporado. También quiero agradecerle a Fernando Soria por toda la ayuda brindada en la inscripción al Máster, Doctorado y lo que respecta al déposito de la tesis doctoral y su defensa.

A mis amigos en Madrid que me ayudan a salir de las matemáticas y ponerme en el mundo real, mostrandome que soy un privilgiado en este mundo por hacer las dos cosas que más me gustan: Viajar e Investigar. Gracias Diego, Lucia, María, Titán, Stef, Fede, Alina, Carlos, Ines, Juan, Lola y Lourdes. Y por supuesto, a mis amigos que supieron trazar puentes para estar cerca mio y mostrarme que no hay distancias que separen nuestras amistades: Ramiro, Joaquín, Laura, Neda y Cecilia. También quiero agradecerle a Marila, Mónika y a toda la familia Lafuente y Martín de Diego por estar siempre conmigo y sentir su apoyo siempre.

Estos años no hubiesen sido los mismos si no hubiese existido Facebook en mi vida. Gracias a esa red social pude estar conectado con todos mis seres queridos, poder vivir en

Argentina y España al mismo tiempo y sentirme apoyado y acompañado por todos mis seres queridos que se encuentran cruzando el charco.

Mi Familia, quién sin su constante apoyo, ayuda y entendimiento me hubiese dado por vencido en ciertos momentos. A mis padres, por ser un pilar fundamental en mi vida y de todas las acciones que hago. Estoy muy orgulloso de tener unos padres como ustedes. A mis hermanos, Adriana y Fabian y sobrinos Valentina y Mateo. A Gabriel y Lorena. A mis tíos y primos, en especial a Julia, Carlos, Nestor, Javier, Karina, Marcela, Ñato, Cuqui y Marie por su amor y compañia estos años. Y aunque ya no estés físicamente con nosotos Nona este trabajo esta hecho gracias a tu amor incondicional durante toda mi vida.

A la luz y amor de mi vida, Fanny, por darme su amor, abrazos y compartir conmigo este camino, rescatarme y llenarme de ilusión. Su valiosa ayuda y compensión en mi profesión, sus maravillosas comidas, desayunos y tiernos despertares me han dado todas las energías para finalizar esta tesis. Así como a mi nueva familia, mi familia política, Mónica, Alex, Pedro, Marina, Boreal, Zenit y Pipo por abrirme sus corazones, apoyarnos, cuidarnos, malcriarnos y compartir con nosotros todas nuestras esperanzas e ilusiones.

¡A Madrid! Ya que el que conoce, sabe que Madrid no es solamente una ciudad con bares, barrios increibles, intelectuales y bohemios, parques preciosos y manifestaciones de todo tipo para atraer nuestra distracción; sino que también posee una zona de sierras y montañas increíble, en donde resido actualmente y he encontrado el mejor ambiente de relajación, naturaleza, aire puro, bosques y pantanos para la escritura de esta tesis. Así que gracias Madrid, Lavapiés y en particular a la urbanización Los Arroyos de Las Zorreras, El Escorial, por darme un ambiente agradable para el transcurso y escritura de esta tesis. A Isifú y Nico, mis gatos, por llenarme de pelos, jugar conmigo, y darme amor incondicional.

Por último, y no por eso menos importante, quiero dar mis últimos dos agradecimientos; primero, agradecerle a mi niño interior por mantener la ilusión siempre en mi vida, patalear y hacerme sentir vivo, llorar y expresarse para hacerme sentir siempre humano y lleno de ilusion, por ser un aliado, ayudarme durante todo este tiempo a poder seguir creciendo felizmente haciendo lo que me gusta y ayudarme a ser amigo de mis sombras, creando puentes, para no tener miedo en esta vida y poder llegar siempre a mis objetivos.

Hay muchas cosas en la vida que son difíciles de obtener: la felicidad, el amor, la sabiduría, teoremas que revolucionen el mundo, ese ejemplar único de nuestro comic favorito, las entradas para un concierto exclusivo de los Rolling Stones... pero si ponemos todo nuestro empeño en ello, ponemos todos los medios, todas nuestras ganas y energias y tenemos un poco de suerte, el universo y todas las energías estelares, dioses y seres de luz (ya tengamos cualquier creencia y/o religión) se alian y conspiran a nuestro favor ayudandonos a conseguir lo que nos propongamos. Las cosas materiales, al final, son más fáciles de obtener que las inmateriales, pero en cualquier caso, siempre es difícil conseguir retos difíciles (ya que si no, no lo serían tanto), pero el premio es mucho mayor, ya que las cosas que son difíciles de obtener, tienen más valor cuando las conseguimos, ya que al habernos costado tanto esfuerzo llegar a ellas, valoramos mucho más el hecho de tenerlas. Por eso, es que quiero agradecerle a la vida y a la fuerza interior que todos llevamos dentro para lograr nuestros objetivos y sueños y sobre todo por enseñarme una cosa muy valiosa...

"el que la persigue, la consigue".

Bibliography

[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] Abrunheiro L, Camarinha M, Cariñena J.F, Clemente Gallardo J, Martínez E and Santos P, *Some applications of quasi-velocities in optimal control.* Int. J. Geom. Methods Mod. Phys. 08, 835 (2011). DOI: 10.1142/S0219887811005427..

[4] Agarwal R. P, *Boundary Value Problems for Higher Order Differential Equations,* World Scientific, Singapore, Philadelphia, (1986).

[5] Arnold V, *Ordinary differential equations*, Springer-Verlag, Berlin, third edition, (1992).

[6] Arnold V, *Dynamical Systems*, Vol. III, SpringerVerlag, New York, Heidelberg, Berlin, (1988).

[7] Aymerich Valls M, *Lagrangian Lie subalgebroids of the canonical Lie algebroid*, Master thesis Universitat Politècnica de Catalunya, 2011.

[8] Baillieul J, *The geometry of controlled mechanical systems.* Mathematical control theory, Springer, New York, 1999, 322-354.

[9] Balseiro P, de León M, Marrero J.C and Martín de Diego D, *The ubiquity of the symplectic Hamiltonian equations in mechanics.* J. Geom. Mech. 1 (2009), no. 1, 134.

[10] Barbero-Liñán M, Echeverría Enríquez A, Martín de Diego D, Muñoz-Lecanda M.C and Román-Roy N, *Skinner-Rusk unified formalism for optimal control systems and applications.* J. Phys. A: Math Theor. 40 (2007), 12071-12093.

[11] Barbero-Liñan M, de León M, Marrero J.C, Martín de Diego D, Muñoz Lecanda M, *Kinematic reduction and the Hamilton-Jacobi equation.* J. Geometric Mechanics, 207–237, Issue 3, 2012.

[12] Barbero-Liñán M, Martín de Diego D and Muñoz-Lecanda M.C, *Lie algebroids and optimal control: abnormality.* Geometry and Physics: XVII International Fall Workshop on Geometry and Physics. AIP Conference Proceedings, **1130** (2009), 113-119.

[13] Bates L and Cushman R, *Global Aspect of Classical Integrable Systems.* Birkhäuser Verlag, Basel (1997).

[14] Benito R, de León M and Martín de Diego D, *Higher-order discrete lagrangian mechanics,* Int. Journal of Geometric Methods in Modern Physics, , 3 (2006), pp. 421–436.

[15] Benito R and Martín de Diego D, *Hidden symplecticity in Hamilton's principle algorithms.* Differential geometry and its applications, 411–419, Matfyzpress, Prague, 2005.

[16] Berndt R, *An introduction to Symplectic Geometry.* Graduate studies in Mathematics, vol. 26, American Mathematical Society.

[17] Bloch A, *Nonholonomic Mechanics and Control.* Interdisciplinary Applied Mathematics Series **24**, Springer-Verlag New-York, (2003).

[18] Bloch A and Crouch P, *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.

[19] Bloch A and Crouch P, *Nonholonomic control systems on Riemannian manifolds*, SIAM J. Control Optim., 33 (1995), pp. 126148.

[20] Bloch A and Crouch P, *On the equivalence of higher order variational problems and optimal control problems*, in Proceedings of the IEEE International Conference on Decision andCon trol, Kobe, Japan, 1996, pp. 16481653.

[21] Bloch A and Crouch P, *Optimal Control, Optimization, and Analytical Mechanics, in Mathematical Control Theory*, J. Baillieul and J.C. Willems, eds., SpringerVerlag, New York, 1998, pp. 268321.

[22] Bloch A and Crouch P, *Constrained variational principles on manifolds*, in Proceedings of the IEEE International Conference on Decision andCon trol, Phoenix, AZ, 1999, pp. 16.

[23] Bloch A and Crouch P, *Newtons law and integrability of nonholonomic systems*, SIAM J. Control Optim., 36 (1998), pp. 20202039.

[24] Bloch A.M and Hussein I, *Optimal Control of Underactuated Nonholonomic Mechanical Systems.*IEEE Transactions on Automatic Control 53, (2008) 668-682.

[25] Bloch A, Baillieul J, Crouch P, and Marsden J, *Nonholonomic Mechanics and Control.* New York, NY: Springer-Verlag, 2003.

[26] Bloch A, Hussein I, Leok M and Sanyal A, *Geometric Structure-Preserving Optimal Control of the Rigid Body*, Journal of Dynamical and Control Systems, 15(3), 307-330, 2009.

[27] Bloch A, Krishnaprasad P, Marsden J and Murray R, *Nonholonomic mechanical systems with symmetry*, Arch. Ration. Mech. Anal., 136 (1996), pp. 2199.

[28] Bobenko AI and Suris YB, *Discrete Lagrangian reduction, discrete Euler-Poincaré equations and semidirect products.* Lett. Math. Phys. **49**, (1999).

[29] Bobenko AI and Suris YB, *Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top.* Comm. Math. Phys., **204(1)**, pp. 147-188, (1999).

[30] 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).

[31] Borisov A and Mamaev I, *On the history of the development of the nonholonomic dynamics.* Regular and Chaotic Dynamics, vol. 7, no. 1, pp. 4347, 2002.

[32] Burnett C, Holm D and Meier D, *Geometric integrators for higher-order mechanics on Lie groups*, Proc. R. Soc. A. 469:20130249 (2013).

[33] Bou–Rabee N and Marsden JE, *Hamilton-Pontryagin integrators on Lie groups I: Introduction and structure-preserving properties. Found. Comput. Math.* **9(2)**, pp. 197-219, (2009).

[34] 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).

[35] Cadzow JA, *Discrete calculus of variations.* Internat. J. Control. **11**, pp. 393-407, (1970).

[36] M. Campos C,de León M and Martín de Diego D, *Constrained variational calculus for higher-order classical fields theory.* J. Phys A, Vol 43, n45 (2010) 455206.

[37] Cannas da Silva A and Weinstein A, *Geometric models for noncommutative algebras.* Amer. Math. Soc., Providence, RI, 1999.

[38] Cantrijn F and Sarlet W, *Higher-order Noether symmetries and constants of the motion.* J. Phys. A 14 (1981), no. 2, 479-492.

[39] Cariñena J.F, Gracia X, Marmo G, Martínez E, Muñoz-Lecanda M.C and Román Roy N, *Geometric Hamilton-Jacobi theory for nonholonomic dynamical systems.* Int. J. Geom. Methods Mod. Phys Vol 7, n3 (2010), 431-454.

[40] Cariñena J and López C, *The time-evolution operator for higher-order singular Lagrangians.* Internat. J. Modern Phys. A 7 (1992), no. 11, 2447-2468.

[41] Cariñena J and López C, *Geometric study of Hamilton's variational principle.* Rev. Math. Phys. **3** (4) (1991), 379-401.

[42] Cariñena J.F and Martínez E, *Lie algebroid generalization of geometric mechanics.* In Lie Algebroids, volume 54 of Banach Center Publications, pages 201-215, 2001.

[43] Cariñena J, Nunes da Costa J and Santos P, *Quasi-coordinates from the point of view of Lie algebroid J. Phys. A: Math. Theor. 40 10031 (2007), doi: 10.1088/1751-8113/40/33/008.*

[44] do Carmo MP, *Riemannian geometry.* Birkhäuser, Boston-Basel- Berlin, (1992).

[45] Cendra H, Marsden J and Ratiu T, *Lagrangian reduction by stages.* Memoirs of the American Mathematical Society, 152 (2001), pp. 1-108.

[46] Chang D and Marsden J, *Asymptotic Stabilization of the Heavy Top Using Controlled Lagrangians.* Proceedings of the 39th IEEE Conference on Decision and Control, (2000) 269 - 273.

[47] Colombo L, de León M, Prieto Martínez P and Román-Roy N, Geometric Hamilton-Jacobi theory for higher-order autonomous systems . Submitted. Preprint available at arXiv:1309.2166 (2013) .

[48] Colombo L, de León M, Prieto Martínez P and Román-Roy N, Unified formalism for the generalized kth-order Hamilton-Jacobi problem . Submitted. Preprint available at arXiv:1310.1071 (2013) .

[49] Colombo L and Martín de Diego D, *Quasivelocities and Optimal Control of Underactuated Mechanical Systems.* Geometry and Physics: XVIII Fall Workshop on Geometry and Physics. AIP Conference Proceedings, no. 1260, 133-140 (2010).

[50] Colombo L, Martín de Diego M and Zuccalli M., *Optimal Control of Underactuated Mechanical Systems: A Geometrical Approach.* Journal Mathematical Physics vol51, (2010) 083519.

[51] Colombo L, Martín de Diego D and Zuccalli M, *On Variational Integrators for Optimal Control of Mechanical Control Systems.* Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas. Volume 106, Issue 1 (2012), Page 161-171.

[52] Colombo L, Martín de Diego D and Zuccalli M, *Higher-order variational problems with constraints.* Journal of Mathematical Physics. Vol 54, 093507 (2013), doi: 10.1063/1.4820817.

[53] Colombo L, Jimenez F and Martín de Diego D, *Discrete Second-Order Euler-Poincar Equations. An application to optimal control.* International Journal of Geometric Methods in Modern Physics. (2012) Vol 9, N4.

[54] Colombo L, Jimenez F and Martín de Diego D, *Variational integrators for underactuated mechanical control systems with symmetries* (submitted). Preprint available at http://arxiv.org/abs/1209.6315 (2013).

[55] Cortés J, de León M, Marrero J.C and Martínez E, Nonholonomic Lagrangian systems on Lie algebroids, Discrete and Continuous Dynamical Systems - Series A, 24 (2), pp. 213271, (2009).

[56] Cortés J, de León M, Martín de Diego D and Martínez S, *Geometric description of vakonomic and nonholonomic dynamics.* SIAM J. Control Optim. **41**, no. 5, 1389–1412, (2002).

[57] Cortés J and Martínez E, *Mechanical control systems on Lie algebroids* SIAM J. Control Optim. **41**, no. 5, 1389–1412, (2002).

[58] J. Cortés J, Martínez S, Ostrowski J.P and Zhang H, *Simple mechanical control systems with constraints and symmetry* SIAM J. Control Optim. 41 (3) (2002),851-874.

[59] Cortés J, *Geometric, control and numerical aspects of nonholonomic systems.* Volume **1793** of Lecture Notes in Mathematics. Springer-Verlag, Berlin, (2002).

[60] Coste, A., P. Dazord, and A. Weinstein, *Groupoïdes symplectiques.* In Publications du Département de Mathématiques. Nouvelle Série. A, Vol. 2, volume 87 of Publ. Dép. Math. Nouvelle Sér. A, pages i-ii, 1,62. Univ. Claude-Bernard, Lyon (1987).

[61] Crampin M, Sarlet W and Cantrijn F, Higher order differential equations and higher order Lagrangian Mechanics. Math. Proc. Camb. Phil. Soc. Vol 99, (1986), 565-587.

[62] Crampin M and Saunders D, *On the geometry of higher-order ordinary differential equations and the Wuenschmann invariant.* Groups, geometry and physics, 79-92, Monogr. Real Acad. Ci. Exact. Fs.-Qum. Nat. Zaragoza, 29, (2006).

[63] Crouch P and Silva-Leite F, *Geometry and the dynamic interpolation problem. American Control Conference*, 1131–1136 (1991).

[64] Crouch P and 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.

[65] Fedorov Y, García-Naranjo L and Marrero J.C, *Unimodularity and preservation of volumes in nonholonomic mechanical* preprint available at http://arxiv.org/abs/1304.1788 (2013).

[66] Ferraro S, Jimenez F and Martín de Diego D, *New developments on the Geometric Nonholonomic Integrator*. Preprint available at http://arxiv.org/abs/1312.1587 (2013).

[67] Fliess M, Lévine J, Martin P and Rouchon P, *Flatness and defect of nonlinear systems,Introductory theory and examples*, International Journal of Control, **61** (6), 1327–1361, (1995).

[68] Gay-Balmaz F, Holm D, Meier D, Ratiu T and Vialard F, *Invariant higher-order variational problems*, Communications in Mathematical Physics, Volume 309, Issue 2, pp 413–458, (2012).

[69] Gay-Balmaz F, Holm D, Meier D, Ratiu T and F.-X Vialard, *Invariant higher-order variational problems II.* Journal of Nonlinear Science **22** (4) 553–597 (2012).

[70] Gay-Balmaz F, Holm D and Ratiu T, *Higher-Order Lagrange-Poincaré and Hamilton-Poincaré Reductions.* Bulletin of the Brazilian Math. Soc. **42** (4) 579–606 (2011).

[71] Godbillon C, *Géometrie diférentielle et mécanique analytique.* Hermann, Paris (1969).

[72] Gotay M, Presymplectic Manifolds, Geometric Constraint Theory and the Dirac-Bergmann Theory of Constraints, Ph. D. Thesis, University of Maryland, 1979.

[73] Gotay M, Nester J and Hinds G, *Presymplectic manifolds and the Dirac-Bergmann theory of constraints.* J. Math. Phys. **19**, no. 11, 2388–2399, (1978).

[74] Grabowska K and Grabowski J, *Variational calculus with constraints on general algebroids.* J. Phys. A: Math. Theor. **41**, pp. 175-204, (2008).

[75] Grabowski J, de Leon M, Marrero J. C and Martín de Diego, D, *Nonholonomic constraints: a new viewpoint.* J. Math. Phys. 50 (2009), no. 1, 013520, 17 pp.

[76] Grabowski J and Urbański P, *Tangent and cotangent lifts and graded Lie algebras associated with Lie algebroids,* Ann. Global Anal. Geom. **15** (1997), 447-486.

[77] Grabowski J and Urbański P, *Algebroids – general differential calculi on vector bundles, J. Geom. Phys.* **31** *(1999), 111–141.*

[78] Gràcia X, Pons J and Román-Roy N, *Higher-order Lagrangian systems: geometric structures, dynamics, and constraints* J. Math. Phys. 32 (1991), no. 10, 2744-2763.

[79] Hairer E, Lubich C and Wanner G, *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations,* Springer Series in Computational Mathematics, Springer-Verlag, Berlin, **31**, (2002).

[80] Higgins P and Mackenzie K, *Algebraic constructions in the category of Lie algebroids,* J. Algebra, **vol 129** (1990) 194-230.

[81] Holm D, *Geometric mechanics. Part I and II,* Imperial College Press, London; distributed by World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008..

[82] Holm D.D, Schamah T and Stoica C, *Geometry Mechanics and Symmetry,* Oxford Text in Applied Mathematics (2009).

[83] Holm D, Marsden J and Ratiu T, *The Euler-Poincaré equations and semidirect products with applications to continuum theories*, Adv. Math. **137**, no. 1 , 1–81, (1998).

[84] Husemoller D, *Fibre Bundles.* Springer series in Graduate Texts in Mathematics. Springer-Verlag, New York, (1994).

[85] Iglesias D, Marrero J.C, Martín de Diego D and Martinez E, *Discrete Nonholonomic Lagrangian Systems on Lie Groupoids.* Journal of Nonlinear Science 18, Number 3 (2008) 221-276.

[86] Iglesias D, Marrero J.C, Martín de Diego D and Padrón E, *Discrete dynamics in implicit form* Discrete and Continuous Dynamical Systems - Series A, 1117 - 1135, Issue 3, 2013.

[87] Iglesias D, Marrero J.C, Martín de Diego D and Sosa D, *Singular lagrangian systems and variational constrained mechanics on Lie algebroids.* Dynamical Systems: an International Journal 23 (2008) Issue 3, 351-397.

[88] Jozwikowski M and Rotkiewicz M, *Bundle-theoretic methods for higher-order variational calculus.* J. Geometric Mechanics Vol 6, n1, 99-120 (2014).

[89] Jozwikowski M and Rotkiewicz M, *Higher algebroids with applications to variational calculus.* Preprint available at http://arxiv.org/abs/1306.3379 (2014).

[90] Jurdjevic V, *Geometric Control Theory* Cambridge Studies in Advanced Mathematics, **52**, Cambridge University Press, 1997.

[91] Jurdjevic V, *Optimal Control, Geometry and Mechanics.* In Mathematical Control Theory, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, 227-267.

[92] Kane C, Marsden JE and Ortiz M, *Symplectic energy-momentum integrators.* J. Math. Phys. 40, pp. 3353-3371, (1999).

[93] Kane C, Marsden JE, Ortiz M and West M, *Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems.* International Journal for Numerical Methods in Engineering, **49(10)** pp. 1295-1325, (2000).

[94] Kelly S and Murray R, *Geometric phases and robotic locomotion*, J. Robotic Systems, 12 (1995), pp. 417431.

[95] Khesin B and Wendt, *The geometry of infinite-dimensional groups* .A Series of Modern Surveys in Mathematics, **51**. Springer-Verlag New-York, (2009).

[96] Klein J, *Espacies variationnels et Mécanique.* Annales de l'institut Fourier, **12**, pp. 1-124, (1962).

[97] Koon W and Marsden J, *Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction*, SIAM J. Control Optim., 35 (1997), pp. 901929.

[98] Kobayashi S and Nomizu K, *Foundations of Differential Geometry. Vol I and II.* Wiley Classics Library. John Wiley and Sons Inc., New York, (1996).

[99] Kobilarov M, *Discrete geometric motion control of autonomous vehicles* PhD thesis, University of Southern California, (2008).

[100] Koiller J, *Reduction of some classical non-holonomic systems with symmetry*, Arch. Ration. Mech. Anal., 118 (1992), pp. 113148.

[101] Konieczna K and Urbański P, *Double vector bundles and duality*, Arch. Math. (Brno) Vol 35 (1999), 59–95.

[102] Lee JM, *Riemannian Manifolds: An Introduction to Curvature.* Springer series in graduate Texts in Mathematics, vol 176. Springer-Verlag, New York, (1997).

[103] Lee T, McClamroch NH and Leok M, *Attitude maneuvers of a rigid spacecraft in a circular orbit.* In American Control Conference, Minneapolis, Minnesota, USA, pp. 1742-1747, (2006).

[104] Lee T, McClamroch N and Leok M, *Optimal control of a rigid body using geometrically exact computations on $SE(3)$.* Proc. IEEE Conf. on Decision and Control, (2006).

[105] Lee T, Leok M and McClamroch N.H, *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).

[106] Leimkmuhler B and Skeel R, *Symplectic Numerical Integrators in Constrained Hamiltonian Systems.* J. Comput. Physics, **112**, (1994).

[107] Leimkuhler B and Reich S, *Simulating Hamiltonian Dynamics.* Cambridge University Press, 2004.

[108] Leimkuhler B and Reich S, *Symplectic integration of constrained Hamiltonian systems.* Math. Comp. 63 (1994), 589-605.

[109] de León M and Lacomba E, *Lagrangian submanifolds and higher-order mechanical systems,* J. Phys. A.:Math. Gen. 22 (1989), 3809-3820.

[110] de León M, Marrero J C and Martín de Diego D, *Linear almost Poisson structures and Hamilton-Jacobi equation. Applications to nonholonomic mechanics.* J. Geom. Mech. 2 (2010), no. 2, 159198.

[111] de León M, Marrero J.C and Martínez E, *Lagrangian submanifolds and dynamics on Lie algebroids,* J. Phys. A: Math. Gen. 38 (2005), 241-308.

[112] de León M and Rodrigues PR, *Methods of Differential Geometry in Analytical Mechanics.* North-Holland, Amsterdam (1989).

[113] Lewis A.D, *Simple mechanical control systems with constraints*, IEEE Trans. Automat. Control, 45 (2000), pp. 14201436.

[114] Lewis A and Murray R, *Variational principles for constrained systems: Theory and experiments*, Internat. J. Non-linear Mech., 30 (1995), pp. 793815.

[115] Lewis A, *Energy-preserving affine connections* Preprint available http://www.mast.queensu.ca/~andrew/papers/abstracts/1997d.html.

[116] Libermann P, *Lie algebroids and mechanics* Arch. Math. (Brno) **32** (1996), 147–162.

[117] Libermann P and Marle C.M, *Symplectic geometry and Analytical Mechanics.* Mathematics and its Applications, **35**. D. Reidel Publishing Co., Dordrecht, (1987).

[118] Mackenzie K, *Lie groupoids and Lie algebroids in differential geometry.* Cambridge University Press, (1987).

[119] Mackenzie K, *General theory of Lie groupoids and Lie algebroids,* volume 213 of London Mathematical Society Lecture Notes Series. Cambridge university Press, Cambridge (2005).

[120] Mackenzie K and Xu P, *Lie bialgebroids and Poisson groupoids,* Duke Math J. **vol 73** (1994) 415-452.

[121] Marle C.M, *From momentum maps and dual pairs to symplectic and Poisson groupoids.* In The breadth of symplectic and Poisson geom- etry, vol. 232 of Progr. Math., pages 493523. Birkhäuser Boston, Boston, MA 2005.

[122] Marmo G, Mendella G and Tulczyjew W, *Constrained Hamiltonian systems as implicit differential equations.* J. Phys. A 30 (1997), no. 1, 277293.

[123] Marrero J.C, *Hamiltonian mechanical systems on Lie algebroids, unimodularity and preservation of volumes,* J. Geometric Mechanics. **vol 2** (2010) 243-263.

[124] Marrero J.C, Martín de Diego D and Martínez E, *Discrete Lagrangians and Hamiltonian mechanics on Lie groupoids* Nonlinearity **19** 1313-1348. (2006).

[125] Marrero J.C, Martín de Diego D, Martínez E, *The local description of discrete mechanics.* To appear in Geometry, mechanics, and dynamics: the legacy of Jerry Marsden, Fields Inst. Commun., Springer (2014). Also available at http://arxiv.org/abs/1303.4047.

[126] Marrero J.C, Martín de Diego D and Stern A, *Symplectic groupoids and discrete constrained Lagrangian systems* arXiv:1103.6250v1 (2011).

[127] Marsden JE and Ostrowski J, *Symmetries in motion: Geometric foundations of motion control.* Nonlinear Sci. Today, (1998).

[128] Marsden JE, Pekarsky S and Shkoller S, *Discrete Euler-Poincaré and Lie-Poisson equations.* Nonlinearity, **12(6)**, pp. 1647-1662, (1999).

[129] Marsden JE, Pekarsky S and Shkoller S, *Symmetry reduction of discrete Lagrangian mechanics on Lie groups.* J. Geom. Phys. **36**, (1999).

[130] Marsden JE and Ratiu TS, *Introduction to mechanics and symmetry.* **17** of Texts in Appied Mathematics, Springer-Verlag, New York, (1999).

[131] Marsden JE and West M, *Discrete Mechanics and variational integrators.* Acta Numerica Vol 10, 357-514, 2001..

[132] Martínez E, *Variational calculus on Lie algebroids. ESAIM: Control, Optimisation and Calculus of Variations. Volume: 14, Issue: 2, page 356-380 (2008).*

[133] Martínez E, *Geometric formulation of Mechanics on Lie algebroids.* In Proceedings of the VIII Fall Workshop on Geometry and Physics, Medina del Campo, (1999), Publicaciones de la RSME, **2**, pp. 209-222, (2001).

[134] Martínez E, *Lagrangian Mechanics on Lie algebroids.* Acta Appl. Math., **67**, pp. 295-320, (2001).

[135] Martínez E, *Reduction in optimal control theory.* Reports on Mathematical Physics. Vol53, pp. 79-90, (2004).

[136] Martínez E and Cortés J., Lie algebroids in classical mechanics and optimal control. SIGMA Symmetry Integrability Geom. Methods Appl. **3** (2007), Paper 050, 17 pp. (electronic).

[137] Maruskin J.M, Bloch A.M, Marsden J.E and Zenkov D.V, A Fiber Bundle Approach to the Transpositional Relations in Nonholonomic Mechanics. Journal of Nonlinear Science, Vol 22, Issue 4, pp 431–461, (2012).

[138] Mestdag T, Crampin M, Anholonomic frames in constrained dynamics. Dynamical Systems. An International Journal **25** 159-187 (2010).

[139] Morimoto A, *Prolongations of G-structures to tangent bundles of higher order.* Nagoya Math. J. Vol 38, (1970) 153-179..

[140] Moser J and Veselov A, *Discrete versions of some classical integrable systems and factorization of matrix polynomials.* Comm. Math. Phys., 139 (2), 217-243 (1991).

[141] Murray R and S.S. Sastry, *Nonholonomic motion planning: Steering using sinusoids*, IEEE Trans. Automat. Control, 38 (1993), pp. 700716.

[142] Murray R, Li Z and Sastry S, *A mathematical introduction to robotic manipulation,* CRC Press, Boca Raton, FL, ISBN 0-8493-7981-4. (1994).

[143] van Nieuwstadt M, Rathinam M and Murray R.M, *Differential Flatness and Absolute Equivalence of Nonlinear Control Systems*, SIAM J. Control Optim. **36** (4), 1225–1239, (1998).

[144] Neimark J and Fufaev N., *Dynamics of Nonholonomic Systems.* Translations of Mathematical Monographs, AMS, 33 (1972).

[145] Nijmeijer H and van der Schaft A, *Nonlinear Dynamical Control Systems.* SpringerVerlag, New York, 1990.

[146] S. Ober-Blöbaum, O. Junge and J.E. Marsden, *Discrete Mechanics and Optimal Control: an Analysis*, Calculus of Variations ESAIM: Control, Optimisation and Calculus of Variations. Volume 17 Issue 02 (2011), pp 322-352.

[147] Oshawa T, Bloch A.M and Leok M, *Nonholonomic Hamilton-Jacobi equation and integrability.* J. Geom. Mech. Vol 1, n4 (2009) 461-481.

[148] Oshawa T and Bloch A.M, *Discrete Hamilton-Jacobi theory* SIAM J. Control and Optim. Vol 49, n4 (2011) 1829-1856.

[149] Ostrowski J.P, *Computing reduced equations for robotic systems with constraints and symmetries*, IEEE Transactions on Robotics and Automation, vol. 15, no. 1, pp. 111-123, February 1999.

[150] Patrick G, *Lagrangian mechanics without ordinary differential equations.* Reports on Mathematical Physics. Vol 57, n3, 437-443, 2006.

[151] Patrick G and Cuell C, *Error analysis of variational integrators of unconstrained Lagrangian systems.* Numer. Math. 113 (2009), no. 2, 243–264.

[152] Pradines J, *Théorie de Lie pour les groupoïdes différentiables.* Calcul différentiel dans la catégorie des groupoïdes infinitésimaux, C. R. Acad. Sci. Paris, **264** Sér. A, pp, 245-248, (1967).

[153] Pradines J, *Fibrés vectoriels doubles et calcul des jets non holonomes.* Amiens (1974).

[154] Prieto-Martínez P.D , Estudi de les estructures geomtriques dels sistemes dinmics d'ordre superior. Universidad Politecnica de Cataluna, 2011.

[155] Prieto-Martínez P.D and Román-Roy N, Lagrangian-Hamiltonian unified formalism for autonomous higher-order dynamical systems, J. Phys. A 44(38) (2011) 385203.

[156] Poincaré H, Sur une forme nouvelle des équations de la méchanique, C. R. Acad. Sci.,**132**, 369-371, (1901).

[157] Pontryagin L.S, Boltyanski V.G, Gamkrelidze R.V and Mischenko E.F, The Mathematical Theory of Optimal Processes, Interscience Publishers, Inc., New York (1962).

[158] Popescu M and Popescu P, *Geometric objects defined by almost Lie structures.* In Proceedings of the Workshop on Lie algebroids and related topics in Differential Geometry, Warsaw 2001, Banach Center Publ. **54**, pp. 217-233, (2001).

[159] Sanz-Serna JM, *Symplectic integrators for Hamiltonian problems: An overview.* Acta Numerica, **1**, Cambridge University Press, pp. 243- 286, (1992).

[160] Sanz-Serna JM and Calvo MP, *Numerical Hamiltonian problems.* Applied Mathematics and Mathematical Computation **7**, Chapman and Hall, London, (1994).

[161] Saunders D, *Prolongations of Lie groupoids and Lie algebroids.* Houston J. Math. 30 (3), pp. 637-655, (2004).

[162] Seto D and Baillieul J, *Control problems in super-articulated mechanical systems.* IEEE Trans. Automat. Control **39**, no. 12 (1994) 2442-2453.

[163] Spindler K, *Optimal attitude control of a rigid body*, Applied Mathematics& Optimization **34** (1), 79-90, (1996).

[164] Suris Y.B, *Hamiltonian methods of Runge-Kutta type and their variational interpretation.* Mat. Model., 2 (4), 78-87 (1990).

[165] Sussman H, *Geometry and Optimal Control.* In Mathematical Control Theory, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, pp. 140-198.

[166] Sussman H and Jurdjevic V, *Contollability of nonlinear systems.* J. Differential Equations, 12:95116, 1972.

[167] Skinner R and Rusk R, *Generalized Hamiltonian dynamics I. Formulation on $T^*Q\oplus TQ$*, Journal of Mathematical Pyhsics, Vol 24 (11), 2589-2594 and 2595-2601, (1983).

[168] Skinner R and Rusk R, *Generalized Hamiltonian dynamics. II. Gauge transformations.* J. Math. Phys. 24 (1983), no. 11, 25952601.

[169] Tulczyjew WM, *Les sous-variétés lagrangiennes et la dynamique hamiltonienne.* C. R. Acad. Sc. Paris **283** Série A, (1976) 15-18 .

[170] Tulczyjew WM, *Les sous-variétés lagrangiennes et la dynamique lagrangienne.* C. R. Acad. Sc. Paris **283** Série A, (1976) 675-678.

[171] Veselov AP, *Integrable discrete-time systems and diference operators.* Funct. Anal. Appl. **22**, pp. 83-93, (1988).

[172] Veselov AP, *Integrable Lagrangian correspondences and the factorization of matrix polynomials.* Funct. Anal. Appl. **25**, pp. 112-122, (1991).

[173] Vitagliano L, *The Lagrangian-Hamiltonian Formalism for Higher Order Field Theories*, J. Geom. Phys. 60 (2010) 857873.

[174] Warner F, *Foundations of differentiable manifolds and Lie groups.* Scott, Foresman, Glenview, Ill, (1973).

[175] Weinstein A, *Symplectic categories.* Port. Math. **67(2)**, pp. 261–278, (2010).

[176] Weinstein A, *Lagrangian Mechanics and groupoids.* Fields Inst. Comm. **7**, (1996).

[177] Weinstein A, *Lectures on symplectic manifolds.* CBMS Regional Conference Series in Mathematics, **29**. American Mathematical Society, Providence, R.I., (1979).

[178] Wendlandt JM and Marsden JE, *Mechanical integrators derived from a discrete variational principle.* Physica D **106**, pp. 223-246, (1997).

[179] WENDLANDT JM AND MARSDEN JE, *Mechanical systems with symmetry, variational principles and integration algorithms.* Current and Future Drections in Applied Mathematics, Birkhäuser, pp. 219-261, (1997).

[180] YOSHIMURA H AND MARSDEN J, *Dirac structures in Lagrangian mechanics. I. Implicit Lagrangian systems.* J. Geom. Phys. Vol 57 (2006), no. 1, 133–156.