Universidad de Zaragoza

Instituto de Ciencias Matemáticas

Facultad de Ciencias

Departamento de Física Teórica

ICMAT-CSIC

UAM-UC3M-UCM

New developments on discrete variational calculus: constrained systems and optimal control

Fernando Jiménez Alburquerque

Zaragoza, Abril 2012

Zaragoza, 18 de Abril de 2012

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 Fernando Jiménez Alburquerque y titulada **New developments on discrete variational calculus: constrained systems and optimal control** ha sido realizada bajo su dirección en el Instituto de Ciencias Matemáticas y tutelada en el Departamento de Física Teórica de la Universidad de Zaragoza. 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 de Zaragoza

Facultad de Ciencias
Departamento de Física Teórica

Instituto de Ciencias Matemáticas

ICMAT-CSIC
UAM-UC3M-UCM

New developments on discrete variational calculus: constrained systems and optimal control

Memoria realizada por
Fernando Jiménez Alburquerque,
presentada ante el Departamento de Física Teórica
de la Universidad de Zaragoza
para la obtención del grado de Doctor en Ciencias.

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.

Zaragoza, Abril 2012.

Perhaps my best years are gone. When there was a chance of happiness. But I wouldn't want them back. Not with the fire in me now.

Samuel Beckett, "Krapp's Last Tape".

Resumen

Durante la década de 1960, poderosas y sofisticadas técnicas provenientes de la geometría diferencial moderna y de la topología fueron introducidas en el estudio de los sistemas dinámicos (incluyendo los sistemas mecánicos). Este nuevo campo de investigación que reformuló la mecánica analítica clásica en lenguaje geométrico y atrajo nuevos métodos topológicos y analíticos es llamado actualmente Mecánica Geométrica.

Por otro lado, uno de los máximos 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 de la década de 1980, y durante todos los 90, 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 es posible la estructura geométrica subyacente.

La Mecánica Discreta, entendida como la confluencia de la Mecánica Geométrica y la Integración Geométrica, es, al mismo tiempo, un área de investigación bien fundamentada y una herramienta poderosa a la hora de entender los sistemas dinámicos y físicos, más concretamente aquéllos relacionados con la mecánica. Una herramienta clave en Mecánica Discreta, ampliamente utilizada en este trabajo, son los integradores variacionales, i.e., integradores geométricos basados en la discretización de los principios variacionales. El trabajo desarrollado en esta tesis se alinea con la Mecánica Discreta y su interralación con la teoría de algebroides y grupoides de Lie (considerados estos últimos como la generalización natural de los espacios sobre los que se define la Mecánica Discreta).

Nuestra intención ha sido desarrollar integradores numéricos con propiedades de preservación geométrica en distintas ramas de la mecánica. Más concretamente, hemos usado técnicas de la Mecánia Geométrica con el objetivo de obtener resultados novedosos en tres aspectos: la relación entre sistemas Lagrangianos con ligaduras y sistemas Hamiltonianos; la integración geométrica de problemas de control óptimo; y la integración geométrica de problemas mecánicos noholónomos.

Abstract

In the 1960's, sophisticated and powerful techniques coming from modern differential geometry and topology have been introduced in the study of dynamical systems (including mechanical ones). This new field that eventually reformulated classical analytic mechanics in geometric language and brought in new methods from topology and analysis is called today Geometric Mechanics.

On the other hand, one of the main goals of numerical analysis and computational mathematics has been rendering physical phenomena into algorithms that produce sufficiently accurate, affordable, and robust numerical approximations. In the late 1980's, and throughout the 1990's, 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 preserve exactly as much of the underlying geometrical structures as possible

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 widely used in this work, are the variational integrators, i.e., geometric integrators for mechanical problems based on the discretization of variational principles. The work developed in this thesis is in line with Discrete Mechanics and its feedback with Lie groupoid and Lie algebroid theory (regarded these as natural generalizations of the spaces where the Discrete Mechanics is defined on).

Our goal has been to develop numerical integrators with geometric preservation properties in several branches of mechanics. More concretely, we have used the techniques from Discrete Mechanics in order to obtain valuable results in three topics: the relationship between constrained Lagrangian systems and Hamiltonian systems, geometric integration of optimal control problems and geometric integration of mechanical nonholonomic problems.

Contents

Introduction

Historically, the research in dynamical systems (including mechanical ones) had a major impact on other areas of mathematics and physics as well as in the development of various engineering technologies. Most of these advances have been based on applied numerical and analytical methods. However, in the 1960's more sophisticated and powerful techniques coming from modern differential geometry and topology have been introduced in their study. This new field that eventually reformulated quantum mechanics (topic which will not be considered in this thesis) and classical analytic mechanics in geometric language and brought in new methods from topology and analysis is called today Geometric Mechanics. It has experienced a spectacular growth in the last 40 years impacting all adjacent mathematical fields as well as mathematical physics and certain areas of engineering. The main guiding idea in this development consists in applying the techniques and methods of differential geometry to the study and description of mechanical systems (classical, field theoretical, or quantum). Symplectic structures and their natural generalizations (like Poisson and Dirac manifolds) turn out to be the natural framework in the description and study of various phenomena that appear in classical (quantum) mechanics, among which we mention the following: symmetry reduction (both for finite and infinite dimensional systems) in a classical and quantum setting, Hamilton-Jacobi theory, mechanical systems that are subjected to external (possibly non-holonomic) constraints, and the modeling of friction. We refer to the following fundamental references ([1, 5, 108, 111, 123]) for additional history, references and background on Geometric Mechanics.

On the other hand, one of the main goals of numerical analysis and computational mathematics has been rendering physical phenomena into algorithms that produce sufficiently accurate, affordable, and robust numerical approximations. Numerical simulations are an invaluable tool for exploring the dynamics of nonlinear differential equations. In the late 1980's, and throughout the 1990's, the field of Geometric Integration arose to design and to analyze numerical methods for ordinary differential equations and, more recently, for partial differential equations (PDEs), that preserve exactly (i.e. up to round-off error) as much of the underlying geometrical structures as possible (see [61]). In this sense, Geometric Integration is concerned with producing numerical approximations preserving the qualitative attributes of the solution to the extent that it is possible (phase space, energy conservation, preservation of integrability under discretization, reversibility, symplecticity, volume preservation, etc) while not disregarding accuracy, affordability, and robustness. In particular, in many problems arising from science and engineering (such as solar system or molecular dynamics) the underlying geometric structure affects the qualitative behavior of solutions, and as such, numerical methods that preserve the geometry of a problem typically yield simulations that are qualitatively more accurate.

As a synergic consequence of the intersection between these two disciplines, i.e. Geometric Mechanics and Geometric Integration, Discrete Mechanics has earned significance in the last years. A flavour of the power of Discrete Mechanics can be given by its capacity to generate geometric variational integrators for mechanical problems.

It is well-known that the equations of motion of a mechanical system described by a Lagrangian function $L : TQ \to \mathbb{R}$ can be obtained applying calculus of variations to the action integral

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

These equations of motion are the so-called **Euler-Lagrange** equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0. \tag{1}$$

For the particular Lagrangian function $L(q, \dot{q}) = \frac{1}{2}\, \dot{q}^T\, M\, \dot{q} - V(q)$, i.e. a Lagrangian of mechanical type, the Euler-Lagrange equations are just

$$M\ddot{q} = -\nabla V(q), \tag{2}$$

which are Newton's equations: mass times acceleration equals force. From the geometric study of Lagrangian problems, it is well-known that the system described by the Euler-Lagrange equations has many special properties. In particular, the flow on state space is symplectic, meaning that it conserves a particular two-form, and if there are symmetry actions on phase space then there are corresponding conserved quantities of the flow, known as momentum maps.

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 [124, 134]). This point of view is a clear consequence of a deeper insight into the geometric structure of numerical methods (provided by the Geometric Integration) and the geometry of the mechanical systems (provided by the Geometric Mechanics) 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 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 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 $\mathcal{A}_L$ using the rectangle rule[1]. In the last expression, $q_0 \approx q(0)$ 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. At the end of the day, this gives the **discrete Euler-Lagrange equations**:

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0,$$

which is the discrete counterpart of (1) 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 (2), 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 (they preserve the same two-form on state space as the true system) and have the property of conserving momentum maps arising from symmetry 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 ([61, 153]). 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, which will be detailed below.

From the point of view of history, the theory of discrete variational mechanics in the form we shall use it has its roots in the optimal control literature of the 1960's: see, for example, Jordan and Polak ([80]), Hwang and Fan ([68]) and Cadzow ([24]). In the context of mechanics early work was done, often independently, by Cadzow ([25]), Logan ([113]), Maeda

[1] More sophisticated quadrature rules lead to higher-accurate integrators.

([115, 116, 117]), and Lee ([97, 98]), by which point the discrete action sum, the discrete EulerLagrange equations and the discrete Noether's theorem were clearly understood. This theory was then pursued further in the context of integrable systems in Veselov ([160, 161]) and Moser and Veselov ([134]), and in the context of quantum mechanics (topic which will not be considered in this thesis) in Jaroszkiewicz and Norton ([74, 75]) and Norton and Jaroszkiewicz ([137]).

The variational view of discrete mechanics and its numerical implementation is further developed in Wendlandt and Marsden ([167, 168]) and then extended in Kane, Marsden and Ortiz ([84]), Marsden, Pekarsky and Shkoller ([121, 122]), Bobenko and Suris ([17, 18]) and Kane, Marsden, Ortiz and West ([85]). A central reference big part of this thesis is based on is the work by Marsden and West ([124]). The beginnings of an extension of these ideas to a nonsmooth framework is given in Kane, Repetto, Ortiz and Marsden ([86]), and is carried further in Fetecau, Marsden, Ortiz and West ([46]).

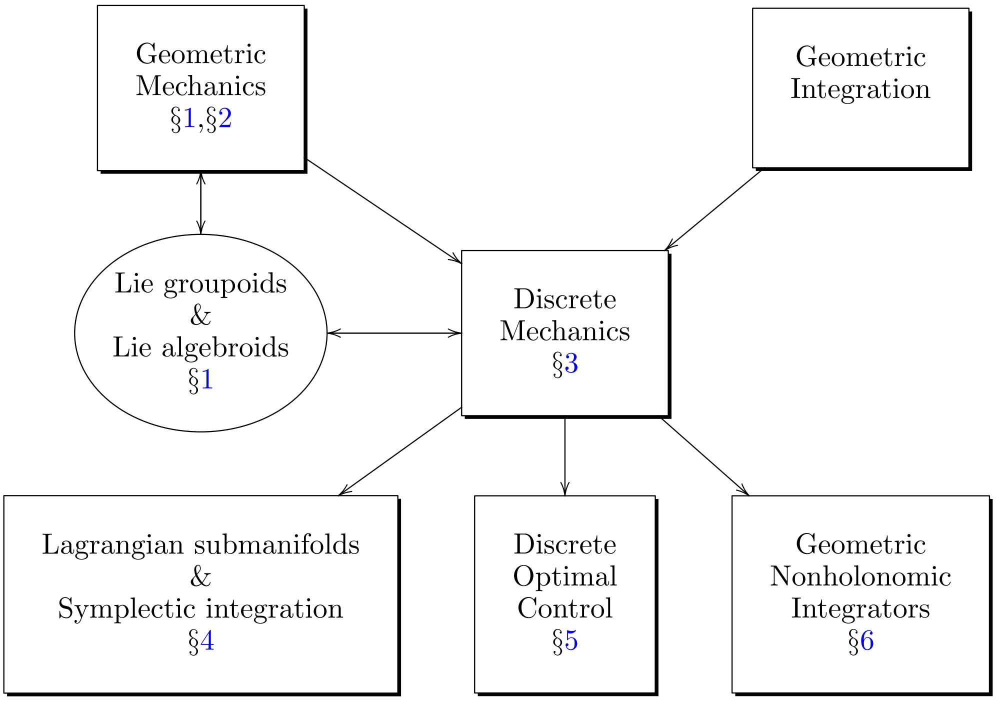
A step further, Alan Weinstein began the study of discrete mechanics on Lie groupoids. His attention was called by the work by Moser and Veselov [134], where the authors study the complete integrability of certain discrete dynamical systems. Moreover the authors describe the Lagrangian and Hamiltonian formalisms for discrete mechanics in two different settings: $Q \times Q$ and a Lie group. Therefore, in [165] 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 [142]. 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 Ch. 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.

We may think a Lie algebroid A over a manifold Q, with projection $\tau : A \to Q$, as a generalized version of the tangent bundle to Q. The geometry and dynamics on Lie algebroids have been extensively studied during the past years. In particular, a geometric formalism of mechanics similar to Klein's formalism ([88]) was developed in [128], while, more recently, a description of the Hamiltonian dynamics on a Lie algebroid was given in [107, 127, 144].

Finally, 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 [118].

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. The work developed in this thesis is therefore in line with Discrete Mechanics and its feedback with Lie groupoid and Lie algebroid theory. For sake of precision, we have used the techniques from Discrete Mechanics in order to obtain valuable results in three different topics: the relationship between constrained Lagrangian systems and Hamiltonian systems, geometric integration of optimal control problems, and geometric integration of mechanical nonholonomic problems. The following diagram shows the master lines followed in our work.



Lagrangian submanifolds and symplectic integration:

A central notion in symplectic geometry is the concept of Lagrangian submanifold (see [164, 166]). This concept arises in several and different interpretations of physical, engineering and geometric phenomena. 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 [157, 158]). Namely, if a mechanical system is defined by a Lagrangian function $L : TQ \to \mathbb{R}$, 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.

A proof of the power of this formalism is its capacity to describe constrained systems. Roughly speaking, a constrained Lagrangian system can be defined by $\ell : C \to \mathbb{R}$, where C is a submanifold of TQ with inclusion $i_C : C \hookrightarrow TQ$. A Lagrangian submanifold $\Sigma_\ell \subset T^*TQ$ can be built as

$$\Sigma_\ell = \{\mu \in T^*TQ \mid i_C^*\,\mu = \mathrm{d}\ell\}.$$

Thus, we can obtain via α_Q a new Lagrangian submanifold of the tangent bundle TT^*Q which completely determines the equations of motion of the constrained dynamics (see [54]), which are, in a regular case, of Hamiltonian type. Taking this into account, it is clear that constrained Lagrangian systems and Hamiltonian systems are closely related (relationship which will be further studied in this work).

Hamiltonian systems form the most important class of ordinary differential equations in the context of Geometric Integration, being one of its outstanding properties the symplecticity of the flow. Hamiltonian theory operates in three different domains (equations of motion, partial differential equations and variational principles) which are all interconnected. Each of these viewpoints leads to the construction of methods preserving the symplecticity. Pioneering work on symplectic integration is due to De Vogelaŕe ([162]), Ruth ([150]) and Feng Kang ([43]). For an overview of symplectic integration, see Sanz-Serna ([152]) and Sanz-Serna and Calvo ([153]). Qualitative properties of symplectic integration of Hamiltonian systems are given in González, Highan and Stuart ([50]) and Cano and Sanz-Serna ([26]). Other interesting references on the subject are [10, 57, 58, 59, 60, 147, 148, 156]. For other references see the large literature on symplectic methods in molecular dynamics, such as [155], and for various applications, see [9, 63, 102].

From the point of view of Geometric Integration and Discrete Mechanics, the connection between constrained Lagrangian systems and Hamiltonian systems, mentioned above, becomes more interesting when dealing with the discrete formalism, which might be interpreted as suitable Lagrangian submanifolds of the Cartesian product of two copies of T^*Q. Since the Hamiltonian flow $F_H^t : T^*Q \to T^*Q$ can be seen as the graph of F_H^t, and consequently as a submanifold of $T^*Q \times T^*Q$, the treatment of this kind of systems from the point of view of Geometric Mechanics, can enlighten the geometric structure of symplectic integrators.

Discrete optimal control:

The optimization and control of physical processes is of crucial importance in all modern technological sciences. In particular, optimal control theory is a mathematical optimization method for deriving control policies such that a certain optimality criterion is achieved. The aim of optimal control is to guide or steer certain processes, arising in nature and engineering, such that a given quantity, for example control effort or maneuver time is minimal. To be more precise, a given **cost functional** has to be optimized by taking into account the dynamics of the process described by **a dynamical system**.

From a purely mathematical point of view, optimal control problems are also variants of a class of problems of the calculus of variations. As in Hamilton's principle, the problem is to minimize an integral (the cost functional) which is now subject to constraints describing the dynamical behavior of the underlying system. These constraints also determine the set of admissible variations and are typically differential equations or, for mechanical systems, given by the Lagrange-d'Alembert principle. Therefore, it seems natural to apply techniques from calculus of variations for a better understanding of optimization processes. Moreover, besides the importance in continuous mechanics, discrete calculus of variations and the corresponding discrete variational principles play an important role in constructing efficient numerical methodologies for the simulation of mechanical systems and for optimizing dynamical systems. Due to their common origin in theory as well as in computational approaches, the combination of these two fields of research, Discrete Mechanics and Optimal Control Theory, provides interesting insight into theoretical and computational issues and provides profitable approaches to specific problems.

More concretely, the application of discrete variational principles already on the dynamical level (namely the discretization of the Lagrange-d'Alembert principle) leads to structure-preserving time-stepping equations which serve as equality constraints for the resulting finite dimensional nonlinear optimization problem. The benefits of variational integrators could be handed down to the optimal control context. For example, in the presence of symmetry groups in the continuous dynamical system, also along the discrete trajectory the change in momentum maps is consistent with the control forces. Choosing the cost function to represent the control effort, which has to be minimized is only meaningful if the system responds exactly according to the control forces.

Applications of discrete optimal control theory were firstly focused on space mission design and formation flying ([81, 82, 83]). There are other applications to robotics and biomechanics ([87, 90, 92, 126, 141, 149]) and to image analysis ([131]). From the theoretical point of view, considering the development of variational integrators, extensions of discrete optimal control to mechanical systems with nonholonomic constraints or to systems with symmetries are quite natural and have already been analyzed in [90, 92]. Extensions for hybrid systems can be found in [126] and for constrained multi-body dynamics in [109, 110]. Discrete optimal control related approaches are presented in [100, 101]. The authors discretize the dynamics by a Lie group variational integrator. Rather than solving the resulting optimization problem numerically, they construct the discrete necessary optimality conditions via the discrete variational principle and solve the resulting discrete boundary value problem (the discrete state and adjoint system). The method is applied to the optimal control of a rigid body and to the computation of attitude maneuvers of a rigid spacecraft.

Other application of discrete optimal control to systems defined on Lie groups can be found in [91]. A wide introduction to optimal control on Lie algebroids and Lie groupoids can be found in [36], while the application of geometric reduction in the context of optimal control in [129].

Geometric nonholonomic integrators:

Nonholonomic constraints have been a subject of deep analysis since the dawn of Analytical Mechanics. The origin of its study is nicely explained in the introduction of the book by Neimark and Fufaev [136],

> "*The birth of the theory of dynamics of nonholonomic systems occurred at the time when the universal and brilliant analytical formalism created by Euler and Lagrange was found, to general amazement, to be inapplicable to the very simple mechanical problems of rigid bodies rolling without slipping on a plane. Lindelöf's error, detected by Chaplygin, became famous and rolling systems attracted the attention of many eminent scientists of the time...*"

Many authors have shown, in the last 25 years, a new interest in that theory and also in its relationship to the new developments in control theory and robotics. The main characteristic of this last period is that nonholonomic systems are studied from a geometric perspective (see L.D. Fadeev and A.M. Vershik [159] as an advanced and fundamental reference). From this perspective, nonholonomic mechanics forms part by its own right of Geometric Mechanics.

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

As mentioned above, the geometric perspective in the study of the nonholonomic systems has been recently introduced (see [14, 16, 27, 34, 94, 104, 159]). On the other hand, recent works (see [38, 42, 70, 132]) have introduced numerical integrators for nonholonomic systems with very good energy behavior and properties such as the preservation of the discrete nonholonomic momentum map. In a similar spirit, the Geometric Nonholonomic Integrator (GNI) and some of its properties were introduced in [44, 45, 93] (some of the results in these references are further developed in this work). Therefore, we can conclude that Nonholonomic Mechanics represents a research field where Discrete Mechanics, as a confluent tool from both Geometric Mechanics and Geometric Integration, can extract new and interesting results.

Outline of the thesis

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

- Chapters 1, 2 and 3 are devoted to introduce the fundamental concepts this thesis is built upon. Specifically those concerning differential geometry, Lie groupoid and Lie algebroid theory, continuous and discrete Lagrangian and Hamiltonian mechanics and, finally, variational integrators. These chapters are included in order to make the thesis project as self-consistent as possible. The ideas and results appearing in these chapters can be found, in more detail and enlarged, in the following references: general texts on differential geometry and other mathematical areas [2, 20, 33, 89, 99, 135], texts on analytical and geometric mechanics [1, 5, 108, 111, 123], groupoid theory [114], fibre bundles [67], texts on symplectic geometry [3, 12], nonholonomic mechanics [14], numerical (geometric) integration [61, 62] and, finally, a fundamental reference on variational integrators [124].

- In chapter 4 we study the relationship between Hamiltonian dynamics and constrained variational calculus (Vakonomic mechanics). The main tools that we employ to describe both are Lagrangian submanifolds (of convenient symplectic manifolds) and the Tulczyjew's triple. The results are also extended to the case of discrete dynamics. In this last sense, our final goal is the capacity of finding interesting applications to geometrical integration of Hamiltonian systems. We also analyze in parallel the case of classical nonholonomic mechanics in the discrete and continuous cases.

  Chapter 4 represents the first central block of the thesis; all its results can be found in [77]. Some of its ideas were previously introduced in [79].

- Chapter 5 accounts for the development of numerical methods for optimal control of mechanical systems in the Lagrangian setting. It extends the theory of discrete mechanics to enable the solutions of optimal control problems through the discretization of variational principles. The key point is to solve the optimal control problem as a variational integrator of a specially constructed higher-dimensional system. The developed framework applies to systems on tangent bundles, Lie groups, underactuated and nonholonomic systems, and can approximate either smooth or discontinuous control inputs. Special attention is paid to Lagrangian systems defined on tangent bundles and Lie groups. The resulting methods inherit the preservation properties (see [118, 124]) of variational integrators. These qualities are associated with numerical stability, robustness and easy implementation, which motivate the development of practical algorithms that can be applied to robotic or aerospace vehicles.

  Chapter 5 is the second central block of this thesis; all its results can be found in [76]. The ideas concerning the extension of discrete optimal control to the Lie groupoid setup were previously introduced in [78].

- Chapter 6 is devoted to develop the ideas concerning the Geometric Nonholonomic Integrator (GNI) presented in [44, 45, 93]. GNI is a discretization scheme adapted to nonholonomic mechanical systems through a discrete geometric approach. This method was designed to account for some of the special geometric structures associated to a nonholonomic motion, namely preservation energy, preservation of constraints or the nonholonomic momentum equation. In chapter 6 we present the GNI extensions of Euler-symplectic methods (see [61]) and discuss some of their convergence properties following the methodology developed in [45]. Additionally, we generalize the method proposed for nonholonomic reduced systems, which represent an important subclass of examples in nonholonomic dynamics. Moreover, we construct extensions of the GNI in the cases of affine constraints and Lie groupoids. Several theoretical examples illustrate the behavior of the proposed method.

Chapter 6 is the third central block of this thesis. Currently its results are a work in progress in collaboration with David Martín de Diego and Sebastián Ferraro (Universidad Nacional del Sur, Argentina).

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

- Appendix A contains the set of Lemmae and proofs, involving the right-trivialized tangent and inverse right-trivialized tangent of a general retraction map (definition 5.2.2), necessary for the derivation of the algorithms obtained in §5. Interesting references in regard to the derivation of the right-trivialized tangent (and its inverse) of a general retraction map are [21, 22].

- Appendix B collects the explicit expression of the Cayley map, as well as its right-trivialized tangent and inverse right-trivialized tangent, for some quadratic matrix Lie groups, namely $SO(3)$, $SE(2)$ and $SE(3)$. These expressions are useful for the practical implementation of the algorithms derived in §5, and which simulations are shown in Figures 5.1, 5.2 and 5.3. Interesting references in regard to matrix expression for Cayley transforms are [61, 91].

Chapter 1

Mathematical background

This chapter gives a brief review of several differential geometric tools used throughout this work. For a more thorough introduction we refer to [1, 2, 89, 108, 112, 123].

1.1 Differentiable manifolds, tangent and cotangent bundles

A minimum knowledge in linear algebra, topology and differential geometry is assumed in the following. For further understanding in this topic, references [1, 2, 89, 135, 163] 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, unless otherwise stated.

Two interesting examples of manifolds which will be extensively used throughout this dissertation are the **tangent** and **cotangent bundles**. Both are discussed next.

1.1.1 Tangent and cotangent bundles

There exist several approaches in order to define the tangent space to an abstract manifold. In this work we will employ the **curves approach**, i.e., to abstract the idea that a tangent vector to a surface is the velocity vector of a curve in the surface, which provides a geometric notion of what the tangent space of a manifold is.

Definition 1.1.1. *Let Q be a manifold and $q \in Q$. A* **curve at** q *is a C^1 map $c : I \to Q$ from an open interval $I \subset \mathbb{R}$ into Q with $0 \in I$ and $c(0) = q$. Let c_1 and c_2 be curves at q and (U, φ) a local chart with $x \in U$. Then we say c_1 and c_2 are* **tangent at** q **with respect to** φ *iff $\varphi \circ c_1$ and $\varphi \circ c_2$ are tangent at 0.*

To understand this definition is necessary to briefly introduce the notion of tangency. Let E, F be two (finite-dimensional) vector spaces with maps $f, g : U \subset E \to F$, being U an open subset of E. We say that f and g are tangent at $u_0 \in U$ iff $\lim_{u \to u_0} \frac{\|f(u)-g(u)\|}{\|u-u_0\|} = 0$, where $\| \cdot \|$ represents any norm on the appropiate space.

Regarding this, two curves are tangent with respect to φ if they have identical tangent vectors (same direction and speed) in the chart φ.

Proposition 1.1.2. *Let c_1 and c_2 be two curves at $q \in Q$. Suppose (U_β, φ_β) are local charts with $q \in U_\beta$, $\beta = 1, 2$. Then c_1 and c_2 are tangent at q with respect to φ_1 iff they are tangent at q with respect to φ_2.*

This proposition (see [1] for the proof) guarantees that the tangency of curves at $q \in M$ is independent of the chart used. Following this idea, is quite evident that tangency at $q \in Q$ is an equivalent relation among curves at q. An equivalent class of such curves is denoted $[c]_q$, where c is a representative of the class.

Definition 1.1.3. *For a manifold Q and $q \in Q$, the tangent space of Q at q is the set of equivalence classes of curves at q, namely*

$$T_qQ = \left\{ [c]_q \middle| \; c \text{ is a curve at } q \right\}.$$

We call $TQ = \cup_{q \in Q} T_qQ$ the **tangent bundle** *of Q. The mapping $\tau_Q : TQ \to Q$ defined by $\tau_Q([c]_q) = q$ is the* **tangent bundle projection** *of M. We will denote by v_q the elements of T_qQ.*

If we have a tangent vector $v_q \in T_q\mathbb{R}^n$, then we can define the notion of **directional derivative** at q, which is a map $D_{v_q} : C^\infty(\mathbb{R}^n) \to \mathbb{R}$, $f \mapsto D_{v_q}(f) = Df(q) \cdot v_q$, where $f \in C^\infty(\mathbb{R}^n)$. In the case of a general manifold Q, we can take the directional derivative to be the map given by $D_{v_q} : C^\infty(Q) \to \mathbb{R}$

$$D_{v_q}(f) = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ c)\Big|_{t=0},$$

where $f \in C^\infty(Q)$ and $v_q = [c]_q$.

If we choose local coordinates (q^i) $i = 1, ..., n$ for $q \in Q$ in the chart (U, φ), the derivative above can be expressed as

$$\left(\frac{\partial f \circ \varphi^{-1}}{\partial q^i} \right) \left(\frac{\mathrm{d}\varphi^i(c(t))}{\mathrm{d}t} \right) \Big|_{t=0}.$$

Denoting $\frac{\partial f \circ \varphi^{-1}}{\partial q^i}$ by $\frac{\partial f}{\partial q^i}$ (which is an abuse of notation) and $X^i = \frac{\mathrm{d}\varphi^i(c(t))}{\mathrm{d}t}\Big|_{t=0}$, we can rewrite the directional derivative like

$$\frac{\mathrm{d}f\,(c(t))}{\mathrm{d}t}\Big|_{t=0} = X^i \frac{\partial f}{\partial q^i} \equiv X[f].$$

All the equivalence classes of curves at $q \in Q$, namely, all the tangent vectors at q, form a **vector space** called the **tangent space** of Q at q and, as previously, denoted by T_qQ. As just mentioned, T_qQ is a vector space with typical local coordinate basis $\frac{\partial}{\partial q^i}\Big|_q$, and $\dim T_qQ = \dim Q$.

Adding up all the tangent spaces corresponding to all the points which belong to Q we get the **tangent bundle**:

$$TQ = \bigcup_{q\in Q} T_qQ.$$

The tangent bundle TQ is a $2n$-dimensional manifold (recall that Q is an n-dimensional manifold).

The canonical tangent projection assigns to each tangent vector its base point. Let define the canonical tangent projection by τ_Q; $\tau_Q : TQ \to Q$. As mentioned just before, for each base point q in Q the tangent space T_qQ is a $\mathbb{R}$-vector space. Thus, we can consider the dual space T_q^*Q, which is called the **cotangent space** at q of Q. T_q^*Q is the space of linear functions from T_qQ to $\mathbb{R}$. In consequence, T_q^*Q has $\mathbb{R}$-vector space structure too. The elements in T_q^*Q are called **covectors** or **momenta** at the point $q \in Q$. As in the tangent case, we can perform the union of all the cotangent spaces to construct the **cotangent bundle** of Q, i.e.:

$$T^*Q = \bigcup_{q\in Q} T_q^*Q.$$

The canonical cotangent projection assigns to each covector its base point, $\pi_Q : T^*Q \to Q$. Since we have introduced some terminology regarding fibre bundles, namely, **tangent bundle, canonical projection**, etc., maybe this is the precise point to refresh some basics about this topic.

Basics on fibre bundles

A **fibre bundle** is, roughly speaking, a sort of generalized product. The notion of fibre bundle will include as particular cases the tangent and cotangent bundles. For more insight in fibre bundles we recommend [67].

Definition 1.1.4 (Bundle). *A* **bundle** *is a triple* (E, p, B)*, where* $p : E \to B$ *is a map. The space* B *is called the* **base space***, the space* E *is called the* **total space***, and the map* p *is called the* **projection of the bundle***. For each* $b \in B$*, the space* $p^{-1}(b) \in E$ *is called the* **fibre** *of the bundle over* $b \in B$*.*

Definition 1.1.5. *A bundle* (E', p', B') *is a subbundle of* (E, p, B) *provided* E' *is a subspace of* E*,* B' *is a subspace of* B *and* $p' = p|_{E'} : E' \to B'$*.*

Definition 1.1.6 (Sections). *A* **section** *of a bundle* (E, p, B) *is a map* $s : B \to E$ *such that* $p \circ s = \mathrm{Id}_B$*, where* Id_B *is the identity map on* B*.*

In other words, a section is a map $s : B \to E$ such that $s(b) \in p^{-1}(b)$, the fibre over b, for each $b \in B$. The set of sections is denoted $\Gamma(p)$ or $\Gamma(E)$ if there is no doubt about the fibre bundle structure.

Let (E', p', B') be a subbundle of (E, p, B), and let $s \in \Gamma(p)$ be a section of (E, p, B). Then $s|_{B'}$ is a cross section of (E', p', B') if and only if $s|_{B'}(b) \in E'$ for each $b \in B'$.

From a more geometrical point of view, we can consider the bundle $p : E \to B$ as a **differentiable fiber bundle** if E and B are differential manifolds and p is a surjective submersion such that it is locally trivial. That is, there exists a manifold F such that, for every $b \in B$, there exists a neighborhood W of b and a diffeomorphism $\varphi : p^{-1}(W) \to W \times F$ such that $pr_1 \circ \varphi = p$, where $pr_1 : W \times F \to W$. Moreover, given two fiber bundle charts (W_1, ϕ_1) and (W_2, ϕ_2) adapted to p, the mapping $\phi_1 \circ \phi_2^{-1}$, called the **transition function**, is a diffeomorphism. Here F is called the **typical fiber**.

Once we have defined what a fibre bundle is, we can establish the concept of bundle morphisms, which, roughly speaking, are fibre preserving maps.

Definition 1.1.7. *Let (E, p, B) and (E', p', B') be two bundles. A* **bundle morphism** *$(u, f) : (E, p, B) \to (E', p', B')$ is a pair of maps $u : E \to E'$, $f : B \to B'$, such that $p' \circ u = f \circ p$.*

The condition $p' \circ u = f \circ p$ is equivalent to the commutativity of the following diagram

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "p"'] & E' \arrow[d, "p'"] \\ B \arrow[r, "f"] & B'. \end{tikzcd}$$

The bundle morphism condition $p' \circ u = f \circ p$ can also be expressed by the relation $u\left(p^{-1}(b)\right) \subset (p')^{-1}\left(f(b)\right)$ for each $b \in B$; that is, the fibre over b is carried into the fibre over $f(b)$ by u. It should be observed that the map f is uniquely determined by u when p is surjective (case we are considering for differentiable fibre bundles).

Definition 1.1.8. *Let (E, p, B) and (E', p', B) be two bundles over B. A* **bundle morphism over** B *$u : (E, p, B) \to (E', p', B)$ is a map $u : E \to E'$ such that $p = p' \circ u$.*

The condition $p = p' \circ u$ is equivalent to the commutativity of the following diagram

$$\begin{tikzcd} E \arrow[rr, "u"] \arrow[dr, "p"'] & & E' \arrow[dl, "p'"] \\ & B. & \end{tikzcd}$$

The bundle morphism condition $p = p' \circ u$ can also be expressed by the relation $u\left(p^{-1}(b)\right) \subset (p')^{-1}(b)$ for each $b \in B$; that is, u is fibre preserving.

In this subsection we have presented two bundles, namely, the tangent bundle $\tau_Q : TQ \to Q$ (in the above notation (TQ, τ_Q, Q)) and the cotangent bundle $\pi_Q : T^*Q \to Q$ (in the above notation (T^*Q, π_Q, Q)). Both are examples of **vector bundles**. A vector bundle is a bundle with an additional vector space structure on each fibre. The concept arose from the study of tangent vector fields to smooth geometric objects, e.g., spheres, projective spaces and, more generically, manifolds.

Definition 1.1.9 (Vector bundle). *A* **vector bundle** *is a fibre bundle $p : E \to B$ such that:*

1. *Each fibre of the bundle is a vector space of fixed dimension, say k. The standard fibre is therefore $\mathbb{R}^k$.*

2. *About each point of B there is a linear trivialization, that is, a local trivialization $\varphi : p^{-1}(U) \to U \times \mathbb{R}^k$ such that for each $x \in U$, the map $\varphi_x : E_x \to \mathbb{R}^k$ defined by $\varphi_x = pr_2 \circ \varphi|_{E_x}$ is a linear isomorphism of the vector spaces E_x and $\mathbb{R}^k$.*

Once we have introduced the concept of **manifold**, **tangent bundle** of a given manifold and **cotangent bundle** of a given manifold, there exist other interesting geometrical concepts, regarding these, which could be useful for the development of this thesis.

Given two manifolds Q and P, one may consider a smooth mapping between them $f : Q \to P$. The set of all these mappings is denoted by $C^\infty(Q, P)$. The tangent map of f is a mapping between the tangent bundles of the previously defined manifolds, i.e., $Df : TQ \to TP$. The tangent map has a pointwise definition as follows: $D_q f : T_q Q \to T_{p=fq)} P$, where $q \in Q$ and $p = f(q) \in P$. When $P = \mathbb{R}$, 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 $X : Q \to TQ$ such that $\tau_Q \circ X = \mathrm{Id}_Q$ (in **bundle** terminology, a vector field is just a section of the vector bundle (TQ, τ_Q, Q)). In other words, it assigns to each $q \in Q$ the tangent vector $X(q) \in T_q Q$. The set of all vector fields on Q is denoted by $\mathfrak{X}(Q)$. An **integral curve of a vector field** is a curve γ on Q, that is $\gamma : I \to Q$, satisfying $\dot\gamma(t) = X(\gamma(t))$ (where $\dot\gamma(t)$ denotes $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)$). Given an initial condition q_0, there always exists a unique integral curve $\phi^X_{q_0} : I \to Q$ of X with that initial condition because of the results about existence and uniqueness of solutions for ordinary differential equations ([33]). The flow of X is a mapping $\phi^X : I \times Q \to Q$, such that $\phi^X(t, q_0) = \phi^X_{q_0}(t)$ and for every $t \in I$, $\phi^X_t : Q \to Q$ is a diffeomorphism on Q given by $\phi^X_t(q) = \phi^X(t, q)$. Observe that $\phi^X_0 = q$ for every $x \in M$ and $\phi^X_{s+r} = \phi^X_s \circ \phi^X_r$ for $s, r \in I$, wherever the composition is defined.

In a similar fashion to the definition of vector fields, a 1-form ω on Q is a mapping $\omega : Q \to T^*Q$ such that $\pi_Q \circ \omega = \mathrm{Id}_Q$ (in **bundle** terminology, a 1-form is just a section of the bundle (T^*Q, π_Q, Q)). In other words, it assigns to each point $q \in Q$ a covector $\omega(q) \in T^*_q Q$. The set of all the 1-forms on 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; $\langle \cdot\,, \cdot \rangle_q : T_q Q \times T^*_q Q \to \mathbb{R}$.

In fact, the vector fields and the 1-forms are particular cases of tensor fields on Q. Given $r, s \in \mathbb{N} \cup \{0\}$, an r-**contravariant and** s-**covariant tensor field** t on Q is a C^∞-section of $T^r_s(Q) = (TQ \otimes \cdots r \cdots \otimes TQ) \otimes (T^*Q \otimes \cdots s \cdots \otimes T^*Q)$; that is, it associates to each point $q \in Q$ an $\mathbb{R}$-multilinear mapping:

$$t(q) : \left(T^*_q Q \times \cdots r \cdots T^*_q Q\right) \times \left(T_q Q \times \cdots s \cdots T_q Q\right) \to \mathbb{R}.$$

Such a geometric element is also called an (r, s)-tensor field on Q. Thus, a vector field is a $(1, 0)$-tensor field and a 1-form is a $(0, 1)$-tensor field. The set of all tensor fields on Q is denoted by $\mathcal{T}(Q)$. The skew-symmetric s-covariant tensor fields are called s-**forms**. The set of all s-forms is denoted by $\Omega^s(Q)$.

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

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

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** 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 $\Omega^k(Q)$, $k = 0, 1, ...$, together with its structure as an infinite-dimensional real vector space and with the multiplication $\wedge$.

When dealing with exterior differential forms, another important geometric object is the exterior derivative, represented byd. It 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, 163]):

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, then $\mathrm{d}(\alpha\big|_U) = \mathrm{d}(\alpha)\big|_U$, where $\alpha \in \Omega^k(Q)$.

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

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

where $v_i \in T_q Q$. Note that the pull-back defines the mapping $f^* : \Omega^k(Q) \to \Omega^k(Q)$. 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 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 $(\mathcal{L}f)\big|_U = \mathcal{L}_X\big|_U(f\big|_U)$, for $f \in C^\infty(Q)$.
2. $\mathcal{L}_X(f\,Y) = (\mathcal{L}_X f)Y + f(\mathcal{L}_X Y)$, for $f \in C^\infty(Q)$.

Indeed, the operator $\mathcal{L}_X$ can be defined on the full tensor algebra of the manifold Q (see [1, 2, 163]).

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 $v_i \in T_qQ$. The operator i_X is an $\wedge-$antiderivation, namely, 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. $i_{[X,Y]}\alpha = \mathcal{L}_X i_Y\alpha - i_Y\mathcal{L}_X\alpha$.

Once we have introduced the concept of bundle, subbundle, tangent bundle and cotangent bundle, is easy to present the concepts of **distribution** and **codistribution**, which will be relevant when dealing with nonholonomic mechanics.

Distributions and codistributions

Definition 1.1.10 (Distribution). *A* k**-dimensional distribution** $\mathcal{D}$ *on a manifold* Q*,* *dim* $Q = n$*, is a class of 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* $X_1, ..., X_k$ *on* U *which span* $\mathcal{D}$ *at each point of* U

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 $\mathcal{D}_q$. The bundle nature of distributions is shown in the following diagram:

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

where $(\mathcal{D}, \tau_{\mathcal{D}}, Q)$ is a subbundle of (TQ, τ_Q, Q) and i represents the inclusion map (we recall that if $A \subset B$, where A and B are general sets, the inclusion map $i : A \hookrightarrow B$ is defined by $i(a) = a$ for any $a \in A$. In the sequel $\hookrightarrow$ will denote an inclusion map.) Some other interesting definitions related to distributions are:

1. A submanifold $S \hookrightarrow Q$ is said to be an **integral manifold** 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 by 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.

Theorem 1.1.11 (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**, that is: *Let* Q *be a manifold. A* **smooth regular codistribution** $\mathcal{D}^*$ **on** T^*Q *is a subbundle of* T^*Q *with* k*-dimensional fiber.* Its bundle nature is shown in the diagram:

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

where $(\mathcal{D}^*, \pi_{\mathcal{D}^*}, Q)$ is a subbundle of (T^*Q, π_Q, Q). Given the concept of codistribution, 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 given by:

$$\mathrm{ann}(\mathcal{D}_x) = \mathcal{D}_x^0 = \left\{ \alpha \in T_q^*Q \mid \alpha(v) = \langle \alpha, v \rangle = 0, \ \forall v \in \mathcal{D}_q \right\}$$

for every $q \in Q$.

1.2 Riemannian geometry

Some good references for more insight in Riemannian Geometry are [20, 30, 99].

Definition 1.2.1 (Riemannian metric). *Let Q be a n-dimensional differentiable manifold. A* **Riemannian metric** *$\mathcal{G}$ on Q is a $(0,2)$ tensor field 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, the 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 **length** or **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

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

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 $^{\flat}\mathcal{G}(X)(Y) = \mathcal{G}(X, Y)$. On the other hand, $^{\sharp}\mathcal{G}$ 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 $\operatorname{grad} f =^{\sharp} \mathcal{G}(\mathrm{d}f) \in \mathfrak{X}(Q)$.

If $(Q, \mathcal{G})$ and $(Q', \mathcal{G}')$ are Riemannian manifolds, a diffeomorphism ϕ from Q to Q' is called an **isometry** if $\phi^* \mathcal{G}' = \mathcal{G}$. If any such isometry exists, then we say that the manifolds are *isometric*. Is easy to check that being isometric is an equivalence relation on the class of Riemannian manifolds.

In a local chart (U, φ) and local coordinates (q^i) for Q, the metric has the form

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

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

$$\begin{array}{rcl} \nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) & \to & \mathfrak{X}(Q) \\ (X, Y) & \mapsto & \nabla(X, Y) = \nabla_X Y, \end{array}$$

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 given by:

$$\nabla\left(\frac{\partial}{\partial q^r}\,,\,\frac{\partial}{\partial q^j}\right) = \nabla_{\frac{\partial}{\partial q^r}}\frac{\partial}{\partial q^j} = \Gamma^i_{r\,j}\frac{\partial}{\partial q^i}.$$

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

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

Given a curve in a manifold Q, we may define the parallel transport of a vector along the curve. Let $c : I \to Q$ be that curve (for sake of simplicity we assume that the image is covered by a single chart (U,φ) with local coordinates q^i). Let X be a vector field defined (at least) along $c(t)$

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

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 is the tangent vector to $c(t)$. In components, the previous condition is written as

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

where q^i are the local components of the curve $c(t)$. 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. In components, the geodesic equations becomes

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

where, as before, q^i are the coordinates of the curve $c(t)$ (see [30]).

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 $f \in \mathcal{C}^\infty(Q)$ is the ordinary directional derivative, namely $\nabla_X f = X[f]$. Then the condition

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

can be exactly rewritten as the Leibniz rule

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

We require this to be true for any product of tensors,

$$\nabla_X(t_1\otimes t_2) = (\nabla_X t_1)\otimes t_2 + t_1\otimes(\nabla_X t_2),$$

where t_1 and t_2 are tensor fields of arbitrary types. With these requirements we can compute the covariant derivative of a one-form $\omega \in \Omega^1(Q)$. Since $\langle \omega, Y\rangle \in \mathcal{C}^\infty(Q)$ for $Y \in \mathfrak{X}(Q)$, we should have

$$X[\langle \omega, Y\rangle] = \nabla_X[\langle \omega, Y\rangle] = \langle \nabla_X \omega, Y\rangle + \langle \omega, \nabla_X Y\rangle.$$

In coordinates we find the expression

$$(\nabla_X \omega)_i = X^j \frac{\partial \omega_i}{\partial q^j} - X^j \Gamma^r_{j\,i} \omega_r,$$

which is easily generalizable to any kind of tensor field.

There exists a natural connection on each Riemannian manifold that is particularly suited to computations in Riemannian geometry. In order to define it is necessary to introduce some extra concepts. Let $(Q, \mathcal{G})$ be a Riemannian (or pseudo-Riemannian) manifold. A affine connection ∇ is said to be **compatible with** $\mathcal{G}$ is it satisfies the following product rule for all vector fields $X, Y, Z \in \mathfrak{X}(Q)$

$$\nabla_X(\mathcal{G}(Y, Z)) = \mathcal{G}(\nabla_X Y, Z) + \mathcal{G}(Y, \nabla_X, Z).$$

Lemma 1.2.2. *The following conditions are equivalent for a connection ∇ on a Riemannian manifold*

a) ∇ *is compatible with* $\mathcal{G}$.

b) $\nabla \mathcal{G} \equiv 0$.

See [20, 99] for the proof. 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 $\tau : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ defined by

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

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

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

Theorem 1.2.3 (**Fundamental Lemma of Riemannian Geometry).** *Let $(Q, \mathcal{G})$ be a Riemannian manifold. There exists a unique connection ∇ on Q that is compatible with $\mathcal{G}$ and symmetric.*

See [20, 30, 99] for the proof. This connection is called the **Riemannian connection** or the **Levi-Civita connection** of $\mathcal{G}$. In the Levi-Civita case, the Christoffel symbols are given in terms of the components of the metric as follows:

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

1.3 Lagrangian submanifolds

Due to its close relation with symplectic manifolds, perhaps is necessary to introduce symplectic algebra and geometry before going deeper into Lagrangian submanifolds. 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, 12, 111].

1.3.1 Symplectic algebra and symplectic geometry

Let V be a finite-dimensional $\mathbb{R}$ vector space with $\dim V = l$. We say that V is a **symplectic vector space** if it is equipped with a symplectic form Ω.

Definition 1.3.1 (Symplectic form). *A symplectic form*

$$\Omega : V \times V \to \mathbb{R}$$

is an skew-symmetric and non-degenerate bilinear form; that is it satisfies $\Omega(v, v) = 0$ *for all* $v \in V$*, and if* $\Omega(v, w) = 0$ *for all* $v \in V$*, then* $w = 0$*,* $w \in V$*.*

In the case of real vector spaces, it is possible to completely describe a skew-symmetric bilinear form.

Proposition 1.3.2. *If* Ω *is an anti-symmetric bilinear form of rank* r*, then* $r = 2n$ *(where* n *is a positive integer) and there is a basis of* V *relative to which*

$$\Omega = \begin{pmatrix} 0 & I_{n\times n} & 0 \\ -I_{n\times n} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $I_{n\times n}$ *is the unit matrix.*

From the definition of symplectic form follows that the skew-symmetric form of a symplectic space is full rank, and consequently the dimension of a symplectic space should be even, i.e., if $\dim V = l$ then $l = 2n$ and

$$\Omega = \begin{pmatrix} 0 & I_{n\times n} \\ -I_{n\times n} & 0 \end{pmatrix}. \tag{1.1}$$

In other words, there exists a basis $\left\{e^i\right\}_{i=1}^{2n}$ of V^* such that $\sum_{i=1}^{n} e^i \wedge e^{i+n}$. The pair (V, Ω) is called **symplectic vector space**. Once the symplectic vector spaces are defined, we can classify their subspaces. Let W be a linear subspace of (V, Ω). We define its skew-orthogonal space by

$$W^{\perp} := \{v \in V \, , \, \Omega(v, w) = 0 \text{ for all } w \in W\} \, .$$

The subspaces of a symplectic manifold are classified in the following definition.

Definition 1.3.3. 1. *A subspace $W \subset V$ such that $\Omega|_W = 0$ is called* **isotropic subspace** *of (V, ω).*

2. *A subspace $W \subset V$ with $\Omega|_W$ non degenerate is called a* **symplectic subspace** *of V.*

3. *A subspace $W \subset V$ with $W^{\perp}$ isotropic is called* **coisotropic**.

4. *A subspace $L \subset V$ which is both isotropic and coisotropic ($L^{\perp} = L$) is called a* **Lagrangian subspace**.

Symplectic geometry arises from the globalization of the symplectic algebra. The central concept is that of a symplectic manifold. In what follows we will assume that P is a smooth manifold of dimension l and assume as well that it is real.

Definition 1.3.4. *P is called a symplectic manifold if there is defined on P a closed nondegenerate 2-form Ω_P; that is an $\Omega_P \in \Omega^2(P)$ such that*

i) $\mathrm{d}\Omega_P = 0$,

ii) on each tangent space T_pP, $p \in P$, if $\Omega_P|_p(X, Y) = 0$ for all $Y \in T_pP$, then $X = 0$.

From now on, if P is a symplectic manifold its associated symplectic 2-form will be denoted by Ω_P. In addition, if the particular symplectic manifold is the cotangent bundle of an arbitrary manifold M, that is $P = T^*M$, then the associated symplectic 2-form will be denoted by Ω_M. Definition 1.3.4 means that the restrictions of Ω_P to each $p \in P$ make the tangent space T_pP into a symplectic vector space. From the fact that the dimension of its tangent space is equal to that of a given manifold, it is already clear that for P a symplectic manifold $\dim P = l = 2n$. It can be shown (see [12]), 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.

Given two symplectic manifolds (P, Ω_P) and $(P', \Omega_{P'})$, let $\phi : P \to P'$ be a smooth map.

Definition 1.3.5. *The map ϕ is called* **symplectic**, *or a* **morphism of symplectic manifolds**, *so long as*

$$\phi^*\Omega_{P'} = \Omega_P.$$

Given a symplectic diffeomorphism ϕ, ϕ^{-1} is also symplectic, and ϕ is called **symplectomorphism**. $Sp(P)$ denotes the **group of symplectomorphisms** from P to itself.

A main result in symplectic geometry is that of Darboux's theorem. In its simplest form, it has the following formulation. To every point p of a symplectic manifold (P, Ω_P) of dimension $2n$, there correspond an open neighborhood U of p and a smooth map

$$\phi : U \to \mathbb{R}^{2n}, \text{ with } \phi^*\Omega = \Omega_P\big|_U,$$

where Ω_0 is the standard symplectic form on $\mathbb{R}^{2n}$ (in other words, the form given in proposition (1.3.2) just by considering $V = \mathbb{R}^{2n}$). It follows immediately that for an appropriate choice of **symplectic coordinates** $p = (x, y)$, $x = (x^1, ..., x^n)$ and $y = (y_1, ..., y_n)$, Ω_P can be written on U in the manner

$$\Omega_P = \sum_{i=1}^{n} \mathrm{d}x^i \wedge \mathrm{d}y_i.$$

Proofs of this theorem are given in [3, 12, 56].

From the definition of symplectic manifold and the classification of the subspaces of a symplectic space given in definition (1.3.3), we arrive to different submanifolds:

Definition 1.3.6. *Let (P, Ω_P) be a symplectic manifold and $i : L \hookrightarrow P$ an immersion. We say L is an* **isotropic (symplectic, coisotropic) immersed submanifold** *of (P, Ω_P) if $(T_x i)(T_x L) \subset T_{i(x)} P$ is an isotropic (symplectic, coisotropic) subspace for each $x \in L$. The same terminology applies for submanifolds of P and for subbundles of TP over submanifolds of P.*

A submanifold $L \subset P$ is called **Lagrangian** *if it is isotropic and there is an isotropic subbundle $E \subset TP\big|_L$ such that $TP\big|_L = TL \oplus E$.*

Remark 1.3.7. *Notice that $i : L \to P$ is isotropic if and only if $i^*\Omega_P = 0$. Also note that if $L \subset P$ is Lagrangian, dim $L = \frac{1}{2}$dim P and $(T_x L)^{\perp} = T_x L$.*

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

Proposition 1.3.8. *Let (P, Ω_P) be a symplectic manifold, $L \subset P$ a submanifold and $i : L \hookrightarrow P$ an immersion. Then L is Lagrangian if and only if $i^*\Omega_P = 0$ and dim $L = \frac{1}{2}$dim P.*

See [1] for the proof.

An interesting kind of Lagrangian submanifolds is the following. Let (P, Ω_P) be a symplectic manifold and $g : P \to P$ a diffeomorphism. Denote by Graph(g) the graph of g, that is Graph$(g) = \{(x, g(x)), \ x \in P\} \subset P \times P$, and by $pr_i : P \times P \to P$, $i = 0, 1.$, the canonical projections. It is well known that $\left(P \times P, \tilde{\Omega}_P\right)$, where $\tilde{\Omega}_P = pr_1^*\,\Omega_P - pr_0^*\,\Omega_P$, is a symplectic manifold. Let $i_g : \text{Graph}(g) \hookrightarrow P \times P$ be the inclusion map, then

$$i_g^* \tilde{\Omega}_P = (pr_0)^* \left(g^*\Omega_P - \Omega_P\right).$$

It is quite clear that dim (Graph g) $= \frac{1}{2}$ dim$(P \times P)$. Moreover, if g is a symplectomorphism besides a diffeomorphism, then $g^*\Omega_P = \Omega_P$ and consequently $i_g^*\tilde{\Omega}_P = 0$. Finally we can conclude that if g is a symplectomorphism then Graph g is a Lagrangian submanifold of $P \times P$.

Let consider now $g = \varphi$, $P = T^*Q$ the cotangent bundle of a given manifold Q and $\Omega_P = \Omega_Q$. As seen in the previous paragraph, every symplectomorphism $\varphi : T^*Q \to T^*Q$ generates a Lagrangian submanifold Graph $\varphi \subset (T^*Q \times T^*Q, \tilde{\Omega}_Q)$, with $\tilde{\Omega}_Q = pr_1^*\,\Omega_Q - pr_0^*\,\Omega_Q$. These Lagrangian submanifolds are generically called **canonical relations** refering to the map φ (this development is again used in §2.2.2 in order to define the **generating function** of a symplectomorphism).

1.4 Lie groups and Lie algebras

In this section, we will recall the definition of Lie group and Lie algebra.

1.4.1 Lie group

Roughly speaking, a Lie group is a manifold on which the group operations, **product** and **inverse**, are defined. Lie groups play an extremely important role in the theory of fibre bundles an also find vast applications in physics, being a crucial notion in the concept of **symmetry** (see §2.1.3 and §2.2 for further details).

Definition 1.4.1. *A* **Lie group** G *is a differentiable manifold which is endowed with a group structure such that the group operations*

1. $\cdot : G \times G \to G;\ (g_1, g_2) \mapsto g_1 \cdot g_2,$
2. $^{-1} : G \to G;\ g \mapsto g^{-1},$

are differentiable.

Remark 1.4.2. *It can be shown that G has a unique analytic structure with which the product and the inverse operations are written as convergent power series.*

The **unit element** of a Lie group is written as e. 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 $g_1 g_2$.

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 $hg' = g'h''$.

1.4.2 Lie algebra

Definition 1.4.3. *A* **Lie algebra** *over* $\mathbb{R}$ *is a real vector space* $\mathfrak{g}$ *together with a bilinear operation* $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *such that, for all* $\xi_1, \xi_2, \xi_3 \in \mathfrak{g}$,

1. $[\xi_1, \xi_2] = -[\xi_2, \xi_1]$.
2. $[\xi_1, [\xi_2, \xi_3]] = [\xi_3, [\xi_1, \xi_2]] = [\xi_2, [\xi_3, \xi_1]]$.

There exists a Lie algebra $\mathfrak{g}$ associated to a Lie group G. We will perform this relationship by means of the following definition:

Definition 1.4.4. *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*

$$R_h(g) = gh, \quad L_h(g) = hg. \tag{1.2}$$

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 $T_gL_h : T_gG \to T_{hg}G$ and $T_gR_h : T_gG \to T_{gh}G$. 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 manner.

Given a Lie group G, there exists a special class of vector fields characterized by an **invariance** under group action (on the usual manifold there is no canonical way of discriminating some vector fields from the others).

Definition 1.4.5. *Let X be a vector field on a Lie group, that is $X \in \mathfrak{X}(G)$. X is said to be a* **left-invariant vector field** *if*

$$T_gL_hX(g) = X(hg).$$

A vector $\xi \in T_eG$ defines a unique left-invariant vector field $\overleftarrow{\xi}$ throughout G by

$$\overleftarrow{\xi}(g) = T_eL_g\xi, \quad g \in G.$$

In fact, it is possible to verify that $\overleftarrow{\xi}(hg) = T_eL_{hg}\xi = T_e(L_h \circ L_g)\xi = (T_eL_h) \circ (T_eL_g)\xi = T_gL_h\overleftarrow{\xi}(g)$. Conversely, a left-invariant vector field $\overleftarrow{\xi}$ defines a unique vector $\xi = \overleftarrow{\xi}(e) \in T_eG$. Let us denote the set of left-invariant vector fields on G by $\mathfrak{g}$. The map $T_eG \to \mathfrak{g}$ 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_eG. 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 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_gL_h to the Lie bracket $[\xi, \eta]$ of $\xi, \eta \in \mathfrak{g}$, we have that

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

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** is defined as the set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket.

Definition 1.4.6. *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 $SO(3)$, which is widely used in §5.

1.4.3 The adjoint representation

Definition 1.4.7. *Take any $h \in G$ and define a homomorphism $\mathrm{Ad}_h : G \to G$ by the conjugation*

$$\mathrm{Ad}_h(g) = L_h \circ R_{h^{-1}} g = hgh^{-1},$$

for $g \in G$. This homomorphism is called the **adjoint action** *of G on G.*

Roughly speaking, the adjoint action 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 adjoint action provides the transition between these two possibilities.

Definition 1.4.8. *The* **adjoint action** *of G on $\mathfrak{g}$ is defined as the map $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ given by*

$$\mathrm{Ad}(h, \xi) = \mathrm{Ad}_h \xi = T_{h^{-1}} L_e (T_e R_{g^{-1}} \xi),$$

for $h \in G$ and $\xi \in \mathfrak{g}$

Definition 1.4.9. *The* **adjoint action** *of $\mathfrak{g}$ on $\mathfrak{g}$ is given by the map $\mathrm{ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined by*

$$\mathrm{ad}(\xi, \eta) = \mathrm{ad}_\xi \eta = [\xi, \eta] = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\mathrm{Ad}_{\exp(t\xi)} \eta),$$

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

1.4.4 Action of a Lie group on a manifold

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 $q \in Q$.

The notion of action of a Lie group on a manifold leads to the definition of **infinitesimal generator**.

Definition 1.4.10. *Suppose $\Phi : G \times Q \to Q$ is an action. For $\xi \in \mathfrak{g}$, the map $\Phi^\xi : \mathbb{R} \times Q \to Q$, defined by $\Phi^\xi(t, q) = \Phi(\exp(t\xi), q)$, where $\exp : \mathfrak{g} \to G$ is the usual exponential map defined below, is an $\mathbb{R}-$action on Q. In other words, $\Phi_{\exp(t\xi)} : Q \to Q$ is a flow on Q. The corresponding vector field $\xi_Q \in \mathfrak{X}(Q)$, given by*

$$\xi_Q(q) := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi(\exp(t\xi), q),$$

is called the **infinitesimal generator** *of the action Φ corresponding to ξ.*

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

Definition 1.4.11. *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.*

1.5 Lie algebroids and Lie groupoids

In this section, we will recall the definition of a Lie algebroid and of the differential calculus associated to them. Moreover, we illustrate the theory with several examples

1.5.1 Lie algebroids

Definition 1.5.1. *A* **Lie algebroid** *A over a manifold Q is a vector bundle $\tau : A \to Q$ together with a Lie bracket $[\![\cdot,\cdot]\!]$ on the space $\Gamma(A)$ of the global cross sections of $\tau : A \to Q$ and a bundle map $\rho : A \to TQ$, called the* **anchor map***, such that if we also denote by $\rho : \Gamma(A) \to \mathfrak{X}(Q)$ the homomorphism of $C^\infty(Q, \mathbb{R})$- modules induced by the anchor map then*

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

for $X, Y \in \Gamma(A)$ and $f \in C^\infty(Q, \mathbb{R})$. The triple $(A, [\![\cdot,\cdot]\!], \rho)$ is called a **Lie algebroid** *over M (see [114, 145] for further details).*

Remark 1.5.2. *If $(A, [\![\cdot,\cdot]\!], \rho)$ is a Lie algebroid over Q then the anchor map $\rho : \Gamma(A) \to \mathfrak{X}(Q)$ is an homomorphism between the Lie algebras $(\Gamma(A), [\![\cdot,\cdot]\!])$ and $(\mathfrak{X}(Q), [\cdot,\cdot])$.*

Given local coordinates (q^i) in the base manifold and a local basis of sections (e_α) of A, then local coordinates of a point $a \in A$ are (q^i, y^α) where $a = y^\alpha e_\alpha(\tau(a))$. In local form, the Lie algebroid structure is determined by the local functions ρ^i_α and $C^\gamma_{\alpha\beta}$ on Q. Both are determined by the relations:

$$\begin{aligned} \rho(e_\alpha) &= \rho^i_\alpha \frac{\partial}{\partial q^i}, \\ [\![e_\alpha, e_\beta]\!] &= C^\gamma_{\alpha\beta} e_\gamma, \end{aligned}$$

and they satisfy the following equations

$$\rho^j_\alpha \frac{\partial \rho^i_\beta}{\partial q^j} - \rho^j_\beta \frac{\partial \rho^i_\alpha}{\partial q^j} = \rho^i_\gamma C^\gamma_{\alpha\beta}, \qquad \sum_{\text{cyclic}\,(\alpha,\beta,\gamma)} \left(\rho_\alpha \frac{\partial C^\nu_{\beta\gamma}}{\partial q^i} + C^\mu_{\beta\gamma} C^\nu_{\alpha\mu} \right) = 0.$$

We present some usual examples of Lie algebroids:

1. Real Lie algebras of finite dimension: Let $\mathfrak{g}$ be a Lie algebra. Then it is clear that $\mathfrak{g}$ is a Lie algebroid over a single point.

2. The tangent bundle: Let TQ be the tangent bundle of a manifold Q. Then, the triple $(TQ, [\cdot,\cdot], \mathrm{Id})$ is a Lie algebroid over Q, where $\mathrm{Id} : TQ \to TQ$ is the identity map.

Lie algebroids morphisms

Let $(A, [\![\cdot,\cdot]\!], \rho)$ (respectively, $(A', [\![\cdot,\cdot]\!]', \rho')$)be a a Lie algebroid over a manifold Q (respectively, Q') and suppose that $\Psi : A \to A'$ is a vector bundle morphism over a map $\Psi_0 : Q \to Q'$. Then, the following diagram is commutative

$$\begin{tikzcd} A \arrow[r, "\Psi"] \arrow[d] & A' \arrow[d] \\ Q \arrow[r, "\Psi_0"] & Q' \end{tikzcd}$$

Now, if $X \in \Gamma(A)$ then

$$\Psi \circ X = \sum_i f_i (X'_i \circ \Psi_0),$$

for suitable $f_i \in C^\infty(Q, \mathbb{R})$ and $X'_i \in \Gamma(A')$. We refer to the previous relation as the $\Psi-$**decomposition of** X.

The pair (Ψ, Ψ_0) is said to be a **Lie algebroid morphism** if

$$\rho' \circ \Psi = T\Psi_0 \circ \rho,$$

$$\begin{aligned} \Psi \circ [\![X, Y]\!] &= \sum_{i,j} f_i g_j ([\![X'_i, Y'_j]\!]) + \sum_j \rho(X)(g_j)(Y'_j \circ \Psi_0) \\ &- \sum_i \rho(Y)(f_i)(X'_i \circ \Psi_0), \end{aligned} \tag{1.3}$$

for $X, Y \in \Gamma(A)$, where $T\Psi_0 : TQ \to TQ'$ is the tangent map of Ψ_0 and

$$\Psi \circ X = \sum_i f_i (X'_i \circ \Psi_0), \quad \Psi \circ Y = \sum_i g_i (Y'_i \circ \Psi_0),$$

are $\Psi-$decompositions of X and Y respectively. The right hand side of (1.3) is independent of the $\Psi-$decompositions of X and Y (for more details, see [64])

If $Q = Q'$, Ψ_0 is the identity map. Then $\Psi \circ X$ is a section of A'. In consequence, (1.3) is equivalent to the condition

$$\Psi \circ [\![X, Y]\!] = [\![\Psi \circ X, \Psi \circ Y]\!]',$$

for $X, Y \in \Gamma(A)$.

The prolongation of a Lie algebroid over a fibration

Let $(A, [\![\cdot,\cdot]\!], \rho)$ be a Lie algebroid over a manifold Q and $\pi : P \to Q$ be a fibration. We consider the subset of $A \times TP$

$$\mathcal{T}^A_p P = \{(b, v) \in A_q \times T_pP \,|\, \rho(b) = T_p\pi(v)\}\,,$$

where $T\pi : TP \to TQ$ is the tangent map to π, $p \in P_q$ and $\pi(p) = q \in Q$. We will frequently use the redundant notation (p, b, v) to denote the element $(b, v) \in \mathcal{T}^A_p P$. $\mathcal{T}^A P$, the union of $\mathcal{T}^A_p P$ over all $p's$, is a vector bundle over P and the vector bundle projection τ^A_P is just the projection over the first factor. The anchor of $\mathcal{T}^A P$ is the projection onto the third factor, that is, the map $\rho^\pi : \mathcal{T}^A P \to TP$ given by $\rho^\pi(p, b, v) = v$. The projection onto the second factor will be denoted by $\mathcal{T}\pi : \mathcal{T}^\pi P \to A$, and it is a morphism of Lie algebroids over π. Explicitly $\mathcal{T}\pi(p, b, v) = b$.

An element $z \in \mathcal{T}^A P$ is said to be **vertical** if it projects to zero, that is $\mathcal{T}\pi(z) = 0$. Therefore it is of the form $(p, 0, v)$, with $\pi-$vertical vector tangent to P at p.

Given local coordinates (q^i, u^E) (here E is just a super(sub)-index) on P and a local basis $\{e_\alpha\}$ of sections of A, we can define a local basis $\{X_\alpha, \Phi_E\}$ of sections of $\mathcal{T}^A P$ by

$$X_\alpha(p) = \left(p, e_\alpha(\pi(p)), \rho^i_\alpha \frac{\partial}{\partial q^i}\Big|_p\right), \quad \Phi_E(p) = \left(p, 0, \frac{\partial}{\partial u^E}\Big|_p\right).$$

If $z = (p, b, v)$ is an element of $\mathcal{T}^A P$, with $b = z^\alpha e_\alpha$, then v is of the form $v = \rho^i_\alpha z^\alpha \frac{\partial}{\partial q^i} + v^E \frac{\partial}{\partial u^E}$, and we can write

$$z = z^\alpha X_\alpha(p) + v^E \Phi_E(p).$$

Vertical elements are linear combinations of $\{\Phi_E\}$.

The anchor map ϕ^π applied to a section Z of $\mathcal{T}^A P$ with local expression $Z = Z^\alpha X_\alpha + V^E \Phi_E$ is the vector field on P whose coordinate expression is

$$\rho^\pi(Z) = \rho^i_\alpha Z^\alpha \frac{\partial}{\partial q^i} + V^E \frac{\partial}{\partial u^E}.$$

Next, we will see that it is possible to induce a Lie bracket structure on the space of sections of $\mathcal{T}^A P$. For that, we say that a section $\tilde{Y}$ of $\tau^A_P : \mathcal{T}^A P \to P$ is **projectable** if there exists a section Y of $\tau : A \to M$ and a vector field $U \in \mathfrak{X}(P)$ which is $\pi-$projectable to the vector field $\rho(X)$ and such that $\tilde{Y}(p) = (Y(\pi(p), U(p)))$, for all $p \in P$. 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 chose a local basis of projectable sections $\Gamma(\tau^A_P)$.

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

$$[\![Z_1, Z_2]\!]^\pi(p) = (p, [\![Y_1, Y_2]\!](x), [U_1, U_2](p)),\ p \in P,\ q = \pi(p).$$

Since any section of $\mathcal{T}^A P$ can be locally written as a linear combination of projectable sections, the definition of a Lie bracket for arbitrary sections of $\mathcal{T}^A P$ follows. In particular, the Lie bracket of the elements of the basis are

$$[\![X_\alpha, X_\beta]\!]^\pi = C^\gamma_{\alpha\beta} X_\gamma,\ [\![X_\alpha, \Phi_E]\!]^\pi = 0,\ [\![\Phi_E, \Phi_B]\!]^\pi = 0,$$

and, therefore, the exterior differential is determined by

$$\mathrm{d}q^i = \rho^i_\alpha X^\alpha, \qquad \mathrm{d}u^E = \Phi^E,$$

$$\mathrm{d}X^\gamma = -\tfrac{1}{2}C^\gamma_{\alpha\beta}X^\alpha \wedge X^\beta, \qquad \mathrm{d}\Phi^E = 0,$$

where $\left\{X^\alpha, \Phi^E\right\}$ is the dual basis of $\{X_\alpha, \Phi_E\}$.

The Lie algebroid $\mathcal{T}^A P$ is called the **prolongation** of A **over** π or the $A-$**tangent bundle to** π. For further details see [64, 105, 127].

1.5.2 Lie groupoids

The global objects corresponding to Lie algebroids are Lie groupoids. We recall the definition of a Lie groupoid and some generalities about them are explained (for more details see [114]). We also discuss some examples which will be interesting in the sequel.

Definition 1.5.3. *A* **groupoid***, denoted by $G \rightrightarrows Q$, consists of two sets G and Q, called respectively the groupoid and the base, together with two maps α and β from G to Q, called respectively the* **source** *and* **target** *projections, a map $\epsilon : Q \to G$, called the* **inclusion***, a* **partial multiplication** *$m : G^{(2)} = \{(g,h) \in G \times G/\alpha(g) = \beta(g)\} \to G$ and a map $i : G \to G$, called the* **inversion***, satisfying the following conditions:*

i) $\alpha(m(g,h)) = \alpha(h)$ *and* $\beta(m(g,h)) = \beta(g)$, *for all* $(g,h) \in G^{(2)}$ *(which is called the* **set of composable pairs***),*

ii) $m(g,m(h,k)) = m(m(g,h),k)$, *for all* $g,h,k \in G$ *such that* $\alpha(g) = \beta(h)$ *and* $\alpha(h) = \beta(k)$,

iii) $\alpha(\epsilon(q))=\beta(\epsilon(q)) = q$, *for all* $q \in Q$,

iv) $m(g,\epsilon(\alpha(g))) = g$ *and* $m(g,\epsilon(\beta(g))) = g$, *for all* $g \in G$,

v) $m(g,i(g)) = \epsilon(\beta(g))$ *and* $m(i(g),g) = \epsilon(\alpha(g))$, *for all* $g \in G$.

If G and Q are manifolds, $G \rightrightarrows Q$ is a **Lie groupoid** if:

i) α and β are differentiable submersions,

ii) m, ϵ, i are differentiable maps.

From the last two conditions, it follows that m is a submersion, ϵ is an immersion and i is a diffeomorphism. In fact, $i^2 = \mathrm{Id}$. From now on, we will usually write gh for $m(g,h)$ and g^{-1} for $i(g)$. Moreover, if $q \in Q$ then $G_q = \alpha^{-1}(q)$ (resp., $G^q = \beta^{-1}(q)$) will be said the $\alpha-$fiber (resp., the $\beta-$fiber) of q. Furthermore, since ϵ is an immersion, we will identify M with $\epsilon(Q)$.

Given a Lie groupoid $G \rightrightarrows Q$ and an element $g \in G$, we can define the **left-translation** $L_g : G_{\beta(g)} \to G_{\alpha(g)}$ and **right-translation** $R_g : G^{\alpha(g)} \to G^{\beta(g)}$ by g defined as

$$L_g(h) = m(g,h) = gh, \;\; R_g(h) = m(h,g) = hg. \tag{1.4}$$

Using these translations, and imitating the case of Lie groups, one may introduce the notion of a left (right)-invariant vector field in a Lie groupoid. Given a Lie groupoid $G \rightrightarrows Q$, a vector field $\xi \in \mathfrak{X}(G)$ is **left-invariant** if ξ is $\alpha-$vertical and $(T_h L_g)(\xi(h)) = \xi(gh)$ for all $(g,h) \in G^{(2)}$. Similarly, ξ is **right-invariant** if ξ is $\beta-$vertical and $(T_h R_g)(\xi(h)) = \xi(hg)$ for all $(h,g) \in G^{(2)}$.

Under these considerations, we may consider the vector bundle $AG \to Q$, whose fiber at a point $q \in Q$ is $A_q G = T_{\epsilon(q)} G$. It is easy to prove that there exists a bijection between the space $\Gamma(AG)$ and the set of left-invariant (respectively, right-invariant) vector fields on G. If X is a section of AG, the corresponding left-invariant (respectively, right-invariant) vector field on G will be denoted by $\overleftarrow{X}$ (respectively, $\overrightarrow{X}$). Using the above facts, we may introduce a Lie algebroid structure $(\llbracket\cdot,\cdot\rrbracket, \rho)$ on AG, which is defined by

$$\overleftarrow{\llbracket X, Y\rrbracket} = [\overleftarrow{X}, \overleftarrow{Y}], \;\; \rho(X)(q) = T_{\epsilon(q)}\alpha(X(q)),$$

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

$$\overrightarrow{\llbracket X, Y\rrbracket} = -[\overrightarrow{X}, \overrightarrow{Y}], \;\; [\overrightarrow{X}, \overleftarrow{Y}] = 0.$$

As a corollary, we can conclude that it is always possible to find a Lie algebroid AG associated to a Lie groupoid $G \rightrightarrows Q$ in the manner we have just showed (the converse is not true [40]).

We introduce now some examples of Lie groupoids:

1. Lie groups: Any Lie group G is a Lie groupoid over $\{e\}$, the identity element of G. The Lie algebroid associated with G is just the Lie algebra $\mathfrak{g}$ of G.

2. The banal groupoid: Let Q be a differentiable manifold. The product manifold $Q \times Q$ is a Lie groupoid over M in the following way: α is the projection onto the second factor and β is the projection onto the first factor; $\epsilon(q,q) = (q,q)$ for all $q \in Q$ and $m((q,q'),(q',q'')) = (q,q'')$. $Q \times Q \rightrightarrows Q$ is called the **banal groupoid**. Its associated Lie algebroid is the tangent bundle TQ of Q.

3. The direct product of Lie groupoids: If $G_1 \rightrightarrows Q_1$ and $G_2 \rightrightarrows Q_2$ are Lie groupoids, then $G_1 \times G_2 \rightrightarrows Q_1 \times Q_2$ is a Lie groupoid in a natural way.

4. Reduced systems: Consider the set $(Q \times Q) \times G$, where G is a Lie group and Q a differentiable manifold. The Lie group structure $(Q \times Q) \times G \rightrightarrows Q$ is given by

$$\begin{aligned}
&\cdot\ \alpha\left(((q,q'),g)\right) = q,\\
&\cdot\ \beta\left(((q,q'),g)\right) = q',\\
&\cdot\ m\left(((q,q'),g),((q',q''),h)\right) = ((q,q''),gh),\\
&\cdot\ \epsilon(q) = ((q,q),e),\\
&\cdot\ i\left(((q,q'),g)\right) = ((q',q),g^{-1}).
\end{aligned}$$

Lie groupoid morphisms

Given two Lie groupoids $G \rightrightarrows Q$ and $G' \rightrightarrows Q'$, a **morphism of Lie groupoids** is a smooth map $\Phi : G \to G'$ such that if $(g,h) \in G^{(2)}$ then $(\Phi(g), \Phi(h)) \in G'^{(2)}$ and $\Phi(gh) = \Phi(g)\Phi(h)$. A morphism of Lie groupoids Φ induces a smooth map $\Phi_0 : Q \to Q'$ in such a way that $\alpha' \circ \Phi = \Phi_0 \circ \alpha$, $\beta' \circ \Phi = \Phi_0 \circ \beta$ and $\Phi \circ \epsilon = \epsilon' \circ \Phi_0$; α, β and ϵ (resp., α', β' and ϵ') being the projections and the inclusion in the Lie groupoid $G \rightrightarrows Q$ (resp., $G' \rightrightarrows M'$). If (Φ, Φ_0) is a morphism between the Lie groupoids $G \rightrightarrows Q$ and $G' \rightrightarrows Q'$, and $AG \to Q$ (resp., $AG' \to Q'$) is the Lie algebroid associated to G (resp., G') then (Φ, Φ_0) induces, in a natural way, a morphism $(A(\Phi), \Phi_0)$ between the Lie algebroids AG and AG' (see [64, 114])

The prolongation of a Lie groupoid over $\tau^* : A^*G \to Q$

An interesting example of Lie groupoid is the **prolongation of a Lie groupoid** over the vector bundle $\tau^* : A^*G \to Q$, which will be used in §5. Given a Lie groupoid $G \rightrightarrows Q$ we may construct the associated Lie algebroid $\tau : AG \longrightarrow Q$, and its dual bundle $\tau^* : A^*G \longrightarrow Q$. Consider the set

$$\mathcal{P}^{\tau^*}G = A^*G \; {}_{\tau^*}\!\times_\alpha G \; {}_\beta\!\times_{\tau^*} A^*G \,.$$

In Reference [118] (see also [154]) this set is called the **prolongation of G over τ^***. $\mathcal{P}^{\tau^*}G$ is a Lie groupoid over A^*G with structure maps

- $\alpha^{\tau^*} : \mathcal{P}^{\tau^*}G \to A^*G; \quad \alpha^{\tau^*}(\mu, g, \mu') = \mu,$
- $\beta^{\tau^*} : \mathcal{P}^{\tau^*}G \to A^*G; \quad \beta^{\tau^*}(\mu, g, \mu') = \mu',$
- $\epsilon^{\tau^*} : G \to \mathcal{P}^{\tau^*}G; \quad \epsilon^{\tau^*}(\mu) = (\mu, \epsilon(\tau^*(\mu)), \mu),$
- $m^{\tau^*} : (\mathcal{P}^{\tau^*}G)^{(2)} \to \mathcal{P}^{\tau^*}G; \quad m^{\tau^*}((\mu, g, \mu'), (\mu', h, \mu'')) = (\mu, gh, \mu''),$
- $i^{\tau^*} : \mathcal{P}^{\tau^*}G \longrightarrow \mathcal{P}^{\tau^*}G; \quad i^{\tau^*}((\mu, g, \mu')) = (\mu', g^{-1}, \mu).$

In the particular case when G is a Lie group we obtain the Lie groupoid $\mathfrak{g}^* \times G \times \mathfrak{g}^* \rightrightarrows \mathfrak{g}^*$.

Now, denote by $\tau : AG \longrightarrow Q$ the Lie algebroid associated to G, and by $A(\mathcal{P}^{\tau^*}G) \to A^*G$ the Lie algebroid associated to $\mathcal{P}^{\tau^*}G$. It is easy to show that

$$\begin{aligned} A_\mu(\mathcal{P}^{\tau^*}G) \;\; &\equiv \;\; \{(a_{\epsilon(\tau^*(\mu))}, Y_\mu) \in A_{\tau^*(\mu)}G \times T_\mu A^*G \mid \\ &\qquad (T_\mu \tau^*)(Y_\mu) = (T_{\epsilon(\tau^*(\mu))}\beta)(a_{\epsilon(\tau^*(\mu))})\}, \end{aligned}$$

where, obviously, $a_{\epsilon(\tau^*(\mu))} \in A_{\tau^*(\mu)}G$ and $\mu \in A^*G$. A section $Z \in \Gamma(A(\mathcal{P}^{\tau^*}G))$ is expressed as

$$Z(\mu) = (X(\tau^*(\mu)), Y(\mu)),$$

where $X \in \Gamma(AG)$ and $Y \in \mathfrak{X}(A^*G)$ verify that $T\beta(X) = T\tau^*(Y)$.

Therefore, it is easy to show, from the definition of m^{τ^*} that

$$\begin{aligned} \overleftarrow{Z}(\mu, g, \mu') \;\; &= \;\; (-Y(\mu), \overleftarrow{X}(g), \mathbf{0}_{\mu'}) \\ \overrightarrow{Z}(\mu, g, \mu') \;\; &= \;\; (\mathbf{0}_\mu, \overrightarrow{X}(g), Y(\mu')) \end{aligned}$$

Chapter 2

Continuous Lagrangian and Hamiltonian mechanics

Classical mechanics deals with the dynamics of particles, rigid bodies, continuous media (fluid, plasma, and solid mechanics), and other fields (such as electromagnetism, gravity, etc.) This theory also plays a crucial role in quantum mechanics, in control theory and other areas of physics, engineering and even chemistry and biology. It begins with a long tradition of qualitative investigation culminating with Kepler and Galileo. Following this is the period of quantitative theory, represented by the works of Newton, Euler, Lagrange, Laplace, Hamilton and Jacobi. The neoqualitative period began with Poincaré and lasts to the present. It consists primarily in the amplification of the geometric methods of Poincaré, the application of these methods to the qualitative questions of the previous period and the consideration of new qualitative questions that could not previously be asked.

Throughout history, mechanics has also played a key role in the development of mathematics. Starting with the creation of calculus stimulated by Newton's mechanics, it continues today with exciting developments in group representations, geometry and topology; these mathematical developments in turn are being applied to interesting problems in physics and engineering. An instance of this role is the mentioned work by Poincaré, which culminated in modern differential geometry and topology.

Mechanics has two main branches, **Lagrangian mechanics** and **Hamiltonian mechanics**. In one sense, Lagrangian mechanics is more fundamental since it is based on variational principles and it is what generalizes most directly to the general relativistic context. In another sense, Hamiltonian mechanics is more fundamental, since it is based directly on the energy concept and it is what is more closely tied to quantum mechanics. Fortunately, in many cases these branches are equivalent.

This chapter is devoted to develop both branches from geometrical and variational perspectives.

2.1 Lagrangian mechanics

The Lagrangian formulation of mechanics is for simplicity set in a finite dimensional manifold (the infinite dimensional case is treated for instance in [65, 123]), which will be denoted by Q, called the **configuration space**, whose tangent bundle TQ describes the states (position and velocity) of the system. Due to the important role that TQ is going to play in this representation of mechanics, we detail in §2.1.1 the basics of its geometry.

2.1.1 The geometry of tangent bundle

Through this subsection, Q denotes an n-dimensional smooth manifold. Local coordinates in Q are denoted $\left(q^i\right)$, $i = 1, ..., n$, and the induced adapted coordinates of TQ and TTQ (we recall here that TQ is also a manifold and consequently we can consider its tangent bundle) are denoted (q^i, v^i) and $(q^i, v^i, \dot{q}^i, \dot{v}^i)$ respectively. If we restrict our system to TQ, along this thesis the coordinates v^i will be also denoted by $\dot{q}^i$ making no difference. According to this, vectors $v_q \in T_qQ$ and $V_q \in T_{v_q}(TQ)$ are of the form

$$v_q = v^i \frac{\partial}{\partial q^i}\Big|_q \quad \text{and} \quad V_q = \dot{q}^i \frac{\partial}{\partial q^i}\Big|_{v_q} + \dot{v}^i \frac{\partial}{\partial q^i}\Big|_{v_q}.$$

As was established before $\tau_Q : v_q \in T_qQ \mapsto q \in Q$ denotes the natural projection of TQ onto Q. Then, given a tangent vector $V_q \in T_{v_q}TQ$, we have that $\tau_{TQ}(V_q) = v_q$. The commutative projection structure is defined by the following diagram:

$$\begin{tikzcd} TTQ \arrow[r, "T\tau_Q"] \arrow[d, "\tau_{TQ}"'] & TQ \arrow[d, "\tau_Q"] \\ TQ \arrow[r, "\tau_Q"'] & Q, \end{tikzcd}$$

which can be expressed in local coordinates as:

$$\tau_Q(q^i, v^i) = (q^i), \;\; \tau_{TQ}(q^i, v^i, \dot{q}^i, \dot{v}^i) = (q^i, v^i) \;\text{ and }\; T\tau_Q(q^i, v^i, \dot{q}^i, \dot{v}^i) = (q^i, \dot{q}^i).$$

Definition 2.1.1. *Let $v_q \in T_qQ$. The* **vertical lift** *of v_q at $w_q \in T_qQ$ is the tangent vector $v^{\vee}_{w_q} \in T_{w_q}TQ$ given by*

$$v^{\vee}_{w_q}(f) = \frac{d}{dt} f(w_q + tv_q)|_{t=0}, \;\; \forall f \in \mathcal{C}^\infty(T_qQ).$$

Given a smooth function $g \in \mathcal{C}^\infty(Q)$,

$$\begin{aligned} (T_w\tau_Q)\left(v^{\vee}_w\right)(f) &= v^{\vee}_w(g \circ \tau_q) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(g \circ \tau_Q)(w + tv)|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} g(q)|_{t=0} \\ &= 0, \end{aligned}$$

where, for convenience, we have omitted the subscript q in the elements belonging to T_qQ. The vertical lift may also be seen as a morphism $X \in \mathfrak{X}(Q) \mapsto X^\vee \in \mathfrak{X}^\vee(TQ)$, where $\mathfrak{X}^\vee(TQ)$ is the module of vector fields over TQ that are vertical with respect to the projection τ_Q. In local coordinates, if $v = (q^i, v^i)$ and $w = (q^i, w^i)$, then

$$v_w^\vee = (q^i, w^i, 0, v^i)$$

for the induced adapted local coordinates of TTQ.

Definition 2.1.2. *The* **vertical endomorphism** *is the linear map $S : TTQ \to TTQ$ that, for any vector $V \in TTQ$, gives the value*

$$S(V) = ((T_v\tau_Q)(V))^\vee ,$$

where $v = \tau_{TQ}(V) \in TQ$.

In adapted coordinates (q^i, v^i) of TQ, the vertical endomorphism has the local expression

$$S = \mathrm{d}q^i \otimes \frac{\partial}{\partial v^i} \quad \text{or} \quad S(q^i, v^i, \dot{q}^i, \dot{v}^i) = (q^i, v^i, 0, \dot{q}^i).$$

Definition 2.1.3. *The* **Liouville** *or* **dilation vector field** *is the vector field $\triangle \in \mathfrak{X}(TQ)$ defined by*

$$\triangle_v = \left(v^\vee\right)_v ,$$

for any $v \in TQ$.

Another way to define the Liouville vector field is as the infinitesimal generator of the 1-parameter group of transformations $\phi_t : v \in TQ \mapsto e^t v \in TQ$. This definition can be easily translated to any vector bundle.

Definition 2.1.4 (SODE). *A* **second order vector field** *or* **differential equation** *(usually abbreviated SODE) is a vector field $X \in \mathfrak{X}(TQ)$ such that $T\tau_Q \circ X = \mathrm{Id}_{TQ}$.*

In adapted coordinates (q^i, v^i) of TQ, a SODE is a vector field

$$X = X^i \frac{\partial}{\partial q^i} + Y^i \frac{\partial}{\partial v^i} \ \text{ such that } \ X^i = v^i.$$

Thus, neither the Liouville vector field nor the vertical lift of a vector field are SODEs. Nevertheless, SODEs are characterized by the equation

$$S(X) = \triangle.$$

Definition 2.1.5. *Given a smooth curve $c : I \to Q$, its first* **lift** *to TQ is the smooth curve $c^{(1)} : I \to TQ$ such that*

$$\left(c^{(1)}(t_0)\right)(f) = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ c)|_{t=t_0}.$$

In local adapted coordinates $c^{(1)} = (c^i, \frac{\mathrm{d}c^i}{\mathrm{d}t})$. Sometimes we will use the notation $c^{(1)} = \dot{c}$.

Proposition 2.1.6. *A vector field $X \in \mathfrak{X}(TQ)$ is a SODE if and only if the integral curves of X are lifts of their own projections to Q; that is, if $\tilde{c}$ is an integral curve of X, then*

$$\tilde{c} = (\tau_Q \circ \tilde{c})^{(1)} .$$

The curve $c = \tau_Q \circ \tilde{c} : I \to Q$ is called a **base integral** *curve of X or a* **solution** *of the SODE given by X. In that sense, the previous equation can be rewritten as $\tilde{c} = c^{(1)}$.*

If $\tilde{c} : I \to TQ$ is an integral curve of a SODE $X \in \mathfrak{X}(TQ)$ locally given by $X = (q^i, v^i, v^i, a^i)$ and $c : I \to Q$ denotes its base integral curve, then along $\tilde{c}(t)$ X satisfies the following:

$$q^i = c^i, \quad v^i = \frac{\mathrm{d}c^i}{\mathrm{d}t} \quad \text{and} \quad a^i = \frac{\mathrm{d}^2 c^i}{\mathrm{d}t^2}.$$

Alternatively, the base integral curve c of $\tilde{c}$ satisfies the system of second order differential equations

$$\frac{\mathrm{d}^2 c^i}{\mathrm{d}t^2} = a^i(c^i, \frac{\mathrm{d}c^i}{\mathrm{d}t}),$$

which intrinsically can be defined by

$$\frac{\mathrm{d}\tilde{c}}{\mathrm{d}t} = X(\tilde{c}(t)),$$

where X is a SODE.

2.1.2 Variational approach to Lagrangian formalism

The main objective of classical mechanics is to seek for trajectories that describe the motion of our system. It is well-known that there exists a variational procedure to obtain these trajectories. Thus, we will consider twice differentiable curves $c : [0, T] \to Q$ joining two fixed points $q_0, q_1 \in Q$. The set of such curves is denoted by

$$\mathcal{C}^2([0,T], Q, q_0, q_1) = \left\{ c \in \mathcal{C}^2([0,T], Q) \mid c(0) = q_0 \, , \, c(T) = q_1 \right\} , \tag{2.1}$$

or $\mathcal{C}^2(q_0, q_1)$ as a shorthand notation. Given $c \in \mathcal{C}^2(q_0, q_1)$, denote by $c^{(1)}$ its lift to TQ (see definition 2.1.5). If (q^i, v^i) are adapted coordinates to TQ, then $c^{(1)}(t) = (q^i(t), v^i(t))$. Locally,

$$c^{(1)}(t) = (c^i(t), \dot{c}^i(t)),$$

where $c^i(t) = (q^i \circ c^{(1)})(t)$ and $\dot{c}^i(t) = (v^i \circ c^{(1)})(t) = (\mathrm{d}c^i/\mathrm{d}t)(t)$.

One of the main ingredients of Lagrangian mechanics is, obviously, the Lagrangian function $L : TQ \to \mathbb{R}$, which is a smooth function defined usually as the kinetic energy minus the potential energy of the system. Given the Lagrangian function, two fixed points $q_0, q_1 \in Q$ and a fixed time interval $[0, T]$, the associated **action integral** is the real valued map $\mathcal{A}_L$ defined on $\mathcal{C}^2([0,T], Q, q_0, q_1)$ given by

$$\mathcal{A}_L(c) = \int_0^T L(c^{(1)}(t)) \, \mathrm{d}t = \int_0^T L(q^i(t), v^i(t)) \, \mathrm{d}t. \tag{2.2}$$

As will be presented soon, the equations of motion of a system described by the Lagrangian function L are obtained by variational methods. Therefore, we must describe how $\mathcal{A}_L$ changes under small variations of c and what these variations are. It can be shown that $\mathcal{C}^2(q_0, q_1)$ may be endowed with a infinite-dimensional smooth manifold structure (see [13] for more details). Namely

$$T_c\mathcal{C}^2(q_0, q_1) = \left\{\delta c \in \mathcal{C}^1([0, T], TQ) \mid \tau_Q \circ \delta c \equiv c,\ \delta c(0) = \delta c(T) = 0\right\}. \tag{2.3}$$

Definition 2.1.7. *Let $c \in \mathcal{C}^2(q_0, q_1)$, a* **variation** *of c is a curve $c_s \in \mathcal{C}^2(q_0, q_1)$, $s \in [-\epsilon, \epsilon]$ for $\epsilon > 0$ belonging to $\mathbb{R}$, such that $c_0 \equiv c$. An* **infinitesimal variation** *of c is a vector field δc along c which vanishes at the endpoints $\delta c(0) = \delta c(T) = 0$.*

Taking into account this definition, the tangent space $T_c\mathcal{C}^2(q_0, q_1)$ at a curve $c \in \mathcal{C}^2(q_0, q_1)$ is the set of infinitesimal variations δc of c, which are induced by variations c_s of c. More precisely,

$$\delta c(t) = \frac{\mathrm{d}c_s(t)}{\mathrm{d}s}\Big|_{s=0},$$

where $t \in [0, T]$ is fixed.

Definition 2.1.8. *Let $\mathcal{F} : \mathcal{C}^2(q_0, q_1) \to \mathbb{R}$ be a functional of class $\mathcal{C}^1$. A* **critical point** *of $\mathcal{F}$ is a curve $c \in \mathcal{C}^2(q_0, q_1)$ such that*

$$\frac{d(\mathcal{F} \circ c_s)}{ds}\Big|_{s=0} = 0,$$

for any variation c_s of c.

Equivalently, c is a critical point of $\mathcal{F}$ if and only if $\mathrm{d}\mathcal{F}(c) \cdot \delta c = 0$ for any infinitesimal variation δc.

Under these considerations, we are able to formulate the variational Hamilton's principle, which states that the dynamics of our physical system is determined from the variational problem related to the integral action $\mathcal{A}_L$:

Theorem 2.1.9 (Continuous Hamilton's principle). *The* **motion** *of a particle in the physical system defined by the Lagrangian function $L : TQ \to \mathbb{R}$ is a critical point of the action integral $\mathcal{A}_L$, that is, a curve $c \in \mathcal{C}^2(q_0, q_1)$ such that $\delta\mathcal{A}_L(c) = 0$.*

An easy calculation helps us to write the Hamilton's principle in terms of the Lagrangian function, giving also the well-known Euler-Lagrange equations.

Theorem 2.1.10. *Consider a given Lagrangian system where $L \in C^2(TQ)$. A twice differentiable curve $c : [0, T] \to Q$ joining $q_0, q_1 \in Q$ is a motion of the system if and only if the lift $c^{(1)}$ of c satisfies the differential equations*

$$\frac{\partial L}{\partial q^i} \circ c^{(1)} - \frac{d}{dt}\left(\frac{\partial L}{\partial v^i} \circ c^{(1)}\right) = 0, \tag{2.4}$$

where (q^i, v^i) are adapted local coordinates in a neighborhood of $c^{(1)}$.

The proof will be shown in §2.4 within the context of constrained systems. See [1, 123] for further details.

2.1.3 Geometric approach to Lagrangian formalism

Definition 2.1.11. *The* **Poincaré-Cartan** *1-form is the pullback of the differential of the Lagrangian function by the vertical endomorphism $\mathcal{S}$ (see definition 2.1.2), namely*

$$\Theta_L = S^*(\mathrm{d}L). \tag{2.5}$$

The Poincaré-Cartan 2-form is then given by

$$\Omega_L = -\mathrm{d}\Theta_L. \tag{2.6}$$

In local and adapted coordinates, the Poincaré-Cartan 1-form reads

$$\Theta_L = \frac{\partial L}{\partial v^i}\mathrm{d}q^i,$$

while the Poincaré-Cartan 2-form

$$\Omega_L = \frac{\partial^2 L}{\partial v^i \partial q^j}\mathrm{d}q^i \wedge \mathrm{d}q^j + \frac{\partial^2 L}{\partial v^i \partial v^j}\mathrm{d}q^i \wedge \mathrm{d}v^j.$$

The Poincaré-Cartan 2-form is exact by definition and hence closed. It is non-degenerate if and only if the Lagrangian function is regular, that is, when the Hessian matrix $\left(\frac{\partial^2 L}{\partial v^i \partial v^j}\right)$ is invertible.

Definition 2.1.12. *The* **Lagrangian energy** *is the smooth function $E_L \in C^\infty(TQ)$ defined by*

$$E_L = \Delta L - L,$$

where Δ denotes the Liouville vector field given in definition 2.1.3.

Definition 2.1.13. *Any vector field $X_L \in \mathfrak{X}(TQ)$ that satisfies the following equation*

$$i_{X_L}\Omega_L = \mathrm{d}E_L, \tag{2.7}$$

is called a **Lagrangian vector field***.*

Theorem 2.1.14. *If the Lagrangian function L is regular, then there exists a unique vector field $X_L \in \mathfrak{X}(TQ)$ which is solution of (2.7). The Lagrangian vector field X_L is a second order differential equation and its integral curves are solutions of the Euler-Lagrange equations (2.4).*

Proof. The existence and uniqueness of a Lagrangian vector field comes out from the fact that Ω_L is non-degenerate when L is regular, hence Ω_L is symplectic. Let X_L be a generic vector field on TQ whose local expression is

$$X = \dot{q}^i\frac{\partial}{\partial q^i} + \dot{v}^i\frac{\partial}{\partial v^i},$$

for adapted coordinates (q^i, v^i) of TQ. Let suppose also that X_L satisfies the equation (2.7). The contraction of the Poincaré-Cartan 2-form Ω_L with X_L reads

$$i_{X_L}\Omega_L = \left(\dot{q}^j \frac{\partial^2 L}{\partial v^j \partial q^i} - \dot{q}^j \frac{\partial^2 L}{\partial v^i \partial q^j} - \dot{v}^j \frac{\partial^2 L}{\partial v^j \partial v^i}\right) \mathrm{d}q^i + \dot{q}^j \frac{\partial^2 L}{\partial v^j \partial v^i} \mathrm{d}v^i,$$

while the differential of the Lagrangian energy E_L is

$$\mathrm{d}E_L = (v^j \frac{\partial^2 L}{\partial v^j \partial q^i} - \frac{\partial L}{\partial q^i})\mathrm{d}q^i + v^j \frac{\partial^2 L}{\partial v^j \partial v^i} \mathrm{d}v^i.$$

Equating coefficients, we have in one hand that

$$\dot{q}^j \frac{\partial^2 L}{\partial v^j \partial v^i} = v^j \frac{\partial^2 L}{\partial v^j \partial v^i}.$$

Thus, if L is regular, $\dot{q}^j = v^j$, which provides that X_L is second order. We will use v^i and $\dot{q}^i$ without distinction in the sequel whenever we consider regularity conditions. On the other hand, since L is regular, we have that

$$\frac{\partial L}{\partial q^i} - v^j \frac{\partial^2 L}{\partial q^j \partial v^i} - \dot{v}^j \frac{\partial^2 L}{\partial v^j \partial v^i} = 0.$$

Let $c : I \to Q$ be a base integral curve of the Lagrangian vector field X_L. Then, $\dot{q}^i = \dot{c}^i = \mathrm{d}c/\mathrm{d}t$ and $\dot{v}^i = \ddot{c}^i = \mathrm{d}^2c/\mathrm{d}t^2$. Replacing this into the previous equation and denoting $c^{(1)} = (c^i, \dot{c}^i)$ the lift of c to TQ, we obtain

$$0 = \frac{\partial L}{\partial q^i} \circ c^{(1)} - \left(\frac{\mathrm{d}c^j}{\mathrm{d}t}\right) \frac{\partial^2 L}{\partial q^j \partial v^i} \circ c^{(1)} - \left(\frac{\mathrm{d}c^j}{\mathrm{d}t}\right) \frac{\partial^2 L}{\partial v^j \partial v^i} \circ c^{(1)} = \frac{\partial L}{\partial q^i} \circ c^{(1)} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^i} \circ c^{(1)}\right),$$

which are precisely the Euler-Lagrange equations (2.4). □

Lagrangian vector fields and flows

The **Lagrangian vector field** $X_L : TQ \to TTQ$ is a second-order vector field on TQ satisfying equation (2.7), and the **Lagrangian flow** $F_L : TQ \times \mathbb{R} \to TQ$ is the flow of X_L. We shall write $F_L^t : TQ \to TQ$ for the map F_L at a frozen time t. A curve $q \in \mathcal{C}^2(q_0, q_1)$ (we recall that $\mathcal{C}^2(q_0, q_1)$ is defined in §2.1.2) is said to be a solution of the Euler-Lagrange equations (2.4) if it is an integral curve of X_L.

Symplecticity of Lagrangian flows

From equation (2.7) we immediately have that $\mathcal{L}_{x_L}\Omega_L = 0$ and therefore $(F_L^t)^*\Omega_L = \Omega_L$, which implies the preservation of the Poincaré-Cartan 2-form by the Lagrangian flow F_L^t. In the following lines we give a variational deduction of this preservation which will be useful in the discrete formalism.

Define the **solution space** $\mathcal{C}_L(Q) \subset \mathcal{C}^2(q_0, q_1)$ to be the set of curves solution of the Euler-Lagrange equations given a regular Lagrangian. As an element $q \in \mathcal{C}_L(Q)$ is an integral curve of X_L, it is uniquely determined by the initial (local) condition $(q^i(0), v^i(0) \in TQ$, where $(q^i(t), v^i(t))$ are the local coordinates of q. Thus, we can identify $\mathcal{C}_L(Q)$ with the space of initial conditions in TQ. Note that, with some abuse of notation, we are using the same symbol for the curve q and its configuration coordinates $q^i(t)$.

Defining the **restricted action sum** $\hat{\mathcal{A}}_L : TQ \to \mathbb{R}$ to be

$$\hat{\mathcal{A}}_L(v_q) = \mathcal{A}_L(q), \quad q \in \mathcal{C}_L(Q) \text{ and } (q^i(0), v^i(0)) \in TQ,$$

where $\mathcal{A}_L$ is defined in (2.2). We see that the Hamilton's principle $\delta\mathcal{A}_L(q)$ reduces to

$$\begin{aligned} \langle \mathrm{d}\hat{\mathcal{A}}_L(v_q), w_{v_q} \rangle &= \langle \Theta_L(v_q(T)), (F^T)_* w_{v_q} \rangle - \langle \Theta_L(v_q(0)), w_{v_q} \rangle \\ &= \langle \left((F_L^T)^* \Theta_L\right)(v_q(0)), w_{v_q} \rangle - \langle \Theta_L(v_q(0)), w_{v_q} \rangle, \end{aligned}$$

for all $w_{v_q} \in T_{v_q}TQ$. Taking a further derivative of this expression, and using the fact that $\mathrm{d}^2\hat{\mathcal{A}}_L = 0$, we obtain the already mentioned preservation property

$$(F_L^T)^* \Omega_L = \Omega_L.$$

Momentum map preservation

Suppose that the Lie group G, with Lie algebra $\mathfrak{g}$, acts on Q by the (left or right) action $\Phi : G \times Q \to Q$. Consider the tangent lift of this action to $\Phi^{TQ} : G \times TQ \to TQ$ given by $\Phi_g^{TQ}(v_q) = T\Phi_g(v_q)$ which locally is

$$\Phi^{TQ}(g, q^i, v^i) = \left(\Phi(g, q^i), \frac{\partial \Phi^i}{\partial q^j}(g, q^i) v^j \right),$$

where $g \in G$ and (q^i, v^i) are the local coordinates of $v_q \in TQ$. For $\xi \in \mathfrak{g}$ define the infinitesimal generators (as is already done in definition 1.4.10) $\xi_Q : Q \to TQ$ and $\xi_{TQ} : TQ \to TTQ$ by

$$\begin{aligned} \xi_Q(q) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi(\exp(t\xi), q), \\ \xi_{TQ}(v_q) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi^{TQ}(\exp(t\xi), v_q), \end{aligned}$$

where $t \in \mathbb{R}$ and $\exp : \mathfrak{g} \to G$ is the usual exponential map (see definition 1.4.10). We now define the **Lagrangian momentum map** $J_L : TQ \to \mathfrak{g}^*$ to be

$$\langle J_L(v_q), \xi \rangle = \langle \Theta_L, \xi_{TQ}(v_q) \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the natural pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$ in the left hand side of the equation, whereas it denotes the natural pairing between T^*TQ and TTQ in the right hand side.

A Lagrangian $L : TQ \to \mathbb{R}$ is said to be **invariant** under the action $\Phi^{TQ} : G \times TQ \to TQ$, that is

$$L \circ \Phi_g^{TQ} = L, \ \ \forall\, g \in G,$$

and in this case the group action is said to be a **symmetry** of the Lagrangian. Differentiating this expression implies that the Lagrangian is **infinitesimally invariant**, which is the statement $\langle \mathrm{d}L, \xi_{TQ} \rangle = 0$ for all $\xi \in \mathfrak{g}$. We will now show that, when the group action is a symmetry of the Lagrangian, then the momentum maps are preserved by the Lagrangian flow. This result was originally due to Noether (1918).

Theorem 2.1.15 (Noether's theorem). *Consider a Lagrangian system $L : TQ \to \mathbb{R}$ which is invariant under the lift of the left (or right) action $\Phi : G \times Q \to Q$. Then, the corresponding Lagrangian momentum map $J_L : TQ \to \mathfrak{g}^*$ is a conserved quantity of the flow, so that $J_L \circ F_L^t = J_L$ for all times t.*

Proof. The action of G on Q induces an action of G on the space of paths $\mathcal{C}^2(q_0, q_1)$ by pointwise action, so that $\Phi_g : \mathcal{C}^2(q_0, q_1) \to \mathcal{C}^2(gq_0, gq_1)$ is given by $\Phi_g(q)(t) = \Phi_g(q(t))$, for $q \in Q$ and $q(t) \in \mathcal{C}^2(q_0, q_1)$. As the action is just the integral of the Lagrangian, invariance of L implies invariance of $\mathcal{A}_L$, and the differential of this gives

$$\langle \mathrm{d}\mathcal{A}_L, \xi_{\mathcal{C}^2(q_0,q_1)}(q) \rangle = \int_0^T \langle \mathrm{d}L, \xi_{TQ} \rangle \mathrm{d}t = 0.$$

Invariance of $\mathcal{A}_L$ also implies that Φ_g maps solution curves to solution curves and thus $\xi_{\mathcal{C}^2(q_0,q_1)}(q) \in T_q\mathcal{C}_L$, which is the corresponding infinitesimal version. We can thus restrict $\langle \mathrm{d}\mathcal{A}_L, \xi_{\mathcal{C}^2(q_0,q_1)} \rangle$ to the space of solutions $\mathcal{C}_L(Q)$ to obtain

$$0 = \langle \mathrm{d}\hat{\mathcal{A}}_L, \xi_{TQ}(v_q) \rangle = \langle \Theta_L(v_q(T)), \xi_{TQ}(v_q(T)) \rangle - \langle \Theta_L(v_q(0)), \xi_{TQ}(v_q(0)) \rangle.$$

Replacing in the definition of J_L shows that this is just

$$0 = \langle J_L(F_L^T(v_q)), \xi \rangle - \langle J_L(v_q), \xi \rangle,$$

which gives the desired result. □

We have thus seen that conservation of momentum maps is a direct consequence of the invariance of the variational principle under a symmetry action.

2.2 Hamiltonian mechanics

Equivalently to the Lagrangian formalism, the Hamiltonian formulation of mechanics is set up in a finite dimensional manifold Q, the **configuration space**, but, in contrast, the positions and momenta of the system under study are described by the cotangent bundle T^*Q of Q. The local coordinates (q^i) of Q induce fiber coordinates (q^i, p_i) on T^*Q.

Let $H : TQ \to \mathbb{R}$ be the **Hamiltonian function** (smooth) of the system, which from the physical point of view is the total energy of the system under study. Typically, the Hamiltonian is defined as the sum of kinetic and potential energies of the system.

Definition 2.2.1. *Given a Hamiltonian function $H : T^*Q \to \mathbb{R}$, the* **Hamiltonian vector field** *associated to H is the unique vector field $X_H \in \mathfrak{X}(T^*Q)$ such that,*

$$i_{X_H}\Omega_Q = \mathrm{d}H, \tag{2.8}$$

where Ω_Q is the canonical symplectic form of T^*Q.

Notice that the nondegeneracy of Ω_Q guarantees that X_H exists. The canonical symplectic form can be geometrically defined in the following way. Define the **canonical 1-form** or **Liouville's 1-form** Θ_Q on T^*Q by

$$\langle \Theta_Q(\alpha_q), v_{\alpha_q} \rangle = \langle \alpha_q, T\pi_Q(v_{\alpha_q}) \rangle,$$

where $\alpha_q \in T^*Q$, $v_{\alpha_q} \in TT^*Q$, $\pi_Q : T^*Q \to Q$ is the standard projection. In local coordinates, $\Theta_Q = p_i \mathrm{d}q^i$. The canonical 2-form Ω_Q is defined to be

$$\Omega_Q = -\mathrm{d}\Theta_Q,$$

which has the local expression $\Omega_Q = \mathrm{d}q^i \wedge \mathrm{d}p_i$ given by the Darboux theorem.

Theorem 2.2.2. *A differentiable curve $c : I \to T^*Q$ is an integral curve of X_H if and only if the* **Hamilton's equations** *hold:*

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \tag{2.9}$$

where $c(t) = (q^i(t), p_i(t))$.

As mentioned above, the Hamiltonian function represents the total energy of the system. Thus, if H is autonomous (i.e. there is no explicit dependence on time) the total energy must be preserved.

Proposition 2.2.3. *Given an ingegral curve $c(t)$ of X_H, we have that $H(c(t))$ is constant.*

Proposition 2.2.4. *Let $F_t \in \mathrm{Diff}(T^*Q)$ be the flow of X_H, then $F_t^*\Omega_Q = \Omega_Q$ for each t, i.e. F_t is a family of symplectomorphisms.*

Proof. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} F_t^*\Omega_Q = \mathcal{L}_{X_H}\Omega_Q = i_{X_H}\mathrm{d}\Omega_Q + \mathrm{d}i_{X_H}\Omega_Q.$$

Since $\mathrm{d}\Omega_Q = 0$ and (2.8) holds, we arrive to

$$\frac{\mathrm{d}}{\mathrm{d}t} F_t^*\Omega_Q = \mathrm{d}\mathrm{d}H = 0.$$

Thus F_t^* is constant in t. Since $F_0 = \mathrm{Id}$, the equation $F_t^*\Omega_Q = \Omega_Q$ results. □

In the previous proof the definition of the Lie derivative and some of its properties, described in §1.1.1, have been used.

For more insight in Hamiltonian mechanics, see [1, 123].

Hamiltonian form of Noether's theorem

Consider a (left or right) action $\Phi : G \times Q \to Q$. The cotangent lift of this action is $\Phi^{T^*Q} : G \times T^*Q \to T^*Q$ given by $\Phi_g^{T^*Q}(p_q) = \Phi^*_{g^{-1}}(p_q)$, where $g \in G$ and $p_q \in T^*Q$. In local coordinates

$$\Phi^{T^*Q}(g,(q^i,p_i)) = \left((\Phi_g^{-1})^i(q), p_j \frac{\partial \Phi_g^j}{\partial q^i}(q) \right),$$

where (q^i, p_i) are the local coordinates of $p_q \in T_q^*Q$. This has its corresponding **infinitesimal generator** $\xi_{T^*Q} : T^*Q \to TT^*Q$ defined by

$$\xi_{T^*Q}(p_q) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi^{T^*Q}(\exp(t\xi), p_q).$$

The **Hamiltonian momentum map** $J_H : T^*Q \to \mathfrak{g}^*$ is defined by

$$\langle J_H(p_q), \xi \rangle = \langle \Theta_Q(p_q), \xi_{T^*Q}(p_q) \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the natural pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$ in the left hand side of the equation, whereas it represents the natural paiting between TT^*Q and T^*T^*Q in the right hand side. For each $\xi \in \mathfrak{g}$ we define $J_H^\xi : T^*Q \to \mathbb{R}$ by $J_H^\xi(p_q) = \langle J_H(p_q), \xi \rangle$, which has the intrinsic expression $J_H^\xi = i_{\xi_{T^*Q}} \Theta_Q$.

A Hamiltonian $H : T^*Q \to \mathbb{R}$ is said to be **invariant** under the cotangent lift of the action $\Phi : G \times Q \to Q$ if

$$H \circ \Phi_g^{T^*Q} = H, \;\; \forall\, g \in G,$$

in which case the action is said a **symmetry** for the Hamiltonian. The derivative of this expression implies that such a Hamiltonian is also **infinitesimally invariant**, which is the requirement $\langle \mathrm{d}H, \xi_{T^*Q} \rangle = 0$ for all $\xi \in \mathfrak{g}$, although the converse is not generally true.

Theorem 2.2.5 (Hamiltonian Noether's theorem). *Let $H : T^*Q \to \mathbb{R}$ be a Hamiltonian which is invariant under the lift of the (left or right) action $\Phi : G \times Q \to Q$. Then the corresponding Hamiltonian momentum map $J_H : T^*Q \to \mathfrak{g}^*$ is a conserved quantity of the flow; that is, $J_H \circ F_H^t = J_H$ for all times t.*

See [1, 124] for the proof. Noether's theorem still holds even if the Hamiltonian is only infinitesimally invariant.

2.2.1 The Legendre transformation

Let us now give the relationship between the Lagrangian formulation on TQ and the Hamiltonian formulation on T^*Q. In fact, they are equivalent in the **hyperregular** case, and are transformed one into the other by the Legendre transform

Definition 2.2.6. *Given a Lagrangian function $L : TQ \to \mathbb{R}$, the* **Legendre transformation** *associated to L is the fibered mapping $\mathbb{F}L : TQ \to T^*Q$ defined implicitly by*

$$\langle \mathbb{F}L(v), w \rangle = \frac{d}{ds}\Big|_{s=0} L(v + sw),$$

where $v, w \in TQ$.

If (q^i, v^i) and (q^i, p_i) denote fibered coordinates on TQ and T^*Q respectively, then the local expression of the Legendre transform is

$$\mathbb{F}L(q^i, v^i) = \left(q^i, p_i = \frac{\partial L}{\partial v^i}\right).$$

Remark 2.2.7. *Note that the* **Legendre transformation** *is a* **fiber derivative** *in the sense shown in §1.1.1. That is, let Q be a manifold and $L : TQ \to \mathbb{R}$ a function on its tangent bundle. Then the map $\mathbb{F}L : TQ \to T^*Q : w_q \mapsto DL_q(w_q) \in \mathrm{Lin}(T_qQ, \mathbb{R}) = T_q^*Q$ is called the* **fiber derivative** *of L. Here, L_q denotes the restriction of L to the fiber over $q \in Q$.*

Proposition 2.2.8. *If L is regular, then $\mathbb{F}L : TQ \to T^*Q$ is a local diffeomorphism.*

Proof. Employing the Inverse Function Theorem, is easy to see that $\mathbb{F}L$ is a local diffeomorphism if

$$\frac{\partial p_i}{\partial v^j} = \frac{\partial^2 L}{\partial v^i \partial v^j}$$

is invertible, which is equivalent to say that L is regular. □

Definition 2.2.9. *A Lagrangian function $L : TQ \to \mathbb{R}$ is said to be hyper-regular whenever $\mathbb{F}L$ is a global diffeomorphism.*

The following theorem shows the relationship of Lagrangian and Hamiltonian formalisms via the Legendre transformation.

Theorem 2.2.10. *Let L be a hyperregular Lagrangian and let $H = E_L \circ (\mathbb{F}L)^{-1} : T^*Q \to \mathbb{R}$, where E_L is the Lagrangian energy. Then the Lagrangian vector field X_L and the Hamiltonian vector field X_H are $(\mathbb{F}L)-$related: $(\mathbb{F}L)_*X_L = X_H$. The integral curves of X_L are mapped by $\mathbb{F}L$ onto integral curves of X_H. Furthermore, X_L and X_H have the same base integral curves*

Proof. It suffices to prove that $(\mathbb{F}L)_*X_L = X_H$. Note that $\tau_Q = \tau_Q^* \circ \mathbb{F}L$, so once the integral curves are $\mathbb{F}L$ related, the base integral curves are deduced to be equal.

Now, writing $v^* = T_v(\mathbb{F}L)(w)$ for $v \in TQ$, $w \in T_vTQ$, we get

$$\begin{aligned}
\Omega_Q\left(T\mathbb{F}L(X_L(v)), v^*\right) &= \Omega_L\left(X_L(v), w\right) \\
&= \langle dE_L(v), w\rangle \\
&= \langle \mathrm{d}(H \circ \mathbb{F}L)(v), w\rangle \\
&= \langle \mathrm{d}H(\mathbb{F}L(v)), v^*\rangle \\
&= \Omega_Q\left(X_H(\mathbb{F}L(v)), v^*\right),
\end{aligned}$$

where, as before, Ω_Q is the standard symplectic form in T^*Q. Since $T_v(\mathbb{F}L)$ is an isomorphism, v^* is arbitrary, so

$$T\mathbb{F}L\left(X_L(v)\right) = X_H\left(\mathbb{F}L(v)\right),$$

that is,

$$(\mathbb{F}L)_*X_L = X_H.$$

□

The transformation $\mathbb{F}L : TQ \to T^*Q$ thus maps the Euler-Lagrange equations into the Hamilton equations.

The following diagram shows the relationship between Lagrangian and Hamiltonian formalisms via the Legendre transformation:

$$\begin{tikzcd}
 & Q & & \\
T^*Q \arrow[ur, "\pi_Q"] \arrow[dr, "H"'] & & TQ \arrow[ul, "\tau_Q"'] \arrow[ll, "\mathbb{F}L"'] \arrow[dl, "E_L"] \arrow[r, "L"] & \mathbb{R} \\
 & \mathbb{R} & &
\end{tikzcd}$$

In §4 a generalization of theorem 2.2.10, showing the relationship between a general Hamiltonian system and its (constrained) Lagrangian counterpart, will be presented.

2.2.2 Generating functions

As with Hamiltonian mechanics, a useful general context for discussing canonical transformations and generating functions is that of symplectic manifolds. Here, we restrict ourselves again to the case of T^*Q with the canonical symplectic form Ω_Q.

Let $F : T^*Q \to T^*Q$ be a transformation from T^*Q to itself and let $\mathrm{Graph}(F) \subset T^*Q \times T^*Q$ be the graph of F. Consider the one-form on $T^*Q \times T^*Q$ defined by

$$\hat{\Theta}_Q = \pi_2^*\Theta_Q - \pi_1^*\Theta_Q,$$

where $\pi_{1,2} : T^*Q \times T^*Q \to T^*Q$ are the projections onto the two components. The corresponding two-form is then

$$\hat{\Omega}_Q = -\mathrm{d}\hat{\Theta} = \pi_2^*\Omega_Q - \pi_1^*\Omega_Q,$$

which is a symplectic 2-form in $T^*Q \times T^*Q$. Denoting the inclusion map by $i_F : \mathrm{Graph}(F) \hookrightarrow T^*Q \times T^*Q$, we see that we have the identities

$$\pi_1 \circ i_F = \pi_1|_{\mathrm{Graph}(F)}, \quad \text{and} \quad \pi_2 \circ i_F = F \circ \pi_1 \text{ on } \mathrm{Graph}(F).$$

Using these relations, we have

$$\begin{aligned}
i_F^*\hat{\Omega}_Q &= i_F^*(\pi_2^*\Omega_Q - \pi_1^*\Omega_Q) \\
&= (\pi_2 \circ i_F)^*\Omega_Q - (\pi_1 \circ i_F)^*\Omega_Q \\
&= (\pi_1|_{\mathrm{Graph}(F)})^*(F^*\Omega_Q - \Omega_Q).
\end{aligned}$$

Using this last equality, it is clear that F is a **canonical transformation** if and only if $i_F^*\hat{\Omega}_Q = 0$ or, equivalently, if and only if $\mathrm{d}(i_F^*\hat{\Theta}_Q) = 0$. By the Poincaré lemma, this last statement is equivalent to there existing, at least locally, a function $S : \mathrm{Graph}(F) \to \mathbb{R}$ such that $i_F^*\hat{\Theta}_Q = \mathrm{d}S$. Such function S is known as the **generating function** of the symplectic transformation F. Note that S is not unique.

The generating function S is specified on the graph $\mathrm{Graph}(F)$, and so can be expressed in any local coordinate system on $\mathrm{Graph}(F)$. The standard choices, for coordinates (q_0, p_0, q_1, p_1) on $T^*Q \times T^*Q$, are any two of the four quantities q_0, p_0, q_1 and p_1; note that $\mathrm{Graph}(F)$ has the same dimension as T^*Q and, moreover, is a Lagrangian submanifold of $T^*Q \times T^*Q$.

Coordinate expression

We will be particulary interested in the choice (q_0, q_1) as local coordinates on $\mathrm{Graph}(F)$, and so we give the coordinate expression for the above general generating function derivation for this particular case . This choice results in generating functions of the so-called **first kind**.

Consider a function $S : Q \times Q \to \mathbb{R}$. Its differential is

$$\mathrm{d}S = \frac{\partial S}{\partial q_0}\mathrm{d}q_0 + \frac{\partial S}{\partial q_1}\mathrm{d}q_1.$$

Let $F : T^*Q \to T^*Q$ be the canonical transformation generated by S. In local coordinates, the quantity $i_F^*\hat{\Theta}_Q$ is

$$i_F^*\hat{\Theta}_Q = -p_0\mathrm{d}q_0 + p_1\mathrm{d}q_1,$$

and so the condition $i_F^*\hat{\Theta}_Q = \mathrm{d}S$ reduces to the equations

$$p_0 = -\frac{\partial S}{\partial q_0}(q_0, q_1), \tag{2.10a}$$

$$p_1 = \frac{\partial S}{\partial q_1}(q_0, q_1), \tag{2.10b}$$

which are an implicit definition of the transformation $F : (q_0, p_0) \to (q_1, p_1)$. From the above general theory, we know that such a transformation is automatically symplectic, and that all symplectic transformations have such a representation, at least locally.

Note that there is not a one-to-one correspondence between symplectic transformations and real-valued functions on $Q \times Q$, because for some functions the above equations either have no solutions or multiple solutions, and so there is no well-defined map $(q_0, p_0) \to (q_1, p_1)$. For example, taking $S(q_0, q_1) = 0$ forces p_0 to be zero, and so there is no corresponding map F. In addition, one has to be careful about the special case of generating the identity transformation, as was noted in [31, 47].

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, where Q is any smooth manifold, as symplectic manifolds. This construction plays a key role in Lagrangian and Hamiltonian mechanics

In [157, 158], 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 [48, 146] for further details) as symplectic manifolds, i.e. $(TT^*Q\,,\, \mathrm{d}_T\,\Omega_Q)$, where $\mathrm{d}_T\,\Omega_Q$ is the tangent lift of Ω_Q, and (T^*TQ, Ω_{TQ}).

Before giving the full picture, we begin with two basic definitions. The **canonical involution** (see [48] 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 $\tilde{\chi}(s,t) := \chi(t,s)$. Note that $\frac{\mathrm{d}}{\mathrm{d}t}\chi(s,t)|_{t=0} : \mathbb{R} \to TQ$. If (q^i) are local coordinates for Q, (q^i, v^i) are the induced coordinates for TQ and $(q^i, v^i, \dot{q}^i, \dot{v}^i)$ for 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 bundles is expressed in the following diagram

$$\begin{tikzcd} TTQ \arrow[r, "\kappa_Q"'] \arrow[d, "\tau_{TQ}"'] & TTQ \arrow[d, "T\tau_Q"] \\ TQ \arrow[r, "\mathrm{Id}"'] & TQ, \end{tikzcd}$$

The **tangent pairing** between TT^*Q and TTQ is the fibered map $\langle\cdot,\cdot\rangle^T : TT^*Q \times_Q TTQ \to \mathbb{R}$ given by

$$\langle\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)\big|_{t=0}, \frac{\mathrm{d}}{\mathrm{d}t}\delta(t)\big|_{t=0}\rangle^T := \frac{\mathrm{d}}{\mathrm{d}t}\langle\gamma(t), \delta(t)\rangle^T\big|_{t=0},$$

where $\gamma : \mathbb{R} \to T^*Q$ and $\delta : \mathbb{R} \to TQ$ are such that $\pi_Q \circ \gamma \equiv \tau_Q \circ \delta$.

Definition 2.3.1. *The* **Tulczyjew's isomorphism** α_Q *is the map* $\alpha_Q : TT^*Q \to T^*TQ$ *given by*

$$\langle\alpha_Q(V), W\rangle := \langle V, \kappa_Q(W)\rangle, \quad V \in TT^*Q, \;\; W \in TTQ.$$

Locally:

$$\alpha_Q(q^i, p_i, \dot{q}^i, \dot{p}_i) = (q^i, \dot{q}^i, \dot{p}_i, p_i)$$

In the following diagram we show the different relationships among the double vector bundles and the α_Q-Tulczyjew's isomorphism:

$$\begin{tikzcd}
 & TT^*Q \arrow[rr, "\alpha_Q"] \arrow[dl, "\tau_{T^*Q}"'] \arrow[dr, "T\pi_Q"] & & T^*TQ \arrow[dl, "\pi_{TQ}"'] \arrow[dr, "T^*\tau_Q"] & \\
T^*Q \arrow[drr, "\pi_Q"'] & & TQ \arrow[d, "\tau_Q"] & & T^*Q \arrow[dll, "\pi_Q"] \\
 & & Q & &
\end{tikzcd}$$

The definition of $T^*\tau_Q$ is given in the following remark.

Remark 2.3.2. *Given a tangent bundle $\tau_N : TN \to N$, for each $y \in T_xN$ we can define*

$$\mathcal{V}_y = ker\{T_y\tau_N : T_yTN \to T_xN\}\,, \quad \tau_N(y) = x.$$

Summing over all y we obtain a vector bundle $\mathcal{V}$ 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 $u^\vee(y)$ (recall the definition 2.1.1). 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 $X^\vee(y) = (X(\tau_N(y)))^\vee$. 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 $T^*\tau_Q : T^*TQ \to T^*Q$ by $\langle T^*\tau_Q(\alpha_u), w\rangle = \langle \alpha_u, w_u^\vee\rangle$; $u, w \in T_qQ$, $\alpha_u \in T_u^*TQ$ and $w_u^\vee \in T_uTQ$.

Definition 2.3.3. *The* **Tulczyjew's isomorphism** *β_Q is the map $\beta_Q : TT^*Q \to T^*T^*Q$ defined by*

$$\beta_Q(V) := i_V\Omega_Q, \quad V \in TT^*Q,$$

where Ω_Q is the canonical symplectic form of T^*Q.

Locally,

$$\beta_Q(q^i, p_i, \dot{q}^i, \dot{p}_i) = (q^i, \dot{q}^i, \dot{p}_i, p_i).$$

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 of Q, denoted by $\mathrm{d}_T\,\Omega_Q$ and which coincides with the previous ones, that is $\mathrm{d}_T\,\Omega_Q = \Omega_{\alpha_Q}$. 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.

$$\begin{tikzcd}
T^*TQ \arrow[dr, "\pi_{TQ}"] & & TT^*Q \arrow[ll, "\alpha_Q"'] \arrow[rr, "\beta_Q"] \arrow[dl, "T\pi_Q"'] \arrow[dr, "\tau_{T^*Q}"] & & T^*T^*Q \arrow[dl, "\pi_{T^*Q}"'] \\
& TQ \arrow[rr, "\mathbb{F}L"] \arrow[dr, "\tau_Q"'] & & T^*Q \arrow[dl, "\pi_Q"] & \\
& & Q & &
\end{tikzcd}$$

2.3.1 Implicit description of mechanics

In middle seventies, W.M. Tulczyjew [157, 158] 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 depict this interpretation we are going to use the notion of Lagrangian submanifold introduced in §1.3.1 and that of Tulczyjew's isomorphisms introduced in the previous subsection §2.3. In addition, it is necessary to introduce the notion of **special symplectic manifold**. In the following definition we are employing an arbitrary symplectic manifold (P, Ω_P); we will particularize to $P = T^*Q$ afterwards.

Definition 2.3.4. *A* **special symplectic manifold** *is a symplectic manifold (P, Ω_P) which is symplectomorphic to a cotangent bundle. More precisely, there exists a fibration $\pi : P \to M$, and a one-form Θ_P on P, such that $\Omega_P = -\mathrm{d}\Theta_P$, and $\alpha : P \to T^*M$ is a diffeomorphism such that $\pi_M \circ \alpha = \pi$ and $\alpha^*\Theta_M = \Theta_P$.*

Here, Q is a smooth manifold, $\pi_Q : T^*Q \to Q$ is the canonical projection and Θ_Q is the canonical one-form in T^*Q.

The following is an important result for our discussion.

Theorem 2.3.5. *Let $(P, \Omega_P = -\mathrm{d}\Theta_P)$ be an special symplectic manifold, let $f : M \to \mathbb{R}$ be a function, and denote by N_f the submanifold of P where $\alpha^*\mathrm{d}f$ and Θ_P coincide. Then N_f is a Lagrangian submanifold of (P, Ω_P) and f is a generating function (see §2.2.2).*

Theorem 2.3.5 applies to the particular case of mechanics. Indeed, if we consider a Lagrangian function $L : TQ \to \mathbb{R}$, we obtain a Lagrangian submanifold N_L to the symplectic manifold $(TT^*Q, \Omega_{\alpha_Q})$ with generating function L.

Now, assume that $H : T^*Q \to \mathbb{R}$ is a Hamiltonian function, with Hamiltonian vector field X_H. We have the following results.

Theorem 2.3.6. *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 2.3.7. *Given a Lagrangian function $L : TQ \to \mathbb{R}$, then the image of $\mathrm{d}L$ is a Lagrangian submanifold $S_{\mathrm{d}L}$ of (T^*TQ, Ω_{TQ}).*

Finally, we relate both Lagrangian submanifolds N_L and $\mathrm{Im}X_H$.

Proposition 2.3.8. *Given a hyper-regular Lagrangian function $L : TQ \to \mathbb{R}$, consider the associated Hamiltonian $H = E_L \circ \mathbb{F}L^{-1}$. We have that $\alpha_Q^{-1}(S_L) = S_{X_H} = \beta_Q^{-1}(S_H)$. In other words, $N_L = S_{X_H}$.*

The results 3.5, 2.3.7, 2.3.8 are summarized in the following diagram

$$\begin{tikzcd}
T^*TQ \arrow[dr, "\pi_{TQ}"] & & TT^*Q \arrow[ll, "\alpha_Q"'] \arrow[rr, "\beta_Q"] \arrow[dl, "T\pi_Q"'] \arrow[dr, "\tau_{T^*Q}"] & & T^*T^*Q \arrow[dl, "\pi_{T^*Q}"'] \\
& TQ \arrow[ul, bend left, "dL"] \arrow[rr, "\mathbb{F}L"] \arrow[dr, "\tau_Q"'] & & T^*Q \arrow[ur, bend right, "dH"'] \arrow[dl, "\pi_Q"] & \\
& & Q & &
\end{tikzcd}$$

In proposition 2.3.8 we derive a Lagrangian submanifold of TT^*Q with a Lagrangian or Hamiltonian system 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 [133], which will be briefly described in §4.1.2.

2.4 Mechanical systems with constraints

Let Q be the configuration manifold of our system. Generically, a constrained system is defined by a $(2n-m)$−dimensional submanifold $\mathcal{M} \subset TQ$, which locally is defined by $\Phi^a = 0$, $1 \le a \le m$, where $\Phi^a : TQ \to \mathbb{R}$. Two distinguished kinds of constraints are the following: **linear** and **affine** (which will be useful in §6):

Definition 2.4.1. *A* **linear constraint** *on Q is defined by a distribution $\mathcal{D}$ on Q. A curve $c : I \to Q$ will be said to satisfy the linear constraint $\mathcal{D}$ if $c^{(1)}(t) \in \mathcal{D}(c(t))$ for all $t \in I$.*

Definition 2.4.2. *An* **affine constraint** *on Q is a pair $(\mathcal{D}, \gamma)$, where $\mathcal{D}$ is a distribution on Q and $\gamma \in \mathfrak{X}(Q)$ is a vector field. A curve $c : I \to Q$ will be said to satisfy the affine constraint $(\mathcal{D}, \gamma)$ if $c^{(1)}(t) - \gamma(c(t)) \in \mathcal{D}(c(t))$ for all $t \in I$.*

In these definitions $c^{(1)}$ is the lift of the curve $c(t)$ (see def.2.1.5). We shall assume that $\mathcal{D}$ has constant rank k for simplicity, and we will use this fact to suppose, at least locally, the existence of $n - k = m$ linearly independent one-forms $\mu^1, ..., \mu^m$, which annihilate the distribution. The previous assertion is equivalent to say the following

$$\mathcal{D} = \ker \left\{\mu^1, ..., \mu^m\right\}.$$

All solutions of the constrained system are required to satisfy the conditions

- $\langle \mu^a(c(t)) \, , \, c^{(1)}(t) \rangle = 0, \quad a = 1, ..., m,$ for **linear constraints**,
- $\langle \mu^a(c(t)) \, , \, c^{(1)}(t) \rangle = \langle \mu^a(c(t)) \, , \, \gamma(c(t)) \rangle, \quad a = 1, ..., m,$ for **affine constraints**,

In the subsections §2.4.2 and §2.4.3, we respectively present nonholonomic and vakonomic methods for deriving the equations of motion for a mechanical system with constraints (which will be linear in both cases). Previously we study the unconstrained case.

2.4.1 Unconstrained mechanics

Let $L : TQ \to \mathbb{R}$ be the Lagrangian function of our system. Theorem 2.1.9 established that the motion of a mechanical system defined by the Lagrangian function $L : TQ \to \mathbb{R}$ is a critical point of the action integral $\mathcal{A}_L$, that is, a curve $c \in \mathcal{C}^2(q_0, q_1)$ (see equation (2.1)) such that $\delta \mathcal{A}_L(c) = 0$. Recall that the action integral is defined by

$$\mathcal{A}_L(c) = \int_0^T L(c^{(1)}(t)) \, \mathrm{d}t, \tag{2.11}$$

where L is a Lagrangian on Q. Note that $\delta \mathcal{A}_L(c) = 0$ if and only if $\langle \mathrm{d}\mathcal{A}_L(c), u \rangle = 0$ for every $u \in T_c\mathcal{C}^2([0, T], Q, q_0, q_1)$ (see equation (2.3)). In other words $u(t) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} c_s(t)$. It is convenient to write

$$\langle \mathrm{d}\mathcal{A}_L(c), u \rangle = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \mathcal{A}_L(c_s(t)),$$

where s is a real parameter. Given (2.11) this can be written as

$$\langle \mathrm{d}\mathcal{A}_L(c), u \rangle = \int_0^T \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L(c_s^{(1)}(t)) \, \mathrm{d}t.$$

We wish to evaluate this expression in local coordinates for Q. By the chain rule we have

$$\begin{aligned}
\langle \mathrm{d}\mathcal{A}(c), u \rangle &= \int_0^T \left(\frac{\partial L}{\partial q^i} \frac{\partial q^i}{\partial s} + \frac{\partial L}{\partial \dot{q}^i} \frac{\partial \dot{q}^i}{\partial s} \right) \Big|_{s=0} \mathrm{d}t \\
&= \int_0^T \left(\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}^i} \right) \delta q^i \mathrm{d}t + \left(\frac{\partial L}{\partial \dot{q}^i} \right) \delta q^i \Big|_0^T
\end{aligned}$$

where $(q^i, \dot{q}^i)$ are the local coordinates for a neighborhood of $c^{(1)}$ and integration by parts has been used. Since $\delta q^i(0) = \delta q^i(T) = 0$ and, moreover, δq^i are independent, the previous expression vanishes if and only if

$$\frac{\partial L(c^{(1)})}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L(c^{(1)})}{\partial \dot{q}^i} = 0.$$

Thus, the previous digression can be considered as a proof for theorem 2.1.10.

2.4.2 Nonholonomic mechanics

In order to depict the nonholonomic setting, we shall start with a configuration space Q and a distribution $\mathcal{D}$ that describes the kinematic constraints of interest. Thus, $\mathcal{D}$ is a collection of linear subspaces denoted by $\mathcal{D}_q \subset T_qQ$, one for each $q \in Q$. As mentioned in definition 2.4.1, a curve $c(t) \in Q$ will be said to **satisfy the constraints** if $c^{(1)}(t) \in \mathcal{D}_{c(t)}$ for all t. This distribution will, in general, be nonintegrable in the sense of Frobeniu's theorem; i.e. the constraints are, in general, nonholonomic.

The Lagrange-d'Alembert Principle.

Consider a given Lagrangian $L : TQ \to \mathbb{R}$. In (generalized) coordinates (q^i), $i = 1, ..., n$, on Q with induced coordinates $(q^i, \dot{q}^i)$ for the tangent bundle, we write, as before, $L(q^i, \dot{q}^i)$. We assume the following principle of Lagrange-d'Alembert.

Definition 2.4.3. *The* **Lagrange-d'Alembert equations of motion** *for the system are those determined by*

$$\delta \int_0^T L(q^i, \dot{q}^i)\,dt = 0,$$

where we choose variations $\delta c(t)$ of the curve $c(t)$ that satisfy $\delta c(t) \in \mathcal{D}_{c(t)}$ for each t, $0 \leq t \leq T$, and $\delta q^i(0) = \delta q^i(T) = 0$. Recall that we are using local coordinates $(q^i, \dot{q}^i)$ for a neighborhood of $c^{(1)}(t)$.

This principle is supplemented by the condition that the curve $c(t)$ itself satisfies the constraint.

As explained before, in such a principle we take the variations δq^i *before* imposing the constraints; that is, we *do not* impose the constraints on the family of curves defining the variation. The usual arguments in the calculus of variations show that the constrained variational principle is equivalent to the equations

$$-\delta L = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i}\right)\delta q^i = 0, \tag{2.12}$$

for all variations δq such that $\delta q \in \mathcal{D}_q$ at each point of the underlying curve $c(t)$.

Structure of the equations of motion

To explore the structure of the equations determined by (2.12), let consider the set of one-forms whose vanishing determines $\mathcal{D}$, that is

$$\mathcal{D}^0 = \operatorname{span}\left\{\mu^1, ..., \mu^m\right\}.$$

Considering this m one-forms independent, it can be proven that one can choose, in a neighborhood of each point, a local coordinate chart such that μ^a, $a = 1, ..., m$, can be written as

$$\mu^a(q) = \mathrm{d}s^a + A^a_\alpha(r, s)\mathrm{d}r^\alpha, \quad a = 1, ..., m,$$

where, locally, $q = (r, s) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$. With this choice, the constraints on $\delta q = (\delta r, \delta s)$ are given in the conditions

$$\delta s^a + A^a_\alpha \delta r^\alpha = 0.$$

Substituting the previous equation into (2.12) and using the fact that δr is arbitrary gives

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha}\right) = A^a_\alpha \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}\right), \quad a = 1, ..., m. \tag{2.13}$$

Equations (2.13) combined with the constraints equation

$$\dot{s}^a + A^a_\alpha \dot{r}^\alpha = 0, \quad a = 1, ..., m, \tag{2.14}$$

give a complete description of the **equations of motion** of the system. Notice that they consist of $n - m$ second-order equations and m first order equations.

In §6 we will give an extension of the nonholonomic equations in terms of affine connections (see §1.2).

The constrained Lagrangian

We now define the **constrained Lagrangian** by substituting the constraints (2.14) into the Lagrangian, i.e:

$$L_c(r^\alpha, s^a, \dot{r}^\alpha) = L(r^\alpha, s^a, -A^a_\alpha(r, s)\dot{r}^\alpha).$$

The equations of motion (2.13) can be written in terms of the constrained Lagrangian in the following way:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta}\dot{r}^\beta,$$

where

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} - A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a}.$$

Letting $\mathrm{d}\mu^b$ be the exterior derivative of μ^b, another straightforward computation using properties of differential forms shows that

$$\mathrm{d}\mu^b(\dot{q}, \cdot) = B^b_{\alpha\beta}\dot{r}^\beta \mathrm{d}r^\alpha,$$

and hence the equations of motion have the form

$$-\delta L_c = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} \right) \delta r^a = -\frac{\partial L}{\partial \dot{s}^b} \mathrm{d}\mu^b(\dot{q}, \delta r). \tag{2.15}$$

This form of the equations isolates the effects of the constraints, and shows that if the constraints are integrable (which is equivalent to $\mathrm{d}\mu^b = 0$, i.e., to $B^b_{\alpha\beta} = 0$), then the correct equations of motion are obtained by substituting the constraints into the Lagrangian and setting the variation of L_c to zero. However, in the nonintegrable case, which is the case of nonholonomic systems, the constraints generates extra forces that must be taken into account.

For a more geometric interpretation of the equations (2.15) see [29]. Particularly, for an interpretation within the Lie algebroid setup see [37, 104].

Nonholonomic equations of motion with Lagrange multipliers

We can obtain the nonholonomic equations of motion with Lagrange multipliers from the Lagrange-d'Alembert principle as follows. Recall that the Lagrange-d'Alembert principle gives us

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0, \;\; i = 1, ..., n, \tag{2.16}$$

for variations $\delta q^i \in \mathcal{D}$, i.e., for variations in the constraint distribution. We shall denote the constraints on δq^i as follows,

$$\mu^a_i \delta q^i = 0, \qquad a = 1, ..., m, \tag{2.17}$$

where we are assuming the Einstein convention. From (2.16) and (2.17) we obtain the set of n equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_a \mu^a_i. \tag{2.18}$$

If, instead of constraints defined by (2.17), the constraints are more generically defined by the vanishing of the set of m independent maps $\Phi^a : TQ \to \mathbb{R}$, that is $\Phi^a(q^i, \dot{q}^i) = 0$, equations (2.18) become

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_a \frac{\partial \Phi^a}{\partial q^i},$$

which are called **Chetaev's equations**.

2.4.3 Vakonomic mechanics

In this variational technique one makes the functional $\mathcal{A}_L$ stationary *after* asking that the solutions satisfy the constraints. Thus, this is a classical constrained minimization problem, and may be solved with techniques from the calculus of variations with constraints. Let consider again a constrained system, with Lagrangian function $L : TQ \to \mathbb{R}$, defined by the submanifold $\mathcal{M} \subset TQ$. This submanifold is $(2n - m)-$dimensional and is determined by $\Phi^a = 0$, $1 \leq a \leq m$, where $\Phi^a : TQ \to \mathbb{R}$.

Now, we introduce a special subset $\tilde{\mathcal{C}}^2(q_0, q_1)$ of $\mathcal{C}^2(q_0, q_1)$ (recall eq.(2.1)) which consists of those curves which are in the submanifold $\mathcal{M}$

$$\tilde{\mathcal{C}}^2(q_0, q_1) = \left\{ c \in \mathcal{C}^2(q_0, q_1) \,|\, c^{(1)} \in \mathcal{M}_{c(t)} = \mathcal{M} \cap \tau_Q^{-1}(c(t)),\ \forall t \in [0, T] \right\}, \tag{2.19}$$

where $\tau_Q : TQ \to Q$ is the usual projection. Let consider the action integral $\mathcal{A}_L$ defined in (2.11), which we want to extremize among the curves satisfying the constraints imposed by $\mathcal{M}$, $c \in \tilde{\mathcal{C}}^2(q_0, q_1)$.

Definition 2.4.4. *A curve $c \in \tilde{\mathcal{C}}^2(q_0, q_1)$ will be a* **solution of the vakonomic problem** *is c is a critical point of $\mathcal{A}_L\big|_{\tilde{\mathcal{C}}^2(q_0,q_1)}$.*

Therefore, following the same procedure in the unconstrained case (§2.4.1) c is a solution of the vakonomic problem if and only if $\langle \mathrm{d}\mathcal{A}_L(c), u\rangle = 0$ for all $u \in T_c\tilde{\mathcal{C}}^2(q_0, q_1)$.

Remark 2.4.5. *We are assuming that the solution curves $c \in \tilde{\mathcal{C}}^2(q_0, q_1)$ admit nontrivial variations in $\tilde{\mathcal{C}}^2(q_0, q_1)$. These solutions are called* **normal** *in the literature, in opposition to the* **abnormal** *ones, which are pathological curves that do not admit nontrivial variations* [5].

The usual way to present the equations of motion of vakonomic mechanics is the following ([35]):

$$\left\{ \begin{array}{l} \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial \dot{q}^i}\right) - \dfrac{\partial L}{\partial q^i} = \dot{\lambda}_a \dfrac{\partial \Phi^a}{\partial \dot{q}^i} + \lambda_a \left[\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial \Phi^a}{\partial \dot{q}^i}\right) - \dfrac{\partial \Phi^a}{\partial q^i}\right], \\[2ex] \Phi^a(q, \dot{q}) = 0,\ 1 \leq a \leq m\,, \end{array} \right. \tag{2.20}$$

where λ_a are the Lagrange multipliers.

Equations (2.20) can be seen as the Euler-Lagrange equations for the extended Lagrangian $\mathcal{L} = L + \lambda_a \Phi^a$. We will not follow this approach here, which has been exploited fruitfully in [41, 95, 130]. Note that if we consider the extended Lagrangian $\lambda_0 L + \lambda_a \Phi^a$, with $\lambda_a = 0$ or 1, then we recover all the solutions, both the normal and the abnormal ones [5].

Assume now that the constraints are written in the following way

$$\dot{q}^a = \Psi(q^i, \dot{q}^\alpha),$$

where $1 \leq a \leq m$, $m+1 \leq \alpha \leq n$ and $1 \leq i \leq n$. Then $(q^i, \dot{q}^\alpha)$ are local adapted coordinates for the submanifold $\mathcal{M}$ of TQ.

Proposition 2.4.6. *A curve $c \in \tilde{\mathcal{C}}^2(q_0, q_1)$ is a normal solution of the vakonomic problem if and only if there exists $\tilde{\lambda} : [0, T] \to \mathbb{R}^m$ such that*

$$\left\{ \begin{array}{l} \dfrac{d}{dt}\left(\dfrac{\partial \tilde{L}}{\partial \dot{q}^\alpha}\right) - \dfrac{\partial \tilde{L}}{\partial q^\alpha} = \tilde{\lambda}_a \left[\dfrac{d}{dt}\left(\dfrac{\partial \Psi^a}{\partial \dot{q}^\alpha}\right) - \dfrac{\partial \Psi^a}{\partial q^\alpha}\right] + \dot{\tilde{\lambda}}_a \dfrac{\partial \Psi^a}{\partial \dot{q}^\alpha}, \\[2ex] \dot{\tilde{\lambda}}_a = \dfrac{\partial \tilde{L}}{\partial q^a} - \tilde{\lambda}_b \dfrac{\partial \Psi^b}{\partial q^a}, \\[2ex] \dot{q}^a = \Psi^a(q^i, \dot{q}^\alpha), \end{array} \right. \tag{2.21}$$

where $\tilde{L} : \mathcal{M} \to \mathbb{R}$ is the restriction of L to $\mathcal{M}$.

See [35] for the proof. Equations (2.21) stress how the information given by L outside $\mathcal{M}$ is irrelevant to obtain the vakonomic equations, contrary to what happens in nonholonomic mechanics. Finally, note that equations (2.20) and (2.21) are related by the transformation $\Phi^a = \Psi^a - \dot{q}^a$ and $\tilde{\lambda}_a = \frac{\partial L}{\partial \dot{q}^a} - \lambda_a$, $1 \leq a \leq m$.

2.5 Extension of mechanics to Lie algebroids

In this section we are going to use the notions of Lie algebroid and prolongation of a Lie algebroid described in §1.5.1.

2.5.1 Lagrangian mechanics

In [128] (see also [105]) a geometric formalism for Lagrangian mechanics on Lie algebroids was introduced. It is developed in the prolongation $\mathcal{T}^A A$ of a Lie algebroid A (see §1.5) over the vector bundle projection $\tau : A \to Q$. The prolongation of the Lie algebroid is playing the same role as TTQ in the standard mechanics, that is mechanics defined in tangent bundles. Following the same program in §2.1.3 to describe geometrically the Lagrangian mechanics for tangent bundles, we specify the canonical geometrical structures defined in $\mathcal{T}^A A$:

- The **vertical lift** $\xi^\vee : \tau^* A \to \mathcal{T}^A A$ given by $\xi^\vee(a, b) = (a, 0, b_a^\vee)$, where $a \in A$ and $b_a^\vee$ is the vector tangent to the curve $a + tb$ at $t = 0$.
- The **vertical endomorphism** $S : \mathcal{T}^A A \to \mathcal{T}^A A$ defined as follows
$$S(a, b, v) = \xi^\vee(a, b) = (a, 0, b_a^\vee).$$
- The **Liouville section**, which is the vertical section corresponding to the Liouville dilation vector field:
$$\Delta(a) = \xi^\vee(a, a) = (a, 0, a_a^\vee).$$

We also mention that the **complete lift** Y^c of a section $Y \in \Gamma(A)$ is the section of $\mathcal{T}^A A$ characterized by the following properties:

i) projects to Y, i.e., $\tau\mathcal{T} \circ Y^c = Y \circ \tau$,

ii) $\mathcal{L}_{(Y^c)}\hat{\mu} = \widehat{\mathcal{L}_Y \mu}$, for all $\mu \in \Gamma(A^*)$,

where by $\hat{\mu} \in C^\infty(A)$ we denote the linear function associated to any $\mu \in \Gamma(A^*)$.

Given a Lagrangian function $L \in C^\infty(A)$ we define the **Cartan 1-section** $\Theta_L \in \Gamma((\mathcal{T}^A A)^*)$ and the **Cartan 2-section** $\Omega_L \in \Gamma(\wedge^2(\mathcal{T}^A A)^*)$ and the **Lagrangian energy** $E_L \in C^\infty(A)$ as
$$\Theta_L = S^*(\mathrm{d}L), \;\; \Omega_L = -\mathrm{d}\Theta_L, \;\; E_L = \mathcal{L}_\Delta L - L.$$

Here, d denotes the differential operation in the prolongation of the Lie algebroid (see §1.5.1) If (q^i, y^α) are local fibered coordinates on A, $(\rho^i_\alpha, C^\gamma_{\alpha\beta})$ are the corresponding local structure functions on E and $\{X_\alpha, \Phi_\alpha\}$ is the corresponding local basis of sections of $\mathcal{T}^A A$ (see §1.5.1) then:

$$S(X_\alpha) = \Phi_\alpha,\ S(\Phi_\alpha) = 0,\ \forall\,\alpha,$$

$$\Delta = y^\alpha \Phi_\alpha,$$

$$\Omega_L = \tfrac{\partial^2 L}{\partial y^\alpha \partial y^\beta} X^\alpha \wedge \Phi^\beta + \tfrac{1}{2}\left(\tfrac{\partial^2 L}{\partial q^i \partial y^\alpha} \rho^i_\beta - \tfrac{\partial^2 L}{\partial q^i \partial y^\beta} \rho^i_\alpha + \tfrac{\partial L}{\partial y^\alpha} C^\gamma_{\alpha\beta} \right) X^\alpha \wedge X^\beta,$$

$$E_L = \tfrac{\partial L}{\partial y^\alpha} y^\alpha - L.$$

From the previous equations it follows that

$$i_{S(X)}\Omega_L = -S^*(i_X \Omega_L),\ \ i_\Delta \Omega_L = -S^*(\mathrm{d}E_L), \tag{2.22}$$

for $X \in \Gamma(\mathcal{T}^A A)$. Here, $\{X^\alpha, \Phi^\alpha\}$ is the dual basis of $\{X_\alpha, \Phi_\alpha\}$.

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

i) c is admissible (that is, $\rho(c(t)) = \dot{m}(t)$, where $m = \tau \circ c$) and

ii) $i_{(c(t),\dot{c}(t))}\Omega_L(c(t)) = \mathrm{d}E_L(c(t))$, for all t.

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

$$\dot{q}^i = \rho^i_\alpha y^\alpha,\ \ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^\alpha} + \frac{\partial L}{\partial y^\gamma} C^\gamma_{\alpha\beta} y^\beta - \rho^i_\alpha \frac{\partial L}{\partial q^i} = 0. \tag{2.23}$$

Note that if A 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, a Lagrangian function L is said to be regular if Ω_L is a symplectic section, that is, if Ω_L is regular at every point as a bilinear form. In such a case, there exists a unique solution Γ_L verifying

$$i_{\Gamma_L}\Omega_L = \mathrm{d}E_L.$$

In addition, using (2.22) it follows that $i_{S(\Gamma_L)}\Omega_L = i_\Delta \Omega_L$ which implies that Γ_L is SODE section, that is,

$$S(\Gamma_L) = \Delta,$$

or alternatively $\mathcal{T}\tau(\Gamma_L(a)) = a$ for all $a \in A$.

Thus, the integral curves of Γ_L (that is, the integral curves of the vector field $\rho^\tau(\Gamma_L)$) are solutions of the Euler-Lagrange equations for L. Γ_L is called the **Euler-Lagrange** section associated with L.

The Lagrangian L is regular if and only if the matrix $\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}$ is regular. Moreover, the local expression of Γ_L is

$$\Gamma_L = y^\alpha X_\alpha + f^\alpha \Phi_\alpha,$$

where the functions f^α satisfy the linear equations

$$\frac{\partial^2 L}{\partial y^\beta \partial y^\alpha} f^\beta + \frac{\partial^2 L}{\partial q^i \partial y^\alpha} \rho^i_\beta y^\beta + \frac{\partial L}{\partial y^\gamma} C^\gamma_{\alpha\beta} y^\beta - \rho^i_\alpha \frac{\partial L}{\partial q^i} = 0,\ \ \forall\,\alpha.$$

As a further topic, an interesting reference on reduction of Lagrangian mechanics within the context of Lie algebroids is [28].

2.5.2 Hamiltonian mechanics

Let $\tau^* : A^* \to Q$ be the vector bundle projection of the dual bundle A^* to A. Consider the prolongation $\mathcal{T}^A A^*$ of A over τ^*

$$\begin{aligned} \mathcal{T}^A A^* &= \{(b, v) \in A \times TA^* \,|\, \rho(b) = T\tau^*(v)\} \\ &= \{(a^*, b, v) \in A^* \times A \times TA^* \,|\, \tau^*(a^*) = \tau(b)\,,\, \rho(b)) = T\tau^*(v)\}\,. \end{aligned}$$

The canonical geometrical structures defined on $\mathcal{T}^A A^*$ are the following:

- The **Liouville section** $\Theta_A \in \Gamma((\mathcal{T}^A A^*)^*)$ defined by

$$\langle \Theta_A(a^*), (b, v)\rangle = \langle a^*, b\rangle.$$

- The **canonical symplectic section** $\Omega_A \in \Gamma(\wedge^2(\mathcal{T}^A A^*)^*)$ is defined by

$$\Omega_A = -\mathrm{d}\Theta_A$$

  where d is the differential on the Lie algebroid $\mathcal{T}^A A^*$.

Take coordinates (q^i, p_α) on A^* and denote by $\left\{\mathcal{Y}_\alpha, \mathcal{P}^\beta\right\}$ the local basis of sections of $\mathcal{T}^A A^*$, with

$$\mathcal{Y}_\alpha(a^*) = \left(a^*, e_\alpha(\tau^*(a^*)), \rho^i_\alpha \frac{\partial}{\partial q^i}\right),\ \mathcal{P}^\beta(a^*) = \left(a^*, 0, \frac{\partial}{\partial p_\alpha}\right).$$

In coordinates the Liouville and canonical sections are written as

$$\Theta_A = p_\alpha \mathcal{Y}^\alpha, \quad \Omega_A = \mathcal{Y}^\alpha \wedge \mathcal{P}_\alpha + \frac{1}{2} p_\gamma C^\gamma_{\alpha\beta} \mathcal{Y}^\alpha \wedge \mathcal{Y}^\alpha.$$

where $\{\mathcal{Y}^\alpha, \mathcal{P}_\beta\}$ is the dual basis of $\left\{\mathcal{Y}_\alpha, \mathcal{P}^\beta\right\}$.

Every function $H \in C^\infty(A^*)$ define a unique section Γ_H of $\mathcal{T}^A A^*$ by the equation

$$i_{\Gamma_H}\Omega_A = \mathrm{d}H,$$

and, therefore, a vector field $\rho^{\tau^*}(\Gamma_H) = X_H$ on A^* which gives the dynamics. In coordinates

$$\Gamma_H = \frac{\partial H}{\partial p_\alpha} \mathcal{Y}_\alpha - \left(\rho^i_\alpha \frac{\partial H}{\partial q^i} + p_\gamma C^\gamma_{\alpha\beta} \frac{\partial H}{\partial p_\beta}\right) \mathcal{P}^\alpha,$$

and therefore

$$X_H = \rho^i_\alpha \frac{\partial H}{\partial p_\alpha} \frac{\partial}{\partial q^i} - \left(\rho^i_\alpha \frac{\partial H}{\partial q^i} + p_\gamma C^\gamma_{\alpha\beta} \frac{\partial H}{\partial p_\beta}\right) \frac{\partial}{\partial p_\alpha}.$$

Thus, the **Hamiltonian equations** are

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \rho^i_\alpha \frac{\partial H}{\partial p_\alpha}, \qquad \frac{\mathrm{d}p_\alpha}{\mathrm{d}t} = -\rho^i_\alpha \frac{\partial H}{\partial q^i} - p_\gamma C^\gamma_{\alpha\beta} \frac{\partial H}{\partial p_\beta}.$$

2.5.3 The Legendre transformation

The relationship between Lagrangian and Hamiltonian mechanics in Lie algebroids is given by the Legendre transformation (see §2.2.1 for the case of mechanics in tangent bundles). Let $L : A \to \mathbb{R}$ be a Lagrangian function and $\Theta_L \in \Gamma((\mathcal{T}^A A)^*)$ be the Poincaré-Cartan 1-section associated with L.

We introduce the **Legendre transformation associated with** L as the smooth map $\mathbb{F}L : A \to A^*$ defined by

$$\langle \mathbb{F}L(a), b\rangle = \frac{\mathrm{d}}{\mathrm{d}t} L(a + tb)\Big|_{t=0},$$

for $a, b \in A_q$, where A_q is the fiber of A over the point $q \in Q$. In other words $\langle \mathbb{F}L(a), b\rangle = \langle \Theta_L(a), z\rangle$, where z is a point in the fiber $\mathcal{T}^A A$ over the point a such that $\mathcal{T}\tau(z) = b$.

The map $\mathbb{F}L$ is well-defined and its local expression in fibered coordinates on A and A^* is

$$\mathbb{F}L(q^i, y^\alpha) = (q^i, \frac{\partial L}{\partial y^\alpha}).$$

From this local expression it is easy to prove that the Lagrangian L is regular if and only if $\mathbb{F}L$ is a local diffeomorphism.

The Legendre transformation induces a map $\mathcal{T}\mathbb{F}L : \mathcal{T}^A A \to \mathcal{T}^A A^*$ defined by

$$\mathcal{T}\mathbb{F}L(b, X_a) = (b, (T_a\mathbb{F}L)(X_a)),$$

for $a, b \in A$ and $(a, b, X_a) \in \mathcal{T}^A_a A \subseteq A_{\tau(a)} \times A_{\tau(a)} \times T_a A$, where $T\mathbb{F}L : TA \to TA^*$ is the tangent map of $\mathbb{F}L$. Note that $\tau^* \circ \mathbb{F}L = \tau$ and thus $\mathcal{T}\mathbb{F}L$ is well-defined.

If we consider local coordinates on $\mathcal{T}^A A$ (resp. $\mathcal{T}^A A^*$) induced by the local basis $\{X_\alpha, \Phi_\alpha\}$ (resp. $\{\mathcal{Y}_\alpha, \mathcal{P}^\alpha\}$) the local expression of $\mathcal{T}\mathbb{F}L$ is

$$\mathcal{T}\mathbb{F}L(q^i, y^\alpha; z^\alpha, v^\alpha) = \left(x^i, \frac{\partial L}{\partial y^\alpha}; z^\alpha, \rho^i_\beta z^\beta \frac{\partial^2 L}{\partial q^i \partial y^\alpha} + v^\beta \frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right).$$

The relationship between Lagrangian and Hamiltonian mechanics is given by the following theorem.

Theorem 2.5.1. [105] *The pair* $(\mathcal{T}\mathbb{F}L, \mathbb{F}L)$ *is a morphism between the Lie algebroids* $(\mathcal{T}^A A, [\![\cdot,\cdot]\!]^\tau, \rho^\tau)$ *and* $(\mathcal{T}^A A^*, [\![\cdot,\cdot]\!]^{\tau^*}, \rho^{\tau^*})$*. Moreover, if* Θ_L *and* Ω_L *(resp.* Θ_A *and* Ω_A*) are the Poincaré-Cartan 1-section and 2-section associated with* L *(respectively, the Liouville 1-section and the canonical symplectic section on* $\mathcal{T}^A A^*$*), then*

$$(\mathcal{T}\mathbb{F}L, \mathbb{F}L)^*\Theta_A = \Theta_L, \quad (\mathcal{T}\mathbb{F}L, \mathbb{F}L)^*\Omega_A = \Omega_L.$$

In addition, in [105], it is proved that if the Lagrangian L is **hyperregular**, that is, $\mathbb{F}L$ is a global diffeomorphism, then $(\mathcal{T}\mathbb{F}L, \mathbb{F}L)$ is a symplectomorphism and the Euler-Lagrange section Γ_L associated with L and the Hamiltonian section Γ_H are $(\mathcal{T}\mathbb{F}L, \mathbb{F}L)-$related, that is

$$\Gamma_H \circ \mathbb{F}L = \mathcal{T}\mathbb{F}L \circ \Gamma_L.$$

Therefore, an admissible curve $a(t)$ on $\mathcal{T}^A A$ is a solution of the Euler-Lagrange equations if and only if the curve $\mu(t) = \mathbb{F}L(a(t))$ is a solution of the Hamilton equations.

Remark 2.5.2. *The extension of nonholonomic mechanics to the Lie algebroid setup was developed in* [37]*, whereas the vakonomic extension was presented in* [72]*, among other references.*

Chapter 3

Discrete Lagrangian and Hamiltonian mechanics

This chapter is mainly based on the work by Marsden and collaborators (see [124] for an introduction to this topic). Other interesting references, as was mentioned in the introduction chapter, are [25, 97, 98, 134]. They used the concept of discrete variational mechanics to derive variational integrators simulating initial value problems in dynamical mechanics. The extension of these ideas to Lie groupoids, presented in the last section of this chapter, were originally developed in the seminal work [118]. Regarding this last topic we follow here the nice survey [36].

3.1 Discrete Lagrangian mechanics

As in the continuous case, we consider the configuration manifold Q, but now we define the **discrete state space** to be $Q \times Q$. That means that rather than taking a position q^i and a velocity $\dot{q}^i$ (as local coordinates of the vector $v_q \in TQ$), we now consider two positions q_0 and q_1 and a time step $h \in \mathbb{R}$. These positions should be thought of as being two points on a curve at time h apart, such that $q_0 \simeq q(0)$ and $q_1 \simeq q(h)$. Roughly speaking, the discrete state space $Q \times Q$ contains the same amount of information as TQ (in other words, both spaces are isomorphic). A **discrete Lagrangian** is a function $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, which we think of as approximating the action integral along the curve segment between q_0 and q_1, namely

$$L_d(q_0, q_1, h) \simeq \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t.$$

In the following we neglect the h dependence except where it is important and consider the discrete Lagrangian as a function $L_d : Q \times Q \to \mathbb{R}$.

We construc the increasing sequence of times $\{t_k = hk \,|\, k = 0, ..., N\} \subset \mathbb{R}$ from the time step h, and define the **discrete path space** to be

$$\mathcal{C}_d(Q) = \mathcal{C}_d(\{t_k\}_{k=0}^N, Q) = \left\{ q_d : \{t_k\}_{k=0}^N \to Q \right\}.$$

We will identify a discrete trajectory $q_d \in \mathcal{C}_d(Q)$ with its image $q_d : \{q_k\}_{k=0}^N$, where $q_k = q_d(t_k)$. The **discrete action map** or **discrete action sum** $\mathcal{A}_{L_d} : \mathcal{C}_d(Q) \to \mathbb{R}$ along this sequence is calculated by summing the discrete Lagrangian on each adjacent pair and defined by

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

As the discrete path space $\mathcal{C}_d(Q)$ is isomorphic to $Q \times \cdots \times Q$ ($N+1$ copies), it can be given a smooth product manifold structure. The discrete action $\mathcal{A}_{L_d}$ inherits the smoothness of the discrete Lagrangian L_d.

The tangent space $T_{q_d}\mathcal{C}_d(Q)$ at q_d is the set of maps $v_{q_d} : \{t_k\}_{k=0}^N \to TQ$ such that $\tau_q \circ v_{q_d} = q_d$, which we will denote $v_{q_d} = \{(q_k, \dot{q}_k)\}_{k=0}^N$.

The discrete object corresponding to TTQ is the set $(Q \times Q) \times (Q \times Q)$. With the **projection operator** π and the **translation operator** σ defined as

$$\begin{aligned}\pi &: ((q_0, q_1), (q_0', q_1')) \mapsto (q_0, q_1),\\ \sigma &: ((q_0, q_1), (q_0', q_1')) \mapsto (q_0', q_1'),\end{aligned}$$

the **discrete second order submanifold** of $(Q \times Q) \times (Q \times Q)$ is defined to be

$$\ddot{Q}_d = \{w_d \in (Q \times Q) \times (Q \times Q) \,|\, \pi_1 \circ \sigma(w_d) = \pi_2 \circ \pi(w_d)\},$$

where $\pi_{1,2} : Q \times Q \to Q$ are the usual projectors of the first and second factors onto Q. The discrete second order submanifold is the set of pairs of the form $((q_0, q_1), (q_1, q_2))$. Now, analogously to the continuous setting, the discrete version of Hamilton's principle describes the dynamics of the discrete mechanical system determined by a discrete Lagrangian L_d on $Q \times Q$.

Theorem 3.1.1 (Discrete Hamilton's principle). *Given a C^k discrete Lagrangian L_d, $k \geq 1$, there exists a unique C^{k-1} mapping $D_{EL}L_d : \ddot{Q}_d \to T^*Q$ and unique C^{k-1} one-forms $\Theta^+_{L_d}$ and $\Theta^-_{L_d}$ on $Q \times Q$, such that for all variations $\delta q_d \in T_{q_d}\mathcal{C}_d(Q)$ of q_d we have*

$$\begin{aligned}\langle \mathrm{d}\mathcal{A}_{L_d}, \delta q_d\rangle = &\sum_{k=0}^{N-1} \langle D_{EL}L_d((q_{k-1}, q_k), (q_k, q_{k+1})), \delta q_k\rangle \\ &+\langle \Theta^+_{L_d}(q_{N-1}, q_N), (\delta q_{N-1}, \delta q_N)\rangle + \langle \Theta^-_{L_d}(q_0, q_1), (\delta q_0, \delta q_1)\rangle. \end{aligned} \tag{3.1}$$

The mapping $D_{EL}L_d$ is called the **discrete Euler-Lagrange map** *and has coordinate expression*

$$D_{EL}L_d((q_{k-1}, q_k), (q_k, q_{k+1})) = D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}).$$

The one-forms $\Theta^+_{L_d}$ and $\Theta^-_{L_d}$ are called the **discrete Lagrangian one-forms** *which local expressions are*

$$\begin{aligned}\Theta^+_{L_d}(q_0, q_1) &= D_2L_d(q_0, q_1)\mathrm{d}q_1 = \frac{\partial L_d}{\partial q_1^i}\mathrm{d}q_1^i,\\ \Theta^-_{L_d}(q_0, q_1) &= -D_1L_d(q_0, q_1)\mathrm{d}q_0 = -\frac{\partial L_d}{\partial q_0^i}\mathrm{d}q_0^i.\end{aligned}$$

Proof. Computing the derivative of the discrete action sum gives

$$\begin{aligned}\langle \mathrm{d}\mathcal{A}_{L_d}, \delta q_d\rangle &= \sum_{k=0}^{N-1} \left(\langle D_1 L_d(q_k, q_{k+1}), \delta q_k\rangle + \langle D_2 L_d(q_{k-1}, q_k), \delta q_{k+1}\rangle\right) \\ &= \sum_{k=0}^{N-1} \langle (D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k)), \delta q_k\rangle \\ &+ \langle D_1 L_d(q_0, q_1), \delta q_0\rangle + \langle D_2 L_d(q_{N-1}, q_N), \delta q_N\rangle,\end{aligned}$$

using a discrete integration by parts (rearrangement of the summation). Identifying the terms with the discrete Euler-Lagrange map and the discrete Lagrangian one-forms gives the desired result. □

Note that two one-forms arise from the boundary terms. However, observe that $\mathrm{d}L_d = \Theta^+_{L_d} - \Theta^-_{L_d}$ and so using that $\mathrm{d}^2 = 0$ we arrive to

$$\mathrm{d}\Theta^+_{L_d} = \mathrm{d}\Theta^-_{L_d}.$$

Thus, although there are two one-form, they give rise to a single discrete two-form, which is important for symplecticity.

Discrete Lagrangian evolution operator and mappings

The **discrete Lagrangian evolution operator** X_{L_d} plays the same role as the continuous Lagrangian vector field, and is defined to be the map $X_{L_d} : Q \times Q \to (Q \times Q) \times (Q \times Q)$ satisfying $\pi \circ X_{L_d} = \mathrm{Id}_{Q\times Q}$ and

$$D_{EL} L_d \circ X_{L_d} = 0.$$

The discrete objet corresponding to the Lagrangian flow is the **discrete Lagrangian map** $F_{L_d} : Q \times Q \to Q \times Q$ defined by

$$F_{L_d} = \sigma \circ X_{L_d}.$$

Since X_{L_d} is of second order, which corresponds to the requirement that $X_{L_d}(Q \times Q) \subset \ddot{Q}_d$, it has the form

$$X_{L_d} : (q_0, q_1) \to ((q_0, q_1), (q_1, q_2)),$$

and so the corresponding Lagrangian map will be $F_{L_d} : (q_0, q_1) \to (q_1, q_2)$.[1]

A discrete path $q_d \in \mathcal{C}_d(Q)$ is said to be a **solution of the discrete Euler-Lagrange equations** if the first term on the right hand side of (3.1) vanishes for all variations $\delta q_d \in T_{q_d}\mathcal{C}_d(Q)$. This means that the points $\{q_k\}$ satisfy $F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1})$ or, equivalently, that they satisfy the **discrete Euler-Lagrange equations**

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \quad \forall\, k = 1, ..., N-1. \tag{3.2}$$

[1]For a regular discrete Lagrangian (see §3.2), these objets are well-defined and the discrete Lagrangian map is invertible.

Example 3.1.2. Let consider the discrete Lagrangian $L_d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$L_d(q_k, q_{k+1}) = \frac{h}{2}\left(\frac{q_{k+1} - q_k}{h}\right)^T M \left(\frac{q_{k+1} - q_k}{h}\right) - hV(q_k),$$

where M is a real symmetric positive-definite $n \times n$ matrix (in other words a mass matrix) and V a potential function. Applying equation (3.2), we find that the discrete Euler-Lagrange equations are

$$M\left(\frac{q_{k+1} - 2q_k + q_{k-1}}{h^2}\right) = -\nabla V(q_k),$$

which is clearly a discretization of Newton's equations

$$M\,\ddot{q} = -V(q),$$

which just are the Euler-Lagrange equations for a continuous Lagrangian system defined by $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T\, M\, \dot{q} - V(q)$.

If, on the other hand, we choose a different discrete Lagrangian, namely

$$L_d(q_k, q_{k+1}) = \frac{h}{2}\left(\frac{q_{k+1} - q_k}{h}\right)^T M \left(\frac{q_{k+1} - q_k}{h}\right) - hV(\frac{q_k + q_{k+1}}{2}),$$

the resulting discrete Euler-Lagrange equations are

$$M\left(\frac{q_{k+1} - 2q_k + q_{k-1}}{h^2}\right) = -\nabla V(q_{k+1/2}) - \nabla V(q_{k-1/2}),$$

where $q_{k+1/2} = \frac{q_{k+1} - q_k}{2}$ and $q_{k-1/2} = \frac{q_k - q_{k-1}}{2}$. Again, this is a discrete analogue of the Newton's law.

3.1.1 Properties of discrete Lagrangian maps

One can show that discrete Lagrangian maps inherit the properties we have presented for continuous Lagrangian flows. That means that the **discrete Lagrangian symplectic form** $\Omega_{L_d} = \mathrm{d}\Theta^+_{L_d} = \mathrm{d}\Theta^-_{L_d}$, with coordinate expression

$$\Omega_{L_d}(q_0, q_1) = \frac{\partial^2 L_d}{\partial q_0^i \partial q_1^j}\mathrm{d}q_0^i \wedge \mathrm{d}q_1^j,$$

is preserved under the discrete Lagrangian map as

$$(F_{L_d})^*\Omega_{L_d} = \Omega_{L_d}.$$

Thus, we say that F_{L_d} is **discretely symplectic** (see [124] for the proof).

There exists a discrete analogue of Noether's theorem which states that momentum maps of symmetries are constants of the motion. To see this, we introduce the action lift to

$Q \times Q$ by the product $\Phi_g^{Q\times Q}(q_0,q_1) = (\Phi_g(q_0), \Phi_g(q_1))$, which has an infinitesimal generator $\xi_{Q\times Q} : Q \times Q \to T(Q \times Q)$ given by

$$\xi_{Q\times Q}(q_0,q_1) = (\xi_Q(q_0), \xi_Q(q_1)),$$

where $\xi_Q : Q \to TQ$ is the infinitesimal generator of the action $\Phi_g : Q \to Q$, both presented in §2.1.3. The two **discrete Lagrangian momentum maps** $J^{\pm}_{L_d} : Q \times Q \to \mathfrak{g}^*$ are

$$\begin{aligned}
\langle J^+_{L_d}(q_0,q_1), \xi\rangle &= \langle \Theta^+_{L_d}, \xi_{Q\times Q}(q_0,q_1)\rangle, \\
\langle J^-_{L_d}(q_0,q_1), \xi\rangle &= \langle \Theta^-_{L_d}, \xi_{Q\times Q}(q_0,q_1)\rangle,
\end{aligned}$$

for all $\xi \in \mathfrak{g}$, or alternatively written as

$$\begin{aligned}
\langle J^+_{L_d}(q_0,q_1), \xi\rangle &= \langle D_2L_d(q_0,q_1), \xi_Q(q_1)\rangle, \\
\langle J^-_{L_d}(q_0,q_1), \xi\rangle &= \langle -D_1L_d(q_0,q_1), \xi_Q(q_0)\rangle.
\end{aligned}$$

The discrete momentum maps $J^+_{L_d}$ and $J^-_{L_d}$ are equal, along solutions of the discrete Euler-Lagrange equations, in the case of a discrete Lagrangian that is invariant under the lifted action, that is

$$L_d \circ \Phi_g^{Q\times Q} = L_d$$

holds for all $g \in G$. Then $J_{L_d} : Q \times Q \to \mathfrak{g}^*$ is the unique **discrete Lagrangian momentum map**.

Theorem 3.1.3 (Discrete Noether's theorem). *Consider a given discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$ which is invariant under the lift of the (left or right) action $\Phi : G \times Q \to Q$. Then, the corresponding discrete Lagrangian momentum map $J_{L_d} : Q \times Q \to \mathfrak{g}^*$ is a conserved quantity of the discrete Lagrangian map $F_{L_d} : Q \times Q \to Q \times Q$, such that $J_{L_d} \circ F_{L_d} = J_{L_d}$.*

See [124] for the proof.

3.2 Discrete Hamiltonian mechanics

Discrete Legendre transforms

Just as the standard Legendre transform maps the Lagrangian state space TQ to the Hamiltonian phase space T^*Q, we can define **discrete Legendre transforms** or **discrete fiber derivatives** $\mathbb{F}L^{\pm}_d : Q \times Q \to T^*Q$, which map the discrete state space $Q \times Q$ to T^*Q. These are given by

$$\begin{aligned}
\langle \mathbb{F}L^+_d(q_0,q_1), \delta q_1\rangle &= \langle D_2L_d(q_0,q_1), \delta q_1\rangle, \\
\langle \mathbb{F}L^-_d(q_0,q_1), \delta q_0\rangle &= \langle -D_1L_d(q_0,q_1), \delta q_0\rangle,
\end{aligned}$$

which can be written as

$$\begin{aligned}
\mathbb{F}L^+_d : (q_0,q_1) &\mapsto (q_1,p_1) = (q_1, D_2L_d(q_0,q_1)), \\
\mathbb{F}L^-_d : (q_0,q_1) &\mapsto (q_0,p_0) = (q_0, -D_1L_d(q_0,q_1)).
\end{aligned}$$

We say that L_d is regular if both discrete fiber derivatives are local isomorphisms (for nearby q_0 and q_1). In general, we assume working with regular discrete Lagrangians. If both discrete fiber derivatives are global isomorphisms we say that L_d is hyperregular. In fact, it is only necessary to have regularity of one of the Legendre maps. Then the regularity of the second is straightforward.

The canonical one- and two-forms and Hamiltonian momentum maps are related to the discrete Lagrangian forms and discrete momentum maps by pullback by the discrete fiber derivatives, such that

$$\Theta^{\pm}_{L_d} = (\mathbb{F}L_d^{\pm})^*\Theta_Q \text{ and } \Omega_{L_d} = (\mathbb{F}L_d^{\pm})^*\Omega_Q,$$

where, as before, Θ_Q is the canonical 1-form and Ω_Q is the symplectic 2-form on T^*Q.

Momentum matching

By introducing the notation

$$\begin{aligned} p^+_{k,k+1} &= p^+(q_k, q_{k+1}) = \mathbb{F}L_d^+(q_k, q_{k+1}), \\ p^-_{k,k+1} &= p^-(q_k, q_{k+1}) = \mathbb{F}L_d^-(q_k, q_{k+1}), \end{aligned}$$

for the momentum at the two endpoints of each interval $[k, k+1]$ the discrete fiber derivatives permit a new interpretation of the discrete Euler-Lagrange equations (3.2) which can be written as

$$\mathbb{F}L_d^+(q_{k-1}, q_k) = \mathbb{F}L_d^-(q_k, q_{k+1}), \tag{3.3}$$

or simply

$$p^+_{k-1,k} = p^-_{k,k+1}.$$

That is, the discrete Euler-Lagrange equations are enforcing the condition that the momentum at time k should be the same when being evaluated from the lower interval $[k-1, k]$ or the upper interval $[k, k+1]$. This means that along a solution curve there is a unique momentum at each time k, which is denoted by

$$p_k = p^+_{k-1,k} = p^-_{k,k+1}.$$

A discrete trajectory $\{q_k\}_{k=0}^N$ in Q can thus also be regarded as either a trajectory $\{(q_k, q_{k+1})\}_{k=0}^{N-1}$ in $Q \times Q$ or, equivalently, as a trajectory $\{(q_k, p_k)\}_{k=0}^N$ in T^*Q.

Note that (3.3) can be also written as

$$\mathbb{F}L_d^+ = \mathbb{F}L_d^- \circ F_{L_d}. \tag{3.4}$$

A consequence of viewing the discrete Euler-Lagrange equations as a matching of momenta is that it gives a condition for when the discrete Lagrangian evolution operator and discrete Lagrangian map are well-defined.

Theorem 3.2.1. *Given a discrete Lagrangian system $L_d : Q \times Q \to \mathbb{R}$, the discrete Lagrangian evolution operator X_{L_d} and the discrete Lagrange map F_{L_d} are well-defined if and only if $\mathbb{F}L_d^-$ is locally and isomorphism. The discrete Lagrange map is well-defined and invertible if and only if the discrete Lagrangian is regular.*

See [124] for the proof.

Discrete Hamiltonian maps

Using the discrete Legendre transformations also enables us to push the discrete Lagrangian map $F_{L_d} : Q \times Q \to Q \times Q$ forward to T^*Q. We define the **discrete Hamiltonian map** $\tilde{F}_{L_d} : T^*Q \to T^*Q$ by $\tilde{F}_{L_d} = \mathbb{F}L_d^{\pm} \circ F_{L_d} \circ (\mathbb{F}L_d^{\pm})$. The fact that the discrete Hamiltonian map can be equivalently defined with either discrete Legendre transform is a consequence of the following theorems.

Theorem 3.2.2. *The following diagram commutes*

$$
\begin{tikzcd}
 & (q_0, q_1) \arrow[rr, mapsto, "F_{L_d}"] \arrow[dl, mapsto, "\mathbb{F}L_d^-"'] \arrow[dr, mapsto, "\mathbb{F}L_d^+"] & & (q_1, q_2) \arrow[dl, mapsto, "\mathbb{F}L_d^-"'] \arrow[dr, mapsto, "\mathbb{F}L_d^+"] & \\
(q_0, p_0) \arrow[rr, mapsto, "\tilde{F}_{L_d}"'] & & (q_1, p_1) \arrow[rr, mapsto, "\tilde{F}_{L_d}"'] & & (q_2, p_2)
\end{tikzcd}
\tag{3.5}
$$

Proof. The central triangle is simply (3.4). Assume that we define the discrete Hamiltonian map by $\tilde{F}_{L_d} = \mathbb{F}L_d^+ \circ F_{L_d} \circ (\mathbb{F}L_d^+)^{-1}$, which gives the right-hand parallelogram. Replicating the right-hand triangle on the left-hand side completes the diagram. If we choose to use the other discrete Legendre transform then the reverse argument applies. □

Corollary 3.2.3. *The following three definitions of the discrete Hamiltonian map,*

$$
\begin{aligned}
\tilde{F}_{L_d} &= \mathbb{F}L_d^+ \circ F_{L_d} \circ (\mathbb{F}L_d^+)^{-1}, \\
\tilde{F}_{L_d} &= \mathbb{F}L_d^- \circ F_{L_d} \circ (\mathbb{F}L_d^-)^{-1}, \\
\tilde{F}_{L_d} &= \mathbb{F}L_d^+ \circ (\mathbb{F}L_d^-)^{-1},
\end{aligned}
$$

are equivalent and have coordinate expression $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$*, where*

$$p_0 = -D_1 L_d(q_0, q_1), \tag{3.6a}$$

$$p_1 = D_2 L_d(q_0, q_1). \tag{3.6b}$$

Proof. The equivalence of the three definitions can be read directly from the diagram in Theorem (3.2.2). The coordinate expression for $\tilde{F}_{L_d} : (q_0, p_0) \mapsto (q_1, p_1)$ can be readily seen from the definition $\tilde{F}_{L_d} = \mathbb{F}L_d^+ \circ (\mathbb{F}L_d^-)^{-1}$. Taking initial condition $(q_0, p_0) \in T^*Q$ and setting $(q_0, q_1) = (\mathbb{F}L_d^-)^{-1}(q_0, p_0)$ implies that $p_0 = -D_1 L_d(q_0, q_1)$, which is (3.6a). Now, letting $(q_1, p_1) = \mathbb{F}L_d^+(q_0, q_1)$ gives $p_1 = D_2 L_d(q_0, q_1)$, which is (3.6b). □

As the discrete momentum map preserves the discrete symplectic form and discrete momentum maps on $Q \times Q$, the discrete Hamiltonian map will preserve the pushforwards of these structures. As we saw above, however, these are simply the canonical symplectic form and momentum map on T^*Q, and so the discrete Hamiltonian map is symplectic and momentum preserving.

We can summarize the relationship between the discrete and continuous systems in the following diagram, where the dashed arrows represent the discretization.

$$\begin{tikzcd} TQ\,,\, F_L \arrow[r, dashed] \arrow[d, "\mathbb{F}L"'] & Q \times Q\,,\, F_{L_d} \arrow[d, "\mathbb{F}L_d"] \\ T^*Q\,,\, F_H \arrow[r, dashed] & T^*Q\,,\, \tilde{F}_{L_d} \end{tikzcd} \tag{3.7}$$

3.2.1 Discrete Lagrangians are generating functions

As we have seen above, a discrete Lagrangian is a real-valued function on $Q \times Q$ which defines a map $\tilde{F}_{L_d} : T^*Q \to T^*Q$. In fact, a discrete Lagrangian is simply a generating function of the first kind for the map $\tilde{F}_{L_d}$, in the sense defined in §2.2.2. This is seen by comparing the coordinate expression (3.6) for the discrete Hamiltonian map with the expression (2.10) for the map generated by a generating function of the first kind.

3.3 Correspondence between discrete and continuous mechanics

We will now define a particular choice of discrete Lagrangian which gives an **exact** correspondence between discrete and continuous systems. To do this, we must firstly recall the following fact

Theorem 3.3.1. *Consider a regular Lagrangian L for a configuration manifold Q, two points $q_0, q_1 \in Q$ and a time step $h \in \mathbb{R}$. If $||q_1 - q_0||$ and h are sufficiently small, then there exists a unique solution $q : \mathbb{R} \to Q$ of the Euler-Lagrange equations for L satisfying $q(0) = q_0$ and $q(h) = q_1$.*

See [123] and [140] for the proof. For a regular Lagrangian L we state the following definition

Definition 3.3.2. *Let the* **exact discrete Lagrangian** *be*

$$L_d^E(q_0, q_1, h) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\, dt,$$

for sufficiently small h and close q_0 and q_1.

Here, $q_{0,1}(t)$ is the unique solution of the Euler-Lagrange equations for L which satisfies the boundary conditions $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$, and whose existence is guaranteed by Theorem 3.3.1.

We will now see that with this exact discrete Lagrangian there is an exact correspondence between the discrete and continuous systems. To do this, we will first establish that there is a special relationship between the Legendre transforms of a regular Lagrangian and its corresponding exact discrete Lagrangian.

Lemma 3.3.3. *A regular Lagrangian L and the corresponding exact discrete Lagrangian L_d^E have Legendre transforms related by*

$$\begin{aligned} \mathbb{F}(L_d^E)^+(q_0, q_1, h) &= \mathbb{F}L(q_{0,1}(h), \dot{q}_{0,1}(h)), \\ \mathbb{F}(L_d^E)^-(q_0, q_1, h) &= \mathbb{F}L(q_{0,1}(0), \dot{q}_{0,1}(0)), \end{aligned}$$

for sufficiently small h and close $q_0, q_1 \in Q$.

See [124] for the proof. Since $(q_{0,1}(h), \dot{q}_{0,1}(h)) = F_L^h(q_{0,1}(0), \dot{q}_{0,1}(0))$, Lemma 3.3.3 is equivalent to the following commutative diagram (recall that F_L^t is the Lagrangian flow, defined in §2.1.3, of the Lagrangian vector field X_L determined by the equation (2.7), that is $i_{X_L}\Omega_L = \mathrm{d}E_L$).

$$\begin{tikzcd}
 & (q_0, q_1) \arrow[dl, mapsto, "\mathbb{F}(L_d^E)^-"'] \arrow[dr, mapsto, "\mathbb{F}(L_d^E)^-"] & \\
(q_0, p_0) & & (q_1, p_1) \\
(q_0, \dot{q}_0) \arrow[u, mapsto, "\mathbb{F}L"] \arrow[rr, mapsto, "F_L^h"'] & & (q_1, \dot{q}_1) \arrow[u, mapsto, "\mathbb{F}L"']
\end{tikzcd}$$

Combining this diagram with Theorem 3.2.2 and with the definition of the Legendre transform gives the following commutative diagram for the exact discrete Lagrangian

$$\begin{tikzcd}
 & (q_0, q_1) \arrow[rr, mapsto, "F_{L_d^E}"] \arrow[dl, mapsto, "\mathbb{F}(L_d^E)^-"'] \arrow[dr, mapsto, "\mathbb{F}(L_d^E)^+"] & & (q_1, q_2) \arrow[dl, mapsto, "\mathbb{F}(L_d^E)^-"'] \arrow[dr, mapsto, "\mathbb{F}(L_d^E)^+"] & \\
(q_0, p_0) \arrow[rr, mapsto, "\tilde{F}_{L_d^E} = F_H^h"'] & & (q_1, p_1) \arrow[rr, mapsto, "\tilde{F}_{L_d^E} = F_H^h"'] & & (q_2, p_2) \\
(q_0, \dot{q}_0) \arrow[u, mapsto, "\mathbb{F}L"'] \arrow[rr, mapsto, "F_L^h"'] & & (q_1, \dot{q}_1) \arrow[u, mapsto, "\mathbb{F}L"'] \arrow[rr, mapsto, "F_L^h"'] & & (q_2, \dot{q}_2) \arrow[u, mapsto, "\mathbb{F}L"']
\end{tikzcd}$$

This proves the following theorem.

Theorem 3.3.4. *Consider a regular Lagrangian L, its corresponding exact discrete Lagrangian L_d^E and the pushforward of both the continuous and discrete systems on T^*Q, yielding a Hamiltonian system with Hamiltonian H and a discrete Hamiltonian map $\tilde{F}_{L_d^E}$, respectively. Then, for a sufficiently small time step $h \in \mathbb{R}$, the Hamiltonian flow map equals the* *pushforward discrete Lagrangian map:*

$$F_H^h = \tilde{F}_{L_d^E}.$$

This theorem is a statement about the time evolution of the system, and can also be interpreted as saying that the diagram (3.7) commutes with the dashed arrows understood as samples at times $\{t_k\}_{k=0}^N$, rather than merely as discretizations.

We can also interpret the equivalence of the discrete and continuous systems as a statement about trajectories. On the Lagrangian side, this gives the following theorem.

Theorem 3.3.5. *Take a series of times $\{t_k = hk\}_{k=0}^N$ for a sufficiently time step $h \in \mathbb{R}$, and a regular Lagrangian L and its corresponding exact discrete Lagrangian L_d^E. Then, solutions $q : [0, t_N] \to Q$ of the Euler-Lagrange equations for L and solutions $\{q_k\}_{k=0}^N$ of the discrete Euler-Lagrange equations for L_d^E are related by*

$$q_k = q(t_k), \quad k = 0, ..., N, \tag{3.8a}$$

$$q(t) = q_{k,k+1}(t), \quad t = [t_k, t_{k+1}]. \tag{3.8b}$$

Here, the curves $q_{k,k+1} : [t_k, t_{k+1}] \to Q$ are the unique solutions of the Euler-Lagrange equations for L satisfying $q_{k,k+1}(kh) = q_k$ and $q_{k,k+1}((k+1)h) = q_{k+1}$.

See [124] for the proof.

3.4 Variational integrators

We now turn our attention to a discrete Lagrangian system as an approximation to a given continuous system. That is, the discrete system is an integrator for the continuous system.

As we have seen, under regularity condition discrete Lagrangian maps preserve the symplectic structure and so, regarded as integrators, they are necessarily symplectic. Furthermore, generating function theory shows that any symplectic integrator for a mechanical system can be regarded as a discrete Lagrangian system, a fact we state here as a theorem

Theorem 3.4.1. *If the integrator $F : T^*Q \times \mathbb{R} \to T^*Q$ is symplectic, then there exists a discrete Lagrangian L_d whose discrete Hamiltonian map $\tilde{F}_{L_d}$ is F.*

In addition, if the discrete Lagrangian inherits the same symmetry groups as the continuous system, then the discrete system will also preserve the corresponding momentum maps. As an integrator, it will thus be a so-called symplectic-momentum integrator.

Just as with continuous mechanics, we have seen that discrete variational mechanics has both a Lagrangian and a Hamiltonian interpretation. These two viewpoints are complementary and both give insight into the behaviour and derivation of useful integrators.

However, the above theorem is not literally used in the construction of variational integrators, but is rather used as the first steps in obtaining inspiration. In this section we will assume that Q, and thus also TQ and T^*Q, is a finite n-dimensional vector space with an inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$.

3.4.1 Implementation of variational integrators

Although the distinction between the discrete Lagrangian map $F_{L_d} : Q \times Q \times \mathbb{R} \to Q \times Q$ and its pushforward $\tilde{F}_{L_d} : T^*Q \times \mathbb{R} \to T^*Q$ is important geometrically, for implementation purposes the two maps are essentially the same. This is because of the observation made in §3.2 that the discrete Euler-Lagrange equations that define F_{L_d} can be interpreted as matching of momenta between adjacent intervals.

In other words, given a trajectory $q_0, q_1, ..., q_{k-1}, q_k$ the map $F_{L_d} : Q \times Q \times \mathbb{R} \to Q \times Q$ calculates q_{k+1} according to

$$D_2 L_d(q_{k-1}, q_k, h) = -D_1 L_d(q_k, q_{k+1}, h).$$

If now we take $p_k = D_2 L_d(q_{k-1}, q_k, h)$ for each k, then this equation is simply

$$p_k = -D_1 L_d(q_k, q_{k+1}, h), \tag{3.9}$$

which together with the next update

$$p_{k+1} = D_2 L_d(q_k, q_{k+1}, h), \tag{3.10}$$

defines the pushforward map $\tilde{F}_{L_d} : T^*Q \times \mathbb{R} \to T^*Q$. Another way to think of this is that the p_k are merely storing the values $D_2 L_d(q_k, q_{k+1}, h)$ from the last step. For this reason it is tipically easier to implement a variational integrator as the single step map $\tilde{F}_{L_d}$, as this also provides a simple method of initialization from initial vales $(q_0, p_0) \in T^*Q$. In the general case when no special form is apparent, the equations (3.9) and (3.10) must be solved directly. The update $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ thus involves first solving the implicit equation (3.9) for q_{k+1} and then evaluating the explicit update (3.10) to give p_{k+1}.

3.4.2 Error analysis

In this section we consider a numerical method $F : T^*Q \times \mathbb{R} \to T^*Q$ which approximates the flow $F_H : T^*Q \times \mathbb{R} \to T^*Q$ of a given Hamiltonian vector field X_H. Error analysis is concerned with difference between and exact trajectory and a discrete trajectory.

Local error and order of the method

An integrator F of X_H is said to be **of order** r if there exist an open set $U \subset T^*Q$ and constants $C_l > 0$ and h_l so that

$$\parallel F(q, p, h) - F_H(q, p, h) \parallel \leq C_l h^{r+1} \tag{3.11}$$

for all $(q, p) \in U$ and $h \leq h_l$ (we are using, with some abuse of notation, (q, p) and (q^i, p_i), without distinction, since T^*Q and T^*Q are vector spaces). The expression on the left hand side of this equation is known as the **local error**, and if the method has order at least 1, then it is said to be **consistent**.

Global error and convergence

Having defined the error after one step, we now consider the error after many steps. The integrator F of F_H is said to be **convergent of order** r if there exist an open set $U \subset T^*Q$ and constants $C_g > 0$, $h_g > 0$ and $T_g > 0$ so that

$$\| (F)^N(q,p,h) - F_H(q,p,T) \| \leq C_g h^r,$$

where $h = T/N$, for all $(q,p) \in U$, $h \leq h_g$ and $T \leq T_g$. The expression on the left hand side is the **global error** at time T.

For one-step methods such as we consider here, convergence follows from a local error bound on the method and a Lipschitz bound on X_H

Theorem 3.4.2. *Suppose that the integrator F for X_H is of order r on the open set $U \subset T^*Q$ with local error constant C_l, and assume that $\ell > 0$ is such that*

$$\left|\left| \frac{\partial X_H}{\partial(p,q)} \right|\right| \leq \ell$$

on U. The method is consistent on U with global error constant C_g given by

$$C_g = \frac{C_l}{\ell}(e^{\ell T_g} - 1).$$

See [62], for instance, for the proof.

Order calculation

Given an integrator F of X_H, the order can be calculated by expanding both the true flow F_H and the integrator F in a Taylor series in h and then comparing terms. If the terms agree up to order r, then the method will be of order r.

3.4.3 Variational error analysis

Rather than considering how closely the trajectory of F matches the exact trajectory given by F_H, we can alternatively consider how closely a discrete Lagrangian matches the ideal discrete Lagrangian given by the action. As we have seen in §3.3, if the discrete Lagrangian is equal to the action, then the corresponding discrete Hamiltonian map $\tilde{F}_{L_d}$ will exactly equal the flow F_H.

The approach taken here is to show that when the discrete Lagrangian approximates a continuous Lagrangian, the discrete integrator approximates the continuous flow and thus the classical theory implies that the global discrete trajectory approximates the continuous trajectory.

Local variational order

Recall that the exact discrete Lagrangian given in definition 3.3.2 is defined by

$$L_d^E(q_0, q_1, h) = \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t,$$

where $q(t)$ is the solution of the Euler-Lagrange equations satisfying $q(0) = q_0$ and $q(h) = q_1$.

We say that given a discrete Lagrangian L_d is of **order r** if there exist an open subset $U_v \subset TQ$ with compact closure and constants $C_v > 0$ and $h_v > 0$ so that

$$\| L_d(q(0), q(h), h) - L_d^E(q(0), q(h), h) \| \leq C_v h^{r+1} \tag{3.12}$$

for all solutions $q(t)$ of the Euler-Lagrange equations with initial conditions $(q, \dot{q}) \in U_v$ and for all $h \leq h_v$.

Discrete Legendre transform order

The discrete Legendre transforms $\mathbb{F}L_d^{\pm}$ of a discrete Lagrangian L_d are said to be of **order r** if there exists an open subset $U_f \subset T^*Q$ with compact closure and constants $C_f > 0$ and $h_f > 0$, so that

$$\| \mathbb{F}L_d^+(q(0), q(h), h) - \mathbb{F}(L_d^E)^+(q(0), q(h), h) \| \leq C_f h^{r+1}, \tag{3.13a}$$

$$\| \mathbb{F}L_d^-(q(0), q(h), h) - \mathbb{F}(L_d^E)^-(q(0), q(h), h) \| \leq C_f h^{r+1}, \tag{3.13b}$$

for all solutions $q(t)$ of the Euler-Lagrange equations with initial condition $(q, \dot{q}) \in U_f$ and for all $h \leq h_f$.

The relationship between the orders of a discrete Lagrangian, its discrete Legendre transforms and its discrete Hamiltonian map is given in the following fundamental theorem.

Theorem 3.4.3. *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian L_d:*

1. *the discrete Hamiltonian map for L_d is of order r,*
2. *L_d is equivalent to a discrete Lagrangian of order r.*

See [124] for the proof. An extension of this theorem is treated in [140].

Variational order calculation

Given a discrete Lagrangian, its order can be calculated by expanding the expression for $L_d(q(0), q(h), h)$ in a Taylor series in h and comparing this to the same expansion for the exact Lagrangian. If the series agree up to r, then the discrete Lagrangian is of order r.

We explicitly evaluate the first terms of the expansion of the exact discrete Lagrangian to give

$$L_d^E(q(0), q(h), h) = hL(q, \dot{q}) + \frac{1}{2}h^2 \left(\frac{\partial L}{\partial q^i}(q, \dot{q})\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q})\ddot{q}^i \right) + \mathcal{O}(h^3),$$

where $q = q(0)$, $\dot{q} = \dot{q}(0)$ and so forth. Higher derivatives of $q(t)$ are determined by the Euler-Lagrange equations (recall that the Euler-Lagrange equations $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(q,\dot{q})}{\partial \dot{q}}\right) - \frac{\partial L(q,\dot{q})}{\partial q} = 0$ determine $\ddot{q}$ in terms of $\dot{q}$ and q).

3.4.4 The adjoint of a method and symmetric methods

For a one-step method $F : T^*Q \times \mathbb{R} \to T^*Q$ the **adjoint method** is $F^* : T^*Q \times \mathbb{R} \to T^*Q$ defined by

$$(F^*)^h \circ F^{-h} = \mathrm{Id} \tag{3.14}$$

that is, $(F^*)^h = (F^{-h})^{-1}$. The method is said to be **self-adjoint** if $F^* = F$. Note that we always have $F^{**} = F$.

Given a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, we define the **adjoint discrete Lagrangian** to be $L_d^* : Q \times Q \times \mathbb{R} \to \mathbb{R}$ defined by

$$L_d^*(q_0, q_1, h) = -L_d(q_1, q_0, -h). \tag{3.15}$$

The discrete Lagrangian L_d is said to be **self-adjoint** if $L_d^* = L_d$. Note that $L_d^{**} = L_d$ for any L_d.

Theorem 3.4.4. *If the discrete Lagrangian L_d has a discrete Hamiltonian map $\tilde{F}_{L_d}$, then the adjoint L_d^* of the discrete Lagrangian has discrete Hamiltonian map equal to the adjoint map, so that $\tilde{F}_{L_d^*} = \tilde{F}_{L_d}^*$. If the discrete Lagrangian is self-adjoint, then the method is self-adjoint. Conversely, if the method is self-adjoint, then the discrete Lagrangian is equivalent to a self-adjoint discrete Lagrangian.*

See [124] for the proof.

Order of adjoint methods

To relate the expressions of L_d and its adjoint in terms of h, it is necessary to work with the modified form

$$L_d^*(q(-h/2), q(h/2), h) = -L_d(q(h/2), q(-h/2), -h),$$

which can be used in the same way as $L_d^*(q(0), q(h), h) = -L_d(q(h), q(0), -h)$. From this it is clear that the expansion of L_d^* is the negative of the expansion of L_d with h replaced by $-h$. In other words, if L_d has the expansion

$$L_d(h) = hL_d' + \frac{1}{2}h^2 L_d'' + \frac{1}{6}h^3 L_d''',$$

where $'$ denotes the derivative with respect to h, then L_d^* will have the expansion

$$L_d^*(h) = hL_d' - \frac{1}{2}h^2 L_d'' + \frac{1}{6}h^3 L_d''',$$

and so the series agree on odd terms and are opposite on even terms.

This shows that the order of the adjoint discrete Lagrangian L_d^* is the same as the order of L_d. Furthermore, if L_d is self-adjoint, then all the even terms in its expansion must be zero, showing that self-adjoint discrete Lagrangians are necessarily of even order (the first nonzero term, which is $r+1$, must be odd).

These same conclusions can be also be reached by working with the discrete Hamiltonian map, and showing that its adjoint has the same order as it, and that it is of even order whenever it is self-adjoint. Theorems 3.4.4 and 3.4.3 then give the corresponding statements for the discrete Lagrangians.

Exact discrete Lagrangian is self-adjoint

It is easy to verify that the exact discrete Lagrangian L_d^E in definition 3.3.2 is self-adjoint. This can be done either directly from $L_d^*(q_0, q_1, h) = -L_d(q_1, q_0, -h)$, or by realizing that the exact flow map F_H generated by L_d^E satisfies equation (3.14) and then using Theorem 3.4.4.

3.5 Discrete mechanics on Lie groupoids

In this section we are going to use the notions of Lie groupoid and prolongation of a Lie groupoid described in §1.5.2. We discuss discrete Lagrangian mechanics on a Lie groupoid $G \rightrightarrows Q$. Instead of the usual Euler-Lagrange equations (2.23) for a Lie algebroid $\tau : A \to Q$ $\tau : A \to Q$ equipped with a Lagrangian function $L : A \to \mathbb{R}$, we obtain a set of difference equations called **discrete Euler-Lagrange** equations for a discrete Lagrangian $L_d : G \to \mathbb{R}$ (see [118] for further details). When the Lie algebroid is precisely $A = AG$ and L_d is a suitable approximation of the continuous Lagrangian $L : AG \to \mathbb{R}$, then we will obtain a geometric integrator for the continuous Euler-Lagrange equations.

Discrete Euler-Lagrange equations

A **discrete Lagrangian system** consists of a Lie groupoid $G \rightrightarrows Q$ (the **discrete space**) and a **discrete Lagrangian** $L_d : G \to \mathbb{R}$. For $g \in G$ fixed, we consider the set of admissible sequences

$$\mathcal{C}_g^N = \left\{ (g_1, ..., g_N) \in G^N \,|\, (g_k, g_{k+1}) \in G^{(2)} \text{ for } k = 1, ..., N-1, \ g_1 \cdots g_N = g \right\}.$$

We may identify the tangent space to $\mathcal{C}_g^N$ with

$$T_{(g_1,...,g_N)}\mathcal{C}_g^N \equiv \{ (v_1, ..., v_{N-1}) \,|\, v_k \in (AG)_{q_k} \text{ and } q_k = \beta(g_k), \ 1 \leq k \leq N-1 \}.$$

An element of $T_{(g_1,...,g_N)}\mathcal{C}_g^N$ is called an **infinitesimal variation**. Now, we define the **discrete action sum** associated to the discrete Lagrangian $L_d : G \to \mathbb{R}$ by

$$\mathcal{S}L_d(g_1, ..., g_N) = \sum_{k=1}^{N} L_d(g_k).$$

Hamilton's principle requires this discrete action sum to be stationary with respect to all infinitesimal variations. This requirement gives the following alternative expression for the **discrete Euler-Lagrange equations**:

$$\overleftarrow{X}(g_k)(L_d) - \overrightarrow{X}(g_{k+1})(L_d) = 0, \tag{3.16}$$

for all sections $X \in \Gamma(AG)$ (see [118] for the proof).

Alternatively, we may rewrite the discrete Euler-Lagrange equations as

$$\mathrm{d}\left(L_d \circ L_{g_k} + L_d \circ R_{g_{k+1}} \circ i\right)(\epsilon(q_k))\Big|_{(AG)_{q_k}} = 0,$$

where $\beta(g_k) = \alpha(g_{(k+1)}) = q_k$ and i represents the inversion map in Lie groupoids defined in §1.5.2. Note that L_{g_k} and $R_{g_{k+1}}$ denote the left-translation by g_k and the right translation by g_{k+1} in the Lie groupoid, respectively. Thus, we may define the **discrete Euler-Lagrange operator**:

$$D_{DEL}L_d : G^{(2)} \to A^*G.$$

where A^*G is the dual algebroid of AG. This operator is given by

$$D_{DEL}L_d(g,h) = \mathrm{d}\left(L_d \circ L_g + L_d \circ R_h \circ i\right)(\epsilon(q))\Big|_{(AG)_q},$$

with $\beta(g) = \alpha(h) = q$.

Discrete Lagrangian evolution operator

Let $\Upsilon_d : G \to G$ be a smooth map such that:

- $\mathrm{Graph}(\Upsilon_d) \subseteq G^{(2)}$, that is, $(g, \Upsilon_d(g)) \subseteq G^{(2)}$, for all $g \in G$ (in other words Υ_d is a **second order operator**).
- $(g, \Upsilon_d(g))$ is a solution of the Euler-Lagrange equations for all $g \in G$, that is $D_{DEL}L_d(g, \Upsilon_d(g)) = 0 \ \forall\, g \in G$.

In such case

$$\overleftarrow{X}(g)(L_d) - \overrightarrow{X}(\Upsilon_d(g))(L_d) = 0,$$

for every section X of AG and every $g \in G$. The map $\Upsilon_d : G \to G$ is called a **discrete flow** or a **discrete Lagrangian evolution operator for L_d**.

3.5.1 Discrete Legendre transformations

Given a discrete Lagrangian $L_d : G \to \mathbb{R}$ we may define two discrete Legendre transformations $\mathbb{F}L_d^{\pm} : G \to A^*G$ by

$$(\mathbb{F}L_d^-)(h)(v_{\epsilon(\alpha(h))}) = -v_{\epsilon(\alpha(h))}(L_d \circ R_h \circ i),$$

for $v_{\epsilon(\alpha(h))} \in (AG)_{\alpha(h)}$ and

$$(\mathbb{F}L_d^+)(g)(v_{\epsilon(\beta(h))}) = v_{\epsilon(\beta(h))}(L_d \circ L_g),$$

for $v_{\epsilon(\beta(h))} \in (AG)_{\beta(h)}$.

A discrete Lagrangian $L_d : G \to \mathbb{R}$ is said to be **regular** if and only if the Legendre transformation $\mathbb{F}L_d^+$ is a local diffeomorphism (equivalently if and only if the Legendre transformation $\mathbb{F}L_d^+$ is a local diffeomorphism). In this case, if $(g_0, h_0) \in G \times G$ is a solution of the discrete Euler-Lagrange equations for L_d, then one may prove (see [118]) that 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 unconstrained Lagrangian evolution operator $\Upsilon_{L_d} : U_0 \to V_0$ such that

1. $\Upsilon_{L_d}(g_0) = h_0$,
2. Υ_{L_d} is a diffeomorphism,
3. Υ_{L_d} is unique, that is, if U_0' is an open subset of G, with $g_0 \in U_0'$, and $\Upsilon'_{L_d} : U_0' \to G$ is a (local) discrete Lagrangian evolution operator, then

$$\Upsilon_{L_d}\big|_{U_0 \cap U_0'} = \Upsilon'_{L_d}\big|_{U_0 \cap U_0'}.$$

Moreover, if $\mathbb{F}L_d^+$ and $\mathbb{F}L_d^-$ are global diffeomorphisms then $\Upsilon_{L_d} = (\mathbb{F}L_d^-)^{-1} \circ (\mathbb{F}L_d^+)$.

If $L_d : G \to \mathbb{R}$ is a regular Lagrangian, then pushing forward to A^*G with the discrete Legendre transformations, we obtain the **discrete Hamiltonian evolution operator**, $\tilde{\Upsilon}_{L_d} : A^*G \to A^*G$ which is given by

$$\tilde{\Upsilon}_{L_d} = (\mathbb{F}L_d^{\pm}) \circ \Upsilon_{L_d} \circ (\mathbb{F}L_d^{\pm})^{-1} = (\mathbb{F}L_d^+) \circ (\mathbb{F}L_d^-)^{-1}.$$

The discrete Hamiltonian evolution operator preserves the Poisson bracket naturally induced by the Lie algebroid structure $\tau : AG \to Q$. In this sense, we are obtaining Poisson preserving numerical methods by considering appropriate discretizations of the continuous reduced Lagrangian.

3.5.2 Examples

Pair or Banal groupoid

We consider the pair (banal) groupoid $G = Q \times Q$ already introduced in §1.5.2, where the structural maps are

$$\alpha(q_1, q_2) = q_1, \quad \beta(q_1, q_2) = q_2, \quad \epsilon(q) = (q, q), \quad i(q_1, q_2) = (q_2, q_1),$$
$$m((q_1, q_2), (q_2, q_3)) = (q_1, q_3).$$

We know that the Lie algebroid of G is isomorphic to the standard Lie algebroid $\tau_Q : TQ \to Q$ and the map

$$\Psi : AG = V_{\epsilon(Q)}\alpha \to TQ, \quad (0_q, v_q) \in T_qQ \times T_qQ \to \Psi_q(0_q, v_q) = v_q, \text{ for } q \in Q,$$

induces an isomorphism (over the identity of Q) between AG and TQ. In the last expression we have that $V_{\epsilon(q)}\alpha = \mathrm{Ker}\,(T_{\epsilon(q)}\alpha)$ (respectively β). If X is a section of $\tau_Q : AG \simeq TQ \to Q$, that is, X is a vector field on Q then $\overrightarrow{X}$ and $\overleftarrow{X}$ are the vector fields on $Q \times Q$ given by

$$\overrightarrow{X}(q_1, q_2) = (-X(q_1), 0_{q_2}) \in T_{q_1}Q \times T_{q_2}Q \quad \text{and} \quad \overleftarrow{X}(q_1, q_2) = (0_{q_1}, X(q_2)) \in T_{q_1}Q \times T_{q_2}Q,$$

for $(q_1, q_2) \in Q \times Q$. On the other hand, if $(q_1, q_2) \in Q \times Q$ we have that the map

$$\begin{array}{rcl} \mathcal{P}^{\tau Q}_{(q_1,q_2)} G \equiv V_{(q_1,q_2)}\beta \oplus V_{(q_1,q_2)}\alpha & \to & T_{(q_1,q_2)}(Q \times Q) \simeq T_{q_1}Q \times T_{q_2}Q, \\ ((v_{q_1}, 0_{q_2}), (0_{q_1}, v_{q_2})) & \to & (v_{q_1}, v_{q_2}) \end{array}$$

induces an isomorphism (over the identity of $Q \times Q$) between the Lie algebroids $\pi^{\tau Q} : \mathcal{P}^{\tau Q} G \equiv V\beta \oplus_G V\alpha \to G = Q \times Q$ and $\tau_{(Q\times Q)} : T(Q \times Q) \to Q \times Q$. Recall that the prolongation of a Lie groupoid over a tangent bundle τ, $\mathcal{P}^\tau G$, was already defined in §1.5.2.

Now, given a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ then the discrete Euler-lagrange equations for L_d are:

$$\overleftarrow{X}(q_1, q_2)(L_d) - \overrightarrow{X}(q_2, q_3)(L_d) = 0, \quad \text{for all } X \in \mathfrak{X}(Q), \tag{3.17}$$

which are equivalent to the classical discrete Euler-Lagrange equations

$$D_2L_d(q_1, q_2) + D_1L_d(q_2, q_3) = 0$$

(see, for instance, [124]). The Poincaré-Cartan 1-sections $\Theta^-_{L_d}$ and $\Theta^+_{L_d}$ on $\pi^{\tau Q} : \mathcal{P}^{\tau Q} G \simeq T(Q \times Q) \to G = Q \times Q$ are the 1-forms on $Q \times Q$ defined by

$$\langle \Theta^-_{L_d}(q_1, q_2), (v_{q_1}, v_{q_2})\rangle = -v_{q_1}(L_d), \quad \langle \Theta^+_{L_d}(q_1, q_2), (v_{q_1}, v_{q_2})\rangle = v_{q_2}(L_d),$$

for $(q_1, q_2) \in Q \times Q$ and $(v_{q_1}, v_{q_2}) \in T_{q_1}Q \times T_{q_2}Q \simeq T_{(q_1,q_2)}(Q \times Q)$.

In addition, if $\xi : G = Q \times Q \to G = Q \times Q$ is a discrete Lagrangian evolution operator then the prolongation of ξ

$$\mathcal{P}^{\tau Q}\xi : \mathcal{P}^{\tau Q} G \simeq T(Q \times Q) \to \mathcal{P}^{\tau Q} G \simeq T(Q \times Q)$$

is just the tangent map to ξ and, thus, we have that

$$\xi^*\Omega_{L_d} = \Omega_{L_d},$$

$\Omega_{L_d} = -\mathrm{d}\Theta^-_{L_d} = -\mathrm{d}\Theta^+_{L_d}$ being the Poincaré-Cartan 2-form on $Q \times Q$. The Legendre transformations $\mathbb{F}L^-_d : G = Q \times Q \to A^*G \simeq T^*Q$ and $\mathbb{F}L^+_d : G = Q \times Q \to A^*G \simeq T^*Q$ associated with L_d are the maps given by

$$\mathbb{F}L^-_d(q_1, q_2) = -D_1L_d(q_1, q_2) \in T^*_{q_1}Q, \quad \mathbb{F}L^+_d(q_1, q_2) = D_2L_d(q_1, q_2) \in T^*_{q_2}Q,$$

for $(q_1, q_2) \in Q \times Q$. The Lagrangian L_d is regular if and only if the matrix $\left(\dfrac{\partial^2 L}{\partial x \partial y}\right)$ is regular, where x and y account for the first and second variables respectively. Finally, a Noether symmetry is a vector field X on Q such that

$$D_1L_d(q_1, q_2)(X(q_1)) + D_2L_d(q_1, q_2)(X(q_2)) = f(q_2) - f(q_1),$$

for $(q_1, q_2) \in Q \times Q$, where $f : Q \to \mathbb{R}$ is a real C^∞-function on Q. If X is a Noether symmetry then

$$q_1 \to f(q_1) = D_1L_d(q_1, q_2)(X(q_1)) - f(q_1)$$

is a constant of the motion.

In conclusion, we recover all the geometrical formulation of the classical discrete mechanics on the discrete state space $Q \times Q$ (see, for instance, [118, 124]).

Lie groups

We consider a Lie group G as a groupoid over one point $M = \{e\}$, the identity element of G. The structural maps are

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

The Lie algebroid associated to G is just the Lie algebra $\mathfrak{g} = T_eG$ of G. Given $\xi \in \mathfrak{g}$ we have the left and right invariant vector fields:

$$\overleftarrow{\xi}(g) = (T_eL_g)(\xi), \quad \overrightarrow{\xi}(g) = (T_eR_g)(\xi), \quad \text{for } g \in G,$$

where, again, L_g and R_g are the left- and right-translations in the Lie group G, respectively. Thus, given a Lagrangian $L_d : G \longrightarrow \mathbb{R}$ its discrete Euler-Lagrange equations are:

$$(T_eL_{g_k})(\xi)(L_d) - (T_eR_{g_{k+1}})(\xi)(L_d) = 0, \quad \text{for all } \xi \in \mathfrak{g} \text{ and } g_k, g_{k+1} \in G,$$

or, $(L^*_{g_k}\mathrm{d}L_d)(e) = (R^*_{g_{k+1}}\mathrm{d}L)(e)$. Denote by $\mu_k = (R^*_{g_k}\mathrm{d}L_d)(e)$ then the discrete Euler-Lagrange equations are written as

$$\mu_{k+1} = \mathrm{Ad}^*_{g_k}\mu_k, \tag{3.18}$$

where $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ is the adjoint action of G on $\mathfrak{g}$ defined in §1.4.3. These equations are known as the **discrete Lie-Poisson equations** (they will be reobtained in a different context in §5). See [18, 121, 122] for further details.

Finally, an infinitesimal symmetry of L_d is an element $\xi \in \mathfrak{g}$ such that $(T_eL_g)(\xi)(L_d) = (T_eR_g)(\xi)(L_d)$, and then the associated constant of the motion is $F(g) = (T_eL_g)(\xi)(L_d) = (T_eR_g)(\xi)(L_d)$. Observe that all the Noether's symmetries are infinitesimal symmetries of L_d.

Chapter 4

Hamiltonian dynamics and constrained variational calculus

The aim of this chapter is to study the relationship between Hamiltonian dynamics and constrained variational calculus or, in other words, vakonomic mechanics (see §2.4.3 for further details). One of our main conclusions is that, under natural regularity conditions, both are equivalent. We describe Hamiltonian and vakonomic mechanics using the notion of Lagrangian submanifolds (see §1.3 for further details) of convenient symplectic manifolds and using the Tulczyjew's triples (see §2.3 for further details). The results are also extended to the case of discrete dynamics. More concretely, our approach allows us to build symplectic integrators and to find out interesting applications to geometrical integration of Hamiltonian systems. Moreover, we analyze in parallel the case of classical nonholonomic mechanics in the discrete and continuous cases: we give a general method to compare common solutions of nonholonomic and vakonomic problems (generalizing the results in [35]) and easily adaptable to more general systems. In order to complete the landscape, some extra geometric notions, not presented in §1, are introduced in §4.1, such as **implicit differential equations**.

4.1 Geometric preliminaries

4.1.1 Lagrangian submanifolds and symplectic structures on the tangent bundle of a symplectic manifold

In §1.3 both symplectic algebra and symplectic geometry were introduced, along with the key notion of **Lagrangian submanifold**. Here we recall the example of Lagrangian submanifold given in §1.3 and, moreover, we introduce some other particular constructions that are interesting for our purposes (see [112, 166]).

An interesting class of Lagrangian submanifolds, which will be useful in this chapter, is the following. Let (P, Ω_P) be a symplectic manifold, where Ω_P is the usual symplectic 2-form, and $g : P \to P$ a diffeomorphism. Denote by $\mathrm{Graph}\,(g)$ the graph of g, that is $\mathrm{Graph}\,(g) = \{(p, g\,(p))\,,\ p \in P\} \subset P \times P$, and by $pr_i : P \times P \to P$, $i = \{0, 1\}$, the canonical projections. Then $\left(P \times P, \tilde{\Omega}_P\right)$, where $\tilde{\Omega}_P = pr_1^*\,\Omega_P - pr_0^*\,\Omega_P$, is a symplectic manifold. Let

$i_g : \mathrm{Graph}\,(g) \hookrightarrow P \times P$ be the inclusion map, then

$$i_g^* \tilde{\Omega}_P = (pr_0)^* \left(g^*\Omega_P - \Omega_P\right).$$

Thus, g is a symplectomorphism (that is, $g^*\Omega_P = \Omega_P$) if and only if Graph g is a Lagrangian submanifold of $P \times P$.

A distinguished symplectic manifold is the cotangent bundle T^*Q of any manifold Q. If we choose local coordinates (q^i), $1 \leq i \leq n$, then T^*Q has induced coordinates (q^i, p_i). Denote by $\pi_Q : T^*Q \to Q$ the canonical projection defined by $\pi_Q(\epsilon_q) = q$, where $\epsilon_q \in T_q^*Q$. Define the Liouville one-form or canonical one-form $\Theta_Q \in \Lambda^1 T^*Q$ by

$$\langle (\Theta_Q)_\epsilon \, , \, X\rangle = \langle \epsilon \, , \, T\pi_Q(X)\rangle, \text{ where } X \in T_\epsilon T^*Q \ , \ \epsilon \in T^*Q.$$

In local coordinates we obtain $\Theta_Q = p_i \, \mathrm{d}q^i$. The canonical two-form Ω_Q on T^*Q is the symplectic form $\Omega_Q = -\mathrm{d}\Theta_Q$ (that is $\Omega_Q = \mathrm{d}q^i \wedge \mathrm{d}p_i$).

Now, we will introduce some special Lagrangian submanifolds of the symplectic manifold (T^*Q, Ω_Q). For instance, the image $\Sigma_\lambda = \lambda(Q) \subset T^*Q$ of a closed one-form $\lambda \in \Lambda^1 Q$ is a Lagrangian submanifold of (T^*Q, Ω_Q), since $\lambda^*\Omega_Q = -\mathrm{d}\lambda$. We then obtain a submanifold diffeomorphic to Q and transverse to the fibers of T^*Q. When λ is exact, that is, $\lambda = \mathrm{d}f$, where $f : Q \to \mathbb{R}$, we say that f is a **generating** function of the Lagrangian submanifold $\Sigma_\lambda = \Sigma_f$. Locally, this is always the case.

A useful extension of the previous construction is the following result due to W.Q. Tulczyjew.

Theorem 4.1.1 ([157],[158])**.** *Let Q be a smooth manifold, $\tau_Q : TQ \to Q$ its tangent bundle projection, $N \subset Q$ a submanifold, and $f\colon N \to \mathbb{R}$. Then*

$$\Sigma_f = \big\{ p \in T^*Q \mid \pi_Q(p) \in N \ \textit{and} \ \langle p, v\rangle = \langle \mathrm{d}f, v\rangle \\ \qquad\qquad \textit{for all } v \in TN \subset TQ \ \textit{such that} \ \tau_Q(v) = \pi_Q(p) \big\}$$

is a Lagrangian submanifold of T^*Q.

Taking f as the zero function we obtain the following Lagrangian submanifold

$$\Sigma_0 = \left\{ p \in T^*Q\big|_N \mid \langle p \, , \, v\rangle = 0 \, , \, \forall \, v \in TN \, \text{with} \, \tau_Q(v) = \pi_Q(p) \right\},$$

which is just the **conormal bundle** of N:

$$\nu^*(N) = \left\{ \xi \in T^*Q\big|_N \, ; \, \xi\big|_{T_{\pi(\xi)}N} = 0 \right\}.$$

Given a symplectic manifold (P, Ω_P), $\dim P = 2n$ it is well-known that its tangent bundle TP is equipped with a symplectic structure denoted by $\mathrm{d}_T\Omega_P$ (see [108]). If we take Darboux coordinates (q^i, p_i) on P, $1 \leq i \leq n$, then $\Omega_P = \mathrm{d}q^i \wedge \mathrm{d}p_i$ and, consequently, we have induced coordinates $(q^i, p_i; \dot{q}^i, \dot{p}_i)$, $(q^i, p_i; a_i, b^i)$ on TP and T^*P, respectively. Thus, $\mathrm{d}_T\Omega_P = \mathrm{d}\dot{q}^i \wedge \mathrm{d}p_i + \mathrm{d}q^i \wedge d\dot{p}_i$ and $\Omega_{P^*} = \mathrm{d}q^i \wedge \mathrm{d}a_i + \mathrm{d}p_i \wedge \mathrm{d}b^i$ (with some abuse of notation Ω_{p^*} denotes the usual symplectic form on T^*P). If we denote by $\flat_{\Omega_P} : TP \to T^*P$ the isomorphism defined by Ω_P, that is $\flat_{\Omega_P}(v) = i_v \Omega_P$, then we have $\flat_{\Omega_P}(q^i, p_i; \dot{q}^i, \dot{p}_i) = (q^i, p_i; -\dot{p}_i, \dot{q}^i)$. Given a function $H : P \to \mathbb{R}$, and its associated Hamiltonian vector field X_H, that is, $i_{X_H}\Omega_{P^*} = \mathrm{d}H$, the image $X_H(P)$ is a Lagrangian submanifold of $(TP, \mathrm{d}_T\Omega_P)$. Moreover, given a vector field $X \in \mathfrak{X}(P)$, it is locally Hamiltonian if and only if its image $X(P)$ is a Lagrangian submanifold of $(TP, \mathrm{d}_T\Omega_P)$. It is interesting to note that $\mathrm{d}_T\Omega_P = -\flat^*_{\Omega_P}\Omega_P$ and $\flat_\omega(X_H(P)) = \mathrm{d}H(P)$.

An important notion in the theory of Lagrangian submanifolds is the concept of generating function. If we have a Lagrangian submanifold N of an exact symplectic manifold $(P, \Omega_P = \mathrm{d}\Theta_P)$, where $\Theta_P \in \Lambda^1 P$, then $0 = i_N^*\Omega_P = i_N^*\mathrm{d}\Theta_P = \mathrm{d}(i_N^*\Theta_P)$. Consequently, applying the Poincaré's lemma, there exists a local function $S : U \to \mathbb{R}$ defined on a open neighborhood U of N such that $i_N^*\Theta_P = \mathrm{d}S$. We say that S is a (local) **generating function** of the Lagrangian submanifold N. The concept of generating function was already introduced in the context of canonical transformations in §2.2.2.

4.1.2 Implicit differential equations

An implicit differential equation on a general smooth manifold Q is a submanifold $E \subset TQ$. A solution of E is any curve $\gamma : I \to Q$, $I \subset \mathbb{R}$, such that the tangent curve $(\gamma(t), \dot{\gamma}(t)) \in E$ for all $t \in I$. The implicit differential equation will be said to be **integrable at a point** if there exists a solution γ of E such that the tangent curve passes through it. Furthermore, the implicit differential equation will be said to be **integrable** if it is integrable at *all points*. Unfortunately, integrability does not mean uniqueness. The integrable part of E is the subset of all integrable points of E. The **integrability problem** consists in identifying such a subset.

Denoting the canonical projection $\tau_Q : TQ \to Q$, a sufficient condition for the integrability of E is

$$E \subset TQ,$$

where $C = \tau_Q(E)$, provided that the projection τ_Q restricted is a submersion onto C.

Extracting the integrable part of E

A recursive algorithm was presented in [133] that allows to extract the integrable part of an implicit differential equation E. We shall define the subsets

$$E_0 = E, \quad C_0 = C,$$

and recursively for every $k \geq 1$,

$$E_k = E_{k-1} \cap TC_{k-1}, \quad C_k = \tau_Q(E_k),$$

then, eventually the recursive construction will stabilize in the sense that $E_k = E_{k+1} = ... = E_\infty$, and $C_k = C_{k+1} = ... = C_\infty$. It is clear by construction that $E_\infty \subset TC_\infty$. Then, provided that the adequate regularity conditions are satisfied during the application of the algorithm, the implicit differential equations E_∞ will be integrable and it will solve the integrability problem.

4.2 Continuous Lagrangian and Hamiltonian mechanics

In §2.3 and §2.3.1 we have shown how the Tulczyjew's triple is used to describe geometrically Lagrangian and Hamiltonian mechanics and its relationship. In this section we will see that it is also possible to adapt this geometric formalism when we introduce constraints into the picture. As was shown in §2.4, there are (at least) two methods that one might use to derive the equations of motion of systems subjected to constraints: **nonholonomic mechanics** (§2.4.2) and constrained variational calculus (or **vakonomic mechanics** §2.4.3). The classical method to derive equations of motion for constrained mechanical systems is the nonholonomic mechanics. The equations derived from the nonholonomic methods *are not* of variational nature, but they describe the correct dynamics of a constrained mechanical system. In order to obtain the nonholonomic equations, if we have linear or affine constraints, is necessary to apply the Lagrange-D'Alembert's principle. When dealing with nonlinear constraints, one should employ the more controversial Chetaev's rule (see [14, 103] for further details). Since the geometrical implementation of the Chetaev's rule is practically equal to the process in the linear case, we shall use it from a pure mathematical perspective.

On the other hand, the equations of motion of constrained variational problems are derivable by using variational techniques (always in the constrained case). These last equations are also known in the literature as vakonomic equations. The terminology vakonomic ("mechanics of variational axiomatic kind") was coined by V.V. Kozlov ([5],[96]). The main applications of the constrained variational calculus appear in problems of mathematical nature (like subriemannian geometry) and in optimal control theory.

4.2.1 Nonholonomic mechanics

A nonholonomic system on a manifold Q consists of a pair $(\mathbb{L}, C)$, where $\mathbb{L} : TQ \to \mathbb{R}$ is the Lagrangian of the mechanical system and C is a submanifold of TQ with canonical inclusion $i_C : C \hookrightarrow TQ$. In the following, we will assume, for sake of simplicity, that $\tau_Q(C) = Q$. Since the motion of the system is forced to take place on the submanifold C, this requires the introduction of some reaction or constraint forces into the system. If $\phi^\alpha(q^i, \dot{q}^i) = 0$, $1 \leq \alpha \leq n$, determine locally the submanifold C, then Chetaev's rule implies that the constrained equations of the system are:

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbb{L}}{\partial \dot{q}^i}\right) - \frac{\partial \mathbb{L}}{\partial q^i} = \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\ &\phi^\alpha(q^i, \dot{q}^i) = 0, \quad 1 \leq \alpha \leq n. \end{aligned} \tag{4.1}$$

Next, we will describe geometrically the nonholonomic equations. First, we need to introduce the vertical endomorphism S which is a $(1,1)$-tensor field on TQ defined by

$$\begin{array}{rccl} S: & TTQ & \longrightarrow & TTQ \\ & W_{v_x} & \longmapsto & \left.\dfrac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (v_x + t\, T\tau_Q(W_{v_x}))\,. \end{array}$$

Its local expression is $S = \frac{\partial}{\partial \dot{q}^i} \otimes \mathrm{d}q^i$.

If we accept Chetaev-type forces, then we define

$$F = S^*(TC)^0.$$

Observe that the vector subbundle F will be generated by the 1-forms $\mu^\alpha = \frac{\partial \phi^\alpha}{\partial \dot{q}^i}\, \mathrm{d}q^i$ because $S^*(\mathrm{d}\phi^\alpha) = \mu^\alpha$.

Now, define the affine subbundle of $T^*_C TQ$ given by

$$\Sigma^{noh} = (\mathrm{d}\mathbb{L}) \circ i_C + F,$$

that is,

$$\begin{aligned} \Sigma^{noh} \quad = \quad & \{(q^i, \dot{q}^i, \mu_i, \tilde{\mu}_i) \in T^*TQ \mid \\ & \mu_i = \frac{\partial \mathbb{L}}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\ & \tilde{\mu}_i = \frac{\partial \mathbb{L}}{\partial \dot{q}^i}, \\ & \phi^\alpha(q, \dot{q}) = 0, \;\; 1 \leq \alpha \leq n\} \,. \end{aligned} \tag{4.2}$$

Therefore, applying the Tulczyjew's isomorphism α_Q we obtain the affine subbundle

$$\begin{aligned} \alpha_Q^{-1}\left(\Sigma^{noh}\right) \quad = \quad & \{(q^i, p_i, \dot{q}^i, \dot{p}_i) \in TT^*Q \mid \\ & p_i = \frac{\partial \mathbb{L}}{\partial \dot{q}^i}, \\ & \dot{p}_i = \frac{\partial \mathbb{L}}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\ & \phi^\alpha(q, \dot{q}) = 0, \;\; 1 \leq \alpha \leq n\} \,. \end{aligned} \tag{4.3}$$

Define now the nonholonomic Legendre transformation $\mathbb{FL}^{noh} : C \to T^*Q$ by

$$\mathbb{FL}^{noh} = \pi_{T^*Q} \circ \alpha_Q^{-1} \circ d\mathbb{L} \circ i_C \,.$$

The solutions for the dynamics given by $\alpha_Q^{-1}\left(\Sigma^{noh}\right)$ are curves $\sigma : I \subset \mathbb{R} \to Q$ such that $\frac{\mathrm{d}\sigma}{\mathrm{d}t}(I) \subset C$ and the induced curve $\gamma : \mathbb{R} \to T^*Q$, $\gamma = \mathbb{FL}^{noh}(\frac{\mathrm{d}\sigma}{\mathrm{d}t})$ verifies that $\frac{\mathrm{d}\gamma}{\mathrm{d}t}(I) \subset \alpha_Q^{-1}\left(\Sigma^{noh}\right)$. Locally, σ must satisfy the system of equations (4.1).

An interesting use of Tulczyjew's triple in order to define Lagrangian submanifolds and generalized Legendre transformations within the nonholonomic framework can be found in [125].

4.2.2 Variational constrained equations (vakonomic equations)

Now, we study the same problem but now using purely variational techniques. As above, let consider a regular Lagrangian $\mathbb{L} : TQ \to \mathbb{R}$, and a set of nonholonomic constraints $\phi^\alpha(q^i, \dot{q}^i)$, $1 \leq \alpha \leq n$, determining a $2m - n$ dimensional submanifold $C \subset TQ$. Now we take the extended Lagrangian $\mathcal{L} = \mathbb{L} + \lambda_\alpha \phi^\alpha$ which includes the Lagrange multipliers λ_α as new extra variables. The equations of motion for the constrained variational problem are the Euler-Lagrange equations for $\mathcal{L}$, that is:

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbb{L}}{\partial \dot{q}^i}\right) - \frac{\partial \mathbb{L}}{\partial q^i} = -\dot{\lambda}_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i} - \lambda_\alpha \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \phi^\alpha}{\partial \dot{q}^i}\right) - \frac{\partial \phi^\alpha}{\partial q^i}\right], \\
&\phi^\alpha(q^i, \dot{q}^i) = 0, \quad 1 \leq \alpha \leq n.
\end{aligned}
\tag{4.4}
$$

From a geometrical point of view, these type of variationally constrained problems are determined by a pair (C, L) where C is a submanifold of TQ, with inclusion $i_C : C \hookrightarrow TQ$, and $L : C \to \mathbb{R}$ a Lagrangian function. Using Theorem 4.1.1 we deduce that Σ_L is a Lagrangian submanifold of (T^*TQ, Ω_{TQ}) (see [54]). Now using the Tulczyjew's symplectomorphism α_Q, we induce a new Lagrangian submanifold $\alpha_Q^{-1}(\Sigma_L)$ of $(TT^*Q, \mathrm{d}_T\Omega_Q)$, which completely determines the constrained variational dynamics. Of course, the case of unconstrained Lagrangian mechanics is generated taking the whole space TQ instead of C and an a Lagrangian function over the tangent bundle $L : TQ \to \mathbb{R}$.

Next we shall prove that, indeed, this procedure gives the correct equations for the constrained variational dynamics. Take an arbitrary extension $\mathbb{L} : TQ \to \mathbb{R}$ of $L : C \to \mathbb{R}$, that is, $\mathbb{L} \circ i_C = L$. As above, assume also that we have fixed local constraints such that locally determines C by their vanishing, i.e: $\phi^\alpha(q, \dot{q}) = 0$, $1 \leq \alpha \leq n$ where $n = 2\dim Q - \dim C$.

Locally

$$
\begin{aligned}
\Sigma_L = \{ & (q^i, \dot{q}^i, \mu_i, \tilde{\mu}_i) \in T^*TQ \mid \\
& \mu_i = \frac{\partial \mathbb{L}}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i}, \\
& \tilde{\mu}_i = \frac{\partial \mathbb{L}}{\partial \dot{q}^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\
& \phi^\alpha(q, \dot{q}) = 0, \ \ 1 \leq \alpha \leq n \} \,.
\end{aligned}
\tag{4.5}
$$

Observe that locally the conormal bundle $\nu^*(C) = \text{span}\ \{\mathrm{d}\phi^\alpha, 1 \leq \alpha \leq n\}$.

Therefore,

$$
\begin{aligned}
\alpha_Q^{-1}(\Sigma_L) = \{ & (q^i, p_i, \dot{q}^i, \dot{p}_i) \in TT^*Q \mid \\
& p_i = \frac{\partial \mathbb{L}}{\partial \dot{q}^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\
& \dot{p}_i = \frac{\partial \mathbb{L}}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i}, \\
& \phi^\alpha(q, \dot{q}) = 0, \ \ 1 \leq \alpha \leq n \} \,.
\end{aligned}
\tag{4.6}
$$

The solutions for the dynamics given by $\alpha_Q^{-1}(\Sigma_L) \subset TT^*Q$ are curves $\gamma : I \subset \mathbb{R} \to T^*Q$ such that $\frac{\mathrm{d}\gamma}{\mathrm{d}t} : I \subset \mathbb{R} \to TT^*Q$ verifies that $\frac{\mathrm{d}\gamma}{\mathrm{d}t}(I) \subset \alpha_Q^{-1}(\Sigma_L)$. Locally, if $\gamma(t) = (q^i(t), p_i(t))$ then it must verify the following set of differential equations:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbb{L}}{\partial \dot{q}^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}\right) - \frac{\partial \mathbb{L}}{\partial q^i} - \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i} &= 0, \\
\phi^\alpha(q^i, \dot{q}^i) &= 0,
\end{aligned}
$$

which clearly coincide with equations (4.4).

Now, we consider adapted coordinates $(q^i, \dot{q}^a)$ to the submanifold C (recall that $\tau_Q(C) = Q$ is now a fibration $C \to Q$), $1 \leq i \leq \dim Q$ and $1 \leq a \leq \dim Q - n$, such that

$$i_C(q^i, \dot{q}^a) = \left(q^i, \dot{q}^a, \Psi^\alpha(q^i, \dot{q}^a)\right) .$$

This means that $\phi^\alpha(q^i, \dot{q}^i) = \dot{q}^\alpha - \Psi^\alpha(q^i, \dot{q}^a) = 0$. Therefore, we have

$$\begin{aligned} \Sigma_L \quad = \quad & \{(q^i, \dot{q}^i, \mu_i, \tilde{\mu}_i) \in T^*TQ | \\ & \mu_i = \frac{\partial L}{\partial q^i} - \tilde{\mu}_\alpha \frac{\partial \Psi^\alpha}{\partial q^i}, \\ & \tilde{\mu}_a = \frac{\partial L}{\partial \dot{q}^a} - \tilde{\mu}_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}, \\ & \dot{q}^\alpha = \Psi^\alpha(q^i, \dot{q}^a), \ \ 1 \leq \alpha \leq n\} . \end{aligned} \tag{4.7}$$

Observe that $(q^i, \dot{q}^a, \tilde{\mu}_\alpha)$ determines a local system of coordinates for Σ_L.

Then,

$$\begin{aligned} \alpha_Q^{-1}(\Sigma_L) \quad = \quad & \{(q^i, p_i, \dot{q}^i, \dot{p}_i) \in TT^*Q \ | \\ & p_a = \frac{\partial L}{\partial \dot{q}^a} - p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}, \\ & \dot{p}_i = \frac{\partial L}{\partial q^i} - p_\alpha \frac{\partial \Psi^\alpha}{\partial q^i}, \\ & \dot{q}^\alpha = \Psi^\alpha(q^i, \dot{q}^a), \ \ 1 \leq \alpha \leq n\} . \end{aligned} \tag{4.8}$$

Consequently, the solutions must verify the following system of differential equations (see [35]):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^a} - p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right) \quad &= \quad \frac{\partial L}{\partial q^a} - p_\alpha \frac{\partial \Psi^\alpha}{\partial q^a} \\ \dot{p}_\beta \quad &= \quad \frac{\partial L}{\partial q^\beta} - p_\alpha \frac{\partial \Psi^\alpha}{\partial q^\beta}, \\ & \quad \dot{q}^\alpha = \Psi^\alpha(q^i, \dot{q}^a), \ \ 1 \leq \alpha \leq n . \end{aligned}$$

The constrained Legendre transformation

Definition 4.2.1. *We define the* constrained Legendre transformation $\mathbb{F}L : \Sigma_L \longrightarrow T^*Q$ *as the mapping* $\mathbb{F}L = \tau_{T^*Q} \circ (\alpha_Q^{-1})|_{\Sigma_L}$.

We will say that the constrained system (L, C) *is* **regular** *if* $\mathbb{F}L$ *is a local diffeomorphism and* **hyperregular** *if* $\mathbb{F}L$ *is a global diffeomorphism.*

Observe that locally, if as above we consider the constraints $\dot{q}^\alpha = \Psi^\alpha(q^i, \dot{q}^a)$ determining C, then

$$\mathbb{F}L(q^i, \dot{q}^a, \tilde{\mu}_\alpha) = (q^i, p_a = \frac{\partial L}{\partial \dot{q}^a} - \tilde{\mu}_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}, p_\alpha = \tilde{\mu}_\alpha) .$$

The constrained system (L, C) is regular if and only if $\left(\frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^b} - \tilde{\mu}_\alpha \frac{\partial^2 \Psi^\alpha}{\partial \dot{q}^a \partial \dot{q}^b}\right)$ is a nondegenerate matrix.

Next, define the **energy function** $E_L : \Sigma_L \to \mathbb{R}$ by

$$E_L(\alpha_u) = \langle \alpha_u, u_u^\vee \rangle - L(u), \quad \alpha_u \in \Sigma_L, u \in C \equiv i_C(C)$$

Locally, we have

$$E_L(q^i, \dot{q}^a, \tilde{\mu}_\alpha) = \dot{q}^a \frac{\partial L}{\partial \dot{q}^a} - \tilde{\mu}_\alpha \frac{\partial \Psi^\alpha(q^i, \dot{q}^a)}{\partial \dot{q}^a} \dot{q}^a + \tilde{\mu}_\alpha \Psi^\alpha(q^i, \dot{q}^a) - L(q^i, \dot{q}^a).$$

Remark 4.2.2. *The constrained Legendre transformation allows us to develop a Lagrangian formalism on Σ_L. Indeed, we can define the 2-form $\Omega_L = (\mathbb{F}L)^*\Omega_Q$ on Σ_L and it is easy to show that the equations of motion of the constrained system are now intrinsically rewritten as*

$$i_X \Omega_L = \mathrm{d}E_L.$$

In consequence, we could develop an intrinsic formalism on the Lagrangian side, that is a Klein formalism ([49, 55, 88, 108]) *for constrained systems without using (at least initially) Lagrangian multipliers. Moreover, notice that the constrained system is regular if and only if Ω_L is a symplectic 2-form on Σ_L*

Then, if the constrained system (L, C) is hyperregular we can define the Hamiltonian function $H : T^*Q \to \mathbb{R}$ by

$$H = E_L \circ (\mathbb{F}L)^{-1} ,$$

and the corresponding Hamiltonian vector field X_H by $i_{X_H}\Omega_Q = \mathrm{d}H$. In this particular case we have that

$$\mathrm{Im} X_H = X_H(T^*Q) = \beta_Q^{-1}(\mathrm{d}H(T^*Q)) = \alpha_Q^{-1}(\Sigma_L).$$

The second equivalence, $X_H(T^*Q) = \alpha_Q^{-1}(\Sigma_L)$, will be studied below. The next diagram summarizes the above discussion:

$$\begin{tikzcd}
TT^*Q \arrow[rr, "\alpha_Q"] \arrow[dd, "\tau_{T^*Q}"'] & & T^*TQ \\
& \alpha_Q^{-1}(\Sigma_L) \arrow[ul, hook] & \\
T^*Q \arrow[dr, "H"'] & & \Sigma_L \arrow[ll, "\mathbb{F}L"] \arrow[ul, "\alpha_Q^{-1}"'] \arrow[uu, hook] \arrow[dl, "E_L"] \\
& \mathbb{R} &
\end{tikzcd}$$

In the singular case, it is necessary to apply the integrability algorithm briefly described in §4.1.2 to find, if it exists, a subset where there are consistent solutions of the dynamics (see [53, 51, 52] for further details).

4.2.3 Comparison of nonholonomic and variational constrained equations. Continuous picture

Let consider a system defined by the Lagrangian function $\mathbb{L} : TQ \to \mathbb{R}$ and an independent set of constraints $\phi^\alpha(q^i, \dot{q}^i) = 0$ determining the submanifold $C \subset TQ$.

As shown in §4.2.1 and §4.2.2, the solutions of the nonholonomic dynamics are geometrically described by the submanifold $\Sigma^{noh} \hookrightarrow TT^*Q$, while the solutions of the constrained variational dynamics are given by the Lagrangian submanifold $\Sigma_L \hookrightarrow TT^*Q$, where $L = \mathbb{L}\big|_C : C \to \mathbb{R}$.

Our aim is to know whether, given a solution of the nonholonomic problem, it is also a solution of the constrained variational problem. In order to capture the common solutions to both problems, we have developed the following geometric algorithm. Consider the fibered product $T^*Q \oplus T^*Q$, where we choose the local coordinates (q^i, p_i, π_i); consider also the tangent bundle $T(T^*Q \oplus T^*Q)$ which can be identified with $TT^*Q \oplus_{T\pi_Q} TT^*Q$, which fibers over TQ. Under these considerations, construct the submanifold $\Sigma^{cons} \hookrightarrow TT^*Q \oplus_{T\pi_Q} TT^*Q$ as follows:

$$
\begin{aligned}
\Sigma^{cons} \quad &= \quad \{(X_{\alpha_q}, Y_{\beta_q}) \in TT^*Q \oplus_{T\pi_Q} TT^*Q \,/\, T_{\alpha_q}\pi_Q(X_{\alpha_q}) = T_{\beta_q}\pi_Q(Y_{\beta_q}), \\
&\text{for} \quad X_{\alpha_q} \in \Sigma^{noh}, Y_{\beta_q} \in \Sigma_L\}.
\end{aligned} \tag{4.9}
$$

It is quite clear that the submanifold Σ^{cons} gathers together both nonholonomic and constrained variational dynamics. Locally, Σ^{cons} is determined by the coordinates $(q^i, p_i, \pi_i, \dot{q}^i, \dot{p}_i, \dot{\pi}_i)$ obeying the nonholonomic and constrained variational conditions respectively presented in (4.2) and (4.5), that is

$$
\begin{aligned}
p_i &= \frac{\partial \mathbb{L}}{\partial \dot{q}^i}, & \pi_i &= \frac{\partial \mathbb{L}}{\partial \dot{q}^i} + \mu_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, \\
\dot{p}_i &= \frac{\partial \mathbb{L}}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}, & \dot{\pi}_i &= \frac{\partial \mathbb{L}}{\partial q^i} + \mu_\alpha \frac{\partial \phi^\alpha}{\partial q^i},
\end{aligned}
$$

subject to $\phi^\alpha(q^i, \dot{q}^i) = 0$. Here, λ_α and μ_α are the nonholonomic and variational constrained Lagrange multipliers, respectively. Finally, in order to find the common solutions we will need to apply the integrability algorithm described in [35, 133].

The following diagram shows the bundle relations:

$$
\begin{tikzcd}
\Sigma^{cons} \arrow[r, hook] & TT^*Q \oplus_{T\pi_Q} TT^*Q \arrow[r, "\widetilde{T\pi_Q}"] \arrow[d, "{(\tau_{T^*Q}, \tau_{T^*Q})}"'] & TQ \arrow[d, "\tau_Q"] \\
& T^*Q \oplus T^*Q \arrow[r, "\widetilde{\pi_Q}"'] & Q
\end{tikzcd}
$$

where $\widetilde{T\pi_Q} : TT^*Q \oplus_{T\pi_Q} TT^*Q \longrightarrow TQ$ and $\widetilde{\pi_Q} : T^*Q \oplus T^*Q \longrightarrow Q$ denote the fibrations of the Whitney sums over TQ and Q, respectively.

As a simple example, consider the case of linear constraints, namely $\phi^\alpha(q^i,\dot q^i) = \dot q^\alpha - \varphi^\alpha_a(q^i)\,\dot q^a = 0$. The relationship determining Σ^{cons} presented in (4.25), as well as the integrability algorithm, implies the following equation:

$$\dot q^a \mu_\alpha R^\alpha_{ab} = 0, \tag{4.10}$$

where

$$R^\alpha_{ab} = \frac{\partial\varphi^\alpha_b}{\partial q^a} - \frac{\partial\varphi^\alpha_a}{\partial q^b} + \varphi^\beta_a\frac{\partial\varphi^\alpha_b}{\partial q^\beta} - \varphi^\beta_b\frac{\partial\varphi^\alpha_a}{\partial q^\beta}$$

can be considered as the curvature of the connection Γ in the local projection $\rho(q^a,q^\alpha) = (q^\alpha)$ such that the horizontal distribution $\mathcal{H}$ is given by prescribing its annihilator to be

$$\mathcal{H}^0 = \{\mathrm{d}q^\alpha - \varphi^\alpha_a \mathrm{d}q^a\,,\, 1\leq\alpha\leq m\}.$$

See more details in [35].

4.2.4 Lagrangian and Hamiltonian mechanics relationship

In this section, we shall discuss the converse case, i.e., starting from a Hamiltonian system we shall show that it is possible to construct a constrained Lagrangian system providing the same dynamics.

Since $\pi_Q : T^*Q \to Q$ is a vector bundle, then it is possible to define the dilation vector field or Liouville vector field $\Delta^* \in \mathfrak{X}(T^*Q)$, which is the generator of the one-parameter group of dilations along the vertical fibres $\mu_q \longrightarrow e^t\mu_q$, $\mu_q \in T^*_qQ$. The dilation vector field is locally expressed by

$$\Delta^* = p_i\frac{\partial}{\partial p_i}.$$

Given a Hamiltonian function $H : T^*Q \to \mathbb{R}$, the Hamilton's equations are written in canonical coordinates by

$$\begin{aligned} \dot q^i &= \frac{\partial H}{\partial p_i}, \\ \dot p_i &= -\frac{\partial H}{\partial q^i}. \end{aligned} \tag{4.11}$$

The solutions of the Hamilton's equations are just the integral curves of the Hamiltonian vector field given by

$$i_{X_H}\Omega_Q = \mathrm{d}H, \tag{4.12}$$

where Ω_Q is a symplectic form on T^*Q .

Given the Hamiltonian function, one defines a function $\mathbb{F}H : T^*Q \to TQ$, i.e. the fiber derivative of H (see [1]), by

$$\langle \mathbb{F}H(\alpha_q)\,,\,\beta_q\rangle = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} H(\alpha_q + t\beta_q),$$

where both $\alpha_q, \beta_q \in T^*_qQ$. In local coordinates,

$$\mathbb{F}H(q^i,p_i) = (q^i, \frac{\partial H}{\partial p_i})\ .$$

Assume that the image of T^*Q under $\mathbb{F}H$ defines a submanifold C of TQ. Mimicking the Gotay and Nester's definition ([51, 52]), we implicitly define the function $L : C \to \mathbb{R}$ by

$$L \circ \mathbb{F}H = \Delta^* H - H. \tag{4.13}$$

The function $L : C \to \mathbb{R}$ will be well-defined if and only if, for any two points $\alpha_q, \beta_q \in T^*Q$ such that $\mathbb{F}H(\alpha_q) = \mathbb{F}H(\beta_q)$, we have that $(\Delta^* H - H)(\alpha_q) = (\Delta^* H - H)(\beta_q)$. Obviously, without additional assumptions there is no reason why this should be true. The following definition states under what conditions such projection L exists.

Definition 4.2.3. *A Hamiltonian $H : T^*Q \to \mathbb{R}$ is* **almost-regular** *if $\mathbb{F}H(T^*Q) = C$ is a submanifold of TQ and $\mathbb{F}H : T^*Q \to C \subset TQ$ is a submersion with connected fibers.*

$$\begin{tikzcd}
T^*Q \arrow[rr, "\mathbb{F}H"] \arrow[drr] & & TQ \\
& & \mathbb{F}H(T^*Q) = C \arrow[u, hook]
\end{tikzcd}$$

Under the assumption that the Hamiltonian H is almost-regular, it is only necessary to show that expression (4.13) defines a single-valued function $L : C \to \mathbb{R}$, or, in other words, that $\Delta^* H - H$ is a constant function in the fibers. Since each fiber of the submersion $\mathbb{F}H$ is connected, it is sufficient to consider the infinitesimal condition, i.e. to show that

$$\mathcal{L}_Z(\Delta^* H - H) = 0, \quad \text{for all} \quad Z \in \ker(\mathbb{F}H_*) \, ,$$

where $\mathcal{L}_Z$ is the Lie derivative in the Z direction. Working in local coordinates, Z will be of the form

$$Z = Z_i \frac{\partial}{\partial p_i}, \text{ with } Z_i \, \frac{\partial^2 H}{\partial p_i \, \partial p_j} = 0, \text{ for all } 1 \leq j \leq n \, .$$

Since $\mathbb{F}H_*(Z) = 0$, the last condition can be obtained taking into account that $\langle \sigma \, , \, \mathbb{F}H_*(Z) \rangle = \langle \mathbb{F}H^*(\sigma) \, , \, Z \rangle = 0$, for σ an arbitrary point of T^*TQ such that $\pi_{TQ}(\sigma) = \tau_{TQ} \left(\mathbb{F}H_*(Z) \right)$. Then

$$\begin{aligned}
\mathcal{L}_Z(\Delta^* H - H) &= Z_i \frac{\partial \left(\Delta^* H - H \right)}{\partial p_i} \\
&= Z_i \frac{\partial H}{\partial p_i} + p_j Z_i \, \frac{\partial^2 H}{\partial p_i \, \partial p_j} - Z_i \, \frac{\partial H}{\partial p_i} = 0.
\end{aligned}$$

In what follows we will assume that H satisfies the almost regularity property.

Theorem 4.2.4. *The following equality holds*

$$\alpha_Q(X_H(T^*Q)) = \Sigma_L \, ,$$

where α_Q is the Tulczyjew's isomorphism.

Proof. Take $W_1 \in \Sigma_{X_H} = X_H(T^*Q) \subset TT^*Q$ and take $\alpha_Q(W_1) \in T^*TQ$. We need to prove that

$$\langle \alpha_Q(W_1), U \rangle = \langle \mathrm{d}L, U \rangle \,,$$

for all $U \in T\,C \subset TTQ$, such that $\tau_{TQ}(U) = T\pi_Q(W_1)$. This is equivalent to the equality

$$\langle \alpha_Q(W_1)\,,\, \mathbb{F}H_*(W_2) \rangle = \langle \mathrm{d}L\,,\, \mathbb{F}H_*(W_2) \rangle \,, \tag{4.14}$$

for all $W_2 \in TT^*Q$ such that $\tau_{TQ}\left(\mathbb{F}H_*(W_2)\right) = T\pi_Q(W_1)$. Therefore, regarding (4.14) the previous equality turns out to be

$$(\mathbb{F}H)^*\left(\alpha_Q(W_1)\right) = (\mathbb{F}H)^*\left(\mathrm{d}L\right) = \mathrm{d}(\Delta^* H - H)\,, \tag{4.15}$$

where the right hand side of the equation comes directly from (4.13). If locally $W_1 = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i}$, in other words $W_1 = (q^i, p_i; \dot{q}^i, \dot{p}_i)$, then

$$\begin{aligned}
(\mathbb{F}H)^*\left(\alpha_Q(W_1)\right) &= \left(\dot{p}_i + p_j \frac{\partial^2 H}{\partial p_j \partial q^i}\right) dq^i + p_j \frac{\partial^2 H}{\partial p_j \partial p_i}\, dp_i \\
&= \left(p_j \frac{\partial^2 H}{\partial p_j \partial q^i} - \frac{\partial H}{\partial q^i}\right) dq^i + p_j \frac{\partial^2 H}{\partial p_j \partial p_i}\, dp_i,
\end{aligned} \tag{4.16}$$

where we consider $\dot{p}_i = -\frac{\partial H}{\partial q^i}$ since we are dealing with a Hamiltonian vector field X_H (see equations (4.11)). From condition, $\tau_{TQ}\mathbb{F}H_*(W_2) = T\pi_Q(W_1)$ we also deduce that $\dot{q}^i = \frac{\partial H}{\partial p_i}$. But this is true since $W_1 \in \Sigma_{X_H}$.

On the other hand, a straightforward computation leads to check that $\mathrm{d}(\Delta^* H - H)$ is exactly (4.16). □

4.3 Discrete equivalence

4.3.1 Discrete nonholonomic mechanics

A discrete nonholonomic system is determined by three ingredients: a discrete lagrangian $\mathbb{L}_d : Q \times Q \to \mathbb{R}$, a constraint distribution $\mathcal{D}$ on Q and a discrete constraint submanifold $C_d \subset Q \times Q$ with canonical inclusion $i_{C_d} : C_d \hookrightarrow Q \times Q$. Notice that discrete mechanics can also be seen within this case, where $\mathcal{D} = TQ$ and $C_d = Q \times Q$. Notice also that, in the discrete version of Lagrangian mechanics, the tangent manifold TQ is substituted by the cartesian product $Q \times Q$.

Define the affine submanifold $\Sigma_d^{noh} \subset T^*(Q \times Q)$ by $\Sigma_d^{noh} = (\mathrm{d}\mathbb{L}_d) \circ i_{C_d} + F_d$, where F_d is the vector subbundle of $T^*_{C_d}(Q \times Q)$ given by

$$F_d = \left(pr_1^* \,\mathcal{D}^0\right)\big|_{C_d}\,.$$

Here, $pr_1 : Q \times Q \to Q$ is the first projection onto Q.

The symplectic manifold $(T^*(Q\times Q)\,,\,\Omega_{Q\times Q})$ is symplectomorphic to $(T^*Q\times T^*Q\,,\,\tilde{\Omega}_Q)$, where $\tilde{\Omega}_Q = pr_1^*\Omega_Q - pr_0^*\Omega_Q$ and $pr_i : T^*Q\times T^*Q \to T^*Q$, $i=0,1$, are the natural projections of $T^*Q\times T^*Q$ onto T^*Q. The symplectomorphism is given by

$$\begin{array}{rcl}\Upsilon : T^*(Q\times Q) & \to & T^*Q\times T^*Q\\ \gamma_{(q_0,q_1)} \equiv (\gamma_{q_0},\gamma_{q_1}) & \mapsto & (-\gamma_{q_0},\gamma_{q_1})\end{array}$$

where $(q_0,q_1)\in Q\times Q$.

Using Υ we induce the affine subbundle $\Upsilon(\Sigma_d^{noh})$ of the symplectic manifold $(T^*Q\times T^*Q,\tilde{\Omega}_Q)$. The dynamics is then determined by the sequences $\gamma_{q_0},\gamma_{q_1},\ldots,\gamma_{q_N}$ such that $(\gamma_{q_i},\gamma_{q_{i+1}})\in\Upsilon(\Sigma_d^{noh})$, $0\leq i\leq N-1$ (see [71]).

We will describe now the equations in terms of local coordinates. Assume that C_d is defined by the vanishing of the following set of independent constraints: $\phi_d^\alpha(q_0,q_1)=0$, $1\leq\alpha\leq n$, where $n = 2\dim\, Q - \dim\, C$ and $\mathcal{D}^0 = \mathrm{span}\{\omega^\alpha = \omega_i^\alpha\,\mathrm{d}q^i\}$. Therefore,

$$\begin{array}{rcl}\Sigma_d^{noh} & = & \{(q_0^i,q_1^i,(\mu_0)_i,(\mu_1)_i)\in T^*(Q\times Q)\mid\\ & & (\mu_0)_i = \dfrac{\partial\mathbb{L}_d}{\partial q_0^i} + (\lambda_0)_\alpha\,\omega_i^\alpha,\\ & & (\mu_1)_i = \dfrac{\partial\mathbb{L}_d}{\partial q_1^i},\\ & & \phi_d^\alpha(q_0,q_1)=0,\ \ 1\leq\alpha\leq n\}\ ,\end{array}\tag{4.17}$$

where $(\lambda_\alpha)_0$ are Lagrange multipliers to be determined. Then,

$$\begin{array}{rcl}\Upsilon(\Sigma_d^{noh}) & = & \{(q_0^i,(p_0)_i,q_1^i,(p_1)_i)\in T^*Q\times T^*Q\mid\\ & & (p_0)_i = -\dfrac{\partial\mathbb{L}_d}{\partial q_0^i} - (\lambda_0)_\alpha\,\omega_i^\alpha,\\ & & (p_1)_i = \dfrac{\partial\mathbb{L}_d}{\partial q_1^i},\\ & & \phi_d^\alpha(q_0,q_1)=0,\ \ 1\leq\alpha\leq n\}\ .\end{array}\tag{4.18}$$

The solutions of the dynamics are therefore given by the equations

$$\begin{array}{l}\dfrac{\partial\mathbb{L}_d}{\partial q_1^i}(q_{k-1},q_k) = -\dfrac{\partial\mathbb{L}_d}{\partial q_0^i}(q_k,q_{k+1}) - (\lambda_k)_\alpha\,\omega_i^\alpha(q_k),\\ \phi_d^\alpha(q_k,q_{k+1})=0.\end{array}$$

These equations are traditionally written in the following manner

$$\begin{array}{rcl}D_2\mathbb{L}_d(q_{k-1},q_k)+D_1\mathbb{L}_d(q_k,q_{k+1}) & = & (\lambda_k)_\alpha\,\omega^\alpha(q_k)\\ \phi_d^\alpha(q_k,q_{k+1}) & = & 0,\end{array}$$

which are the expression of the discrete nonholonomic equations (see [38] for more details).

4.3.2 Constrained discrete Lagrangian mechanics

A discrete constrained system [119] is determined by a pair (C_d, L_d) where C_d is a submanifold of $Q\times Q$, with inclusion $i_{C_d}: C_d \hookrightarrow Q\times Q$, and $L_d: C_d \to \mathbb{R}$ a discrete Lagrangian function.

Using again Theorem 4.1.1 we deduce that Σ_{L_d} is a Lagrangian submanifold of $(T^*(Q\times Q)\,,\,\Omega_{Q\times Q})$.

Using Υ we induce the Lagrangian submanifold $\Upsilon(\Sigma_{L_d})$ of the symplectic manifold $(T^*Q\times T^*Q, \tilde{\Omega}_Q)$ (see the diagram below).

$$\begin{tikzcd}
T^*(Q\times Q) \arrow[r, "\Upsilon"] & T^*Q\times T^*Q \\
\Sigma_{L_d} \arrow[u, hook] \arrow[r, "\Upsilon"] & \Upsilon(\Sigma_{L_d}) \arrow[u, hook]
\end{tikzcd}$$

The dynamics is determined by the sequences $\gamma_{q_0}, \gamma_{q_1}, \ldots, \gamma_{q_N}$ such that $(\gamma_{q_i}, \gamma_{q_{i+1}}) \in \Upsilon(\Sigma_{L_d})$, $0 \leq i \leq N-1$ (see [71, 119]). Observe that

$$\gamma_{q_k} \in T^*_{q_k}Q \cap pr_0(\Upsilon(\Sigma_{L_d})) \cap pr_1(\Upsilon(\Sigma_{L_d})), 1 \leq k \leq N-1. \tag{4.19}$$

After determining intrinsically the dynamics, as in the continuous case, we now consider local expressions. Take an arbitrary extension $\mathbb{L}_d : Q\times Q \to \mathbb{R}$ of $L_d : C_d \to \mathbb{R}$, that is, $\mathbb{L}_d \circ i_{C_d} = L_d$. Assume also that we have fixed local constraints such that determines the submanifold C_d. This definition is performed by the vanishing of the following set of independent constraints: $\phi^\alpha_d(q_0, q_1) = 0$, $1 \leq \alpha \leq n$ where $n = 2\dim\, Q - \dim\, C$.

Locally

$$\begin{aligned}
\Sigma_{L_d} \quad = \quad & \{(q_0^i, q_1^i, (\mu_0)_i, (\mu_1)_i) \in T^*(Q\times Q)\mid \\
& (\mu_0)_i = \frac{\partial \mathbb{L}_d}{\partial q_0^i} + (\lambda_1)_\alpha \frac{\partial \phi^\alpha_d}{\partial q_0^i}, \\
& (\mu_1)_i = \frac{\partial \mathbb{L}_d}{\partial q_1^i} + (\lambda_1)_\alpha \frac{\partial \phi^\alpha_d}{\partial q_1^i}, \\
& \phi^\alpha_d(q_0, q_1) = 0,\;\; 1 \leq \alpha \leq n\}\ ,
\end{aligned} \tag{4.20}$$

where $(\lambda_1)_\alpha$ are Lagrange multipliers to be determined.

Therefore,

$$\begin{aligned}
\Upsilon(\Sigma_{L_d}) \quad = \quad & \{(q_0^i, (p_0)_i, q_1^i, (p_1)_i) \in T^*Q\times T^*Q\mid \\
& (p_0)_i = -\frac{\partial \mathbb{L}_d}{\partial q_0^i} - (\lambda_1)_\alpha \frac{\partial \phi^\alpha_d}{\partial q_0^i}, \\
& (p_1)_i = \frac{\partial \mathbb{L}_d}{\partial q_1^i} + (\lambda_1)_\alpha \frac{\partial \phi^\alpha_d}{\partial q_1^i}, \\
& \phi^\alpha_d(q_0, q_1) = 0,\;\; 1 \leq \alpha \leq n\}\ .
\end{aligned} \tag{4.21}$$

The solutions of the dynamics come from (4.19) and are given by the equations

$$\frac{\partial \mathbb{L}_d}{\partial q_1^i}(q_{k-1},q_k) + (\lambda_k)_\alpha \frac{\partial \phi_d^\alpha}{\partial q_1^i}(q_{k-1},q_k) = -\frac{\partial \mathbb{L}_d}{\partial q_0^i}(q_k,q_{k+1}) - (\lambda_{k+1})_\alpha \frac{\partial \phi_d^\alpha}{\partial q_0^i}(q_k,q_{k+1}),$$
$$\phi_d^\alpha(q_{k-1},q_k) = 0,$$
$$\phi_d^\alpha(q_k,q_{k+1}) = 0.$$

These equations are traditionally written as

$$D_2\mathbb{L}_d(q_{k-1},q_k) + D_1\mathbb{L}_d(q_k,q_{k+1}) + (\lambda_k)_\alpha D_2\phi_d^\alpha(q_{k-1},q_{k+1}) + (\lambda_{k+1})_\alpha D_1\phi_d^\alpha(q_k,q_{k+1}) = 0,$$
$$\phi_d^\alpha(q_{k-1},q_k) = 0,$$
$$\phi_d^\alpha(q_k,q_{k+1}) = 0,$$

which are the expression of the discrete vakonomic equations (see [11] for more details).

Now, as a particular case, we assume that we can choose adapted coordinates (q_0^i, q_1^a), $1 \leq i \leq \dim Q$ and $1 \leq a \leq \dim Q - n$, on C_d in such a way the inclusion is written as

$$i_{C_d}(q_0^i, q_1^a) = (q_0^i, q_1^a, \Psi_d^\alpha(q_0^i, q_1^a)). \tag{4.22}$$

In other words, we can write the constraints as $\phi_d^\alpha(q_0,q_1) = q_1^\alpha - \Psi_d^\alpha(q_0^i, q_1^a) = 0$.

Thus, locally we have that

$$\begin{aligned} \Sigma_{L_d} \;\; = \;\; & \{(q_0^i, q_1^i, (\mu_0)_i, (\mu_1)_i) \in T^*(Q\times Q) | \\ & (\mu_0)_i = \frac{\partial L_d}{\partial q_0^i} - (\mu_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_0^i}, \\ & (\mu_1)_a = \frac{\partial L_d}{\partial q_1^a} - (\mu_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_1^a}, \\ & q_1^\alpha = \Psi_d^\alpha(q_0^i, q_1^a), \;\; 1 \leq \alpha \leq n\}\;. \end{aligned} \tag{4.23}$$

Observe that $(q_0^i, q_1^a, (\mu_1)_\alpha)$ gives a local coordinate system for Σ_{L_d}.

Therefore, we obtain the following expression of the Lagrangian submanifold $\Upsilon(\Sigma_{L_d})$:

$$\begin{aligned} \Upsilon(\Sigma_{L_d}) \;\; = \;\; & \{(q_0^i, (p_0)_i, q_1^i, (p_1)_i) \in T^*Q \times T^*Q \mid \\ & (p_0)_i = -\frac{\partial L_d}{\partial q_0^i} + (p_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_0^i}, \\ & (p_1)_a = \frac{\partial L_d}{\partial q_1^a} - (p_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_1^a}, \\ & q_1^\alpha = \Psi_d^\alpha(q_0^i, q_1^a), \;\; 1 \leq \alpha \leq n\}\;. \end{aligned} \tag{4.24}$$

Consequently, the solutions must verify the following system of difference equations:

$$(D_2)_a(L_d - (p_k)_\alpha \Psi_d^\alpha)(q_{k-1},q_k) + (D_1)_a(L_d - (p_{k+1})_\alpha \Psi_d^\alpha)(q_k,q_{k+1}) = 0,$$
$$(p_k)_\beta + (D_1)_\beta L_d(q_k,q_{k+1}) - (p_{k+1})_\alpha (D_1)_\beta \Psi_d^\alpha(q_k,q_{k+1}) = 0,$$
$$q_{k+1}^\alpha = \Psi_d^\alpha(q_k^i, q_{k+1}^a)\;,$$

where $(D_j)_l$ just means $\frac{\partial}{\partial q_j^l}$, being $j = \{1,2\}$ and $l = \{a, \alpha\}$.

The constrained discrete Legendre transformations

Definition 4.3.1. [119] *We define the* **constrained discrete Legendre transformations** $\mathbb{F}L_d^{\pm} : \Sigma_{L_d} \longrightarrow T^*Q$ *as the mappings*

$$
\begin{aligned}
\mathbb{F}L_d^- &= pr_0 \circ (\Upsilon)\big|_{\Sigma_{L_d}}, \\
\mathbb{F}L_d^+ &= pr_1 \circ (\Upsilon)\big|_{\Sigma_{L_d}}.
\end{aligned}
$$

We will say that the constrained system (L_d, C_d) is **regular** *if $\mathbb{F}L_d^-$ is a local diffeomorphism and* **hyperregular** *if $\mathbb{F}L_d^-$ is a diffeomorphism.*

Remark 4.3.2. *It is easy to prove that $\mathbb{F}L_d^-$ is a local diffeomorphism if and only if $\mathbb{F}L_d^+$ is a local diffeomorphism; therefore, it is possible to characterize the regularity of the constrained system using any of the two Legendre transformations (see [119]).*

Observe that if we consider the local constraints $q_1^\alpha = \Psi_d^\alpha(q_0^i, q_1^a)$ determining C_d, then

$$
\begin{aligned}
\mathbb{F}L_d^-(q_0^i, q_1^a, (\mu_1)_\alpha) &= (q_0^i, (p_0)_i = -\frac{\partial L_d}{\partial q_0^i} + (\mu_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_0^i}) \\
&= (q_0^i, -(D_1)_i(L_d - (\mu_1)_\alpha \Psi_d^\alpha)(q_0, q_1))\ . \\
\mathbb{F}L_d^+(q_0^i, q_1^a, (\mu_1)_\alpha) &= (q_1^a\ ,\ q_1^\alpha = \Psi_d^\alpha(q_0^i, q_1^a), \\
&\quad (p_1)_a = \frac{\partial L_d}{\partial q_1^a} - (\mu_1)_\alpha \frac{\partial \Psi_d^\alpha}{\partial q_1^a}\ ,\ (p_1)_\alpha = (\mu_1)_\alpha) \\
&= (q_1^a\ ,\ q_1^\alpha = \Psi_d^\alpha(q_0^i, q_1^a), \\
&\quad (p_1)_a = (D_2)_a(L_d - \mu_\alpha^1 \Psi_d^\alpha)(q_0, q_1)\ ,\ (p_1)_\alpha = (\mu_1)_\alpha)\ .
\end{aligned}
$$

So, the constrained system (L_d, C_d) is regular if and only if the matrix

$$
(A_{ia}, A_{i\alpha}) = \left(\frac{\partial^2 L_d}{\partial q_0^i \partial q_1^a} - \mu_\alpha^1 \frac{\partial^2 \Psi_d^\alpha}{\partial q_0^i \partial q_1^a}, \frac{\partial \Psi_d^\alpha}{\partial q_0^i} \right)
$$

is nondegenerate.

Remark 4.3.3. *The constrained Legendre transformations allow us to define a univocally presymplectic 2-form on Σ_{L_d}*

$$
\Omega_{L_d} = (\mathbb{F}L_d^-)^* \Omega_Q = (\mathbb{F}L_d^+)^* \Omega_Q
$$

which is symplectic if the system is regular.

Then, if the constrained system (L_d, C_d) is hyperregular, we can define the discrete dynamics determined by $\Upsilon^{-1}(\Sigma_{L_d})$ as the graph of the canonical transformation $(\mathbb{F}L_d^+) \circ (\mathbb{F}L_d^-)^{-1} : T^*Q \to T^*Q$, that is,

$$
\text{Graph } (\mathbb{F}L_d^+) \circ (\mathbb{F}L_d^-)^{-1} = \Upsilon(\Sigma_{L_d})\ .
$$

4.3.3 Comparison of nonholonomic and variational constrained equations. Discrete picture

Let consider the system defined by the discrete Lagrangian function $\mathbb{L}_d : Q \times Q \to \mathbb{R}$ and a set of constraints $\phi_d^\alpha(q_0, q_1) = 0$ determining the submanifold $C_d \subset Q \times Q$.

As shown in §4.3.1 and §4.3.2, the solutions of the discrete nonholonomic dynamics are geometrically described by the affine subbundle $\Upsilon(\Sigma_d^{noh}) \subset T^*Q \times T^*Q$, while the solutions of the discrete constrained variational dynamics are given by the Lagrangian submanifold $\Upsilon(\Sigma_{L_d}) \subset T^*Q \times T^*Q$, where $L_d = \mathbb{L}_d\big|_{C_d} : C_d \to \mathbb{R}$.

Given a solution of the discrete nonholonomic problem, we want to know when is also a solution of the associated discrete constrained variational problem. To capture the set of common solutions to both problems, we have developed the following geometric integrability algorithm similar to the continuous one.

As in the continuous case, take the Whitney sum $T^*Q \oplus T^*Q$, and the cartesian product

$$(T^*Q \oplus T^*Q) \times (T^*Q \oplus T^*Q) \equiv (T^*Q \times T^*Q) \oplus_{\pi_Q \times \pi_Q} (T^*Q \times T^*Q).$$

Construct the submanifold $\Sigma_d^{cons} \hookrightarrow (T^*Q \oplus T^*Q) \times (T^*Q \oplus T^*Q)$ as follows:

$$\begin{aligned} \Sigma_d^{cons} = \{ \quad & (\gamma_{q_0}, \gamma_{q_1}, \widetilde{\gamma}_{q_0}, \widetilde{\gamma}_{q_1}) \in (T^*Q \times T^*Q) \oplus_{\pi_Q \times \pi_Q} (T^*Q \times T^*Q) \, / \\ & (\gamma_{q_0}, \gamma_{q_1}) \in \Sigma_d^{noh}, (\widetilde{\gamma}_{q_0}, \widetilde{\gamma}_{q_1}) \in \Sigma_{L_d} \}. \end{aligned} \tag{4.25}$$

It is quite clear that the submanifold Σ_d^{cons} gathers together both nonholonomic and constrained variational dynamics and applying the discrete version of the integrability algorithm developed in [71] we will find the set where there are common solutions for both dynamics.

The following diagram shows the bundle relations:

$$\begin{tikzcd}
\Sigma_d^{cons} \arrow[r, hook] & (T^*Q \oplus T^*Q) \times (T^*Q \oplus T^*Q) \arrow[rr, "\widetilde{\pi_Q} \times \widetilde{\pi_Q}"'] \arrow[d, shift left] \arrow[d, shift right] & & Q \times Q \arrow[d, shift left] \arrow[d, shift right] \\
& T^*Q \oplus T^*Q \arrow[rr, "\widetilde{\pi_Q}"'] & & Q
\end{tikzcd}$$

where the vertical arrows represent the projections onto the first and second factor of the respective cartesian products.

4.3.4 Construction of a discrete constrained problem from a Lagrangian submanifold of $T^*Q \times T^*Q$

Given a Lagrangian submanifold Λ of $(T^*Q \times T^*Q, \tilde{\Omega}_Q)$ we will construct, under some regularity conditions, a constrained Lagrangian problem given by a submanifold $C_d \subset Q \times Q$ and a function $L_d : C_d \to \mathbb{R}$.

We first construct the Lagrangian submanifold $\Upsilon^{-1}(\Lambda)$ of $(T^*(Q \times Q), \Omega_{Q\times Q})$ and we assume that the restriction of $\Omega_{Q\times Q}$ to $\Upsilon^{-1}(\Lambda)$ is exact, that is, we have a generating function $S : \Upsilon^{-1}(\Lambda) \to \mathbb{R}$. In addition, we suppose that the image of $\Upsilon^{-1}(\Lambda)$ by $\pi_{Q\times Q}$ is a submanifold C_d of $Q\times Q$, and, also, that $(\pi_{Q\times Q})\big|_{\Upsilon^{-1}(\Lambda)}$ is a submersion with connected fibers (see the diagrams below).

$$\begin{tikzcd}
T^*(Q\times Q) \arrow[r, "\Upsilon"] & T^*Q\times T^*Q \\
\Upsilon^{-1}(\Lambda) \arrow[u, hook] & \Lambda \arrow[u, hook] \arrow[l, "\Upsilon^{-1}"]
\end{tikzcd}$$

$$\begin{tikzcd}
 & T^*(Q\times Q) \arrow[dl, "\pi_{Q\times Q}"'] & \\
Q\times Q\supset C_d \arrow[dr, "L_d"'] & & \Upsilon^{-1}(\Lambda) \arrow[ul, hook', "i_{\Upsilon^{-1}(\Lambda)}"'] \arrow[dl, "S"] \\
 & \mathbb{R} &
\end{tikzcd}$$

Theorem 4.3.4. *Under the previous conditions the function $S : \Upsilon^{-1}(\Lambda) \to \mathbb{R}$ is $(\pi_{Q\times Q})\big|_{\Upsilon^{-1}(\Lambda)}$-projectable onto a function $L_d : C_d \to \mathbb{R}$. Moreover, the following equation holds*

$$\Upsilon^{-1}(\Lambda) = \Sigma_{L_d}\ .$$

Proof. The submanifold $\Upsilon^{-1}(\Lambda)$ is defined as

$$\Upsilon^{-1}(\Lambda) = \{(\gamma_{q_0}, \gamma_{q_1}) \in T^*_{(q_0,q_1)}(Q\times Q) \mid i^*_{\Upsilon^{-1}(\Lambda)}\Theta_{Q\times Q}(\gamma_{q_0}, \gamma_{q_1}) = \mathrm{d}S(\gamma_{q_0}, \gamma_{q_1})\}. \tag{4.26}$$

By definition of the Liouville 1-form $\Theta_{Q\times Q}$ we have that

$$\langle \Theta_{Q\times Q}\,,\, \ker T\,\pi_{Q\times Q}\rangle = 0.$$

By applying the chain rule $T(\pi_{Q\times Q}\circ i_{\Upsilon^{-1}(\Lambda)}) = T\,\pi_{Q\times Q}\circ T\,i_{\Upsilon^{-1}(\Lambda)}$, is easy to check that $T\,i_{\Upsilon^{-1}(\Lambda)}\left(\ker T\,\pi_{Q\times Q}\big|_{\Upsilon^{-1}(\Lambda)}\right) \subset \ker T\,\pi_{Q\times Q}$. Thus, we finally deduce that S is projectable into $L_d : C_d \to \mathbb{R}$, i.e.

$$S = (\pi_{Q\times Q}\big|_{\Upsilon^{-1}(\Lambda)})^* L_d \tag{4.27}$$

since $\pi_{Q\times Q}(\Upsilon^{-1}(\Lambda)) = C_d$.

On the other hand, we can define a tangent vector $X = (X_{\gamma_{q_0}}, X_{\gamma_{q_1}}) \in T\,\Upsilon^{-1}(\Lambda) \subset TT^*(Q\times Q)$, such that $T_{(\gamma_{q_0},\gamma_{q_1})}\pi_{Q\times Q}(X) = (u_{q_0}, u_{q_1}) \in T_{(q_0,q_1)}C_d$, by

$$\langle(\gamma_{q_0}, \gamma_{q_1})\,,\,(u_{q_0}, u_{q_1})\rangle = \langle \mathrm{d}S(\gamma_{q_0}, \gamma_{q_1})\,,\, X\rangle,$$

where $(\gamma_{q_0}, \gamma_{q_1}) \in \Upsilon^{-1}(\Lambda)$. The equation above comes directly from the definitions of both the Liouville one-form and the Lagrangian submanifold $\Upsilon^{-1}(\Lambda)$ in (4.26). Regarding equation

(4.27) and taking into account that the pullback and the exterior derivative commute, we arrive to

$$\begin{aligned}\langle \mathrm{d}S(\gamma_{q_0}, \gamma_{q_1}) \,,\, X\rangle &= \langle \mathrm{d}L_d(q_0, q_1) \,,\, T\pi_{Q\times Q}\big|_{\Upsilon^{-1}(\Lambda)}(X)\rangle \\ &= \langle \mathrm{d}L_d(q_0, q_1) \,,\, (u_{q_0}, u_{q_1})\rangle.\end{aligned}$$

In the last line of the expression just above, we recognize the definition given in Theorem 4.1.1 of Σ_{L_d}, that is

$$\Sigma_{L_d} = \big\{(\gamma_{q_0}, \gamma_{q_1}) \in T^*_{(q_0,q_1)}(Q\times Q) \,|\, \langle(\gamma_{q_0}, \gamma_{q_1}) \,,\, (u_{q_0}, u_{q_1})\rangle = \\ = \langle \mathrm{d}L_d(q_0, q_1) \,,\, (u_{q_0}, u_{q_1})\rangle \;\text{ for all }\; (u_{q_0}, u_{q_1}) \in T_{(q_0,q_1)}C_d\big\}.$$

In consequence, we deduce that $\Upsilon^{-1}(\Lambda) = \Sigma_{L_d}$. □

Remark 4.3.5. *We would like to point out that the proof of Theorem 4.3.4 can be easily extended to the continuous case. Namely, let consider $\tilde\Lambda \subset TT^*Q$ a Lagrangian submanifold. Under some regularity conditions, it is possible to construct a constrained Lagrangian problem given by a submanifold $C \subset TQ$ and a Lagrangian function $L : C \to \mathbb{R}$. If we consider the Lagrangian submanifold $\alpha_Q(\tilde\Lambda) \subset (T^*TQ \,,\, \Omega_{TQ})$, where α_Q is the Tulczyjew's isomorphism, we can build an analogy with the discrete case by assuming that the restriction of Ω_{TQ} to $\alpha_Q(\tilde\Lambda)$ is exact, that is, we have a generating function $\tilde S : \alpha_Q(\tilde\Lambda) \to \mathbb{R}$;*

$$\alpha_Q(\tilde\Lambda) = \Big\{\gamma \in T^*_{v_q}TQ \,|\, i^*_{\alpha_Q(\tilde\Lambda)}\Theta_{TQ}(\gamma) = \mathrm{d}\tilde S(\gamma)\Big\}\,,$$

where $v_q \in TQ$ such that $\tau_Q(v_q) = q \in Q$. In addition, we suppose that $\pi_{TQ}\Big(\alpha_Q(\tilde\Lambda)\Big)$ is a submanifold $C \subset TQ$, and that $\pi_{TQ}\big|_{\alpha_Q(\tilde\Lambda)}$ is a submersion with connected fibers.

Again, by the definition of the Liouville one-form Θ_{TQ} we have that $\langle\Theta_{TQ} \,,\, kerT\pi_{TQ}\rangle = 0$, and consequently that $\tilde S$ is projectable into L, that is $\tilde S = (\pi_{TQ}\big|_{\alpha_Q(\tilde\Lambda)})^*L$.

Following similar arguments that in the proof of Theorem 4.3.4 we deduce that $\alpha_Q(\tilde\Lambda) = \Sigma_L$.

Since we are not fixing the Lagrangian submanifold $\tilde\Lambda$, this is a more general result than that one provided by Theorem 4.2.4, i.e. $\alpha_Q(X_H(T^*Q)) = \Sigma_L$. Nevertheless, among all the Lagrangian submanifolds $\tilde\Lambda$ of T^*TQ, we choose $X_H(T^*Q)$, that is the image of the cotangent bundle T^*Q by the Hamiltonian vector field provided by the equations $i_{X_H}\Omega_Q = \mathrm{d}H$, in order to stress the relationship between Hamiltonian and constrained Lagrangian systems.

4.4 Examples

4.4.1 Linear Constraints

Let consider a dynamical system denoted by the Lagrangian

$$\mathbb{L}(v_q) = \frac{1}{2}\,\mathcal{G}(v_q, v_q) - V(q) = \frac{1}{2}\mathcal{G}_{ij}\dot q^i\dot q^j - V(q) \tag{4.28}$$

where $v_q \in T_qQ$, with local coordinates $v_q = (q^i, \dot{q}^i)$, $\mathcal{G}$ is a Riemannian metric with components $(\mathcal{G}_{ij})$ (see §1.2). Moreover, $V : Q \to \mathbb{R}$ is a potential function. Additionally, the system is subject to the linear constraints

$$\phi^\alpha(v_q) = \dot{q}^\alpha - \Gamma^\alpha_a(q)\,\dot{q}^a, \tag{4.29}$$

where $\dot{q}^i = \{\dot{q}^a, \dot{q}^\alpha\}$. Locally, the constraints define a submanifold $C \subset TQ$. Moreover, we have the restriction of $\mathbb{L}$ to C, $L : C \to \mathbb{R}$. In local coordinates,

$$L(q^i, \dot{q}^a) = \frac{1}{2}\gamma_{ab}\dot{q}^a\dot{q}^b - V(q)\ ,$$

where

$$\gamma_{ab}\,(q) = \mathcal{G}_{ab} + \mathcal{G}_{a\alpha}\Gamma^\alpha_b(q) + \mathcal{G}_{b\alpha}\Gamma^\alpha_a(q) + \mathcal{G}_{\alpha\beta}\Gamma^\alpha_a(q)\Gamma^\beta_b(q).$$

Observe that (γ_{ab}) is invertible since $\mathcal{G}$ is a Riemannian metric.

Using expression (4.8), we can find local coordinates for $\alpha_Q^{-1}(\Sigma_L)$:

$$\begin{aligned}
p_a &= \gamma_{ab}(q)\,\dot{q}^b - p_\alpha\,\Gamma^\alpha_a(q)\ ,\\
p_i &= \frac{1}{2}\frac{\partial\gamma_{ab}}{\partial q^i}\dot{q}^a\dot{q}^b - \frac{\partial V}{\partial q^i} - p_\alpha\frac{\partial\Gamma^\alpha_a}{\partial q^i}(q)\,\dot{q}^a\ ,\\
\dot{q}^\alpha &= \Gamma^\alpha_a(q)\,\dot{q}^a\ .
\end{aligned}$$

The Legendre transformation is defined by $\mathbb{F}L = \tau_{T^*Q}\circ(\alpha_Q^{-1})\big|_{\Sigma_L}$, or locally by:

$$\mathbb{F}L(q^i, \dot{q}^a, \tilde{\mu}_\alpha) = (q^i, \gamma_{ab}\dot{q}^a - \tilde{\mu}_\alpha\,\Gamma^\alpha_a(q), \tilde{\mu}_\alpha)\ .$$

Since $\left(\frac{\partial^2 L}{\partial\dot{q}^a\partial\dot{q}^b} - \tilde{\mu}_\alpha\frac{\partial^2\Psi^\alpha}{\partial\dot{q}^a\partial\dot{q}^b}\right) = (\gamma_{ab})$ the constrained system (L, C) is regular.

Moreover, the energy function $E_L : \Sigma_L \to \mathbb{R}$ is precisely

$$E_L(q^i, \dot{q}^a, \tilde{\mu}_\alpha) = \frac{1}{2}\gamma_{ab}\dot{q}^a\dot{q}^b - V(q)\ .$$

Therefore, the Hamiltonian function can be expressed by $H = E_L\circ(\mathbb{F}L)^{-1} : T^*Q \to \mathbb{R}$

$$H(q\,,\,p) = \frac{1}{2}\,\gamma^{ab}(q)\,P_a\,P_b + V(q), \tag{4.30}$$

where $\gamma_{a\,b}\,\gamma^{b\,c} = \delta^c_a$ and $P_a = p_a + p_\alpha\,\Gamma^\alpha_a(q)$.

Discretization: symplectic Euler method

Taking into account equations (4.28) and (4.29), we define the discrete Lagrangian $\mathbb{L}_d : Q\times Q \to \mathbb{R}$ and the set of independent constraints in the following way (see [124] for more details):

$$\begin{aligned}
&\mathbb{L}_d(q_0, q_1) = h\mathbb{L}(q_0, \frac{q_1 - q_0}{h}) = \frac{1}{2h}\mathcal{G}_{i\,j}(q_0)(q_1^i - q_0^i)(q_1^j - q_0^j) - hV\,(q_0)\,,\\
&\left(\frac{q_1^\alpha - q_0^\alpha}{h}\right) = \Gamma^\alpha_a(q_0)\left(\frac{q_1^a - q_0^a}{h}\right).
\end{aligned} \tag{4.31}$$

Next, we can explicitly obtain the coordinates for the submanifold $\Upsilon(\Sigma_{L_d})$ given in equations (4.21), namely

$$
\begin{aligned}
(p_0)_a &= \frac{1}{h}\mathcal{G}_{a\,j}(q_0)(q_1^j - q_0^j) + h\partial_a V(q_0) - \frac{1}{2h}\partial_a \mathcal{G}_{i\,j}(q_0)(q_1^i - q_0^i)(q_1^j - q_0^j) \\
&\quad + (\lambda_1)_\beta\, \partial_a \Gamma_b^\beta(q_0)\left(q_1^b - q_0^b\right) - (\lambda_1)_\alpha \Gamma_a^\alpha(q_0), \qquad (4.32)
\end{aligned}
$$

$$
\begin{aligned}
(p_0)_\alpha &= \frac{1}{h}\mathcal{G}_{\alpha\,j}(q_0)(q_1^j - q_0^j) + h\partial_\alpha V(q_0) - \frac{1}{2h}\partial_\alpha \mathcal{G}_{i\,j}(q_0)(q_1^i - q_0^i)(q_1^j - q_0^j) \\
&\quad + (\lambda_1)_\alpha + (\lambda_1)_\beta \partial_\alpha \Gamma_b^\beta(q_0)(q_1^b - q_0^b), \qquad (4.33)
\end{aligned}
$$

$$
(p_1)_a = \frac{1}{h}\mathcal{G}_{a\,j}(q_0)(q_1^j - q_0^j) - (\lambda_1)_\alpha \Gamma_a^\alpha(q_0), \qquad (4.34)
$$

$$
\begin{aligned}
(p_1)_\alpha &= \frac{1}{h}\mathcal{G}_{\alpha\,j}(q_0)(q_1^j - q_0^j) + (\lambda_1)_\alpha, \\
(q_1^\alpha - q_0^\alpha) &= \Gamma_a^\alpha(q_0)(q_1^a - q_0^a),
\end{aligned}
$$

where ∂_a, ∂_α mean $\frac{\partial}{\partial q^a}$ and $\frac{\partial}{\partial q^\alpha}$, respectively. It is important to note that (4.32) is a set of $2m+n$ equations with $2m+n$ unknowns, which are $(q_1)^a, (q_1)^\alpha, (p_1)_a$, $(p_1)_\alpha$ and $(\lambda_1)_\alpha$.

Alternatively, we can apply the so-called Euler symplectic method (see [61])

$$
p_1 = p_0 - h\frac{\partial H}{\partial q}(q_0, p_1), \quad q_1 = q_0 + h\frac{\partial H}{\partial p}(q_0, p_1) \qquad (4.35)
$$

to the Hamiltonian function $H : T^*Q \to \mathbb{R}$ defined in (4.30). We deduce the following set of equations:

$$
\begin{aligned}
(q_1)^a &= (q_0)^a + h\gamma^{ab}(P_1)_b, \\
(q_1)^\alpha &= (q_0)^\alpha + h\Gamma_a^\alpha \gamma^{ab}(P_1)_b, \\
(p_1)_a &= (p_0)_a - h\partial_a V - h\left\{\frac{1}{2}\left(\partial_a \gamma^{bc}\right)(P_1)_b (P_1)_c + \gamma^{bc}(P_1)_b \partial_a (P_1)_c\right\}, \\
(p_1)_\alpha &= (p_0)_\alpha - h\partial_\alpha V - h\left\{\frac{1}{2}\left(\partial_\alpha \gamma^{ab}\right)(P_1)_a (P_1)_b + \gamma^{ab}(P_1)_a \partial_\alpha (P_1)_b\right\},
\end{aligned} \qquad (4.36)
$$

where $(P_1)_a = (p_1)_a + (p_1)_\alpha \Gamma_a^\alpha$, and V, γ^{ab}, Γ_a^α, $(P_1)_a$ are evaluated at q_0. Regarding equations (4.32), is easy to express λ_1 in terms of p and q. Since

$$
0 = \partial_a \delta_d^c = \partial_a\left(\gamma^{cb}\gamma_{bd}\right) = \left(\partial_a \gamma^{cb}\right)\gamma_{bd} + \gamma^{cb}(\partial_a \gamma_{bd}), \qquad (4.37)
$$

and $\left(\partial_a \gamma^{cb}\right)\gamma_{bd} = -\gamma^{cb}(\partial_a \gamma_{bd})$, is easy to check, after a straightforward calculation, that equations (4.32) reduce to (4.36).

Discretization: Midpoint rule

Define the discrete Lagrangian and the discrete constraints using the **midpoint rule** (see [124] for more details), that is:

$$
\begin{aligned}
&\mathbb{L}_d(q_0,q_1) = h\mathbb{L}(\frac{q_1+q_0}{2}, \frac{q_1-q_0}{h}) = \\
&\frac{1}{2h}\mathcal{G}_{ij}(\frac{q_1+q_0}{2})(q_1-q_0)^i(q_1-q_0)^j - hV(\frac{q_1+q_0}{2}), \\
&\left(\frac{q_1-q_0}{h}\right)^{\alpha} = \Gamma^{\alpha}_{a}(\frac{q_1+q_0}{2})\left(\frac{q_1-q_0}{h}\right)^{a}.
\end{aligned}
\tag{4.38}
$$

Now, we can explicitly obtain the corresponding equations derived from the submanifold $\Upsilon(\Sigma_{L_d})$. A straightforward computation shows that these equations are equivalent to the corresponding ones derived from the so-called midpoint rule

$$
q_1 = q_0 + h\frac{\partial H}{\partial p}\left(\frac{q_1+q_0}{2}, \frac{p_1+p_0}{2}\right), \quad p_1 = p_0 - h\frac{\partial H}{\partial q}\left(\frac{q_1+q_0}{2}, \frac{p_1+p_0}{2}\right), \tag{4.39}
$$

which is a symplectic method of order 2.

$$
\begin{tikzcd}
\Sigma_L \arrow[r, dashed, "\varphi_d"] \arrow[d, "\phi_L"'] & \Sigma_{L_d} \arrow[rr, "\mathbb{F}L_d^-"] \arrow[d, "\phi_{L_d}"'] & & T^*Q \arrow[d, shift right, "\phi_{\mathbb{F}L_d^-}"'] \arrow[d, shift left, "(\phi_H)_d"] & T^*Q \arrow[l, dashed, "\hat{\varphi}_d"'] \arrow[d, "\phi_H"] \\
\Sigma_L \arrow[r, dashed, "\varphi_d"'] & \Sigma_{L_d} \arrow[rr, "\mathbb{F}L_d^-"] & & T^*Q & T^*Q \arrow[l, dashed, "\hat{\varphi}_d"]
\end{tikzcd}
\tag{4.40}
$$

$$
\begin{tikzcd}
\Sigma_L \arrow[r, dashed, "\varphi_d"] \arrow[d, "\phi_L"'] & \Sigma_{L_d} \arrow[rr, "\mathbb{F}L_d^+"] \arrow[d, "\phi_{L_d}"'] & & T^*Q \arrow[d, shift right, "\phi_{\mathbb{F}L_d^+}"'] \arrow[d, shift left, "(\phi_H)_d"] & T^*Q \arrow[l, dashed, "\hat{\varphi}_d"'] \arrow[d, "\phi_H"] \\
\Sigma_L \arrow[r, dashed, "\varphi_d"'] & \Sigma_{L_d} \arrow[rr, "\mathbb{F}L_d^+"] & & T^*Q & T^*Q \arrow[l, dashed, "\hat{\varphi}_d"]
\end{tikzcd}
\tag{4.41}
$$

The results of the previous examples can be summarized in the diagrams (4.40) and (4.41), which are explained in the following lines.

- **From right to left**: ϕ_H is the Hamiltonian flow derived from (4.11) applied to a Hamiltonian function H, $\hat{\varphi}_d$ is the discretization of that flow (concretely the symplectic Euler method (4.35) and the midpoint rule (4.39)), $(\phi_H)_d$ is the discrete flow in T^*Q provided by $\hat{\varphi}_d$.
- **From left to right**: ϕ_L is the constrained Lagrangian flow resulting from equations (4.4) applied to the continuous Lagrangian, φ_d is the discretization applied to that

Lagrangian ((4.31) and (4.38)), ϕ_{L_d} is the discrete flow within Σ_{L_d} (4.23) due to φ_d, $\phi_{\mathbb{F}L_d^{\pm}}$ is the discrete flow in $T^*Q \times T^*Q$ (4.24).

In order to be more explicit, $\hat{\varphi}_d : T^*Q \to T^*Q$ represents the discrete flow generated by applying a symplectic method to the Hamiltonian equations (e.g. (4.35) and (4.39)). On the other hand, φ_d represents the discretization mapping, that is $\varphi_d : TQ \to Q \times Q$, e.g. $\varphi_d(q, \dot{q}) = (q_k, \frac{q_{k+1}-q_k}{h})$ or $\varphi_d(q, \dot{q}) = (\frac{q_{k+1}+q_k}{2}, \frac{q_{k+1}-q_k}{h})$.

As expected from Theorems 4.2.4 and 4.3.4, what we explicitly show is that $\phi_{\mathbb{F}L_d^{\pm}} = (\phi_H)_d$ using some discretizations (we have depicted the particular cases of the symplectic Euler methods and the midpoint rule). In other words, the diagrams (4.40) and (4.41) are commutative in those particular cases, as also is when the discretization $\hat{\varphi}_d$ corresponds to the exact discrete Lagrangian's (see [124]). Therefore, using a discrete variational integrator for the constrained continuous Lagrangian system or applying a symplectic integrator to the associated continuous Hamiltonian problem are equivalent approaches.

4.4.2 The Martinet case: symplectic integrators for sub-Riemannian geometry

Let us consider the Hamiltonian function $H : T^*\mathbb{R}^3 \to \mathbb{R}$

$$H(q,p) = \frac{1}{2}\left(\left(p_x + p_z\frac{y^2}{2}\right)^2 + \frac{p_y^2}{(1+\beta\, x)^2}\right), \tag{4.42}$$

where $q = (x, y, z)^T \in \mathbb{R}^3$ and $p = (p_x, p_y, p_z) \in (\mathbb{R}^3)^* \simeq \mathbb{R}^3$. From (4.42) we can locally define $X_H(T^*M)$, particulary $X_H(T^*\mathbb{R}^3)$, through the Hamiltonian equations, i.e:

$$\begin{array}{rclrcl} \dot{x} & = & p_x + p_z\frac{y^2}{2}, & \dot{p}_x & = & \frac{\beta\, p_y^2}{(1+\beta\, x)^3}, \\ \dot{y} & = & \frac{p_y}{(1+\beta\, x)^2}, & \dot{p}_y & = & -\left(p_x + p_z\frac{y^2}{2}\right)\, p_z\, y, \\ \dot{z} & = & \left(p_x + p_z\frac{y^2}{2}\right)\, \frac{y^2}{2}, & \dot{p}_z & = & 0. \end{array} \tag{4.43}$$

The associated Legendre transform $\mathbb{F}H$ is in this particular case written as

$$\mathbb{F}H(x, y, z; p_x, p_y, p_z) = (x, y, z\,;\, (p_x + p_z\frac{y^2}{2}), \frac{p_y}{(1+\beta\, x)^2}, \left(p_x + p_z\frac{y^2}{2}\right)\, \frac{y^2}{2}). \tag{4.44}$$

Looking at (4.43) and (4.44) is easy to realize that

$$C \subset T\mathbb{R}^3 = \left\{(x, y, z\,;\, \dot{x}, \dot{y}, \dot{z}) \;\text{ s.t. }\; \dot{z} = \frac{y^2}{2}\,\dot{x}\right\}.$$

Next, we will obtain the Lagrangian function $L : C \to \mathbb{R}$ using the implicit equation given in (4.13):

$$L \circ \mathbb{F}H = \frac{1}{2}\left(\left(p_x + p_z\frac{y^2}{2}\right)^2 + \frac{p_y^2}{(1+\beta\, x)^2}\right).$$

Finally, using $\mathbb{F}H$ we arrive to

$$L(x,y,z,\dot{x},\dot{y}) = \frac{1}{2}\left(\dot{x}^2 + (1+\beta\, x)^2\, \dot{y}^2\right). \tag{4.45}$$

Consequently, our approach allows us to conclude that the Hamiltonian system (4.42) is equivalent to the Lagrangian one (4.45) subject to the constraints $\dot{z} = \frac{y^2}{2}\,\dot{x}$. We clearly recognize in (4.45) a Martinet sub-Riemannian structure ([4, 19, 32]), which is described by the triple $(U\,,\,\Delta\,,\,\mathcal{G})$. In this triple, U is an open neighborhood of the origin in $\mathbb{R}^3$, Δ is a distribution corresponding to $\Delta = \ker\alpha$ for $\alpha = \mathrm{d}z - \frac{y^2}{2}\mathrm{d}x$ and $\mathcal{G}$ is a Riemannian metric. In the particular case $\mathcal{G} = \mathrm{d}x^2 + (1+\beta\, x)^2\mathrm{d}y^2$, we deduce that (4.45) corresponds to $L = \mathbb{L}\big|_C$ where $\mathbb{L}(q,\dot{q}) = \frac{1}{2}\,\mathcal{G}\,(\frac{\partial}{\partial q}\,,\,\frac{\partial}{\partial q})$, where $\frac{\partial}{\partial q} = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{z}\frac{\partial}{\partial z}$. In addition, the constraints are given by $\alpha\,(\frac{\partial}{\partial q}) = 0$.

Discrete Case

Let us consider the symplectic Euler method (4.35). Is easy to see that a type-2 generating function of the approximated Hamiltonian flow is

$$H^+(q_0,p_1) = q_0\, p_1 + h\, H(q_0,p_1), \tag{4.46}$$

where H is the Hamiltonian function $H : T^*\mathbb{R}^3 \to \mathbb{R}$. In other words:

$$p_0 = \frac{\partial\, H^+(q_0,p_1)}{\partial q_0}, \qquad q_1 = \frac{\partial\, H^+(q_0,p_1)}{\partial p_1}.$$

Under these considerations, we can define the local coordinates for $\Upsilon^{-1}(\Lambda)$:

$$\Upsilon^{-1}(\Lambda) = \left\{q_0\;,\;\frac{\partial\, H^+(q_0,p_1)}{\partial q_0}\;,\;\frac{\partial\, H^+(q_0,p_1)}{\partial p_1}\;,\;p_1\right\}.$$

Now, projecting $\Upsilon^{-1}(\Lambda)$ onto $\mathbb{R}^3\times\mathbb{R}^3$, we obtain that

$$\begin{aligned} x_1 &= x_0 + h\left((p_1)_x + (p_1)_z\frac{y_0^2}{2}\right), \\ y_1 &= y_0 + h\frac{(p_1)_y}{(1+\beta\, x_0)^2}, \\ z_1 &= z_0 + h\left((p_1)_x + (p_1)_z\frac{y_0^2}{2}\right)\frac{y_0^2}{2}. \end{aligned} \tag{4.47}$$

From the last equations we obtain the constraint $(z_1 - z_0) = \frac{y_0^2}{2}(x_1 - x_0)$, which defines the submanifold

$$C_d = \{(x_0,y_0,z_0;x_1,y_1,z_1) \in \mathbb{R}^3\times\mathbb{R}^3 \mid (z_1 - z_0) = \frac{y_0^2}{2}(x_1 - x_0)\}.$$

The next step to completely determine the discrete constrained Lagrangian system is to obtain L_d. In that sense, we take the usual generating function S mentioned in Theorem

4.3.4. Transforming S into a **type-2 generating function** ([106]) we arrive to the implicitly defined expression: $S(q_0, p_1) = p_1\, q_1 - H^+(q_0, p_1)$, which, taking into account (4.46) leads to

$$S(q_0, p_1) = h\left(p_1 \frac{\partial H(q_0, p_1)}{\partial p_1} - H(q_0, p_1)\right). \tag{4.48}$$

We have shown in Theorem 4.3.4 that S is $\pi_{Q\times Q}\big|_{\Upsilon^{-1}(\Lambda)}$-projectable onto $L_d : C_d \to \mathbb{R}$. Thus, from (4.48) and according to equations (4.47) we finally arrive to

$$L_d(q_0, q_1) = \frac{1}{2h}\left((x_1 - x_0)^2 + \frac{y_0^2}{2}(y_1 - y_0)^2\right), \tag{4.49}$$

for

$$C_d \subset Q \times Q = \left\{(q_0\,,\, q_1) \mid (z_1 - z_0) = \frac{y_0^2}{2}(x_1 - x_0)\right\}. \tag{4.50}$$

This is the expected result as is easily seen taking into account the continuous Lagrangian (4.45) and the submanifold $C \subset TQ$ defined by the continuous constraint $\dot z = \frac{y^2}{2}\,\dot x$. If we define both the discrete Lagrangian and the discrete submanifold C_d in the usual symplectic Euler discretization ([124]), i.e. $L_d(q_0, q_1) = h\, L(q_0, \frac{q_1 - q_0}{h})$, we easily obtain (4.49) and (4.50).

4.4.3 Symplectic Störmer-Verlet method

Due to its well-behaved features, namely reversibility, symplecticity, volume preservation and conservation of first integrals, the Störmer-Verlet method is one of the most important examples in geometric numerical integration (see [61]). For a Hamiltonian system determined by $H : T^*Q \to \mathbb{R}$, the Störmer-Verlet method reads

$$\begin{aligned} p_{k+1/2} &= p_k - \frac{h}{2} H_q(q_k, p_{k+1/2}),\\ q_{k+1} &= q_k + \frac{h}{2}\left(H_p(q_k, p_{k+1/2}) + H_p(q_{k+1}, p_{k+1/2})\right),\\ p_{k+1} &= p_{k+1/2} - \frac{h}{2} H_q(q_{k+1}, p_{k+1/2}), \end{aligned} \tag{4.51}$$

where $q_k \in \mathbb{R}^n$, $p_k \in (\mathbb{R}^n)^*$ and H_q, H_p are the derivatives of the Hamiltonian function respect q and p, respectively. As we did for the momenta, we can fix an intermediate configuration point $q_{k+1/2} = q_k + \frac{h}{2} H_p(q_k, p_{k+1/2})$ and consider (4.51) as a two step integrator:

$$\begin{aligned} q_{k+1/2} &= q_k + \frac{h}{2} H_p(q_k, p_{k+1/2}),\\ p_{k+1/2} &= p_k - \frac{h}{2} H_q(q_k, p_{k+1/2}), \end{aligned} \tag{4.52}$$

$$\begin{aligned} q_{k+1} &= q_{k+1/2} + \frac{h}{2} H_p(q_{k+1}, p_{k+1/2}),\\ p_{k+1} &= p_{k+1/2} - \frac{h}{2} H_q(q_{k+1}, p_{k+1/2}). \end{aligned} \tag{4.53}$$

Equations (4.52) and (4.53) show the well-known fact that the Störmer-Verlet method is the composition of two different symplectic Euler schemes. In addition, it is clear that they are respectively generated by the 2- and 3-type generating functions

$$\begin{aligned} H^{+}(q_k, p_{k+1/2}) &= p_{k+1/2}\, q_k + \frac{h}{2} H(q_k, p_{k+1/2}), \\ H^{-}(q_{k+1}, p_{k+1/2}) &= p_{k+1/2}\, q_{k+1} - \frac{h}{2} H(q_{k+1}, p_{k+1/2}), \end{aligned}$$

which, taking into account that $p_1\, \mathrm{d}q_1 - p_0\, \mathrm{d}q_0 = \mathrm{d}S(q_0, q_1)$, lead to

$$S^{+}(q_k, p_{k+1/2}) = p_{k+1/2}\, q_{k+1/2} - p_{k+1/2}\, q_k - \frac{h}{2} H(q_k, p_{k+1/2}), \tag{4.54}$$

$$S^{-}(p_{k+1/2}, q_{k+1}) = -p_{k+1/2}\, q_{k+1/2} + p_{k+1/2}\, q_{k+1} - \frac{h}{2} H(q_{k+1}, p_{k+1/2}). \tag{4.55}$$

Now, as shown in [106], we can construct a 1-type generating function $S(q_k, q_{k+1})$ by

$$\begin{aligned} S(q_k, q_{k+1}) &= S^{+}(q_k, p_{k+1/2}) + S^{-}(p_{k+1/2}, q_{k+1}) \\ &= p_{k+1/2}\, q_{k+1} - p_{k+1/2}\, q_k - \frac{h}{2}\left(H(q_k, p_{k+1/2}) + H(q_{k+1}, p_{k+1/2})\right), \end{aligned} \tag{4.56}$$

and a extremal condition in the intermediate variable $p_{k+1/2}$. That is,

$$\begin{aligned} \mathrm{d}S &= \frac{\partial S}{\partial q_k}\mathrm{d}q_k + \frac{\partial S}{\partial q_{k+1}}\mathrm{d}q_{k+1} \\ &= \left\{ \frac{\partial S^{+}}{\partial q_k} + \left(\frac{\partial S^{+}}{\partial p_{k+1/2}} + \frac{\partial S^{-}}{\partial p_{k+1/2}} \right) \frac{\partial p_{k+1/2}}{\partial q_k} \right\} \mathrm{d}q_k \\ &+ \left\{ \frac{\partial S^{-}}{\partial q_{k+1}} + \left(\frac{\partial S^{+}}{\partial p_{k+1/2}} + \frac{\partial S^{-}}{\partial p_{k+1/2}} \right) \frac{\partial p_{k+1/2}}{\partial q_{k+1}} \right\} \mathrm{d}q_{k+1}, \end{aligned}$$

which leads to

$$\frac{\partial S^{+}}{\partial p_{k+1/2}} + \frac{\partial S^{-}}{\partial p_{k+1/2}} = 0, \tag{4.57}$$

(put in another way, $S(q_k, q_{k+1})$ is not a function of $p_{k+1/2}$ and consequently its partial derivative with respect to this variable should vanish). In other words, we obtain the first equation in (4.51) by $-p_k = \frac{\partial S(q_k, q_{k+1})}{\partial q_k}$, the third one by $p_{k+1} = \frac{\partial S(q_k, q_{k+1})}{\partial q_{k+1}}$, and the second one by $\frac{\partial S(q_k, q_{k+1})}{\partial p_{k+1/2}} = 0$.

As shown in Theorem 4.3.4, $S(q_k, q_{k+1})$ is projectable onto L_d, while condition (4.57) provides $C_d \subset Q \times Q$.

Regular systems

Consider the usual mechanical Hamiltonian

$$H(q, p) = \frac{1}{2}\, p\, M^{-1}\, p^{T} + V(q),$$

where M is a symmetric regular $n \times n$ matrix. From the second equation in (4.51) is easy to check that $p^T_{k+1/2} = M(\frac{q_{k+1}-q_k}{h})$. Hence, projecting (4.56) onto $Q \times Q$ we arrive to

$$L_d(q_k, q_{k+1}) = \frac{1}{2h}(\frac{q_{k+1}-q_k}{h})^T M \, (\frac{q_{k+1}-q_k}{h}) - \frac{h}{2}\,(V(q_k) + V(q_{k+1}))\,.$$

In the expression just above, we clearly recognize the discretization

$$L_d(q_k, q_{k+1}) = \frac{h}{2} L(q_k, \frac{q_{k+1}-q_k}{h}) + \frac{h}{2} L(q_{k+1}, \frac{q_{k+1}-q_k}{h})$$

for the usual mechanical Lagrangian $L(q, \dot q) = \frac{1}{2}\,\dot q^T \, M \, \dot q - V(q)$.

Martinet structure

Consider again the sub-Riemannian Martinet structure in (4.42). Recall that $q = (x, y, z)^T \in \mathbb{R}^3$ and $p = (p_x, p_y, p_z) \in \mathbb{R}^3$. From (4.57) we obtain

$$\begin{aligned}
x_{k+1} &= x_k + \frac{h}{2}\left\{\left((p_{k+1/2})_x + (p_{k+1/2})_z \frac{y_k^2}{2}\right) + \left((p_{k+1/2})_x + (p_{k+1/2})_z \frac{y_{k+1}^2}{2}\right)\right\},\\
y_{k+1} &= y_k + \frac{h}{2}\left\{\frac{(p_{k+1/2})_y}{(1+\beta\, x_k)^2} + \frac{(p_{k+1/2})_y}{(1+\beta\, x_{k+1})^2}\right\},\\
z_{k+1} &= z_k + \frac{h}{2}\left\{\left((p_{k+1/2})_x + (p_{k+1/2})_z \frac{y_k^2}{2}\right)\frac{y_k^2}{2} + \left((p_{k+1/2})_x + (p_{k+1/2})_z \frac{y_{k+1}^2}{2}\right)\frac{y_{k+1}^2}{2}\right\}.
\end{aligned}$$

As above, we consider a **two steps** (of $h/2$ size) interpretation of this sub-Riemannian system through equations (4.52) and (4.53) in the following manner (we denote $p_{k+1/2}$ by p for sake of simplicity):

$(q_k, p_k) \to (q_{k+1/2}, p_{k+1/2})$	$(q_{k+1/2}, p_{k+1/2}) \to (q_{k+1}, p_{k+1})$
$x_{k+1/2} = x_k + \frac{h}{2}\left(p_x + p_z \frac{y_k^2}{2}\right)$	$x_{k+1} = x_{k+1/2} + \frac{h}{2}\left(p_x + p_z \frac{y_{k+1}^2}{2}\right)$
$y_{k+1/2} = y_k + h\left\{\frac{p_y}{(1+\beta\, x_k)^2}\right\}$	$y_{k+1} = y_{k+1/2} + h\left\{\frac{p_y}{(1+\beta\, x_{k+1})^2}\right\}$
$z_{k+1/2} = z_k + \frac{h}{2}\left(p_x + p_z \frac{y_k^2}{2}\right)\frac{y_k^2}{2}$	$z_{k+1} = z_{k+1/2} + \frac{h}{2}\left(p_x + p_z \frac{y_{k+1}^2}{2}\right)\frac{y_{k+1}^2}{2}$
$\Downarrow$	$\Downarrow$
$C_d : \left(\frac{z_{k+1/2}-z_k}{h/2}\right) = \frac{y_k^2}{2}\left(\frac{x_{k+1/2}-x_k}{h/2}\right)$	$C_d : \left(\frac{z_{k+1}-z_{k+1/2}}{h/2}\right) = \frac{y_{k+1}^2}{2}\left(\frac{x_{k+1}-x_{k+1/2}}{h/2}\right)$
$L_d^+(q_k, q_{k+1/2})$	$L_d^-(q_{k+1/2}, q_{k+1})$

As shown in the table above, both discrete submanifolds are respectively defined by the constraints $\left(\frac{z_{k+1/2}-z_k}{h/2}\right)=\frac{y_k^2}{2}\left(\frac{x_{k+1/2}-x_k}{h/2}\right)$ and $\left(\frac{z_{k+1}-z_{k+1/2}}{h/2}\right)=\frac{y_{k+1}^2}{2}\left(\frac{x_{k+1}-x_{k+1/2}}{h/2}\right)$. Taking into account equations (4.54) and (4.55) in §4.4.3, define the two discrete Lagrangian functions (implicitly expressed) as

$$L_d^+(q_k,q_{k+1/2})=p_{k+1/2}\,q_{k+1/2}-p_{k+1/2}\,q_k-\frac{h}{2}H(q_k,p_{k+1/2}),$$

with $q_{k+1/2}=q_k+\frac{h}{2}H_p(q_k,p_{k+1/2})$ (recall that this expression corresponds to the generating function after projecting onto $Q\times Q$). On the other hand

$$L_d^-(q_{k+1/2},q_{k+1})=-p_{k+1/2}\,q_{k+1/2}+p_{k+1/2}\,q_{k+1}-\frac{h}{2}H(q_{k+1},p_{k+1/2}),$$

with $q_{k+1}=q_{k+1/2}+\frac{h}{2}H_p(q_{k+1},p_{k+1/2})$.

From the previous expressions and following the discussion in [124] §2.5.1, we divide each step (q_k,q_{k+1}) into 2 substeps $(q_k=q_k^0\,,\,q_k^1=q_{k+1/2})$ and $(q_k^1=q_{k+1/2}\,,\,q_{k+1}^0=q_{k+1})$. Take the discrete action sum

$$\begin{aligned}\mathfrak{S}_d(\{q_k^0,q_k^1\}_0^{N-1}) &= \sum_{k=0}^{N-1}\left(L_d^+(q_k^0,q_k^1)+L_d^-(q_k^1,q_{k+1}^0)\right)\\ &= \sum_{k=0}^{N}\left(L_d^+(q_k,q_{k+1/2})+L_d^-(q_{k+1/2},q_{k+1})\right).\end{aligned}$$

The corresponding Euler-Lagrange equations, resulting from requiring this action to be stationary, pair both neighboring discrete Lagrangians together to give

$$\begin{aligned}D_2\,L_d^+(q_k^0.q_k^1) &+ D_1\,L_d^-(q_k^1,q_k^2)=0,\\ D_2\,L_d^-(q_k^1,q_k^2) &+ D_1\,L_d^+(q_{k+1}^0,q_{k+1}^1)=0.\end{aligned}$$

The equations just above completely determine the discrete dynamics in equations (4.52) and (4.53) for the Martinet sub-Riemannian system (4.42). In other words, the discrete scheme in (4.52) and (4.53) could be understood as two discrete Hamiltonian flows $F_{L_d^+}$ and $F_{L_d^-}$, each one of half-step $h/2$, respectively generated by the generating functions L_d^+ and L_d^-. The map over the entire time-step h is thus the composition of the maps $F_{L_d^-}\circ F_{L_d^+}$.

Chapter 5

Discrete mechanics and optimal control

The goal of this chapter is to develop, from a geometric point of view, numerical methods for optimal control of Lagrangian mechanical systems. We will employ the theory of discrete mechanics and variational integrators presented in §3 to derive both an integrator for the dynamics and an optimal control algorithm in a unified manner. The proposed framework is general and applies to unconstrained systems, as well as systems with symmetries, underactuation, and nonholonomic constraints. We pay special attention to Lagrangian systems defined on tangent bundles and Lie groups. The extension to principal bundles and nonholonomic mechanics is carefully studied in §5.3.

The main idea is the following: we take an approximation of the Lagrange-d'Alembert principle for forced Lagrangian systems, which models control inputs and external forces such as gravity or drag forces. In principle, we admit the possibility of piecewise continuous control forces, as happens in real applications. We observe that the discrete equations of motion for this type of systems are interpreted as the discrete Euler-Lagrange equations of a new Lagrangian defined in an augmented discrete phase space. Next, we apply discrete variational calculus techniques to derive the discrete optimality conditions. After this, we recover two sequences of discrete controls modeling a piecewise control trajectory.

Moreover, since we are reducing the optimality conditions to discrete Euler-Lagrange equations, the geometric preservation properties like symplectic-momentum preservation in the standard case or Poisson bracket and momentum preservation for reduced systems are automatically guaranteed using the results in [118, 124].

Several theoretical and a practical examples, e.g. the control of an underwater vehicle, will illustrate the application of the proposed approach.

5.1 Optimal control of a mechanical system defined in TQ

5.1.1 Continuous Lagrangian picture

Let Q be the configuration manifold, with (q^i) local coordinates $i = 1, ..., n$. We consider a mechanical system described by a regular Lagrangian $L : TQ \to \mathbb{R}$. The induced local coordinates on TQ are $(q^i, \dot{q}^i)$. Additionally, there are **external-control forces** present defined by the map

$$f : TQ \times U \to T^*Q,$$

such that $f(v_q, u) \in T_q^*Q$ for $v_q \in T_qQ$.

To define these control forces we have introduced the **control manifold** $U \subset \mathbb{R}^m$ $(m \leq n)$ for a given interval $I = [0, T]$. The **control path space** is defined by

$$\mathcal{D}(U) = \mathcal{D}([0, T], U) = \{u : [0, T] \to U \,|\, u \in L^\infty\},$$

with $u(t) \in U$ also called the **control parameter**. $L^p(\mathbb{R}^m)$ denotes the usual Lebesgue space of measurable functions $x : [0, T] \to \mathbb{R}^m$ with $|x(\cdot)|^p$ integrable, equipped with its standard norm

$$||x||_{L^p} = \left(\int_0^T |x(t)|^p \mathrm{d}t\right)^{\frac{1}{p}},$$

where $|\cdot|$ is the Euclidean norm. 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 coordinates as

$$f(u) : (q^i, \dot{q}^i) \mapsto (q^i, f(u)(q^i, \dot{q}^i)).$$

Remark 5.1.1. *Note that the definition of a control force also includes forces that are independent on the control parameter. Thus, in the following, we restrict to a formulation with control forces which gives us the opportunity to include friction or dissipative forces as well.*

The optimal control problem can be defined as follows: during the time interval $I = [0, T]$ the mechanical system described by L moves on a curve $q(t) \in Q$ from an initial state $(q(0), \dot{q}(0))$ to a final state $(q(T), \dot{q}(T))$. The motion is influenced via the external forces f with chosen control parameter $u(t)$ such that a given **cost functional**

$$\mathcal{J}(q, \dot{q}, u) = \int_0^T C(q(t), \dot{q}(t), u(t))\mathrm{d}t, \tag{5.1}$$

where $C : TQ \times U \to \mathbb{R}$ is the **cost function** (continuously differentiable).

The initial state is fixed in the following way: $q(0) = q^0$ and $\dot{q}(0) = \dot{q}^0$, where $(q^0, \dot{q}^0) \in TQ$ is a fixed value for the initial state, that is input data coming from experiments or any theoretical assumption.

On the other hand, the final state is given by the time constraint $r(q(T), \dot{q}(T), q^T, \dot{q}^T) = 0$ with $r : TQ \times TQ \to \mathbb{R}^{n_r}$ (n_r is an integer), where $(q^T, \dot{q}^T) \in TQ$ is a fixed value for the desired final state. This boundary condition allows the simplest velocity constraint $\dot{q}(T) = \dot{q}^T$, that is $r(q(T), \dot{q}(T), q^T, \dot{q}^T) = \dot{q}(T) - \dot{q}^T$, or more involved ones.

At the same time, the motion $q(t)$ of the system has to satisfy the **Lagrange-d'Alembert principle**, which requires that

$$\delta \int_0^T L(q(t), \dot{q}(t))\, \mathrm{d}t + \int_0^T f(q(t), \dot{q}(t), u(t))\, \delta q(t)\ \mathrm{d}t = 0, \tag{5.2}$$

where we consider arbitrary variations $\delta q(t) \in T_{q(t)}Q$ with $\delta q(0) = 0$ and $\delta q(T) = 0$.

The optimal control problem for a Lagrangian system can be now formulated as follows

Problem 5.1.2 (**Lagrangian optimal control problem).**

$$\min_{(q(\cdot),\dot{q}(\cdot),u(\cdot),T)} \quad \int_0^T C(q(t), \dot{q}(t), u(t))\mathrm{d}t, \tag{5.3a}$$

subject to

$$\delta \int_0^T L(q(t), \dot{q}(t))\, \mathrm{d}t + \int_0^T f(q(t), \dot{q}(t), u(t))\, \delta q(t)\ \mathrm{d}t = 0, \tag{5.3b}$$

$$q(0) = q^0, \ \ \dot{q}(0) = \dot{q}^0, \tag{5.3c}$$

$$r(q(T), \dot{q}(T), q^T, \dot{q}^T) = 0. \tag{5.3d}$$

The final time T may either be fixed or appear as a degree of freedom in the optimization problem.

Under these conditions it is necessary to apply the Pontryaguin maximum principle (see [14, 143] for further details) in order to derive the equations of motion. Generally, it is not possible to explicitly integrate these equations and, consequently, it is necessary to apply a numerical method.

5.1.2 Discrete Lagrangian picture

To obtain a discrete formulation, we replace each expression in the previous subsection by its discrete counterpart in terms of discrete variational mechanics, that is following the prescriptions in §3. Briefly, we replace the state space TQ of the system by $Q \times Q$ and a path $q : [0, T] \to Q$ by a discrete path $q_d : \{0, h, 2h, ..., Nh = T\} \to Q$, $N \in \mathbb{N}$, with $q_k = q_d(hk) \approx q(t_k)$. Analogously, the continuous control path $u : [0, T] \to U$ is replaced by a discrete control path $u_d : \Delta\tilde{t} \to U$. In order to establish a discrete version of the Lagrange-d'Alembert principle, is necessary to clarify the properties of this discrete control path and, in addition, the discrete control forces.

Discrete control forces

For the replacement of the control space by a discrete one we introduce a new time grid $\Delta\tilde{t}$. This time grid is generated by an increasing sequence of intermediate control points

$c = \{c_l \,|\, 0 \leq c_l \leq 1,\, l = 1, ..., s\}$ as $\Delta\tilde{t} = \{t_{kl} \,|\, k \in \{0, ..., N-1\}\,,\; l \in \{1, ..., s\}\}$, where $t_{lk} = t_k + c_l h$. With this notation the **discrete control path space** is defined to be

$$\mathcal{D}_d(U) = \mathcal{D}_d(\Delta\tilde{t}, U) = \left\{u_d : \Delta\tilde{t} \to U\right\}.$$

We define the **intermediate control samples** u_k on $[t_k, t_{k+1}]$ as $u_k = (u_{k1}, ..., u_{ks}) \in U^s$ to be the values of the control parameters guiding the system from $q_k = q_d(t_k)$ to $q_{k+1} = q_d(t_{k+1})$, where $u_{kl} = u_d(t_{kl})$ for $l \in \{1, ..., s\}$.

The set of discrete controls U^s can be viewed as a finite dimensional subspace of the control path space $\mathcal{D}([0, h], U)$.

There are several approaches to discrete optimal control theory in the mathematical literature. An inspiring approach for our work is the one in [139], where the authors consider the following discretization of the control forces: take two **discrete control forces** $f_k^{\pm} : Q \times Q \times U^s \to T^*Q$:

$$f_k^-(q_k, q_{k+1}, u_k) \in T^*_{q_k}Q, \tag{5.4a}$$
$$f_k^+(q_k, q_{k+1}, u_k) \in T^*_{q_{k+1}}Q, \tag{5.4b}$$

also called **left and right forces**[1]. Analogously to the continuous case, the two discrete control forces are interpreted as two parameter-dependent discrete fiber-preserving forces $f_k^{\pm}(u_k) : Q \times Q \to T^*Q$ in the sense that $\pi_Q \circ f_k^{\pm} = \pi_Q^{\pm}$, where $\pi_Q : T^*Q \to Q$ is the usual projection and $\pi_Q^{\pm} : Q \times Q \to Q$ are projection operators defined by $\pi_Q^-(q_k, q_{k+1}) = q_k$ and $\pi_Q^+(q_k, q_{k+1}) = q_{k+1}$. The two discrete control forces are combined to give a single one-form $f_k(u_k) : Q \times Q \to T^*(Q \times Q)$ defined by

$$f_k(u_k)(q_k, q_{k+1})(\delta q_k, \delta q_{k+1}) = f_k^+(u_k)(q_k, q_{k+1})\delta q_{k+1} + f_k^-(u_k)(q_k, q_{k+1})\delta q_k. \tag{5.5}$$

The left discrete force f_{k-1}^+ is interpreted as the force resulting from the continuous control force acting during the time span $[t_{k-1}, t_k]$ on the configuration node q_k. The right discrete force f_k^- is the force acting on q_k resulting from the continuous control force during the time span $[t_k, t_{k+1}]$.

Discrete Lagrange-d'Alembert principle

As with discrete Lagrangian functions, the discrete control forces also depend on the time step h, which is important when relating discrete and continuous mechanics. Given such forces, the discrete Hamilton's principle is modified, following [85], to the **discrete Lagrange-d'Alembert principle**, which seeks discrete curves $\{q_k\}_{k=0}^N$ that satisfy

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) + \sum_{k=0}^{N-1} \left(f_k^+(u_k)(q_k, q_{k+1})\delta q_{k+1} + f_k^-(u_k)(q_k, q_{k+1})\delta q_k\right) = 0, \tag{5.6}$$

for all variations $\{\delta q_k\}_{k=0}^N$ vanishing at the endpoints. This is equivalent to the **forced discrete Euler-Lagrange equations**

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) + f_{k-1}^+(q_{k-1}, q_k, u_{k-1}) + f_k^-(q_k, q_{k+1}, u_k) = 0, \tag{5.7}$$

[1] Observe, that the discrete control force is now dependent on the discrete control path.

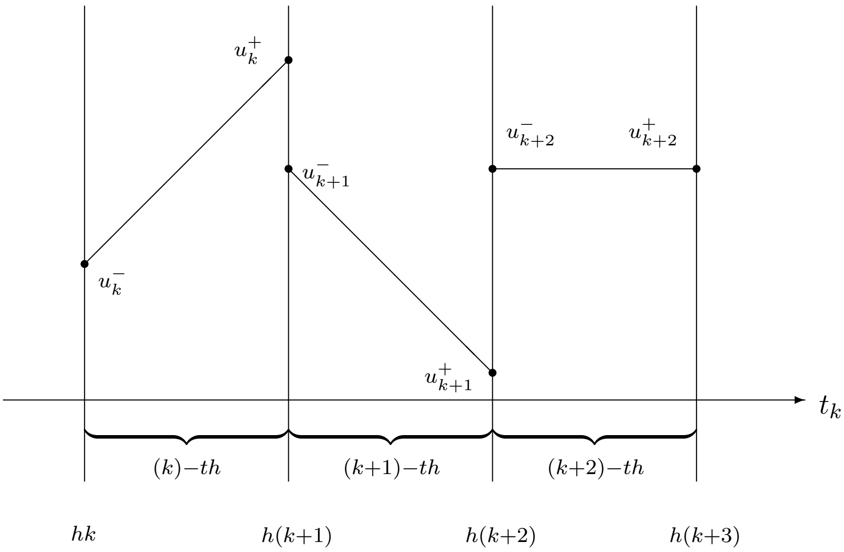
which are the same as the standard discrete Euler-Lagrange equations with the discrete forces added. These implicitly define the **forced discrete Lagrangian map** $F_{L_d}(u_{k-1}, u_k) : Q \times Q \to Q \times Q$. For further details see [85, 124, 138, 139].

5.1.3 Discrete optimal control problem

Now, we propose another discretization of the control forces which allow us to consider a more general class of controls (see [76]). Namely, we allow in the sequel two different sequences of discrete controls $\{u_k^+\}$ and $\{u_k^-\}$ in (5.4). That is

$$f_k^-(q_k, q_{k+1}, u_k^-) \in T^*_{q_k}Q, \tag{5.8a}$$
$$f_k^+(q_k, q_{k+1}, u_k^+) \in T^*_{q_{k+1}}Q. \tag{5.8b}$$

In the notation followed through the rest of this chapter, the time interval between $[k, k+1]$ is denoted as the k-th interval. This choice allows us to model piecewise continuous controls, admitting discrete jumps at the time steps $t_k = hk$. Our notation is completely depicted in the following figure:



Under these definitions we take the following approximation of the control forces in (5.2):

$$f_k^-(q_k, q_{k+1}, u_k^-)\,\delta q_k + f_k^+(q_k, q_{k+1}, u_k^+)\,\delta q_{k+1} \approx \int_{kh}^{(k+1)h} f(q(t), \dot{q}(t), u(t))\delta q(t)\;\mathrm{d}t$$

where $\left(f_k^-(q_k, q_{k+1}, u_k^-), f_k^+(q_k, q_{k+1}, u_k^+)\right) \in T^*_{q_k}Q \times T^*_{q_{k+1}}Q$. Consequently, the discrete Lagrange-d'Alembert principle given in (5.6) is modified in the following way

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) + \sum_{k=0}^{N-1} \left(f_k^-(q_k, q_{k+1}, u_k^-)\, \delta q_k + f_k^+(q_k, q_{k+1}, u_k^+)\, \delta q_{k+1}\right) = 0,$$

for all variations $\{\delta q_k\}_{k=0,\dots N}$ with $\delta q_k \in T_{q_k}Q$ such that $\delta q_0 = \delta q_N = 0$. From this principle is easy to derive the system of difference equations:

$$D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) + f_{k-1}^+(q_{k-1}, q_k, u_{k-1}^+) + f_k^-(q_k, q_{k+1}, u_k^-) = 0, \tag{5.9}$$

where $k = 1, \dots, N-1$. Equations (5.9) are the new **forced discrete Euler-Lagrange equations**.

We can also approximate the cost functional (5.1) in a single time step h by

$$C_d(q_k, u_k^-, q_{k+1}, u_k^+) \approx \int_{kh}^{(k+1)h} C(q(t), \dot q(t), u(t))\, \mathrm{d}t,$$

yielding the **discrete cost functional**:

$$\mathcal{J}_d(q_{0:N-1}, u^{\pm}_{0:N-1}) = \sum_{k=0}^{N-1} C_d(q_k, u_k^-, q_{k+1}, u_k^+)\ .$$

Observe that $C_d : Q \times U \times Q \times U \to \mathbb{R}$.

Boundary conditions

In the next step, we need to incorporate the boundary conditions $q(0) = q^0$, $\dot q(0) = \dot q^0$ and $r(q(T), \dot q(T), q^T, \dot q^T) = 0$ into the discrete description. Those on the configuration level can be used as constraints in a straightforward what as $q(0) = q_0 = q^0$. However, since in the present formulation velocities are approximated in a time interval $[t_k, t_{k+1}]$, the velocity conditions have to be transformed to conditions on the conjugate momenta. These are defined at each and every time node using the discrete Legendre transform. The presence of forces at the time nodes has to be incorporated into that transformation leading to the forced discrete Legendre transforms $\mathbb{F}^f L_d^{\pm}(u) : Q \times Q \to T^*Q$ defined by

$$\mathbb{F}^f L_d^+(u) : (q_0, q_1) \mapsto (q_1, p_1) = (q_1, D_2L_d(q_0, q_1) + f_0^+(q_0, q_1, u)), \tag{5.10a}$$

$$\mathbb{F}^f L_d^-(u) : (q_0, q_1) \mapsto (q_0, p_0) = (q_0, -D_1L_d(q_0, q_1) - f_0^-(q_0, q_1, u)). \tag{5.10b}$$

Using the standard Legendre transform $\mathbb{F}L : TQ \to T^*Q$, $(q, \dot q) \to (q, p) = (q, D_2L(q, \dot q))$ defined in §2.2.1 and the two different sequences of discrete controls $\{u_k^{\pm}\}_{k=0}^N$ leads to the discrete initial constraint on the conjugate momentum:

$$D_2L(q^0, \dot q^0) + D_1L_d(q_0, q_1) + f_0^-(q_0, q_1, u_0^-) = 0. \tag{5.11}$$

We can transform the boundary condition from a formulation with configuration and velocity to a formulation with configuration and conjugate momentum. Thus, instead of considering a discrete version of the final time constraint r on TQ we use a discrete version of the final time constraint $\tilde{r}$ on $T^*Q \times T^*Q$. We define the **discrete boundary condition** on the configuration level to be

$$r_d : Q \times Q \times U^s \times TQ \to \mathbb{R}^{n_r},$$
$$r_d(q_{N-1}, q_N, u^+_{N-1}, q^T, \dot{q}^T) = \tilde{r}_d(\mathbb{F}^f L^+_d(q_{N-1}, q_N, u^+_{N-1}), \mathbb{F}L(q^T, \dot{q}^T)),$$

with $(q_N, p_N) = \mathbb{F}^f L^+_d(q_{N-1}, q_N, u^+_{N-1})$ and $(q^T, p^T) = \mathbb{F}L(q^T, \dot{q}^T)$, which can be translated in the final boundary conditions

$$p_N = D_2 L_d(q_{N-1}, q_N) + f^+_{N-1}(q_{N-1}, q_N, u^+_{N-1}), \tag{5.12a}$$
$$p^T = D_2 L(q^T, \dot{q}^T). \tag{5.12b}$$

Remark 5.1.3. *For the simple velocity constraint $r(q(T), \dot{q}(T), q^T, \dot{q}^T) = \dot{q}(T) - \dot{q}^T$, we obtain for the transformed condition on the momentum level $\tilde{r}(q(T), p(T), q^T, p^T) = p(T) - p^T$ the discrete constraint*

$$-D_2 L(q^T, \dot{q}^T) + D_2 L_d(q_{N-1}, q_N) + f^+_{N-1}(q_{N-1}, q_N, u^+_{N-1}) = 0.$$

Finally, after performing the above discretization steps, one is faced with the following **discrete optimal control problem**:

Problem 5.1.4 (Discrete optimal control problem).

$$\min_{(q_{0:N}, u^{\pm}_{0:N-1}, h)} \sum_{k=0}^{N-1} C_d(q_k, u^-_k, q_{k+1}, u^+_k), \tag{5.13a}$$

subject to

$$\delta \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}) + \sum_{k=0}^{N-1} \left(f^-_k(q_k, q_{k+1}, u^-_k)\, \delta q_k + f^+_k(q_k, q_{k+1}, u^+_k)\, \delta q_{k+1} \right) = 0, \tag{5.13b}$$
$$q(0) = q_0 = q^0, \tag{5.13c}$$
$$D_2 L(q^0, \dot{q}^0) + D_1 L_d(q_0, q_1) + f^-_0(q_0, q_1, u^-_0) = 0, \tag{5.13d}$$
$$r_d(q_{N-1}, q_N, u^+_{N-1}, q^T, \dot{q}^T) = 0. \tag{5.13e}$$

Recall that $f^{\pm}_k$ are dependent on $u_k \in U^s$. To incorporate a free final time T as in the continuous setting, the step size h appears as a degree of freedom within the optimization problem. However, in the following formulations and considerations we restrict ourselves to the case of fixed final time T and thus fixed step size h.

In the sequel we develop our new method to approach (5.13), which consists in solving the discrete optimal control problem as a variational integrator of a specially constructed higher-dimensional system.

5.1.4 Fully-actuated case

We say that a system is **fully-actuated** when the dimension of the control space equals the number of degrees of freedom of the system under study. We perform the fully actuation by means of the following definition:

Definition 5.1.5. (Fully-actuated discrete system) *We say that the discrete mechanical control system is fully-actuated if the mappings*

$$f_k^-\big|_{(q_k,q_{k+1})} : U \to T^*_{q_k}Q, \qquad f_k^-\big|_{(q_k,q_{k+1})}(u) = f_k^-(q_k, q_{k+1}, u),$$
$$f_k^+\big|_{(q_k,q_{k+1})} : U \to T^*_{q_{k+1}}Q, \qquad f_k^+\big|_{(q_k,q_{k+1})}(u) = f_k^+(q_k, q_{k+1}, u),$$

are both diffeomorphisms.

Define the momenta

$$p_k = -D_1L_d(q_k, q_{k+1}) - f_k^-(q_k, q_{k+1}, u_k^-), \tag{5.14}$$
$$p_{k+1} = D_2L_d(q_k, q_{k+1}) + f_k^+(q_k, q_{k+1}, u_k^+). \tag{5.15}$$

Since both $f_k^\pm\big|_{(q_k,q_{k+1})}$ are diffeomorphisms we can work out $u_k^\pm$ in terms of $(q_k, p_k, q_{k+1}, p_{k+1})$ from (5.14) and (5.15). Now, we are in situation to define a new Lagrangian $\mathcal{L}_d : T^*Q\times T^*Q \to \mathbb{R}$ by

$$\begin{aligned}&\mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1}) = \\ &= C_d(q_k\,,\, (f_k^-\big|_{(q_k,q_{k+1})})^{-1}(-D_1L_d - p_k)\,,\, q_{k+1}\,,\, (f_k^+\big|_{(q_k,q_{k+1})})^{-1}(-D_2L_d + p_{k+1})).\end{aligned}$$

The system is fully-actuated, consequently the Lagrangian $\mathcal{L}_d$ is well defined on the entire discrete space $T^*Q\times T^*Q$. Now the discrete phase space is the Cartesian product $T^*Q\times T^*Q$ of two copies of the cotangent bundle. The definition (5.14), (5.15) gives us a matching of momenta (see [124]) which automatically implies

$$D_2L_d(q_{k-1}, q_k) + f_{k-1}^+(q_{k-1}, q_k, u_{k-1}^+) = -D_1L_d(q_k, q_{k+1}) - f_k^-(q_k, q_{k+1}, u_k^-),$$

$k = 1, \ldots, N-1$, which are the forced discrete Euler-Lagrange equations (5.9). In other words, the matching condition enforces that the momentum at time k should be the same when evaluated from the lower interval $[k-1, k]$ or the upper interval $[k, k+1]$. Consequently, along a solution curve there is a unique momentum at each time t_k, which can be called p_k.

Once we have the new discrete Lagrangian $\mathcal{L}_d$ it is possible to define the discrete action sum

$$\mathcal{S}_d = \sum_{k=0}^{N-1} \mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1}).$$

Applying the Hamilton's principle we obtain the discrete Euler-Lagrange equations of motion for the Lagrangian $\mathcal{L}_d : T^*Q \times T^*Q \to \mathbb{R}$:

$$D_3\mathcal{L}_d(q_{k-1}, p_{k-1}, q_k, p_k) + D_1\mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1}) = 0, \tag{5.16}$$
$$D_4\mathcal{L}_d(q_{k-1}, p_{k-1}, q_k, p_k) + D_2\mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1}) = 0\,. \tag{5.17}$$

In conclusion, we have obtained the discrete equations of motion for a fully-actuated mechanical optimal control problem as the discrete Euler-Lagrange equations for a Lagrangian defined on the Cartesian product of two copies of the cotangent bundle.

Example 5.1.6 (Optimal control problem for a mechanical Lagrangian with configuration space $\mathbb{R}^n$). *Let consider $x \in \mathbb{R}^n$, M a $n \times n$ constant and symmetric matrix and the mechanical Lagrangian $L : \mathbb{R}^{2n} \to \mathbb{R}$ defined by $L(x, \dot{x}) = \frac{1}{2}\dot{x}^T M \dot{x} - V(x)$, where $V : \mathbb{R}^n \to \mathbb{R}$ is the potential function and $\dot{x}$ represents the time derivative of x. Our boundary constraints are $(x(0), \dot{x}(0))$ and $(x(T), \dot{x}(T))$. The system is fully actuated and there exist no velocity constraints. Note that in the continuous setting we can define the momentum by the continuous Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, $(q, \dot{q}) \mapsto (q, p)$: $p = \frac{\partial L}{\partial \dot{x}}$, i.e. $p(t) = \dot{x}^T(t)\, M$. In consequence, we can define boundary constraints also in the phase space: $(x(0)\,,\, p(0) = \dot{x}(0)^T\, M)$ and $(x(T)\,,\, p(T) = \dot{x}(T)^T\, M)$.*

We set the Trapezoidal discretization for the Lagrangian (see [61]), that is,

$$L_d(x_k, x_{k+1}) = \frac{h}{2}\, L(x_k, \frac{x_{k+1} - x_k}{h}) + \frac{h}{2}\, L(x_{k+1}, \frac{x_{k+1} - x_k}{h})$$

where, as above, h is the fixed time step and $x_1, x_2, \ldots, x_N$ is a sequence of elements on $\mathbb{R}^n$. Our concrete discrete Lagrangian is

$$L_d(x_k, x_{k+1}) = \frac{1}{2h}(x_{k+1} - x_k)^T M (x_{k+1} - x_k) - \frac{h}{2}\,(V(x_k) + V(x_{k+1}))\,.$$

The control forces are $f_k^-(x_k, x_{k+1}, u_k^-) \in T^*_{x_k}\mathbb{R}^n$ and $f_k^+(x_k, x_{k+1}, u_k^+) \in T^*_{x_{k+1}}\mathbb{R}^n$. For sake of clarity, we are going to fix the control forces in the following manner $f^{\pm}(x_k, x_{k+1}, u_k^{\pm}) = u_k^{\pm}$. Looking at equations (5.14) and (5.15) is easy to obtain the associated momenta p_k and p_{k+1}, namely

$$\begin{aligned} p_k &= \frac{1}{h}\,(x_{k+1} - x_k)^T\, M + \frac{h}{2}\, V_x(x_k)^T - u_k^-, \\ p_{k+1} &= \frac{1}{h}\,(x_{k+1} - x_k)^T\, M - \frac{h}{2}\, V_x(x_{k+1})^T + u_k^+. \end{aligned}$$

Let

$$C_d = \frac{h}{4}\sum_{k=0}^{N-1}\left[(u_k^-)^2 + (u_k^+)^2\right]$$

be a discrete cost function. Consequently, the Lagrangian over $T^*\mathbb{R} \times T^*\mathbb{R}$ is

$$\begin{aligned} &\mathcal{L}_d(x_k, p_k, x_{k+1}, p_{k+1}) = \\ &= \frac{1}{4}\sum_{k=0}^{N-1}\left(p_k - \left(\frac{x_{k+1} - x_k}{h}\right)^T M - \frac{h}{2}\, V_x(x_k)^T\right)^2 \\ &+ \frac{1}{4}\sum_{k=0}^{N-1}\left(p_{k+1} - \left(\frac{x_{k+1} - x_k}{h}\right)^T M + \frac{h}{2}\, V_x(x_{k+1})^T\right)^2, \end{aligned}$$

where V_x represents the derivative of V with respect to the variable x. Applying equations (5.16) and (5.17) to $\mathcal{L}_d$ we obtain the following equations:

$$p_k - \left(\frac{x_{k+1} - x_{k-1}}{2h}\right)^T M = 0, \tag{5.18}$$

$$\begin{aligned}&\left(p_k - (\frac{x_{k+1} - x_k}{h})^T M - \frac{h}{2} V_x(x_k)^T\right)\left(M - \frac{h^2}{2} V_{xx}(x_k)^T\right) \\ &- \left(p_k - (\frac{x_k - x_{k-1}}{h})^T M + \frac{h}{2} V_x(x_k)^T\right)\left(M - \frac{h^2}{2} V_{xx}(x_k)^T\right) = 0,\end{aligned} \tag{5.19}$$

where both set of equations are defined for $k = 1, ..., N-1$. It is quite clear that we could remove the p_k dependence in equation (5.19). However, we prefer to keep it in order to stress that the discrete variational Euler-Lagrange equations (5.16) and (5.17) are defined in $T^*Q \times T^*Q$ ($T^*\mathbb{R} \times T^*\mathbb{R}$ in the particular case we are considering in this example).

Expressions (5.18) and (5.19) mean $2(N-1)n$ equations for the $2(N+1)n$ unknowns $\{x_k\}_{k=0}^N$, $\{p_k\}_{k=0}^N$. Nevertheless, we translate the boundary conditions (5.13c), (5.13d) and (5.13e) (in its Hamiltonian version $\tilde{r}_d(x_N, p_N, x^T, p^T) = 0$) into

$$\begin{aligned} x_0 &= x(0), \quad p_0 = p(0), \\ x_N &= x(T), \quad p_N = p(T), \end{aligned}$$

which contribute $4n$ extra equations and convert eqs. (5.18) and (5.19) in a nonlinear root finding problem of $2(N-1)n$ and the same amount of unknowns.

5.1.5 Under-actuated case

We say that a system is **under-actuated** when the dimension of the control space is fewer than the number of degrees of freedom of the system under study. We perform the under-actuation by means of the following definition:

Definition 5.1.7. (Under-actuated discrete system) *We say that the discrete mechanical control system is underactuated if the mappings*

$$\begin{aligned} f_k^-\big|_{(q_k,q_{k+1})} &: U \to T^*_{q_k}Q, \qquad & f_k^-\big|_{(q_k,q_{k+1})}(u) = f_k^-(q_k, q_{k+1}, u), \\ f_k^+\big|_{(q_k,q_{k+1})} &: U \to T^*_{q_{k+1}}Q, \qquad & f_k^+\big|_{(q_k,q_{k+1})}(u) = f_k^+(q_k, q_{k+1}, u), \end{aligned}$$

are both embeddings.

Under the embedding requirement we ensure $f_k^{\pm}(U) \subset T^*Q$ to has a submanifold structure. Furthermore, under this hypothesis we deduce that $\mathcal{M}^-_{(q_k,q_{k+1})} = f_k^-\big|_{(q_k,q_{k+1})}(U)$, $\mathcal{M}^+_{(q_k,q_{k+1})} = f_k^+\big|_{(q_k,q_{k+1})}(U)$ are submanifolds of $T^*_{q_k}Q$ and $T^*_{q_{k+1}}Q$, respectively. Therefore, $f_k^{\pm}\big|_{(q_k,q_{k+1})}$ are diffeomorphisms onto its image. Moreover, $\dim \mathcal{M}^-_{(q_k,q_{k+1})} = \dim \mathcal{M}^+_{(q_k,q_{k+1})} = \dim U$.

As in the fully-actuated case, we define the Lagrangian function

$$\begin{aligned}&\mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1})\\ &= C_d(q_k\,,\,(f_k^-|_{(q_k,q_{k+1})})^{-1}(-D_1L_d - p_k)\,,\,q_{k+1}\,,\,(f_k^+|_{(q_k,q_{k+1})})^{-1}(-D_2L_d + p_{k+1})),\end{aligned}$$

where we have defined the momenta

$$p_k = -D_1L_d(q_k, q_{k+1}) - f_k^-(q_k, q_{k+1}, u_k^-), \tag{5.20}$$

$$p_{k+1} = D_2L_d(q_k, q_{k+1}) + f_k^+(q_k, q_{k+1}, u_k^+). \tag{5.21}$$

To define $\mathcal{L}_d$ in the way we have just done, is necessary to consider that $(q_k\,,\,-D_1L_d(q_k, q_{k+1}) - p_k\,,\,q_{k+1}\,,\,-D_2L_d(q_k, q_{k+1}) + p_{k+1})$ is a point of $\mathcal{M}^-_{(q_k,q_{k+1})} \times \mathcal{M}^-_{(q_k,q_{k+1})}$ in order to calculate its image by the inverse functions $\left(f_k^\pm|_{(q_k,q_{k+1})}\right)^{-1}$. That is, the Lagrangian function $\mathcal{L}_d$ is defined for points $(q_k, p_k, q_{k+1}, p_{k+1})$ satisfying

$$\begin{aligned}(q_k\,,\,-D_1L_d(q_k, q_{k+1}) - p_k) &\in \mathcal{M}^-_{(q_k,q_{k+1})} \subset T^*_{q_k}Q,\\ (q_{k+1}\,,\,-D_2L_d(q_k, q_{k+1}) + p_{k+1}) &\in \mathcal{M}^+_{(q_k,q_{k+1})} \subset T^*_{q_{k+1}}Q.\end{aligned}$$

In many examples of interest, these conditions are performed by means of constraint functions $\Phi^-_\alpha, \Phi^+_\alpha : T^*Q \times T^*Q \to \mathbb{R}$, $1 \le \alpha \le n - \dim U$ and therefore the solutions of the optimal control problem are now viewed as the solutions of the discrete constrained problem (in other words Vakonomic) determined by the Lagrangian $\mathcal{L}_d$ and the constraints $\Phi^\pm_\alpha$. Since $f^\pm|_{(q_k,q_{k+1})}$ are embeddings, as established in definition (5.1.7), the number of constraints is determined by n minus the dimension of U. Note that the total number of constraints, $\Phi^\pm_\alpha$, is therefore $2(n - \dim U)$.

To solve this problem we introduce Lagrange multipliers $(\lambda_k^-)^\alpha$,$(\lambda_k^+)^\alpha$ and the extended Lagrangian

$$\begin{aligned}\widetilde{L}_d(q_k, p_k, \lambda_k^-, q_{k+1}, p_{k+1}, \lambda_k^+) =&\mathcal{L}_d(q_k, p_k, q_{k+1}, p_{k+1})\\ &+ (\lambda_k^-)^\alpha \Phi^-_\alpha(q_k, p_k, q_{k+1}, p_{k+1})\\ &+ (\lambda_k^+)^\alpha \Phi^+_\alpha(q_k, p_k, q_{k+1}, p_{k+1}).\end{aligned}$$

Observe that, in spite the constraints are functions of the cartesian product of two copies of the cotangent bundle i.e. $\Phi^\pm_\alpha : T^*Q \times T^*Q \to \mathbb{R}$, neither Φ^-_α depends on p_{k+1} nor Φ^+_α on p_k. The discrete Euler-Lagrange equations gives us the solutions of the underactuated problem.

Typically, the underactuated systems appear in an affine way that is

$$\begin{aligned}f_k^-(q_k, q_{k+1}, u_k^-) &= A_k^-(q_k, q_{k+1}) + B_k^-(q_k, q_{k+1})(u_k^-)\\ f_k^+(q_k, q_{k+1}, u_k^+) &= A_k^+(q_k, q_{k+1}) + B_k^+(q_k, q_{k+1})(u_k^+)\end{aligned}$$

where $A_k^-(q_k, q_{k+1}) \in T^*_{q_k}Q$, $A_k^+(q_k, q_{k+1}) \in T^*_{q_{k+1}}Q$. Moreover $B_k^-(q_k, q_{k+1}) \in \mathrm{Lin}(U, T^*_{q_k}Q)$ and $B_k^+(q_k, q_{k+1}) \in \mathrm{Lin}(U, T^*_{q_{k+1}}Q)$ are linear maps (we assume that U is a vector space and $\mathrm{Lin}(E_1\,,\,E_2)$ is the set of all linear maps between E_1 and E_2). In consequence $B_k^-(q_k, q_{k+1})(u_k^-) \in T^*_{q_k}Q$ and $B_k^+(q_k, q_{k+1})(u_k^+) \in T^*_{q_{k+1}}Q$.

Then the constraints are deduced using the compatibility conditions:

$$\begin{aligned}\operatorname{rank} B_k^- &= \operatorname{rank}\left(B_k^-\,;\,-D_1L_d(q_k,q_{k+1})-p_k-A_k^-(q_k,q_{k+1})\right),\\ \operatorname{rank} B_k^+ &= \operatorname{rank}\left(B_k^+\,;\,-D_2L_d(q_k,q_{k+1})+p_{k+1}-A_k^+(q_k,q_{k+1})\right),\end{aligned}$$

which imply constraints in (q_k, q_{k+1}, p_k) and (q_k, q_{k+1}, p_{k+1}) respectively. The fact that $f_k^{\pm}\big|_{(q_k,q_{k+1})}$ are both embeddings implies furthermore that $\operatorname{rank} B_k^- = \operatorname{rank} B_k^+ = \dim U$.

5.2 Optimal control of a mechanical system defined in a Lie group

5.2.1 Continuous Lagrangian picture

We consider the optimal control of a mechanical system on a finite dimensional Lie group with Lagrangian that is left invariant under group actions. 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**.

The standard way to solve such problems is to first derive the continuous equations of motion of the system. Among the trajectories satisfying these equations one can find extremal (cost function minimizing) by solving a variational problem.

Let the configuration space be a n-dimensional Lie group G with Lie algebra $\mathfrak{g}$ and a Lagrangian function $L : TG \to \mathbb{R}$ which is left invariant under the action of G. Using this invariance we can left-trivialize such systems (i.e., we can consider the following isomorphism $TG \simeq G \times \mathfrak{g}$) by introducing the **body fixed** velocity $\xi \in \mathfrak{g}$ defined by the left-translation to the origin $\xi = T_eL_{g^{-1}}\dot{g}$ (T_eL is the left translation to the origin, not to be confused with the Lagrangian function L). This last expression will be called henceforth the **reconstruction equation**:

$$\dot{g} = T_eL_g\xi = g\xi. \tag{5.22}$$

The reduced Lagrangian $l : TG/G \simeq \mathfrak{g} \to \mathbb{R}$ is:

$$l(\xi) = L(g^{-1}g, g^{-1}\dot{g}) = L(e, \xi), \tag{5.23}$$

where the invariance under G and (5.22) have been used.

In addition, the control forces are defined by the map $f : \mathfrak{g} \times U \to \mathfrak{g}^*$. As in §5.1.1, the **control manifold** U is a subspace of $\mathbb{R}^m$, with $m \leq n$. Since the motion of the system is influenced by control forces, the dynamics will be governed by the **Lagrange-d'Alembert principle**, which in this case looks like:

$$\delta \int_0^T l(\xi(t))\mathrm{d}t + \int_0^T \langle f(\xi(t), u(t)), \eta(t)\rangle \mathrm{d}t = 0, \tag{5.24}$$

for all variations $\delta\xi(t)$ of the form $\delta\xi(t) = \dot{\eta}(t) + [\xi(t), \eta(t)]$, where $\eta(t) = g^{-1}(t)\delta g(t)$ is an arbitrary curve on the Lie algebra $\mathfrak{g}^*$ with $\eta(0) = \eta(T) = 0$ (see [123] for further details), The derivation of $\delta\xi$ is straightforward by taking usual variations in (5.22), moreover $[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ is the usual Lie bracket of $\mathfrak{g}$. Finally, $\langle\cdot,\cdot\rangle$ represents the usual pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$. Taking variations in (5.24) we arrive to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial l}{\partial \xi}\right)=\mathrm{ad}^*_\xi\left(\frac{\partial l}{\partial \xi}\right)+f. \tag{5.25}$$

These equations are called the **forced Euler-Poincaré** equations (see [66] for further details). $\mathrm{ad}^*_\xi\lambda$ is defined by $\langle \mathrm{ad}^*_\xi\lambda,\eta\rangle=\langle\lambda,\mathrm{ad}_\xi\eta\rangle$, where $\lambda\in\mathfrak{g}^*$ and $\mathrm{ad}_\xi\eta=[\xi,\eta]$, $\eta\in\mathfrak{g}$. Together with the reconstruction equation (5.22), the Euler-Poincaré equations provide the continuous trajectory $(g(t),\xi(t))$ of a system evolving on a Lie group under the action of control forces. If we take the Legendre transform $\mu=\frac{\partial l}{\partial\xi}$, $\mu\in\mathfrak{g}^*$, the Euler-Poincaré equations become

$$\begin{aligned}\dot\mu &= \mathrm{ad}^*_\xi\mu+f,\\ \mu &= \frac{\partial l}{\partial \xi},\\ \dot g &= g\xi,\end{aligned}\tag{5.26}$$

which are called the **Lie-Poisson equations**.

The optimal control problem can be defined as follows: during the time interval $I=[0,T]$ the mechanical system described by l moves on a curve $g(t)\in G$ from an initial state $(g(0),\xi(0))$ to a final state $(g(T),\xi(T))$. The motion is influenced via the external foces f with chosen control parameter $u(t)$ such that a given **cost functional**

$$\mathcal{J}(\xi,u)=\int_0^T C(\xi(t),u(t))\mathrm{d}t, \tag{5.27}$$

where $C:\mathfrak{g}\times U\to\mathbb{R}$ is the **cost function** (continuously differentiable).

As in §5.1.1, the initial state is fixed by $(g(0),\xi(0))=(g^0,\xi^0)$, for $(g^0,\xi^0)\in G\times\mathfrak{g}$ a input value for the initial state, while the final condition is given by the time constraint $r(g(T),\xi(T),g^T,\xi^T)=0$, with $r:G\times\mathfrak{g}\times G\times\mathfrak{g}\to\mathbb{R}^{n_r}$, where $(g^T,\xi^T)\in G\times\mathfrak{g}$ is a fixed value for the desired final state.

Equivalently to the optimal control problem on TQ defined in (5.3), the optimal control problem for a Lagrangian system evolving on a Lie group can be defined as follows:

Problem 5.2.1 (Lagrangian optimal control problem).

$$\min_{(\xi(\cdot),u(\cdot),T)}\int_0^T C(\xi(t),u(t))\mathrm{d}t, \tag{5.28a}$$

subject to

$$\tag{5.28b}$$

$$\delta\int_0^T l(\xi(t))\mathrm{d}t+\int_0^T\langle f(\xi(t),u(t)),\eta(t)\rangle\mathrm{d}t=0, \tag{5.28c}$$

$$\dot g=g\xi, \tag{5.28d}$$

$$g(0)=g^0,\ \ \xi(0)=\xi^0, \tag{5.28e}$$

$$r(g(T),\xi(T),g^T,\xi^T)=0. \tag{5.28f}$$

The final time T may either be fixed or appear as a degree of freedom in the optimization problem.

5.2.2 Discrete Lagrangian picture

Departing from a continuous picture where the Lagrangian is defined in the tangent bundle of a Lie group, i.e. $L : TG \to \mathbb{R}$, the prescription shown in §3.1 enforces a discrete Lagrangian $L_d : G \times G \to \mathbb{R}$. We assume L_d invariant in the following sense:

$$L_d(g_k, g_{k+1}) = L_d(\bar{g}g_k, \bar{g}g_{k+1})$$

for any element $\bar{g} \in G$ and $(g_k, g_{k+1}) \in G \times G$. According to this, we can define a reduced discrete Lagrangian $l_d : G \to \mathbb{R}$ by

$$l_d(W_k) = L_d(e, g_k^{-1}g_{k+1}),$$

just by choosing $\bar{g} = g^{-1}$. In the last e is the identity of the Lie group G. On the other hand,

$$W_k = g_k^{-1}g_{k+1} \tag{5.29}$$

is the discrete counterpart of (5.22) and is called the **discrete reconstruction equation**. By taking usual variations in (5.29) we arrive to

$$\delta W_k = -\eta_k W_k + W_k \eta_{k+1}, \tag{5.30}$$

where $\mathfrak{g} \ni \eta_k = g_k^{-1}\delta g_k$ is an arbitrary element of the algebra.

Put in another way, we can construct the reduced discrete Lagrangian $l_d : G \to \mathbb{R}$ as an approximation of the action integral, that is

$$l_d(W_k) \approx \int_{kh}^{(k+1)h} l(\xi(t))\, \mathrm{d}t.$$

when we deal with a reduced continuous problem $l : \mathfrak{g} \to \mathbb{R}$ as shown in (5.23). Thus, we are replacing the Lie algebra $\mathfrak{g}$ by the Lie group G and the continuous curves $\xi(t)$ by by sequences $(W_0, W_1, ..., W_{N-1}) \in G^N$ (since the Lie algebra is the infinitesimal version of a Lie group, its proper discretization is consequently that Lie group [122, 124]) that will reproduce the discrete trajectory $\{g_k\}_{k=0}^N$ by means of the discrete reconstruction equation (5.29).

Let define the discrete external forces in the following way: $f_k^{\pm} : G \times U \to \mathfrak{g}^*$, where $U \subset \mathbb{R}^m$ for $m \leq n = \dim \mathfrak{g}$. As in §5.1.3 and sections therein, we are going to allow two different sequences of discrete controls $\{u_k^+\}_{k=0}^N$ and $\{u_k^-\}_{k=0}^N$ in order to describe discrete jumps of the controls in the time nodes t_k. In consequence, we are going to use the following quadrature rule:

$$\langle f_k^-(W_k, u_k^-)\,,\, \eta_k\rangle + \langle f_k^+(W_k, u_k^+)\,,\, \eta_{k+1}\rangle \approx \int_{kh}^{(k+1)h} \langle f(\xi(t), u(t)), \eta(t)\rangle\, \mathrm{d}t,$$

where $(f_k^-(W_k, u_k^-), f_k^+(W_k, u_k^+)) \in \mathfrak{g}^* \times \mathfrak{g}^*$ and $\eta_k \in \mathfrak{g}$, for all k. In addition $\eta_0 = \eta_N = 0$. For sake of simplicity we might omit the dependence on $G \times U$ of both f_k^+ and f_k^- in the following.

Taking all the previous into account, we derive a **discrete version of the Lagrange-d'Alembert principle for Lie groups**:

$$\delta\sum_{k=0}^{N-1} l_d(W_k)+\sum_{k=0}^{N-1}\left(\langle f_k^-,\eta_k\rangle+\langle f_k^+,\eta_{k+1}\rangle\right)=0, \tag{5.31}$$

for all variations $\{\delta W_k\}_{k=0}^{N-1}$ verifying the relation (5.30). $\{\eta_k\}_{k=0}^{N-1}$ is an arbitrary sequence of elements of $\mathfrak{g}$ which satisfies $\eta_0=\eta_N=0$ (see [17, 18, 121] for further details).

From this principle is easy to derive the system of difference equations:

$$L^*_{W_{k-1}}\mathrm{d}l_d(W_{k-1})-R^*_{W_k}\mathrm{d}l_d(W_k)+f^+_{k-1}(W_{k-1},u^+_{k-1})+f^-_k(W_k,u^-_k)=0, \tag{5.32}$$

$k=1,\dots,N-1$, which are called the **controlled discrete Euler-Poincaré equations**. In (5.32), $L:G\times G\to G$ is the left-translation of the Lie group G (which shall not be confused with the continuous Lagrangian function L) and $R:G\times G\to G$ is the right-translation. Clearly, the controlled discrete Euler-Poincaré equations are the discrete counterpart of (5.25).

5.2.3 Discrete optimal control problem

We will relate now the discrete changes in the group configuration (5.29) to elements in the Lie algebra ξ_k by means of a retraction map.

A **retraction map** $\tau:\mathfrak{g}\to G$ is an analytic local diffeomorphism around the identity such that $\tau(\xi)\tau(-\xi)=e$, where $\xi\in\mathfrak{g}$. Two standard choices for $\tau-$ the exponential map and the Cayley map-are employed in this chapter.

The variational principle will now be expressed in terms of the chosen map τ. The resulting discrete mechanics will thus involve the derivatives of the map which we define next (see also [22, 73, 91]):

Definition 5.2.2. *Given a map $\tau:\mathfrak{g}\to G$, its* **right trivialized tangent** $\mathrm{d}\tau_\xi:\mathfrak{g}\to\mathfrak{g}$ *and is* **inverse** $\mathrm{d}\tau_\xi^{-1}:\mathfrak{g}\to\mathfrak{g}$*, are such that for $g=\tau(\xi)\in G$ and $\eta\in\mathfrak{g}$, the following holds*

$$\begin{aligned}\partial_\xi\tau(\xi)\,\eta&=\mathrm{d}\tau_\xi\,\eta\,\tau(\xi),\\ \partial_\xi\tau^{-1}(g)\,\eta&=\mathrm{d}\tau_\xi^{-1}(\eta\,\tau(-\xi)).\end{aligned}$$

Using these definitions, variations $\delta\xi$ and δg are constrained by

$$\delta\xi_k=\mathrm{d}\tau^{-1}_{h\xi_k}(-\eta_k+\mathrm{Ad}_{\tau(h\xi_k)}\eta_{k+1})/h, \tag{5.33}$$

where $\eta_k=g_k^{-1}\delta g_k$, which is obtained by straightforward differentiation of $\xi_k=\tau^{-1}(g_k^{-1}g_{k+1})/h$.

Regarding ξ_k as the velocity along the segment between g_k and g_{k+1}, we set the discrete Lagrangian $l_d:G\to\mathbb{R}$ to be

$$l_d(W_k)=hl(\xi_k)=\tilde{l}_d(\xi_k), \tag{5.34}$$

where $\xi_k = \tau^{-1}(g_k^{-1}g_{k+1})/h = \tau^{-1}(W_k)/h$. 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 in order to enable unconstrained optimization in the linear space $\mathfrak{g}$ for optimal control purposes. Note that the new Lagrangian is well-defined only on $\mathfrak{U}$, where $\mathfrak{U} \subset \mathfrak{g}$ is an open neighborhood around the identity for which τ is a diffeomorphism. To make the notation as simple as possible we retain the Lagrangian definition to the full space $\tilde{l}_d : \mathfrak{g} \to \mathbb{R}$.

With these new elements, the discrete version of the Lagrange-d'Alembert principle (5.31) can be rewritten as

$$\delta \sum_{k=0}^{N-1} h l(\xi_k) + \sum_{k=0}^{N-1} \left(\langle f_k^-, \eta_k \rangle + \langle f_k^+, \eta_{k+1} \rangle \right) = 0. \tag{5.35}$$

Taking variations in the last equation, taking into account the definition (5.2.2) for the right-trivialized (and inverse) retraction map, equation (5.33), lemmae defined in Appendix A and the definition of the discrete conjugate momenta

$$\mu_k := (\mathrm{d}\tau^{-1}_{h\,\xi_k})^* \frac{\partial l(\xi_k)}{\partial \xi_k}, \tag{5.36}$$

we arrive to the equations

$$\begin{aligned} &\mu_k - \mathrm{Ad}^*_{\tau(h\xi_{k-1})}\mu_{k-1} = \tilde{f}_k^-(\xi_k, u_k^-) + \tilde{f}_{k-1}^+(\xi_{k-1}, u_{k-1}^+), \\ &k = 1, .., N-1, \\ \\ &\mu_k = (\mathrm{d}\tau^{-1}_{h\,\xi_k})^* \partial_\xi l(\xi_k), \;\; k = 0, ..., N-1, \\ &g_k^{-1} g_{k+1} = \tau(h\xi_k), \qquad k = 0, ..., N-1, \end{aligned} \tag{5.37}$$

where $\tilde{f}_k^\pm$ and $f_k^\pm$ are related by the retraction map. Equations (5.37) are the discrete counterpart of the Lie-Poisson equations (5.26).

In addition, using arguments similar to those given in [124] (forced discrete mechanics section), the discrete forced Noether theorem yields the following boundary conditions

$$\mu_0 - \partial_\xi l(\xi(0)) = \tilde{f}_0^-(\xi_0, u_0^-), \tag{5.38a}$$

$$\partial_\xi l(\xi(T)) - \mathrm{Ad}^*_{\tau(h\xi_{N-1})}\mu_{N-1} = \tilde{f}_N^+(\xi_N, u_N^+), \tag{5.38b}$$

where $\xi(0)$ and $\xi(T)$ are the (given) initial and final velocities. These boundary conditions account for what happens in the initial and final points. Note the distinction between ξ_0 and $\xi(0)$ on one hand, and between ξ_N and $\xi(T)$ on the other. These quantities are not necessary the same since $\xi(t)$ refers to the point on the continuous curve at time t while ξ_k can be thought of as an average velocity along the k-th trajectory segment resulting from the discretization.

The exact form of (5.37) and (5.38) depends on the choice of τ . It is important to point out that this choice will influence the computational efficiency of the optimization framework when the equalities above are enforced as constraints. As mentioned above, there are several choices commonly used for integration on Lie groups. Among them, the more usual are the exponential map and the Cayley map:

a) The exponential map $\exp : \mathfrak{g} \to G$, defined by $\exp(\xi) = \gamma(1)$, with $\gamma : \mathbb{R} \to G$ in the integral curve through the identity of the vector field associated with $\xi \in \mathfrak{g}$ (hence, with $\dot{\gamma}(0) = \xi$). The right trivialized derivative and its inverse are defined by

$$\begin{aligned} \mathrm{dexp}_x\, y &= \sum_{j=0}^{\infty} \frac{1}{(j+1)!}\,\mathrm{ad}_x^j\, y, \\ \mathrm{dexp}_x^{-1}\, y &= \sum_{j=0}^{\infty} \frac{B_j}{j!}\,\mathrm{ad}_x^j\, y, \end{aligned}$$

where B_j are the Bernoulli numbers (see [61]). Typically, these expressions are truncated in order to achieve a desired order of accuracy.

b) 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 (see [61]) that include the groups of interest in this chapter (e.g. $SO(3)$, $SE(2)$ and $SE(3)$. The exact form of the Cayley map for these groups is given in Appendix B). Its right trivialized derivative and inverse are defined by

$$\begin{aligned} \mathrm{dcay}_x\, y &= (e - \frac{x}{2})^{-1}\, y\, (e + \frac{x}{2})^{-1}, \\ \mathrm{dcay}_x^{-1}\, y &= (e - \frac{x}{2})\, y\, (e + \frac{x}{2}). \end{aligned}$$

To end up with the discrete optimization problem, we have to take an approximation the cost functional (5.64):

$$C_d(u_k^-, W_k, u_k^+) \approx \int_{kh}^{(k+1)h} C(\xi(t), u(t))\,\mathrm{d}t, \tag{5.39}$$

yielding the **discrete cost functional**:

$$\mathfrak{J}_d(W_{0:N-1}, u_{0:N-1}^{\pm}) = \sum_{k=0}^{N-1} C_d(u_k^-, W_k, u_k^+)\,. \tag{5.40}$$

Observe that now $C_d : U \times G \times U \to \mathbb{R}$. By means of the retraction map we can define also a discrete cost function defined in the space $U \times \mathfrak{g} \times U$, that is:

$$\tilde{\mathfrak{J}}_d(\xi_{0:N-1}, u_{0:N-1}^{\pm}) = \sum_{k=0}^{N-1} \tilde{C}_d(u_k^-, \xi_k, u_k^+)\,. \tag{5.41}$$

To be precise, the cost function is well-defined only on $U \times \mathfrak{U} \times U$, where $\mathfrak{U} \subset \mathfrak{g}$ is an open neighborhood around the identity for which τ is a diffeomorphism. To make the notation as simple as possible we retain the cost function definition to the full space $U \times \mathfrak{g} \times U$.

Discrete optimal control problem

The discrete optimal control problem for a system with reduced Lagrangian $l : \mathfrak{g} \to \mathbb{R}$ and fixed initial and final states $(g(0), \xi(0))$ and $(g(T), \xi(T))$ respectively can be directly formulated as

Problem 5.2.3 (Discrete optimal control problem).

$$\min_{\xi_{0:N-1}, u^{\pm}_{0:N-1}} \sum_{k=0}^{N-1} \tilde{C}_d(u_k^-, \xi_k, u_k^+), \tag{5.42a}$$

subject to

$$\mu_0 - \partial_\xi l(\xi(0)) = \tilde{f}_0^-(\xi_0, u_0^-), \tag{5.42b}$$
$$\mu_k - \mathrm{Ad}^*_{\tau(h\xi_{k-1})}\mu_{k-1} = \tilde{f}_k^-(\xi_k, u_k^-) + \tilde{f}_{k-1}^+(\xi_{k-1}, u_{k-1}^+), \quad k = 1, ..., N-1, \tag{5.42c}$$
$$\partial_\xi l(\xi(T)) - \mathrm{Ad}^*_{\tau(h\xi_{N-1})}\mu_{N-1} = \tilde{f}_N^+(\xi_N, u_N^+), \tag{5.42d}$$
$$\mu_k = (\mathrm{d}\tau^{-1}_{h\,\xi_k})^*\partial_\xi l(\xi_k), \quad k = 0, ..., N-1, \tag{5.42e}$$
$$g_0 = g(0), \tag{5.42f}$$
$$g_{k+1} = g_k \tau(h\xi_k), \quad k = 0, ..., N-1, \tag{5.42g}$$
$$\tau^{-1}(g_N^{-1} g(T)) = 0, \tag{5.42h}$$

where (5.42h) enforces the condition $g(T) = g_N$. Is clear that (5.42c) are the discrete Lie-Poisson equations coming from the discrete Lagrange-d'Alembert principle (5.35), which in problem 5.2.3 are enforced as constraints. The variables denoted ξ_N and μ_N have no effect on the trajectory $g_{0:N}$ so we can treat these last points as irrelevant to the optimization. This is coherent with thinking of the velocities ξ_k as the average body-fixed velocity along the k-th path segment between configurations g_k and g_{k+1}.

As we did in the case of tangent bundles, in the sequel we develop our new method to approach prob.5.2.3, which consists in solving the discrete optimal control problem as a variational integrator of a specially constructed higher-dimensional system

5.2.4 Fully-actuated case

As in §5.1.4, we perform the fully actuation by means of the following definition:

Definition 5.2.4. (Fully-actuated discrete system) *We say that the discrete mechanical control system is fully-actuated if the mappings*

$$f_k^-|_{W_k} : U \to \mathfrak{g}^*, \qquad f_k^-|_{W_k}(u_k^-) = f_k^-(W_k, u_k^-),$$
$$f_k^+|_{W_k} : U \to \mathfrak{g}^*, \qquad f_k^+|_{W_k}(u_k^+) = f_k^+(W_k, u_k^+),$$

are both diffeomorphisms.

Therefore, we can construct the Lagrangian

$$\mathcal{L}_d : \mathfrak{g}^* \times G \times \mathfrak{g}^* \longrightarrow \mathbb{R}$$

by

$$\begin{aligned}&\mathcal{L}_d(\nu_k, W_k, \nu_{k+1}) = \qquad (5.43)\\ &= C_d\left((f_k^-|_{W_k})^{-1}(R^*_{W_k}\mathrm{d}l_d(W_k) - \nu_k), W_k, (f_k^+|_{W_k})^{-1}(-L^*_{W_k}\mathrm{d}l_d(W_k) + \nu_{k+1})\right),\end{aligned}$$

where we have the equivalence

$$\begin{aligned}\nu_k &= R^*_{W_k}\mathrm{d}l_d(W_k) - f_k^-(W_k, u_k^-),\\ \nu_{k+1} &= L^*_{W_k}\mathrm{d}l_d(W_k) + f_k^+(W_k, u_k^+).\end{aligned} \qquad (5.44)$$

Clearly ν_k and ν_{k+1} belong to $\mathfrak{g}^*$. Observe now that our discrete phase space is $\mathfrak{g}^* \times G \times \mathfrak{g}^*$ which is a space not considered in the previous cases, i.e. two copies of $\mathfrak{g}^*$ and a Lie group G, but in some sense a mixture of both: a Lie groupoid ([118]).

We realize that the discrete optimal control problem defined in (5.31) and (5.39) has been reduced to a Lagrangian one, with Lagrangian function $\mathcal{L}_d : \mathfrak{g}^* \times G \times \mathfrak{g}^* \to \mathbb{R}$, as shown just above. In consequence, we are able to apply discrete variational calculus to obtain the discrete equations of motion of the variables in the phase space $\mathfrak{g}^* \times G \times \mathfrak{g}^*$.

Let us show how to derive these equations from a variational point of view (see [118] for further details). Define first the discrete action sum

$$\mathcal{S}_d = \sum_{k=0}^{N-1} \mathcal{L}_d(\nu_k, W_k, \nu_{k+1}).$$

Consider sequences of the type $\{(\nu_k, W_k, \nu_{k+1})\}_{k=0,\ldots,N-1}$ with boundary conditions: ν_0, ν_N and the composition $\bar{W} = W_0\, W_1 \cdots W_{N-2}\, W_{N-1}$ fixed. Therefore an arbitrary variation of this sequence has the form

$$\{\nu_k(\epsilon)\,,\, h_k^{-1}(\epsilon)\, W_k\, h_{k+1}(\epsilon)\,,\, \nu_{k+1}(\epsilon)\}_{k=0,\ldots,N-1},$$

with $\epsilon \in (-\delta, \delta) \in \mathbb{R}$ (both ϵ and $\delta > 0$ are real parameters) and $\nu_0(\epsilon) = \nu_0$, $\nu_k(0) = \nu_k$, $\nu_N(\epsilon) = \nu_N$, $h_k(\epsilon) \in G$ and $h_0(\epsilon) = h_N(\epsilon) = e$, for all ϵ. Additionally $h_k(0) = e$ for all k.

The critical points of the discrete action sum subjected to the previous boundary conditions are characterized by

$$\begin{aligned}0 &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0}\left(\sum_{k=0}^{N-1}\mathcal{L}_d(\nu_k(\epsilon)\,,\, h_k^{-1}(\epsilon)\, W_k\, h_{k+1}(\epsilon)\,,\, \nu_{k+1}(\epsilon))\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0}\big\{\mathcal{L}_d(\nu_0\,,\, W_0\, h_1(\epsilon),\, \nu_1(\epsilon)) + \mathcal{L}_d(\nu_1(\epsilon)\,,\, h_1^{-1}(\epsilon)\, W_1\, h_2(\epsilon)\,,\, \nu_2(\epsilon))\\ &\quad + \ldots + \mathcal{L}_d(\nu_{N-2}(\epsilon)\,,\, h_{N-2}^{-1}(\epsilon)\, W_{N-2}\, h_{N-1}(\epsilon)\,,\, \nu_{N-1}(\epsilon))\\ &\quad + \mathcal{L}_d(\nu_{N-1}(\epsilon)\,,\, h_{N-1}^{-1}(\epsilon)\, W_{N-1}\,,\, \nu_N)\big\}\,.\end{aligned}$$

Taking derivatives we obtain

$$\begin{aligned} 0 &= \sum_{k=1}^{N-1}\left[L^*_{W_{k-1}}\mathrm{d}\mathcal{L}_d\big|_{(\nu_{k-1},\nu_k)}(W_{k-1}) - R^*_{W_k}\mathrm{d}\mathcal{L}_d\big|_{(\nu_k,\nu_{k+1})}(W_k)\right]\delta h_k \\ &\quad + \sum_{k=1}^{N-1}\left[D_2\mathcal{L}_d\big|_{(W_{k-1})}(\nu_{k-1},\nu_k) + D_1\mathcal{L}_d\big|_{(W_k)}(\nu_k,\nu_{k+1})\right]\delta\nu_k, \end{aligned}$$

where $\mathcal{L}_d\big|_{(W)} : \mathfrak{g}^* \times \mathfrak{g}^* \to \mathbb{R}$ and $\mathcal{L}_d\big|_{(\nu,\nu')} : G \to \mathbb{R}$ are defined by $\mathcal{L}_d\big|_{(W)}(\nu,\nu') = \mathcal{L}_d\big|_{(\nu,\nu')}(W) = \mathcal{L}_d(\nu, W, \nu')$, where $W \in G$ and $\nu, \nu' \in \mathfrak{g}^*$. Since δh_k (which is defined as $\frac{d h_k}{d\epsilon}|_{\epsilon=0}$) and $\delta\nu_k$ (which is defined as $\frac{d\nu_k}{d\epsilon}|_{\epsilon=0}$), $k = 1, \ldots, N-1$ are arbitrary, we deduce the following discrete equations of motion:

$$\begin{aligned} L^*_{W_{k-1}}\mathrm{d}\mathcal{L}_d\big|_{(\nu_{k-1},\nu_k)}(W_{k-1}) - R^*_{W_k}\mathrm{d}\mathcal{L}_d\big|_{(\nu_k,\nu_{k+1})}(W_k) &= 0, \\ D_2\mathcal{L}_d\big|_{(W_{k-1})}(\nu_{k-1},\nu_k) + D_1\mathcal{L}_d\big|_{(W_k)}(\nu_k,\nu_{k+1}) &= 0, \end{aligned} \tag{5.45}$$

for $k = 1, \ldots, N-1$. Similarly to §5.1.4 we obtain the control inputs u_k^- and u_k^+ using (5.44).

5.2.5 Under-actuated case

The under-actuated case can now be considered by adding constraints. Similarly to §5.1.5 underactuation restricts the control forces to lie in a subspace spanned by vectors $\{e^s\}$ of the basis $\{e^s, e^\sigma\}$ of $\mathfrak{g}^*$, where $\{s, \sigma\} = 1, ..., n$. Then

$$\begin{aligned} f_k^-(W_k, u_k^-) &= a_k^-(W_k) + (b_k^-(W_k, u_k^-))_s e^s, \\ f_k^+(W_k, u_k^+) &= a_k^+(W_k) + (b_k^+(W_k, u_k^+))_s e^s, \end{aligned}$$

where $a_k^-(W_k), a_k^+(W_k) \in \mathfrak{g}^*$ and $(b_k^-(W_k, u_k^-))_s, (b_k^+(W_k, u_k^+))_s \in \mathbb{R}$, for all s. Additionally, the embedding condition implies that $\operatorname{rank} b_k^- = \operatorname{rank} b_k^+ = \dim U$. Then, taking the dual basis $\{e_s, e_\sigma\}$, we induce the following constraints:

$$\Phi_\sigma^-(\nu_k, W_k, \nu_{k+1}) = \langle R^*_{W_k}\mathrm{d}l_d(W_k) - \nu_k - a_k^-(W_k), e_\sigma\rangle = 0, \tag{5.46a}$$

$$\Phi_\sigma^+(\nu_k, W_k, \nu_{k+1}) = \langle \nu_{k+1} - L^*_{W_k}\mathrm{d}l_d(W_k) - a_k^+(W_k), e_\sigma\rangle = 0. \tag{5.46b}$$

Observe in (5.46) that, even though the constraints are functions $\Phi_\sigma^\pm : \mathfrak{g}^* \times G \times \mathfrak{g}^* \to \mathbb{R}$, neither Φ_σ^- depends on ν_{k+1} nor Φ_σ^+ on ν_k. Once we have defined the constraints we can implement the Lagrangian multiplier rule in order to solve the underactuated problem. Namely, we define the extended Lagrangian as:

$$\begin{aligned} \tilde{\mathcal{L}}_d(\nu_k, \lambda_k^-, W_k, \nu_{k+1}, \lambda_k^+) &= \mathcal{L}_d(\nu_k, W_k, \nu_{k+1}) + \\ &+ (\lambda_k^-)^\sigma \Phi_\sigma^-(\nu_k, W_k, \nu_{k+1}) + (\lambda_k^+)^\sigma \Phi_\sigma^+(\nu_k, W_k, \nu_{k+1}). \end{aligned}$$

Defining the discrete action sum

$$\mathcal{S}_d^{\text{under}} = \sum_{k=0}^{N-1} \tilde{\mathcal{L}}_d(\nu_k, \lambda_k^-, W_k, \nu_{k+1}, \lambda_k^+),$$

we obtain the underactuated discrete equations of motion

$$
\begin{aligned}
& L^*_{W_{k-1}}\, \mathrm{d}\mathcal{L}_d\big|_{(\nu_{k-1},\nu_k)}(W_{k-1}) - R^*_{W_{k-1}}\, \mathrm{d}\mathcal{L}_d\big|_{(\nu_k,\nu_{k+1})}(W_k) \\
& + L^*_{W_{k-1}} \left((\lambda^-_{k-1})^\sigma\, \mathrm{d}\,\Phi^-_\sigma\big|_{(\nu_{k-1},\nu_k)}(W_{k-1}) + (\lambda^+_{k-1})^\sigma\, \mathrm{d}\,\Phi^+_\sigma\big|_{(\nu_{k-1},\nu_k)}(W_{k-1}) \right) \\
& - R^*_{W_{k-1}} \left((\lambda^-_k)^\sigma\, \mathrm{d}\,\Phi^-_\sigma\big|_{(\nu_k,\nu_{k+1})}(W_k) + (\lambda^+_k)^\sigma\, \mathrm{d}\,\Phi^+_\sigma\big|_{(\nu_k,\nu_{k+1})}(W_k) \right) = 0, \\
& D_2\, \mathcal{L}_d\big|_{(W_{k-1})}(\nu_{k-1},\nu_k) + D_1\, \mathcal{L}_d\big|_{(W_k)}(\nu_k,\nu_{k+1}) + \left[(\lambda^+_{k-1})^\sigma - (\lambda^-_k)^\sigma \right] e_\sigma = 0, \\
& \Phi^-_\sigma(\nu_k, W_k, \nu_{k+1}) = 0, \\
& \Phi^+_\sigma(\nu_k, W_k, \nu_{k+1}) = 0,
\end{aligned}
\tag{5.47}
$$

where the subscripts (W_{k-1}), (W_k), (ν_{k-1},ν_k), (ν_k,ν_{k+1}) denoted variables that are fixed.

5.2.6 Numerical methods for systems on Lie groups

We now put the discrete optimal control equations (5.45) and (5.47) into a form suitable for algorithmic implementation. The numerical methods are construsted using the following guidelines:

1. good approximation of the dynamics and optimality,
2. globally valid parametrization,
3. guarantee for numerical robustness and convergence,
4. numerical efficiency.

The discrete mechanics approach provides an accurate approximation of the dynamics (requirement 1) through momentum and symplectic form preservation and good energy behavior. In addition, we will satisfy requirement 2 for systems on Lie groups by lifting the optimization to the Lie algebra through a retraction map that will be defined in this section. The resulting algorithms are numerically robust in the sense that there are no issues with coordinate singularities and the dynamics and optimality conditions remain close to their continuous counterparts even at big time steps. Yet, as with any other nonlinear optimization scheme it is difficult to formally claim that the algorithm will always converge (requirement 3). Nevertheless, in practice there are only isolated cases for underactuated systems that fail to converge. A remedies for such cases has been suggested in [91]. In general, the resulting algorithms require a small number of iterations, e.g. between 10 and 20 to converge (requirement 4).

The optimization variables W_k are regarded as small displacements on the Lie group. Thus, it is possible to express each term through a Lie algebra element that can be though of the averaged velocity of this displacement. This is accomplished using the **retraction maps** defined in §5.2.3 and considering the discrete Lagrangian l_d to be $l_d(W_k) = hl(\xi_k)$.

Next, the discrete forces and cost function are defined through a trapezoidal approximation, i.e.

$$\tilde{f}_k^{\pm}(\xi_k, u_k^{\pm}) = \frac{h}{2}\, f(\xi_k, u_k^{\pm}),$$

and

$$\tilde{C}_d(u_k^-, \xi_k, u_k^+) = \frac{h}{2}\, C(\xi_k, u_k^-) + \frac{h}{2}\, C(\xi_k, u_k^+),$$

respectively. With the choice of a retraction map and the trapezoidal rule the equations of motion (5.37) become

$$\begin{aligned}
&\mu_k - \mathrm{Ad}^*_{\tau(h\xi_{k-1})}\mu_{k-1} = \frac{h}{2}\, f(\xi_k, u_k^-) + \frac{h}{2}\, f(\xi_{k-1}, u_{k-1}^+), && k = 1, ..., N-1,\\
&\mu_k = (\mathrm{d}\tau^{-1}_{h\xi_k})^*\partial_\xi l(\xi_k), && k = 0, ..., N-1,\\
&g_{k+1} = g_k \tau(h\xi_k), && k = 0, ..., N-1,
\end{aligned}$$

while the momenta defined in (5.44) take the form

$$\nu_k = \mu_k - \frac{h}{2}\, f(\xi_k, u_k^-), \tag{5.48}$$

$$\nu_{k+1} = \mathrm{Ad}^*_{\tau(h\xi_k)}\mu_k + \frac{h}{2}\, f(\xi_k, u_k^+). \tag{5.49}$$

Finally, define the Lagrangian $\ell_d : \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \to \mathbb{R}$ such that

$$\ell_d(\nu, \xi, \nu') = \mathcal{L}_d(\nu, \tau(h\xi), \nu').$$

Note that the Lagrangian is well-defined only on $\mathfrak{g}^* \times \mathfrak{U} \times \mathfrak{g}^*$, where $\mathfrak{U} \subset \mathfrak{g}$ is an open neighborhood around the identity for which τ is a diffeomorphism. To make the notation as simple as possible we retain the Lagrangian definition to the full space $\mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^*$.

The optimality conditions corresponding to (5.45) become

$$(\mathrm{d}\tau^{-1}_{-h\xi_{k-1}})^*\,\mathrm{d}\,\ell_d\big|_{(\nu_{k-1},\nu_k)}(\xi_{k-1}) - (\mathrm{d}\tau^{-1}_{h\xi_k})^*\,\mathrm{d}\,\ell_d\big|_{(\nu_k,\nu_{k+1})}(\xi_k) \;=\; 0, \tag{5.50}$$

$$D_2\,\ell_d\big|_{(\xi_{k-1})}(\nu_{k-1}, \nu_k) + D_1\,\ell_d\big|_{(\xi_k)}(\nu_k, \nu_{k+1}) \;=\; 0, \tag{5.51}$$

for $k = 0, ..., N-1$. Here, $\ell_d\big|_{(\xi)}(\nu, \nu') = \ell_d\big|_{(\nu,\nu')}(\xi) = \ell_d(\nu, \xi, \nu')$. Equations (5.50) and (5.51) can be also obtained from (5.45) employing Lemma 6.6.3 and Lemma 6.6.4 in Appendix A.

In the underactuated case we define

$$\begin{aligned}
\tilde{\ell}_d(\nu, \xi, \nu', \lambda^-, \lambda^+) =&\mathcal{L}_d(\nu, \tau(h\xi), \nu')\\
&+ (\lambda^-)^\sigma \Phi_\sigma^-\big|_{(\nu,\nu')}(\tau(h\xi)) + (\lambda^+)^\sigma \Phi_\sigma^+\big|_{(\nu,\nu')}(\tau(h\xi)),
\end{aligned} \tag{5.52}$$

and from (5.47) obtain the equations

$$\begin{aligned}
&(\mathrm{d}\tau^{-1}_{-h\xi_{k-1}})^*\,\mathrm{d}\,\tilde{\ell}_d\big|_{(\nu_{k-1},\nu_k,\lambda^{\pm}_{k-1})}(\xi_{k-1}) - (\mathrm{d}\tau^{-1}_{h\xi_k})^*\,\mathrm{d}\,\tilde{\ell}_d\big|_{(\nu_k,\nu_{k+1},\lambda^{\pm}_k)}(\xi_k) = 0,\\
&D_2\,\mathcal{L}_d\big|_{\tau(h\xi_{k-1})}(\nu_{k-1}, \nu_k) + D_1\,\mathcal{L}_d\big|_{\tau(h\xi_k)}(\nu_k, \nu_{k+1}) + \lambda^+_{k-1} - \lambda^-_k = 0,\\
&\Phi^-_\sigma(\nu_k, \tau(h\xi_k), \nu_{k+1}) = 0,\\
&\Phi^+_\sigma(\nu_k, \tau(h\xi_k), \nu_{k+1}) = 0,
\end{aligned} \tag{5.53}$$

where we employed the notation $\lambda^{\pm} := (\lambda^{\pm})^\sigma e_\sigma$ and used the definitions (5.46).

Boundary conditions

Establishing the exact relationship between the discrete and continuous momenta, μ_k and $\mu(t) = \partial_\xi l(\xi(t))$, respectively, is particularly important for properly enforcing boundary conditions that are given in terms of continuous quantities (see equations (5.38)). The following equations relate the momenta at the initial and final times $t = 0$ and $t = T$ and are used to transform between the continuous and discrete representations:

$$\begin{aligned} \mu_0 - \partial_\xi l(\xi(0)) &= \frac{h}{2} f(\xi(0), u_0^-), \\ \partial_\xi l(\xi(T)) - \mathrm{Ad}^*_{\tau(h\xi_{N-1})} \mu_{N-1} &= \frac{h}{2} f(\xi(T), u_N^+). \end{aligned}$$

which also corresponds to the relations $\nu_0 = \partial_\xi l(\xi(0))$ and $\nu_N = \partial_\xi l(\xi(T))$. These equations can also be regarded as structure-preserving **velocity boundary conditions**, i.e., for given fixed velocities $\xi(0)$ and $\xi(T)$.

The exact form of the previous equations depends on the choice of τ. This choice will also influence the computational efficiency of the optimization framework when the above equalities are enforced as constraints. The numerical procedure to compute the trajectory is summarized as follows:

Algorithm 5.2.5. Optimal control

Data: group G; mechanical Lagrangian l; control functions a, b; cost function C; final time T; number of segments N.

1. *Input: boundary conditions $(g(0), \xi(0))$ and $(g(T), \xi(T))$.*
2. *Set momenta $\nu_0 = \partial_\xi l(\xi(0))$ and $\nu_N = \partial_\xi l(\xi(T))$*
3. *Solve for $(\xi_0, ..., \xi_{N-1}, \nu_1, ..., \nu_{N-1}, \lambda_1^\pm, ..., \lambda_{N-1}^\pm)$ the relations:*
$$\left\{ \begin{array}{l} \text{equations (5.53) for all } k = 1, ..., N-1, \\ \tau^{-1}\left(\tau(h\xi_{N-1})^{-1} ... \tau(h\xi_0)^{-1} g(0)^{-1} g(T)\right) = 0 \end{array} \right.$$
4. *Output: optimal sequence of velocities $\xi_0, ..., \xi_{N-1}$.*
5. *Reconstruct path $g_0, ..., g_N$ by $g_{k+1} = g_k \tau(h\xi_k)$ for $k = 0, ..., N-1$.*

The solution is computed using root-finding procedure such as Newton's method.

5.2.7 Example: optimal control effort

Consider a Lagrangian consisting of the kinetic energy only

$$l(\xi) = \frac{1}{2} \langle \mathbb{I}(\xi), \xi \rangle,$$

full unconstrained actuation, no potential or external forces and no velocity constraint. The map $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$ is called the inertia tensor and is assumed full rank.

In the fully actuated case we have $f(\xi_k, u_k^{\pm}) \equiv u_k^{\pm}$. We consider a minimum effort control problem, i.e.

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

The optimal control problem 5.2.3 for fixed initial and final states $(g(0)\,,\xi(0))$ and $(g(T)\,,\xi(T))$ can now be summarized as:

Compute: $\xi_{0:N-1}\ , u_{0:N}^{\pm}$,

minimizing: $\frac{h}{4}\sum_{k=0}^{N-1}\left(\| u_k^- \|^2 + \| u_k^+ \|^2\right),$

subject to:

$$\begin{aligned}
&\mu_0 - \mathbb{I}(\xi(0)) = \tfrac{h}{2}\, u_0^-, \\
&\mu_k - \mathrm{Ad}^*_{\tau(h\xi_{k-1})}\,\mu_{k-1} = h(u_k^- + u_{k-1}^+), && k = 1, ..., N-1, \\
&\mathbb{I}(\xi(T)) - \mathrm{Ad}^*_{\tau(h\xi_{N-1})}\,\mu_{N-1} = \tfrac{h}{2}\, u_N^+, \\
&\mu_k = (\mathrm{d}\tau^{-1}_{h\,\xi_k})^*\,\mathbb{I}(\xi_k), \\
&g_{k+1} = g_k\,\tau(h\xi_k), && k = 0, ...N-1, \\
&\tau^{-1}(g_N^{-1}\,g(T)) = 0.
\end{aligned}$$

On the other hand, the optimality conditions for this problem are derived as follows in the approach developed in this chapter. The Lagrangian becomes

$$\ell_d(\nu_k, \xi_k, \nu_{k+1}) = \frac{1}{4h}\sum_{k=0}^{N-1}\left(\| \nu_k - (\mathrm{d}\,\tau^{-1}_{h\xi_k})^*\mathbb{I}(\xi_k) \|^2 + \| \nu_{k+1} - (\mathrm{d}\,\tau^{-1}_{-h\xi_k})^*\mathbb{I}(\xi_k) \|^2\right),$$

where the momentum has been computed according to

$$\nu_k = \frac{1}{2}\left((\mathrm{d}\,\tau^{-1}_{h\xi_k})^*\mathbb{I}(\xi_k) + (\mathrm{d}\,\tau^{-1}_{-h\xi_{k-1}})^*\mathbb{I}(\xi_{k-1})\right), \tag{5.54}$$

Thus the optimality conditions become

$$\begin{aligned}
&(\mathrm{d}\,\tau^{-1}_{h\xi_k})^*\,\mathrm{d}\ell_d\big|_{(\nu_k,\nu_{k+1})}(\xi_k) - (\mathrm{d}\,\tau^{-1}_{-h\xi_{k-1}})^*\,\mathrm{d}\ell_d\big|_{(\nu_{k-1},\nu_k)}(\xi_{k-1}) = 0, \\
&k = 1, ..., N-1, \\
\\
&\tau^{-1}\left(\tau(h\xi_{N-1})^{-1}...\tau(h\xi_0)^{-1}\,g_0^{-1}g(T)\right) = 0.
\end{aligned}$$

It is important to note that these last two equations define $N\cdot n$ equations in the $N\dot{n}$ unknowns $\xi_{0:N-1}$. A solution can be found using nonlinear root finding. Once $\xi_{0:N}$ have been computed, is possible to obtain the final configuration g_N by reconstructing the curve by these velocities. Beside, the boundary condition $g(T)$ is enforced through the relation $\tau^{-1}(g_N^{-1}\,g(T)) = 0$ without the need to optimize over any of the configurations g_k.

5.2.8 Extension: the configuration-dependent case

The developed framework can be extended to a configuration-dependent Lagrangian $L : G \times \mathfrak{g} \to \mathbb{R}$, for instance defined in terms of a kinetic energy $K : \mathfrak{g} \to \mathbb{R}$ and potential energy $V : G \to \mathbb{R}$ according to

$$L(g, \xi) = K(\xi) - V(g),$$

where $g \in G$ and $\xi \in \mathfrak{g}$. The controlled Lie-Poisson equations are in this case

$$\begin{aligned}
&\dot{\mu} - \mathrm{ad}^*_{\xi}\mu = -L^*_g\, \partial_g\, V(g) + f,\\
&\mu = \partial_\xi K(\xi),\\
&\dot{g} = g\,\xi,
\end{aligned}$$

(recall that L_g denotes the left-translation by g) where the external forces are defined as $f : G \times \mathfrak{g} \times U \to \mathfrak{g}^*$. Our discretization choice $L_d : G \times G \to \mathbb{R}$ will be (recall that $\xi_k = \tau^{-1}(g_k^{-1} g_{k+1})/h$)

$$\begin{aligned}
L_d(g_k, g_{k+1}) &= \frac{h}{2}L(g_k, \xi_k) + \frac{h}{2}L(g_{k+1}, \xi_k)\\
&= h\,K(\xi_k) - h\,\frac{V(g_k) + V(g_{k+1})}{2},
\end{aligned}$$

while the G-dependent discrete forces now become

$$f_k^-(g_k, \xi_k, u_k^-) = \frac{h}{2}\, f(g_k, \xi_k, u_k^-), \qquad f_k^+(g_{k+1}, \xi_k, u_k^+) = \frac{h}{2}\, f(g_{k+1}, \xi_k, u_k^+).$$

This leads to the discrete equations

$$\begin{aligned}
&\mu_k - \mathrm{Ad}^*_{\tau(h\xi_{k-1})}\mu_{k-1} = -h\, L^*_{g_k}\partial_g V(g_k)\\
&+\frac{h}{2}\, f(g_k, \xi_k, u_k^-) + \frac{h}{2}\, f(g_k, \xi_{k-1}, u_{k-1}^+), \qquad k = 1, ..., N,\\
&\\
&\mu_k = (\mathrm{d}\tau^{-1}_{h\xi_k})^*\partial_\xi K(\xi_k), \qquad k = 0, ..., N-1,\\
&g_{k+1} = g_k \tau(h\xi_k), \qquad k = 0, ..., N-1.
\end{aligned}$$

The momenta become

$$\begin{aligned}
\nu_k &= \mu_k + \frac{h}{2}\, L^*_{g_k}\, \partial_g\, V(g_k) - \frac{h}{2}\, f(g_k, \xi_k, u_k^-),\\
\nu_{k+1} &= \mathrm{Ad}^*_{\tau(h\xi_k)}\mu_k - \frac{h}{2}\, L^*_{g_{k+1}}\, \partial_g\, V(g_{k+1}) + \frac{h}{2}\, f(g_{k+1}, \xi_k, u_k^+).
\end{aligned}$$

In consequence, we can define a discrete Lagrangian

$$\mathfrak{L}_d : \mathfrak{g}^* \times\ G \times \mathfrak{g} \times \mathfrak{g}^* \to \mathbb{R},$$

depending on the variables $(\nu_k, g_k, \xi_k, \nu_{k+1})$ which discrete equations of motion will be a mixture between (5.45) and (5.50), (5.51), namely

$$\begin{aligned}
&D_2\,\mathfrak{L}_d\big|_{(g_{k-1},\xi_{k-1})}(\nu_{k-1},\nu_k) + D_1\,\mathfrak{L}_d\big|_{(g_k,\xi_k)}(\nu_k,\nu_{k+1}) = 0,\\
&\Big(L^*_{g_{k-1}}\,\mathrm{d}\,\mathfrak{L}_d\big|_{(\nu_{k-1},\xi_{k-1},\nu_k)}(g_{k-1}) + R^*_{g_k}\,\mathrm{d}\,\mathfrak{L}_d\big|_{(\nu_k,\xi_k,\nu_{k+1})}(g_k)\Big)\\
&+\Big((\mathrm{d}\tau^{-1}_{-h\xi_{k-1}})^*\,\mathrm{d}\,\mathfrak{L}_d\big|_{(\nu_{k-1},g_{k-1},\nu_k)}(\xi_{k-1}) - (\mathrm{d}\tau^{-1}_{h\xi_k})^*\,\mathrm{d}\,\mathfrak{L}_d\big|_{(\nu_k,g_k,\nu_{k+1})}(\xi_k)\Big) = 0,
\end{aligned}$$

for $k = 1, ..., N-1$.

5.2.9 Application 1: underwater vehicle

We illustrate the developed algorithm with an application to a simulated unmanned underwater vehicle. Figure (5.1) shows the model equipped with five thrusters which produce forces and torques in all directions but the body-fixed “y”-axis. Since the input directions span only a five-dimensional subspace the problem is solved through the underactuated framework.

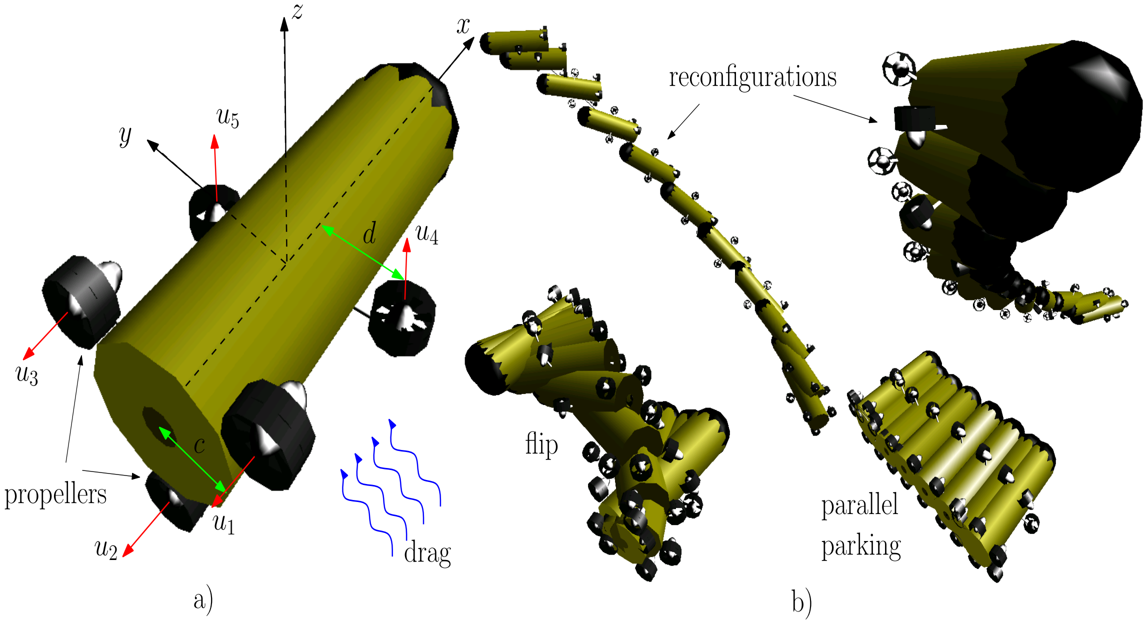


Figure 5.1: An underwater vehicle model (a) and a various computed optimal trajectories between chosen states (b). Only a few frames along the path are shown for clarity.

The vehicle configuration space is $G = SE(3)$. We make the identification $SE(3) \sim SO(3) \times \mathbb{R}^3$ using elements $R \in SO(3)$ and $x \in \mathbb{R}^3$. The exact form of a group element $g \in SE(3)$ and an algebra element $\xi \in \mathfrak{se}(3)$ are described in Appendix B, together with the analytic form of the Cayley map. Elements of the Lie algebra are identified with body-fixed angular and linear velocities denoted $\omega \in \mathbb{R}^3$ and $v \in \mathbb{R}^3$, respectively, through

$$\xi = \begin{pmatrix} \hat{\omega} & v \\ 0_{3\times 3} & 0 \end{pmatrix},$$

where the hat map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined is defined also in Appendix B. The algorithm is thus implemented in terms of vectors in $\mathbb{R}^6$ rather than matrices in $\mathfrak{se}(3)$.

The map $\tau = \mathrm{cay} : \mathfrak{se}(3) \to SE(3)$ is chosen, instead of the exponential, since it results in more computationally efficient implementation. The matrix representation of the right-trivialized tangent inverse $\mathrm{dcay}^{-1}_{(\omega,v)} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3$ becomes

$$[\mathrm{dcay}^{-1}_{(\omega,v)}] = \begin{bmatrix} \mathbf{I}_3 - \frac{1}{2}\widehat{\omega} + \frac{1}{4}\omega\omega^T & \mathbf{0}_3 \\ -\frac{1}{2}\left(\mathbf{I}_3 - \frac{1}{2}\widehat{\omega}\right)\widehat{v} & \mathbf{I}_3 - \frac{1}{2}\widehat{\omega} \end{bmatrix}. \tag{5.55}$$

The vehicle inertia tensor $\mathbb{I}$ is computed assuming cylindrical mass distribution with mass $m = 3$kg. The control basis vectors are $\{e_s\}_{s=1}^5 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\}$, while the non-actuated direction is $e_\sigma = \mathbf{e}_6$, where $\mathbf{e}_i$ is the i-th standard basis vector of $\mathbb{R}^6$. The control functions take the form

$$\begin{aligned}
b(W,u)_1 &= d(u_5 - u_4), \\
b(W,u)_2 &= c((u_1 + u_2)/2 - u_3), \\
b(W,u)_3 &= (c \sin\frac{\pi}{3})(u_2 - u_1), \\
b(W,u)_4 &= u_1 + u_2 + u_3, \\
b(W,u)_5 &= u_4 + u_5, \\
a(W) &= H\tau^{-1}(W),
\end{aligned}$$

here H is a negative definite viscous drag matrix and the constants c, d are the lengths of the thrusting torque moment arms (see Figure 5.1).

We are interested in computing a minimum control effort trajectory between two given boundary states, i.e. conditions on both the configurations and velocities. Such a cost function is defined in §5.2.7. The optimal control problem is solved using equations (5.53). The computation is performed using Algorithm 5.2.5. Figure 5.2 shows the computed velocities and controls for the "reconfiguration" trajectory shown in Figure 5.1. The algorithms requires between 10-20 iterations depending on the boundary conditions and when applied to $N = 32$ segments.

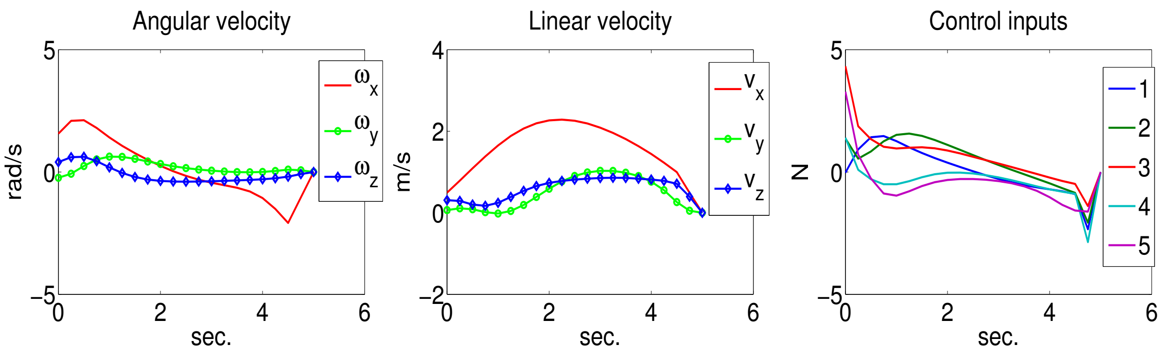
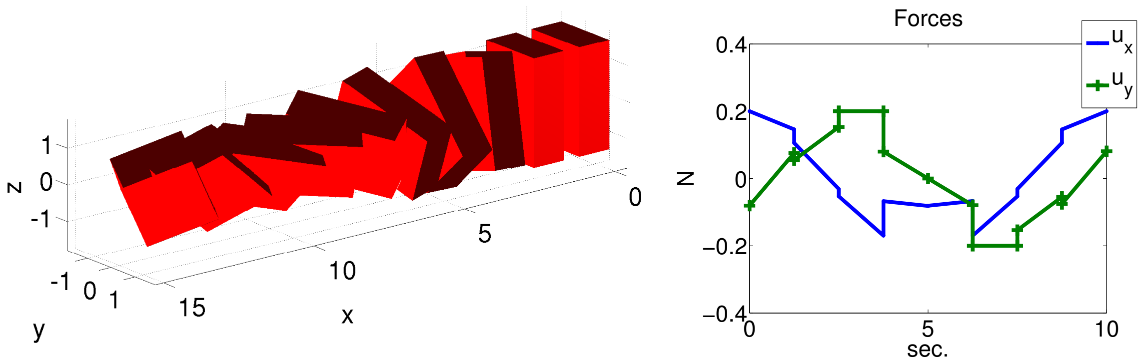


Figure 5.2: Details of the computed optimal path for the reconfiguration maneuver given in Figure (5.1).

5.2.10 Application 2: discontinuous control

One of the advantages of employing the discrete variational framework is the treatment of discontinuous control inputs as illustrated in §5.1. The nature of the control curve depends on the cost function. In the standard squared control effort case (i.e. L_2 control curve norm employed in §5.2.9) the resulting control is smooth. Another cost function of interest is $\int_0^T \|u(t)\|\mathrm{d}t$ (i.e. the L_1 control curve norm) which is typically imposed along with the constraints $u_{min} \leq u(t) \leq u_{max}$. This case results in a discontinuous optimal control curve. Our formulation can handle such problems easily since the terms u_k^- and u_k^+ are regarded as the forces before and after time t_k, respectively. A computed scenario of a rigid body actuated with two control torques around its principles axes of inertia (Fig. 5.3) illustrates the discontinuous case.



a) b)

Figure 5.3: An optimal trajectory of an underactuated rigid body on $SO(3)$ (a). The body is controlled using two force inputs around the body-fixed x and y axes. An L_1-control cost function results in a discontinuous optimal trajectory (b) which our algorithm can handle.

5.3 Extensions

The methods developed in the previous sections are easily adapted to other cases which are of great interest in real applications. In particular, this section will be devoted to the discussion of two important extensions: the case of optimal control problems for Lagrangians of the type $l : TM \times \mathfrak{g} \to \mathbb{R}$ (that is, reduction by symmetries on a trivial principal fiber bundle) and the case of nonholonomic systems. Here, M denotes a smooth manifold. Observe that the phase space $TM \times \mathfrak{g}$ unifies the previously studied cases of a tangent bundle and a Lie algebra.

The notion of principal fiber bundle is present in many locomotion and robotic systems [16, 23, 120]. When the configuration manifold is $Q = M \times G$, there exists a canonical splitting between variables describing the position and variables describing the orientation of the mechanical system. Then, we distinguish the pose coordinates $g \in G$ (the elements in the Lie algebra will be denoted by $\xi \in \mathfrak{g}$), and the variables describing the internal shape of the system, that is $x \in M$ (in consequence $(x, \dot{x}) \in TM$). Observe that the Lagrangians of the type $l : TM \times \mathfrak{g} \to \mathbb{R}$ mainly appears as reduction of Lagrangians of the type $L : T(M \times G) \to \mathbb{R}$, which are invariant under the action of the Lie group G. Under the identification $T(M \times G)/G \equiv TM \times \mathfrak{g}$ we obtain the reduced Lagrangian l. We first develop the discrete optimal control problem for systems in an unconstrained principle bundle setting in §5.3.1. Nonholonomic constraints are then added to treat the more general case of locomotion systems in §5.3.2.

5.3.1 Discrete optimal control on principle bundles

The discrete case is modeled by a Lagrangian $l_d : M \times M \times G \to \mathbb{R}$ which is an approximation of the action integral in one time step

$$l_d(x_k, x_{k+1}, W_k) \simeq \int_{hk}^{h(k+1)} l\left(x(t), \dot{x}(t), \xi(t)\right) \mathrm{d}t,$$

where $(x_k, x_{k+1}) \in M \times M$ and $W_k \in G$. Again, we make an election for the discrete control forces $f_k^{\pm} : M \times M \times G \times U \to T^*M \times \mathfrak{g}^*$, where $U \subset \mathbb{R}^m$:

$$\begin{aligned} f_k^-(x_k, x_{k+1}, W_k, u_k^-) &= \left(\bar{f}_k^-(x_k, x_{k+1}, W_k, u_k^-), \hat{f}_k^-(x_k, x_{k+1}, W_k, u_k^-)\right), \\ f_k^+(x_k, x_{k+1}, W_k, u_k^+) &= \left(\bar{f}_k^+(x_k, x_{k+1}, W_k, u_k^+), \hat{f}_k^+(x_k, x_{k+1}, W_k, u_k^+)\right), \end{aligned}$$

here $f_k^- \in T^*_{x_k}M \times \mathfrak{g}^*$ and $f_k^+ \in T^*_{x_{k+1}}M \times \mathfrak{g}^*$ (more concretely $\bar{f}_k^- \in T^*_{x_k}M$, $\bar{f}_k^+ \in T^*_{x_{k+1}}M$, $\hat{f}_k^- \in \mathfrak{g}^*$, $\hat{f}_k^+ \in \mathfrak{g}^*$).

Similarly to the developments in §5.1 and §5.2.4 we can formulate the **discrete Lagrange-D'Alembert principle**:

$$\begin{aligned} \delta \sum_{k=0}^{N-1} l_d(x_k, x_{k+1}, W_k) \quad &+ \quad \sum_{k=0}^{N-1} \langle f_k^-, (\delta x_k, \eta_k)\rangle \\ &+ \quad \sum_{k=0}^{N-1} \langle f_k^+, (\delta x_{k+1}, \eta_{k+1})\rangle = 0, \end{aligned}$$

which can be rewritten as

$$\begin{aligned} \delta \sum_{k=0}^{N-1} l_d(x_k, x_{k+1}, W_k) \quad &+ \quad \sum_{k=0}^{N-1} \bar{f}_k^- \delta x_k + \sum_{k=0}^{N-1} \bar{f}_k^+ \delta x_{k+1} \\ &+ \quad \sum_{k=0}^{N-1} \langle \hat{f}_k^-, \eta_k\rangle + \sum_{k=0}^{N-1} \langle \hat{f}_k^+, \eta_{k+1}\rangle = 0, \end{aligned}$$

for all variations $\{\delta x_k\}_{k=0}^N$ with $\delta x_k \in T_{x_k}M$ and $\delta x_0 = \delta x_N = 0$; also $\{\delta W_k\}_{k=0}^N$ with $\delta W_k \in T_{g_k}G$, such that $\delta W_k = -\eta_k W_k + W_k \eta_{k+1}$, being $\{\eta_k\}_{k=0}^N$ a sequence of independent elements of $\mathfrak{g}$ such that $\eta_0 = \eta_N = 0$.

Applying variations in the last expression and rearranging the sum, we finally obtain the complete set of **forced discrete Euler-Lagrange equations**:

$$D_1 l_d(x_k, x_{k+1}, W_k) + D_2 l_d(x_{k-1}, x_k, W_{k-1}) + \bar{f}_k^- + \bar{f}_{k-1}^+ = 0, \tag{5.56}$$

$$l^*_{W_{k-1}} D_3 l_d(x_{k-1}, x_k, W_{k-1}) - r^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) + \hat{f}_k^- + \hat{f}_{k-1}^+ = 0, \tag{5.57}$$

with $k = 1, \ldots, N-1$. Since we are dealing with an optimal control problem, we introduce a discrete cost function $C_d : M \times G \times M \times U \times U \to \mathbb{R}$. As in previous cases, our objective is to extremize the following sum

$$\sum_{k=0}^{N-1} C_d(x_k, W_k, x_{k+1}, u_k^-, u_k^+),$$

subjected to equations (5.56) and (5.57). Let us initially restrict our attention to the case of fully actuated systems.

Definition 5.3.1. (Fully actuated discrete system) *We say that the discrete mechanical control system is fully actuated if the mappings*

$$f_k^-\big|_{(x_0,x_1,W_1)} : U \to T^*_{x_0} M \times \mathfrak{g}^*, \quad f_k^-\big|_{(x_0,x_1,W_1)}(u) = f_k^-(x_0, x_1, W_1, u),$$
$$f_k^+\big|_{(x_0,x_1,W_1)} : U \to T^*_{x_1} M \times \mathfrak{g}^*, \quad f_k^+\big|_{(x_0,x_1,W_1)}(u) = f_k^+(x_0, x_1, W_1, u)$$

are both diffeomorphisms.

According to equations (5.56) and (5.57), we can introduce the momenta by means of the following discrete Legendre transforms:

$$p_k = -D_1 l_d(x_k, x_{k+1}, W_k) - \bar{f}_k^-,$$
$$p_{k+1} = D_2 l_d(x_k, x_{k+1}, W_k) + \bar{f}_k^+,$$

$$\mu_k = R^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) - \hat{f}_k^-,$$
$$\mu_{k+1} = L^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) + \hat{f}_k^+,$$

where, we recall, L and R denote the left and right translations, respectively, in the Lie group. In the fully actuated case, is possible to find the value of all control forces in terms of $x_k, x_{k+1}, W_k, p_k, p_{k+1}, \mu_k, \mu_{k+1}$, that is:

$$u_k^- = u_k^-(x_k, x_{k+1}, W_k, p_k, \mu_k), \tag{5.58}$$

$$u_k^+ = u_k^+(x_k, x_{k+1}, W_k, p_{k+1}, \mu_{k+1}). \tag{5.59}$$

Replacing (5.58) and (5.59) into C_d, we finally obtain the discrete Lagrangian that completely describes our system:

$$\mathcal{L}_d : T^*M \times \mathfrak{g}^* \times G \times \mathfrak{g}^* \times T^*M \longrightarrow \mathbb{R}.$$

The associated **discrete cost functional** is

$$\mathcal{J}_d(x_{0:N}, p_{0:N}, \mu_{0:N}, W_{0:N-1}) = \sum_{k=0}^{N-1} \mathcal{L}_d(x_k, p_k, \mu_k, W_k, \mu_{k+1}, x_{k+1}, p_{k+1}). \tag{5.60}$$

As usual, we take now variations in (5.60) in order to obtain the discrete Euler-Lagrange equations for our optimal control problem (with some abuse of notation we denote $\hat{Q}_k = (x_k, p_k, \mu_k, W_k, \mu_{k+1}, x_{k+1}, p_{k+1})$ the whole set of coordinates in the new phase space):

$$\begin{aligned} D_6\mathcal{L}_d(\hat{Q}_{k-1}) \quad &+ \quad D_1\mathcal{L}_d(\hat{Q}_k) = 0\,, \\ D_7\mathcal{L}_d(\hat{Q}_{k-1}) \quad &+ \quad D_2\mathcal{L}_d(\hat{Q}_k) = 0\,, \\ D_5\mathcal{L}_d(\hat{Q}_{k-1}) \quad &+ \quad D_3\mathcal{L}_d(\hat{Q}_k) = 0\,, \\ L^*_{W_{k-1}}D_4\mathcal{L}_d(\hat{Q}_{k-1}) \quad &- \quad R^*_{W_k}D_4\mathcal{L}_d(\hat{Q}_k) = 0, \end{aligned}$$

together with the forced discrete Euler-Lagrange equations (5.56) and (5.57).

Typically, actuation is achieved by controlling only a subset of the shape variables. In our setting this is can be regarded as **underactuation** – the mappings in definition 5.3.1 become embeddings. If this is the case, it is necessary to introduce constraints and apply constrained variational calculus as in §5.1.7 and §5.2.4.

5.3.2 Discrete optimal control of nonholonomic systems

This subsection is devoted to add nonholonomic constraints to the picture. Holonomic constraints might be considered as a particular case of the nonholonomic ones (see [110] for further details). With this extension it would be possible consider examples of optimal control of robotic vehicles. In the following we will expose the theoretical framework, leaving for future research the application to concrete examples.

A controlled discrete nonholonomic system on $M \times M \times G$ is given by the following quadruple (see [70, 93]):

i) A **regular discrete Lagrangian** $l_d : M \times M \times G \to \mathbb{R}$.

ii) A **discrete constraint embedded submanifold** $\mathcal{M}_c$ of $M \times M \times G$.

iii) A **constraint distribution**, $\mathcal{D}_c$, which is a vector subbundle of the vector bundle $\tau_{TM\times\mathfrak{g}} : TM \times \mathfrak{g} \to M$, such that $\dim \mathcal{M}_c = \dim \mathcal{D}_c$. Typically, there is a relation between the constraint distribution and the discrete constraint, since from $\mathcal{M}_c$ we induce for every $x \in M$, the subspace $\mathcal{D}_c(x)$ of $T_xM \times \mathfrak{g}$ given by

$$\mathcal{D}_c(x) = T_{(x,x,e)}\mathcal{M}_c \cap (T_xM \times \mathfrak{g})\,,$$

where we are identifying $T_xM \times \mathfrak{g} \equiv 0_x \times T_xM \times T_eG$, with e being the identity element of the Lie group G.

iv) The discrete control forces $f_k^\pm : \mathcal{M}_c \times U \to T^*M \times \mathfrak{g}^*$ where $U \subset \mathbb{R}^m$ (again, forces $f_k^\pm$ split into $\bar{f}_k^\pm$ and $\hat{f}_k^\pm$ as in the previous section).

We have the following **discrete version of the Lagrange-D'Alembert principle for**

controlled nonholonomic systems:

$$\delta \sum_{k=0}^{N-1} l_d(x_k, x_{k+1}, W_k) \;+\; \sum_{k=0}^{N-1} \langle f_k^-, (\delta x_k, \eta_k)\rangle$$
$$+\; \sum_{k=0}^{N-1} \langle f_k^+, (\delta x_{k+1}, \eta_{k+1})\rangle = 0,$$

for all variations $\{\delta x_k\}_{k=0}^N$, with $\delta x_0 = \delta x_N = 0$; and $\{\delta W_k\}_{k=0}^N$, such that $\delta W_k = -\eta_k W_k + W_k \eta_{k+1}$, being $\{\eta_k\}_{k=0}^N$, verifying $(\delta x_k, \eta_k) \in \mathcal{D}_c(x_k) \subseteq T_{x_k}M \times \mathfrak{g}$ such that $\eta_0 = \eta_N = 0$. Moreover, $(x_k, x_{k+1}, W_k) \in \mathcal{M}_c$, $k = 0, \ldots, N-1$ (see [70]).

Take a basis of sections $\{(X^a, \tilde{\eta}^a)\}$ of the vector bundle $\tau_{\mathcal{D}_c} : \mathcal{D}_c \longrightarrow M$, where $X^a \in \mathfrak{X}(M)$ and $\tilde{\eta}^a : M \to \mathfrak{g}$ for $a = 1, ..., \mathrm{rank}(\mathcal{D}_c)$. Hence, the equations of motion derived from the discrete Lagrange-D'Alembert principle for controlled nonholonomic systems are:

$$\begin{aligned} 0 &= \langle D_1 l_d(x_k, x_{k+1}, W_k) + D_2 l_d(x_{k-1}, x_k, W_{k-1}) + \bar{f}_k^- + \bar{f}_{k-1}^+ \,,\, X^a(x_k)\rangle \\ &+\langle L^*_{W_{k-1}} D_3 l_d(x_{k-1}, x_k, W_{k-1}) - R^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) + \hat{f}_k^- + \hat{f}_{k-1}^+ \,,\, \tilde{\eta}^a(x_k)\rangle, \end{aligned} \tag{5.61}$$

$$0 = \Psi^\alpha(x_k, x_{k+1}, W_k), \tag{5.62}$$

where $\Psi^\alpha(x_k, x_{k+1}, W_k) = 0$ are the constraints which locally determine $\mathcal{M}_d$.

In a more geometric way, we can write equations (5.61) and (5.62) as follows

$$\begin{aligned} 0 = (i_{\mathcal{D}_c})^* \Big(D_1 l_d(x_k, x_{k+1}, W_k) \;&+\; D_2 l_d(x_{k-1}, x_k, W_{k-1}) + \bar{f}_k^- + \bar{f}_{k-1}^+, \\ L^*_{W_{k-1}} D_3 l_d(x_{k-1}, x_k, W_{k-1}) \;&-\; R^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) + \hat{f}_k^- + \hat{f}_{k-1}^+ \Big), \end{aligned}$$

where$(x_k, x_{k+1}, W_k) \in \mathcal{M}_c$ and $i_{\mathcal{D}_c} : \mathcal{D}_c \hookrightarrow TM \times \mathfrak{g}$ is the canonical inclusion.

Given a discrete cost function $C_d : U \times \mathcal{M}_c \times U \longrightarrow \mathbb{R}$ and the optimal control problem is to minimize the action sum

$$\sum_{k=0}^{N-1} C_d(u_k^-, x_k, W_k, x_{k+1}, u_k^+)$$

subject to equations (5.61) and (5.62) and to some given boundary conditions. We next distinguish between the fully and under–actuated case using the following definition:

Definition 5.3.2. (Fully-actuated nonholonomic discrete system) *We say that the discrete nonholonomic mechanical control system is fully-actuated if the mappings*

$$\begin{aligned} F_k^-\big|_{(x_0,x_1,W_1)} : U \to \mathcal{D}_c^*, &\quad F_k^-\big|_{(x_0,x_1,W_1)}(u) = (i_{\mathcal{D}_c})^*(f_k^-(x_0, x_1, W_1, u)), \\ F_k^+\big|_{(x_0,x_1,W_1)} : U \to \mathcal{D}_c^*, &\quad F_k^+\big|_{(x_0,x_1,W_1)}(u) = (i_{\mathcal{D}_c})^*(f_k^+(x_0, x_1, W_1, u)), \end{aligned}$$

are both diffeomorphisms for all $(x_0, x_1, W_1) \in \mathcal{M}_c$.

Regarding equation (5.61) and its geometric redefinition just below, let introduce the following momenta:

$$\pi_k = (i_{\mathcal{D}_c})^* \left(-D_1 l_d(x_k, x_{k+1}, W_k) - \bar{f}_k^-, R^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) - \hat{f}_k^- \right),$$
$$\pi_{k+1} = (i_{\mathcal{D}_c})^* \left(D_2 l_d(x_k, x_{k+1}, W_k) + \bar{f}_k^+, L^*_{W_k} D_3 l_d(x_k, x_{k+1}, W_k) + \hat{f}_k^+ \right),$$

where both π_k and π_{k+1} belong to $\mathcal{D}_c^*$. In the fully actuated case, the value of all control forces can be completely determined in terms of $x_k, x_{k+1}, W_k, \pi_k, \pi_{k+1}$, where the coordinates (x_k, x_{k+1}, W_k) always belong to $\mathcal{M}_c$. Therefore we can re-express the cost function in terms of these variables and, in consequence, derive the discrete Lagrangian

$$\mathcal{L}_d : (\mathcal{D}_c^*)\ {}_{\tau_{\mathcal{D}_c^*}}\!\times_{pr_1} (\mathcal{M}_c)\ {}_{pr_2}\!\times_{\tau^*_{\mathcal{D}_c}} (\mathcal{D}_c^*) \to \mathbb{R},$$

where $pr_i : \mathcal{M}_d \subseteq M \times M \times G \to M$ are the projections onto the first and second arguments and $\tau_{\mathcal{D}_c^*} : \mathcal{D}_c^* \to M$ the vector bundle projection.

Observe that we can consider this case as a constrained discrete variational problem taking an extension

$$\widetilde{\mathcal{L}_d} : \mathcal{D}_c^* \times G \times \mathcal{D}_c^* \to \mathbb{R}$$

of $\mathcal{L}_d$ subjected to the constraints $\Psi^\alpha(x_k, x_{k+1}, W_k) = 0$.

Therefore, denoting $\hat{Q}_k = (x_k, \pi_k, W_k, x_{k+1}, \pi_{k+1})$ as the whole set of coordinates of the new phase space $\mathcal{D}_c^* \times G \times \mathcal{D}_c^*$, we deduce that the equations of motion are

$$\begin{aligned}
D_4\widetilde{\mathcal{L}_d}(\hat{Q}_{k-1}) + D_1\widetilde{\mathcal{L}_d}(\hat{Q}_k) &= \lambda_\alpha^{k-1} D_2\Psi^\alpha(x_{k-1}, x_k, W_{k-1}) \\
&\quad + \lambda_\alpha^k D_1\Psi^\alpha(x_k, x_{k+1}, W_k), \\
D_5\widetilde{\mathcal{L}_d}(\hat{Q}_{k-1}) + D_2\widetilde{\mathcal{L}_d}(\hat{Q}_k) &= 0\,, \\
L^*_{W_{k-1}} D_3\widetilde{\mathcal{L}_d}(\hat{Q}_{k-1}) - R^*_{W_k} D_3\widetilde{\mathcal{L}_d}(\hat{Q}_k) &= \lambda_\alpha^{k-1} L^*_{W_{k-1}} D_3\Psi^\alpha(x_{k-1}, x_k, W_{k-1}) \\
&\quad - \lambda_\alpha^k R^*_{W_k} D_3\Psi^\alpha(x_k, x_{k+1}, W_k), \\
\Psi^\alpha(x_k, x_{k+1}, W_k) &= 0\,,
\end{aligned}$$

where λ_α^k are the Lagrange multipliers of the new constrained problem. The underactuated case can be handled by adding new constraints and applying discrete constrained variational calculus similarly to §5.2.2.

A natural framework that simplifies the previous construction is based on discrete mechanics on Lie groupoids [118]. The Lie groupoid structure generalizes the case of $Q \times Q$, the Lie group G and also many intermediate situations. In particular, many of the examples studied in this chapter can be modeled using Lie groupoid techniques adapted to our formalism (see [78]).

5.3.3 Extension to Lie algebroids

All the previous cases can be considered as particular examples of optimal control systems of mechanical type modeled on a Lie algebroid. The ideas developed in this subsection were firstly introduced in [78]. Consider the reduced Lagrangian $l : AG \longrightarrow \mathbb{R}$, where AG is the Lie algebroid associated to the Lie groupoid G (see §1.5). The external forces are modeled, in this case, by a mapping $f : AG \times U \to A^*G$, where U is the control space and A^*G is the dual algebroid of AG, such that $\tau = \tau^* \circ f|_U$ (recall that $\tau : AG \to Q$ and $\tau^* : A^*G \to Q$, where Q is the base manifold of the vector bundle that defines the algebroid).

Thus, is possible to adapt the derivation of the Lagrange-d'Alembert's principle to the case of mechanical systems defined on Lie algebroids. With that purpose, fix two points q_0, q_1 in the configuration manifold Q, then we look for admissible curves $\xi : I \subset \mathbb{R} \to AG$ (that is, curves such that $\rho(\xi(t)) = \dfrac{\mathrm{d}}{\mathrm{d}t}\tau(\xi(t))$) which satisfy

$$\delta \int_0^T l(\xi(t))\,\mathrm{d}t + \int_0^T \langle f(\xi(t), u(t)), \eta(t)\rangle\,\mathrm{d}t = 0, \tag{5.63}$$

where the infinitesimal variations to be considered are $\delta\xi = \eta^c$, for all $\eta(t) \in \Gamma(\tau)$, where $\Gamma(\tau)$ is the set of sections of the vector bundle τ, with $\eta(0) = 0$ and $\eta(T) = 0$. Here, η^c is a time-dependent vector field on AG along $\xi(t)$, namely the complete lift, locally given by

$$\eta^c = \rho^i_\alpha \eta^\alpha \frac{\partial}{\partial q^i} + (\dot{\eta}^\alpha - C^\alpha_{\beta\gamma}\eta^\beta y^\gamma)\frac{\partial}{\partial y^\alpha}$$

where we have chosen coordinates (q^i) on Q and we have fixed a basis of sections $\{e_\alpha\}$ of $\tau : AG \to Q$, inducing local coordinates (q^i, y^α) on AG. Recall from §1.5.1 that $C^\alpha_{\beta\gamma}$ and ρ^i_α are the local functions that determine the Lie algebroid structure on Q. From this principle we derive the equations of motion for solution curve:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial l}{\partial y^\alpha}\right) = \rho^i_\alpha \frac{\partial l}{\partial q^i} - C^\gamma_{\alpha\beta} y^\beta \frac{\partial l}{\partial y^\gamma} + f_\alpha(q^i, y^\beta, u), \quad f_\alpha = \langle f, e_\alpha\rangle,\\
&\frac{\mathrm{d}q^i}{\mathrm{d}t} = \rho^i_\alpha y^\alpha\,.
\end{aligned}$$

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

$$\mathcal{J}(\xi, u) = \int_0^T C(\xi(t), u(t))\,\mathrm{d}t \tag{5.64}$$

where now $C : AG \times U \to \mathbb{R}$.

We now consider the associated discrete problem. Let construct the discrete lagrangian $l_d : G \to \mathbb{R}$ as an approximation of the continuous action.

$$l_d(g_k) \approx \int_{kh}^{(k+1)h} l(\xi(t))\,\mathrm{d}t, \tag{5.65}$$

where $\{g_k\}_{k=0}^N \in G^{N+1}$, G is the Lie groupoid. The discrete external forces are defined as

$$\langle f_k^-(g_k, u_k^-), \eta_k\rangle + \langle f_k^+(g_k, u_k^+), \eta_{k+1}\rangle \approx \int_{kh}^{(k+1)h} \langle f(\xi(t), u(t)), \eta(t)\rangle\,\mathrm{d}t,$$

where again we are allowing two different sequences of discrete controls $\{u_k^{\pm}\}_{k=0}^{N}$. Moreover, $(f_k^-(g_k, u_k^-), f_k^+(g_k, u_k^+)) \in A^*_{\alpha(g_k)}G \times A^*_{\beta(g_k)}G$ and $\eta_k \in A^*_{\alpha(g_k)}G$, for all k and $\eta_0 = \eta_N = 0$. Note that α and β are the submersions defining the Lie groupoid $G \rightrightarrows Q$ (see §1.5.2).

Therefore, we derive a **discrete version of the Lagrange-D'Alembert principle for Lie groupoids**:

$$\delta \sum_{k=0}^{N-1} l_d(g_k) + \sum_{k=0}^{N-1} \big(\langle f_k^-(g_k, u_k^-), X_k(g_k)\rangle + \langle f_k^+(g_k, u_k^+), X_{k+1}(g_k)\rangle\big) = 0 \tag{5.66}$$

for all variations $\{\delta g_k\}_{k=0}^{N-1}$ verifying the relation $\delta g_k = \overleftarrow{X_{k+1}}(g_k) - \overrightarrow{X_k}(g_k)$ with $\{X_k\}_{k=0}^{N}$ an arbitrary sequence of sections of $\tau : AG \to Q$ with $X_0, X_N = 0$. Recall from §1.5.2 that $\overleftarrow{X}$ and $\overrightarrow{X}$ belong to $\mathfrak{X}(G)$ (they are left-invariant and right-invariant vector fields respectively).

Therefore, we deduce that

$$\begin{aligned} 0 &= \sum_{k=0}^{N-1} \langle \mathrm{d}l_d(g_k), \delta g_k\rangle + \sum_{k=0}^{N-1} \langle f_k^-(g_k, u_k^-), X_k(g_k)\rangle + \langle f_k^+(g_k, u_k^+), X_{k+1}(g_k)\rangle \\ &= \sum_{k=0}^{N-1} \langle \mathrm{d}l_d(g_k), \overleftarrow{X_{k+1}}(g_k) - \overrightarrow{X_k}(g_k)\rangle + \sum_{k=0}^{N-1} \langle (f_k^-(g_k, u_k^-), X_k(g_k)\rangle + \langle f_k^+(g_k, u_k^+), X_{k+1}(g_k)\rangle. \end{aligned}$$

From this last expression we obtain the following discrete equations of motion:

$$\overleftarrow{X}(g_k)(l_d) - \overrightarrow{X}(g_{k+1})(l_d) + \langle f_{k+1}^-(g_{k+1}, u_{k+1}^-) + f_k^+(g_k, u_k^+), X(g_k)\rangle = 0, \qquad 0 \le k \le N-1, \tag{5.67}$$

for all $X \in \Gamma(\tau)$.

As we did in §5.1.3 (tangent bundles case) and §5.2.3 (Lie algebras case), we take now an approximation of the cost functional in the following manner

$$C_d(u_k^+, g_k, u_k^-) \simeq \int_{kh}^{(k+1)h} C(\xi(t), u(t))\,\mathrm{d}t, \tag{5.68}$$

where g_k belongs to the groupoid G. This leads to the discrete cost functional:

$$\mathcal{J}_d(g_{0:N-1}, u^{\pm}_{0:N-1}) = \sum_{k=0}^{N-1} C_d(u_k^+, g_k, u_k^-).$$

Next, as we already did in §5.1.4 (tangent bundles case) and §5.2.4 (Lie algebras case), we perform the fully-actuation by means of the following definition

Definition 5.3.3 (Fully-actuated discrete systems). *We say that the discrete control system is fully-actuated if the mappings*

$$\begin{aligned} f_k^-\big|_{g_k} &: U \to A^*_{\alpha(g_k)}G, \qquad & f_k^-\big|_{g_k}(u_k^-) &= f_k^-(g_k, u_k^-), \\ f_k^+\big|_{g_k} &: U \to A^*_{\beta(g_k)}G, \qquad & f_k^+\big|_{g_k}(u_k^+) &= f_k^+(g_k, u_k^+), \end{aligned}$$

are both diffeomorphisms

Therefore, we can construct the new discrete Lagrangian over the new phase space

$$\mathcal{P}^{\tau^*}G = A^*G \,{}_{\tau^*}\!\times_\alpha G \,{}_\beta\!\times_{\tau^*} A^*G,$$

which is the prolongation of the Lie groupoid $G \rightrightarrows Q$ over $\tau^* : A^*G \to Q$ (the definition of all structure maps and sections can be found in §1.5.2). Therefore, taking into account equations (3.16) in §3.5 (see [118]), the equations of motion of a discrete Lagrangian $\mathcal{L}_d : \mathcal{P}^{\tau^*}G \to \mathbb{R}$ are

$$\overrightarrow{Z}(\nu, g, \nu')(\mathcal{L}_d) - \overleftarrow{Z}(\nu', h, \nu'')(\mathcal{L}_d) = 0,$$

where (ν, g, ν') and (ν', h, ν'') belong to $\mathcal{P}^{\tau^*}G$, $(g, h) \in G^{(2)}$ and $Z \in \Gamma(A(\mathcal{P}^{\tau^*}G))$. The previous equations, with some abuse of notation, can be written as

$$\langle D_2\mathcal{L}_d\big|_g(\nu, \nu') - D_1\mathcal{L}_d\big|_h(\nu', \nu''), Y(\nu')\rangle + \overleftarrow{X}(g)(\mathcal{L}_d) - \overrightarrow{X}(h)(\mathcal{L}_d) = 0, \tag{5.69}$$

where $D_2\mathcal{L}_d\big|_g(\nu, \nu') = D_3\mathcal{L}_d(\nu, g, \nu')$, $D_1\mathcal{L}_d\big|_h(\nu', \nu'') = D_1\mathcal{L}_d(\nu', h, \nu'')$. These equations hold for every $X \in \Gamma(\tau)$ and $Y \in \mathfrak{X}(A^*G)$ verifying that $T\beta(X) = T\tau^*(Y)$.

When G is just a Lie group, $\mathcal{P}^{\tau^*}G = \mathfrak{g}^* \times G \times \mathfrak{g}^* \rightrightarrows \mathfrak{g}^*$, which is the space where the discrete Lagrangian (5.43) is defined. The associated Lie algebroid is $A(\mathfrak{g}^* \times G \times \mathfrak{g}^*) = \mathfrak{g}^* \times T\mathfrak{g}^* \to \mathfrak{g}^*$. The sections of this Lie algebroid are of the form $a^* \to (\xi, Y(a^*))$, where $a^* \in \mathfrak{g}^*$, $\xi\mathfrak{g}$ and $Y \in \mathfrak{X}(\mathfrak{g}^*)$. Therefore, for a discrete Lagrangian $\mathcal{L}_d : \mathfrak{g}^* \times G \times \mathfrak{g}^* \to \mathbb{R}$, we obtain from (5.69):

$$\begin{aligned} L^*_{g_{k-1}}\mathrm{d}\mathcal{L}_d\big|_{(\nu_{k-1},\nu_k)}(g_{k-1}) - R^*_{g_k}\mathrm{d}\mathcal{L}_d\big|_{(\nu_k,\nu_{k+1})}(g_k) &= 0, \\ D_2\mathcal{L}_d\big|_{g_{k-1}}(\nu_{k-1}, \nu_k) + D_1\mathcal{L}_d\big|_{g_k}(\nu_k, \nu_{k+1}) &= 0, \end{aligned}$$

where $k = 1, ..., N-1$, and $\mathcal{L}_d\big|_g(\nu, \nu') = \mathcal{L}_d\big|_(\nu, \nu')(g) = \mathcal{L}_d(\nu, g, \nu')$. In the last expression we recognize the equations (5.45). Take the equivalence

$$\begin{aligned} \nu_k &= -R^*_{g_k}\mathrm{d}l_d(g_k) - f_k^-(g_k, u_k^-), \\ \nu_{k+1} &= l^*_{g_k}\mathrm{d}l_d(g_k) + f_k^+(g_k, u_k^+), \end{aligned}$$

where l_d is defined in (5.65). We recover, just by considering the composability condition (see §1.5.2) of the groupoid $\mathfrak{g}^* \times G \times \mathfrak{g}^* \to \mathfrak{g}^*$, the forced discrete Euler-Poincaré equations (5.32):

$$L^*_{g_{k-1}}\mathrm{d}l_d(g_{k-1}) - R^*_{g_k}\mathrm{d}l_d(g_k) + f^+_{k-1}(g_{k-1}, u^+_{k-1}) + f_k^-(g_k, u_k^+) = 0,$$

with $k = 1, \ldots, N-1$.

Thus, by choosing a concrete instance of Lie groupoid G (namely a Lie group) in the formalism presented in this subsection, we are able to revisit the case of discrete optimal control on Lie groups developed in §5.2.4.

Chapter 6

Geometric Nonholonomic Integrator (GNI)

As was mentioned in the introduction, the geometric perspective has been introduced in the study of nonholonomic systems in the last decades. This new perspective allows us to use techniques from differential geometry to handle problems in nonholonomic constraints. This chapter accounts for new developments in the line of [44, 45, 93], where the Geometric Nonholonomic Integrator (GNI) and some of its properties were presented. More concretely, we study the GNI extensions of Euler-symplectic methods (see [61]) and discuss some of their convergence properties following the methodology developed in [45]. Additionally, we generalize the method proposed for nonholonomic reduced systems, which represent an important subclass of examples in nonholonomic dynamics. Moreover, we construct extensions of the GNI in the cases of affine constraints and Lie groupoids.

6.1 Nonholonomic mechanical systems

In §2.4.2 nonholonomic dynamics has been widely treated. As a summary, let us recall some basics. Our configuration manifold will be denoted by Q, with local constraints q^i, $i = 1, ..., n$, while the local coordinates for the tangent bundle will be $(q^i, \dot{q}^i)$. The Lagrangian function will be denoted $L : TQ \to \mathbb{R}$. Let consider the nonholonomic distribution $\mathcal{D}$ defined by the set of linear constraints

$$\phi^a\left(q^i, \dot{q}^i\right) = \mu_i^a\left(q\right)\dot{q}^i = 0, \;\; 1 \leq a \leq m, \tag{6.1}$$

where $\operatorname{rank}\left(\mathcal{D}\right) = n - m$. The annihilator $\mathcal{D}^\circ$ is locally given by

$$\mathcal{D}^\circ = \operatorname{span}\left\{\mu^a = \mu_i^a(q)\,dq^i;\; 1 \leq a \leq m\right\},$$

where the 1-forms μ^a are independent.

As was mentioned in §2.4.2, a way to establish the nonholonomic dynamics is the Lagrange-d'Alembert principle, namely:

$$\delta\int_0^T L\left(q\left(t\right), \dot{q}\left(t\right)\right) \mathrm{d}t = 0$$

for all variations such that $\delta q\,(t) \in \mathcal{D}_{q(t)}$, $0 \leq t \leq T$, and if the curve itself satisfies the constraints. From this principle we arrive to the nonholonomic equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a, \tag{6.2a}$$

$$\mu_i^a(q)\,\dot{q}^i = 0, \tag{6.2b}$$

where λ_a, $a = 1, ..., m$ is a set of Lagrange multipliers. The right-hand side of equation (6.2a) represents the force induced by the constraints, and equations (6.2b) represent the constraints themselves.

Now we restrict ourselves to the case of nonholonomic mechanical systems where the Lagrangian is of mechanical type. Now, there is a Riemannian metric $\mathcal{G}$ defined on the configuration space Q. Thus, the Lagrangian function will be defined by

$$L(v_q) = \frac{1}{2}\mathcal{G}(v_q, v_q) - V(q), \quad v_q \in T_q Q,$$

where $V : Q \to \mathbb{R}$ is a potential function. Locally, the metric is determined by the matrix $M = (\mathcal{G}_{ij})_{1 \leq i,j \leq n}$ where $\mathcal{G}_{ij} = \mathcal{G}(\partial/\partial q^i, \partial/\partial q^j)$.

Using some basic tools of Riemannian geometry (see §1.2 for further details), we may write the equations of motion of the unconstrained system as

$$\nabla_{\dot{c}(t)}\dot{c}(t) = -\mathrm{grad}\; V(c(t)), \tag{6.3}$$

where ∇ is the Levi–Civita connection associated to $\mathcal{G}$. Observe that if $V \equiv 0$ then the Euler–Lagrangian equations are the equations of the geodesics for the Levi-Civita connection.

When the system is subjected to nonholonomic constraints, the equations become

$$\nabla_{\dot{c}(t)}\dot{c}(t) = -\mathrm{grad}\; V(c(t)) + \lambda(t), \quad \dot{c}(t) \in \mathcal{D}_{c(t)},$$

where λ is a section of $\mathcal{D}^\perp$ along c. Here $\mathcal{D}^\perp$ stands for the orthogonal complement of $\mathcal{D}$ with respect to the metric $\mathcal{G}$. Obviously, $\mathcal{D}$ stands for the distribution determining the nonholonomic constraints.

In coordinates, by defining the n^3 functions Γ_{ij}^k (Christoffel symbols for ∇) by

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

we may rewrite the nonholonomic equations of motion as

$$\ddot{q}^k(t) + \Gamma_{ij}^k(c(t))\dot{q}^i(t)\dot{q}^j(t) = -\mathcal{G}^{ki}(c(t))\frac{\partial V}{\partial q^i} + \lambda_a(t)\mathcal{G}^{ki}(c(t))\mu_i^a(c(t))$$

$$\mu_i^a(c(t))\dot{q}^i(t) = 0$$

where $t \mapsto (q^1(t), \ldots, q^n(t))$ is the local representative of c and $(\mathcal{G}^{ij})$ is the inverse matrix of M.

Since $\mathcal{G}$ is a Riemannian metric, the $m \times m$ matrix $(C^{ab}) = (\mu_i^a \mathcal{G}^{ij} \mu_j^b)$ is symmetric and regular. Define now the vector fields Z^a, $1 \leq a \leq m$ on Q by

$$\mathcal{G}(Z^a, Y) = \mu^a(Y), \quad \text{for all vector fields } Y,\ 1 \leq a \leq m;$$

that is, Z^a is the gradient vector field of the 1-form μ^a. Thus, $\mathcal{D}^\perp$ is spanned by Z^a, $1 \leq a \leq m$. In local coordinates, we have

$$Z^a = \mathcal{G}^{ij} \mu_i^a \frac{\partial}{\partial q^j}.$$

We can construct two complementary projectors

$$\begin{aligned} \mathcal{P}&\colon TQ \to \mathcal{D}, \\ \mathcal{Q}&\colon TQ \to \mathcal{D}^\perp, \end{aligned}$$

which are orthogonal with respect to the metric $\mathcal{G}$. The projector $\mathcal{Q}$ is locally described by

$$\mathcal{Q} = C_{ab} Z^a \otimes \mu^b = C_{ab} \mathcal{G}^{ij} \mu_i^a \mu_k^b \frac{\partial}{\partial q^j} \otimes \mathrm{d}q^k.$$

Using these projectors we may rewrite the equations of motion as follows. A curve $c(t)$ is a motion for the nonholonomic system if it satisfies the constraints, i.e., $\dot{c}(t) \in \mathcal{D}_{c(t)}$, and, in addition, the "projected equation of motion"

$$\mathcal{P}(\nabla_{\dot{c}(t)} \dot{c}(t)) = -\mathcal{P}(\text{grad } V(c(t))) \tag{6.4}$$

is fulfilled.

Summarizing, we have obtained the dynamics of the nonholonomic system (6.4) applying the projector $\mathcal{P}$ to the dynamics of the free system (6.3). Next, we will use $\mathcal{P}$ and $\mathcal{Q}$ to obtain a geometric integrator for nonholonomic systems.

6.2 The Geometric Nonholonomic Integrator (GNI)

In [44, 45, 93] a numerical method for the integration of nonholonomic systems is proposed and developed (GNI henceforth). It is not truly variational; however, it is geometric in nature. It is shown that GNI preserves the discrete nonholonomic momentum map in the presence of horizontal symmetries. Moreover, the energy of the system is preserved under certain symmetry conditions.

Now, we employ the notions concerning discrete mechanics developed in §3. Consider a discrete Lagrangian $L_d\colon Q \times Q \to \mathbb{R}$. The proposed discrete nonholonomic equations are

$$\mathcal{P}^*_{q_k}(D_1 L_d(q_k, q_{k+1})) + \mathcal{P}^*_{q_k}(D_2 L_d(q_{k-1}, q_k)) = 0, \tag{6.5a}$$

$$\mathcal{Q}^*_{q_k}(D_1 L_d(q_k, q_{k+1})) - \mathcal{Q}^*_{q_k}(D_2 L_d(q_{k-1}, q_k)) = 0, \tag{6.5b}$$

where the subscript q_k emphasizes the fact that the projections take place in the fiber over q_k. The first equation is the projection of the discrete Euler–Lagrange equations to the constraint distribution $\mathcal{D}$, while the second one can be interpreted as an elastic impact of the system against $\mathcal{D}$ (see [69]). This is what will provide the preservation of energy. Note that we can combine both equations into

$$D_1L_d(q_k, q_{k+1}) + (\mathcal{P}^* - \mathcal{Q}^*)D_2L_d(q_{k-1}, q_k) = 0,$$

from which we see that the system defines a unique discrete evolution operator if and only if the matrix $(D_{12}L_d)$ is regular, that is, if the discrete Lagrangian is regular. Locally, the method can be written as

$$D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) = (\lambda_k)_b\, \mu^b(q_k) \tag{6.6a}$$

$$\mathcal{G}^{ij}(q_k)\mu^a_i(q_k)\left(\frac{\partial L_d}{\partial q^j_0}(q_k, q_{k+1}) - \frac{\partial L_d}{\partial q^j_1}(q_{k-1}, q_k)\right) = 0. \tag{6.6b}$$

Using the discrete Legendre transformations defined in §3.2, define the pre- and post-momenta, which are covectors at q_k, by

$$p^+_{k-1,k} = p^+(q_{k-1}, q_k) = \mathbb{F}L^+_d(q_{k-1}, q_k) = D_2L_d(q_{k-1}, q_k)$$
$$p^-_{k,k+1} = p^-(q_k, q_{k+1}) = \mathbb{F}L^-_d(q_k, q_{k+1}) = -D_1L_d(q_k, q_{k+1}).$$

In these terms, equation (6.6b) can be rewritten as

$$\mathcal{G}^{ij}(q_k)\mu^a_i(q_k)\left(\frac{(p^-_{k,k+1})_j + (p^+_{k-1,k})_j}{2}\right) = 0$$

which means that the average of post- and pre-momenta satisfies the constraints. In this sense the proposed numerical method also preserves the nonholonomic constraints.

We may rewrite the discrete nonholonomic equations as

$$p^-_{k,k+1} = (\mathcal{P} - \mathcal{Q})^*_{q_k}\,(p^+_{k-1,k}). \tag{6.7}$$

We interpret this equation as a jump of momenta during the nonholonomic evolution. Compare this with the condition $p^-_{k,k+1} = p^+_{k-1,k}$ imposed by the discrete Euler–Lagrange equations (that is, for unconstrained systems). In our method, the momenta are related by a reflection with respect to the image of the projector $\mathcal{P}^*\colon T^*Q \to (\mathcal{D}^\perp)^o$.

6.2.1 Left-invariant discrete Lagrangians on Lie groups

Consider a discrete nonholonomic Lagrangian system on a Lie group G, with a discrete Lagrangian $L_d\colon G \times G \to \mathbb{R}$ that is invariant with respect to the left diagonal action of G on $G \times G$ (see [18, 121]) We do not impose yet any invariance conditions on the distribution $\mathcal{D}$. If we write $W_k = g_k^{-1}g_{k+1}$, then we can define the reduced discrete Lagrangian $l_d\colon G \to \mathbb{R}$ as $l_d(W_k) = L_d(g_k, g_{k+1})$. Note that $\mathrm{d}l_d(W_k) \in T^*_{W_k}G$.

Computing the derivative, we obtain

$$p^-_{k,k+1} = -D_1L_d(g_k, g_{k+1}) = L^*_{g_k^{-1}}R^*_{W_k}\mathrm{d}l_d(W_k),$$

where L^* and R^* are the mappings on T^*G induced by left and right multiplication on the group, respectively (this shall not be confused with the Lagrangian L). We use this to write

$$p^+_{k,k+1} = D_2L_d(g_k, g_{k+1}) = L^*_{g_k^{-1}}\mathrm{d}l_d(W_k) = L^*_{g_k^{-1}}R^*_{W_k^{-1}}L^*_{g_k}p^-_{k,k+1} = R^*_{W_k^{-1}}p^-_{k,k+1}.$$

Therefore, the discrete nonholonomic equations (6.7) become

$$p^-_{k,k+1} = (\mathcal{P} - \mathcal{Q})^*\left(R^*_{W_{k-1}^{-1}}p^-_{k-1,k}\right). \tag{6.8}$$

The relationships between the pre- and post-momenta are depicted in the figure 6.1:

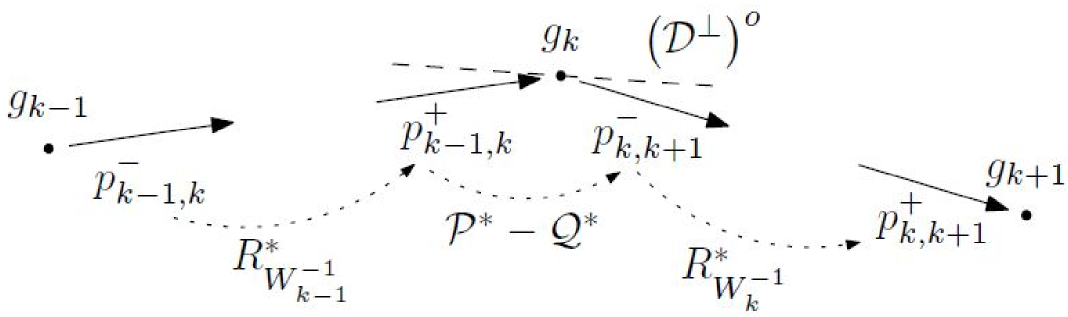


Figure 6.1: Evolution of momenta, depicted here as solid arrows. The right translations are a consequence of the left-invariance of L_d, and the reflection at g_k is the proposed method.

Note that we do not need here that the metric used to build the projectors is the metric giving the kinetic energy in the Lagrangian.

6.2.2 Properties

- *Preserving energy on Lie groups:*

  Let us now consider the case where Q is a Lie group G, the nonholonomic distribution $\mathcal{D}$ is not necessarily G-invariant, and L is regular and bi-invariant.

  Since we are restricting ourselves to Lagrangians of mechanical type, the potential energy is necessarily zero. The left-invariance of L implies that it must be of the form

$$L(v_g) = \frac{1}{2}\langle \mathbb{I}(g^{-1}v_g), g^{-1}v_g\rangle, \tag{6.9}$$

  where $\mathbb{I}\colon \mathfrak{g} \to \mathfrak{g}^*$ is a symmetric non-singular inertia tensor. The bi-invariance, however, imposes the equivariance condition $\mathrm{Ad}^*_{g^{-1}} \circ \mathbb{I} = \mathbb{I} \circ \mathrm{Ad}_g$ for all $g \in G$, as is straightforward to check. We remark that in this section, the metric used to build the projectors will be the same that defines the Lagrangian. If we take a discretization $L_d\colon G \times G \to \mathbb{R}$ (which needs to be **left**-invariant only), the equations of motion (6.8) hold. Then the following result can be proven (see [44] for the proof):

Theorem 6.2.1. *Consider a nonholonomic system on a Lie group with a regular, bi-invariant Lagrangian and with an arbitrary distribution $\mathcal{D}$, and take a discrete Lagrangian that is left-invariant. Then the proposed discrete nonholonomic method (6.5) is energy-preserving.*

- *The average momentum*

  Take a discrete nonholonomic system on G as in the previous section, but add the condition that $\mathcal{D}$ is right-invariant. Since the metric on the group is right-invariant, so is the projector $\mathcal{P}$. Take a trajectory of the system and define at each g_k the average momentum
$$\widetilde{p}_k = \frac{1}{2}\left(p^+_{k-1,k} + p^-_{k,k+1}\right). \tag{6.10}$$
  Using (6.7), (6.8) and the fact that $(\mathcal{P}-\mathcal{Q})^*$ is its own inverse, we have
$$\begin{aligned}\widetilde{p}_k &= \frac{1}{2}\left((\mathcal{P}-\mathcal{Q})^*(p^-_{k,k+1}) + p^-_{k,k+1}\right) = \mathcal{P}^*(p^-_{k,k+1}) = \mathcal{P}^*(R^*_{W^{-1}_{k-1}} p^-_{k-1,k})\\ &= R^*_{W^{-1}_{k-1}}\mathcal{P}^*(p^-_{k-1,k}) = R^*_{W^{-1}_{k-1}}\widetilde{p}_{k-1}.\end{aligned}$$
  Since the norm $\|\cdot\|_{\mathbb{I}}$ on each fiber of T^*G defined in the proof of Theorem 6.2.1 is right-invariant, we obtain $\|\widetilde{p}_k\|_{\mathbb{I}} = \|\widetilde{p}_{k-1}\|_{\mathbb{I}}$, so
$$H(g_k,\widetilde{p}_k) = H(g_{k-1},\widetilde{p}_{k-1}).$$
  In addition, by equation (6.7), we have that $\mathcal{Q}^*(\widetilde{p}_k) = 0$, so $\widetilde{p}_k$ satisfies the constraints.

- *Preservation of the nonholonomic momentum map*

  Let us recall some concepts regarding symmetries of nonholonomic systems. Suppose that a Lie group G acts on the configuration manifold Q. Define, for each $q \in Q$, the vector subspace $\mathfrak{g}^q$ consisting of those elements of $\mathfrak{g}$ whose infinitesimal generators (see definition 1.4.10) at q satisfy the nonholonomic constraints, i.e.,
$$\mathfrak{g}^q = \{\xi \in \mathfrak{g} \,|\, \xi_Q(q) \in \mathcal{D}_q\}.$$
  The (generalized) bundle over Q whose fiber at q is $\mathfrak{g}^q$ is denoted by $\mathfrak{g}^{\mathcal{D}}$.

  A horizontal symmetry is an element $\xi \in \mathfrak{g}$ such that $\xi_Q(q) \in \mathcal{D}_q$ for all $q \in Q$. Note that a horizontal symmetry is related naturally to a constant section of $\mathfrak{g}^{\mathcal{D}}$.

  Now consider a discrete Lagrangian $L_d\colon Q\times Q \to \mathbb{R}$, and define the discrete nonholonomic momentum map $J^{\mathrm{nh}}_d\colon Q\times Q \to (\mathfrak{g}^{\mathcal{D}})^*$ as in [38] by
$$\begin{aligned}J^{\mathrm{nh}}_d(q_{k-1},q_k)\colon \mathfrak{g}^{q_k} &\to \mathbb{R}\\ \xi &\mapsto \langle D_2L_d(q_{k-1},q_k), \xi_Q(q_k)\rangle.\end{aligned}$$
  For any smooth section $\widetilde{\xi}$ of $\mathfrak{g}^{\mathcal{D}}$ we have a function $(J^{\mathrm{nh}}_d)_{\widetilde{\xi}}\colon Q\times Q \to \mathbb{R}$, defined as $(J^{\mathrm{nh}}_d)_{\widetilde{\xi}}(q_{k-1},q_k) = J^{\mathrm{nh}}_d(q_{k-1},q_k)\left(\widetilde{\xi}(q_k)\right)$. Taking this into account, the following theorem and corollary can be proven, ensuring nonholonomic momentum preservation (see [44] for the proof):

Theorem 6.2.2. *Assume that L_d is G-invariant, and let $\widetilde{\xi}$ be a smooth section of $\mathfrak{g}^{\mathcal{D}}$. Then, under the proposed nonholonomic integrator, $(J_d^{\mathrm{nh}})_{\widetilde{\xi}}$ evolves according to the equation*

$$(J_d^{\mathrm{nh}})_{\widetilde{\xi}}(q_k, q_{k+1}) - (J_d^{\mathrm{nh}})_{\widetilde{\xi}}(q_{k-1}, q_k) = \left\langle D_2 L_d(q_k, q_{k+1}), (\xi_{k+1} - \xi_k)_Q\,(q_{k+1})\right\rangle$$

where $\xi_k, \xi_{k+1} \in \mathfrak{g}$ are the result of dropping the base points of $\widetilde{\xi}(q_k)$ and $\widetilde{\xi}(q_{k+1})$ respectively.

Corollary 6.2.3. *If L_d is G-invariant and ξ is a horizontal symmetry, then the proposed nonholonomic integrator preserves $(J_d^{\mathrm{nh}})_\xi$.*

6.3 GNI extensions of symplectic-Euler methods

Consider the autonomous dynamical system

$$\begin{aligned} \dot{x} &= f(x), \\ x(0) &= x_0, \end{aligned}$$

where $x(t), x_0 \in \mathbb{R}^n$, $\dot{x}$ is the time derivative of x and $f : \mathbb{R}^n \to \mathbb{R}^n$. In Hairer's terminology [61] there are two types of Euler integrators: explicit and implicit.

$$\begin{aligned} x_{k+1} &= x_k + hf(x_k) \qquad \text{Euler explicit}, \\ x_{k+1} &= x_k + hf(x_{k+1}) \qquad \text{Euler implicit}, \end{aligned}$$

where, as usual in numerical integration of ordinary differential equations (see [62]), $x_k \approx x(t_k)$ with $t_k = hk$ (h is the time step and k is a positive integer).

Let consider now the tangent TQ and cotangent T^*Q bundles of the configuration manifold $Q = \mathbb{R}^n$ and its local coordinates, $(q^i, \dot{q}^i)$ and (q^i, p_i), $i = 1, ..., n$, respectively. Moreover, let consider the mechanical Lagrangian system $L(q, \dot{q}) = \frac{1}{2}\,\dot{q}^T\,M\,\dot{q} - V(q)$, where M is a $n \times n$ constant regular matrix and $V : Q \to \mathbb{R}$ the potential function. On the other hand, the function $H(q, p) = \frac{1}{2}\,p^T\,M^{-1}\,p + V(q)$ is its Hamiltonian counterpart.

It is well-known that that explicit and implicit Euler methods applied to the Hamilton's equations (2.9) are not symplectic. Nevertheless, the so-called symplectic-Euler methods indeed are. In the sequel, we will denote these methods by **Euler A** and **Euler B**. This result corresponds to the following theorem:

Theorem 6.3.1. *The so-called symplectic-Euler methods*

$$q_{k+1} = q_k + h\frac{\partial H}{\partial p}(p_k, q_{k+1}), \qquad p_{k+1} = p_k - h\frac{\partial H}{\partial q}(p_k, q_{k+1}), \quad \textit{Euler A},$$

$$q_{k+1} = q_k + h\frac{\partial H}{\partial p}(p_{k+1}, q_k), \qquad p_{k+1} = p_k - h\frac{\partial H}{\partial q}(p_{k+1}, q_k), \quad \textit{Euler B},$$

are symplectic methods of order 1.

See [61] for the proof and §3.4.2 for more details regarding the order of accuracy. In addition, Euler A and Euler B methods are adjoint of each other (see §3.4.4). Applied to the specific mechanical Hamiltonian $H(q,p) = \frac{1}{2}\,p^T\,M^{-1}\,p + V(q)$, Euler A and B methods look like

$$\begin{array}{cc} \text{Euler A} & \text{Euler B} \\ q_{k+1} = q_k + hM^{-1}p_k, & q_{k+1} = q_k + hM^{-1}p_{k+1}, \\ p_{k+1} = p_k - h\frac{\partial V}{\partial q}\left(q_k\right), & p_{k+1} = p_k - h\frac{\partial V}{\partial q}\left(q_{k+1}\right). \end{array}$$

As variational integrators (see §3.4) they correspond to the following discrete Lagrangians:

$$L_d^A(q_k, q_{k+1}) = hL(q_k, \frac{q_{k+1}-q_k}{h}), \;\; L_d^B(q_k, q_{k+1}) = hL(q_{k+1}, \frac{q_{k+1}-q_k}{h}). \tag{6.11}$$

Applying the GNI equations (6.6) to the Lagrangians in (6.11) we obtain the following numerical schemes:

- **Euler A**:

$$q_{k+1} - 2q_k + q_{k-1} = -h^2M^{-1}\left(V_q(q_k) + \mu^T\left(q_k\right)\tilde{\lambda}_k\right), \tag{6.12a}$$

$$0 = \mu(q_k)\left(\frac{q_{k+1}-q_{k-1}}{2h} + \frac{h}{2}M^{-1}V_q(q_k)\right). \tag{6.12b}$$

- **Euler B**:

$$q_{k+1} - 2q_k + q_{k-1} = -h^2M^{-1}\left(V_q(q_k) + \mu^T\left(q_k\right)\tilde{\lambda}_k\right), \tag{6.13a}$$

$$0 = \mu(q_k)\left(\frac{q_{k+1}-q_{k-1}}{2h} - \frac{h}{2}M^{-1}V_q(q_k)\right), \tag{6.13b}$$

where $\tilde{\lambda}_k = \lambda_k/h$ and $V_q = \partial V/\partial q$. By means of the momentum relations $\tilde{p}_k = M(q_{k+1} - q_{k-1})/2h$ and $p_{k+1/2} = M(q_{k+1} - q_k)/h$, we can rewrite equations (6.12) and (6.13) as:

- **Euler A**:

$$p_{k+1/2} = \tilde{p}_k - \frac{h}{2}\left(V_q(q_k) + \mu^T(q_k)\tilde{\lambda}_k\right), \tag{6.14a}$$

$$q_{k+1} = q_k + hM^{-1}p_{k+1/2}, \tag{6.14b}$$

$$\mu(q_k)M^{-1}\left(\tilde{p}_k + \frac{h}{2}V_q(q_k)\right) = 0, \tag{6.14c}$$

$$\tilde{p}_{k+1} = p_{k+1/2} - \frac{h}{2}\left(V_q(q_{k+1}) + \mu^T(q_{k+1})\tilde{\lambda}_{k+1}\right), \tag{6.14d}$$

$$\mu(q_{k+1})M^{-1}\left(\tilde{p}_{k+1} + \frac{h}{2}V_q(q_{k+1})\right) = 0. \tag{6.14e}$$

- **Euler B**:

$$p_{k+1/2} = \tilde{p}_k - \frac{h}{2}\left(V_q(q_k) + \mu^T(q_k)\tilde{\lambda}_k\right), \tag{6.15a}$$

$$q_{k+1} = q_k + hM^{-1}p_{k+1/2}, \tag{6.15b}$$

$$\mu(q_k)M^{-1}\left(\tilde{p}_k - \frac{h}{2}V_q(q_k)\right) = 0, \tag{6.15c}$$

$$\tilde{p}_{k+1} = p_{k+1/2} - \frac{h}{2}\left(V_q(q_{k+1}) + \mu^T(q_{k+1})\tilde{\lambda}_{k+1}\right), \tag{6.15d}$$

$$\mu(q_{k+1})M^{-1}\left(\tilde{p}_{k+1} - \frac{h}{2}V_q(q_{k+1})\right) = 0. \tag{6.15e}$$

These numerical schemes provides $k+1$ values through an intermediate momentum step $k+1/2$, i.e:

$$(q_k, \tilde{p}_k, \tilde{\lambda}_k) \to (q_{k+1}, p_{k+1/2}, \tilde{\lambda}_k) \to (q_{k+1}, \tilde{p}_{k+1}, \tilde{\lambda}_{k+1}).$$

We recognize in (6.14c) and (6.15c) a Hamiltonian version for the discretization of the nonholonomic constraints (6.12b) and (6.13b) (Lagrangian version). These constraints are provided by the GNI equations (6.5b) or (6.6b).

Remark 6.3.2. *Method (6.12) (and the corresponding B version) clearly reminds the extension of the SHAKE method (see [151]) proposed by R. McLachlan and M. Permuttler [132] as a reversible method for nonholonomic systems* **not based** *on the discrete Lagrange-d'Alembert principle. Namely*

$$\begin{aligned} q_{k+1} - 2q_k + q_{k+1} &= -h^2M^{-1}\left(V_q(q_k) + \mu^T(q_k)\tilde{\lambda}_k\right), \\ 0 &= \mu(q_k)\left(\frac{q_{k+1} - q_{k-1}}{2h}\right). \end{aligned}$$

At the same time, SHAKE method is an extension of the classical Störmer-Verlet method in the presence of holonomic constraints. The RATTLE method is algebraically equivalent to SHAKE (this equivalence is shown in [102]). Its nonholonomic extension, that is:

$$\begin{aligned} &p_{k+1/2} = \tilde{p}_k - \frac{h}{2}\left(V_q(q_k) + \mu^T(q_k)\tilde{\lambda}_k\right), \\ &q_{k+1} = q_k + hM^{-1}p_{k+1/2}, \\ &\mu(q_k)M^{-1}\tilde{p}_k = 0, \\ &\tilde{p}_{k+1} = p_{k+1/2} - \frac{h}{2}\left(V_q(q_{k+1}) + \mu^T(q_{k+1})\tilde{\lambda}_{k+1}\right), \\ &\mu(q_{k+1})M^{-1}\tilde{p}_{k+1} = 0, \end{aligned}$$

clearly reminds (6.14).

As shown in [44], the nonholonomic SHAKE extension can be obtained by applying GNI equations to the discrete Lagrangian

$$L_d(q_k, q_{k+1}) = \frac{h}{2}L(q_k, \frac{q_{k+1} - q_k}{h}) + \frac{h}{2}L(q_{k+1}, \frac{q_{k+1} - q_k}{h}), \tag{6.16}$$

which also provides the Störmer-Verlet method in the variational integrators sense. Moreover, as shown in [45], the nonholonomic RATTLE method is globally second-order convergent.

Next we present one of the main results of this chapter

Theorem 6.3.3. *The nonholonomic extension of the Euler A (B) method is globally first-order convergent.*

The scheme of the proof is equivalent to that followed in [45] to show that the nonholonomic RATTLE method is second-order convergent. Therefore, it will be useful in the sequel to give a Hamiltonian version of nonholonomic equations (6.2) when $H(q,p) = \frac{1}{2}\, p^T M^{-1} p + V(q)$, namely

$$\begin{aligned} \dot q &= M^{-1}p, \\ \dot p &= -V_q - \mu^T\lambda, \\ \mu(q)\, M^{-1}p &= 0. \end{aligned}$$

Since the constraints are satisfied along the solutions, we can differentiate them with respect to time in order to obtain the actual values of the Lagrange multipliers, i.e.

$$\lambda = \mathcal{C}^{-1}\left(\mu_q[M^{-1}p, M^{-1}p] - \mu M^{-1}V_q\right),$$

where $\mathcal{C}(q) = \mu(q)M^{-1}\mu^T(q)$ is a regular matrix and $\mu_q[M^{-1}p, M^{-1}p]$ is the $m\times 1$ matrix $\frac{\partial \mu_i^\alpha}{\partial q^j}\left(M^{-1}\right)^{jj'} p_{j'} \left(M^{-1}\right)^{ii'} p_{i'}$. Taking this into account, the Hamiltonian nonholonomic system becomes

$$\dot q = M^{-1}p, \tag{6.17a}$$
$$\dot p = -V_q - \mu^T\mathcal{C}^{-1}\left(\mu_q[M^{-1}p, M^{-1}p] - \mu M^{-1}V_q\right), \tag{6.17b}$$

with initial condition satisfying $\mu(q)M^{-1}p = 0$.

Proof. We present the proof for Euler A method, the corresponding to Euler B is equivalent.

Consider the unconstrained problem

$$\begin{aligned} \dot q &= M^{-1}p, \\ \dot p &= \phi\,(q,p)\,, \end{aligned}$$

with a smooth enough function $\phi : \mathbb{R}^{2n} \to \mathbb{R}$. These equations can be discretized by

$$q_{k+1} = q_k + hp_{k+1/2}, \tag{6.18a}$$
$$p_{k+1/2} = p_{k-1/2} + h\phi\left(q_k, p_{k+1/2}\right), \tag{6.18b}$$

which is a first-order global convergent method, using standard arguments of Taylor expansions. Therefore, taking into account equations (6.17), from (6.18) we deduce the following first-order method for the nonholonomic system

$$q_{k+1} = q_k + hM^{-1}p_{k+1/2}, \tag{6.19a}$$

$$\begin{aligned} p_{k+1/2} = p_{k-1/2} &- hV_q\,(q_k) + h\mu^T(q_k)\,\mathcal{C}^{-1}(q_k)\,\mu\,(q_k)\,M^{-1}V_q\,(q_k) \\ &- h\mu^T(q_k)\,\mathcal{C}^{-1}(q_k)\,\mu_q[M^{-1}p_{k+1/2}, M^{-1}p_{k+1/2}]. \end{aligned} \tag{6.19b}$$

The next step is to prove that the nonholonomic Euler A method (6.14) reproduces (6.19). From equations (6.14) we see that the nonholonomic Euler A method assumes the form

$$\begin{aligned} q_{k+1} &= q_k + hM^{-1}p_{k+1/2}, \\ p_{k+1/2} &= p_{k-1/2} - hV_q(q_k) - h\mu^T(q_k)\tilde{\lambda}_k, \\ 0 &= \mu(q_k)M^{-1}\left(\frac{p_{k+1/2}+p_{k-1/2}}{2} + \frac{h}{2}V_q(q_k)\right) \end{aligned}$$

or, after some computations

$$q_{k+1} = q_k + hM^{-1}p_{k+1/2}, \tag{6.20a}$$

$$p_{k+1/2} = p_{k-1/2} - hV_q(q_k) - 2\mu^T(q_k)\mathcal{C}^{-1}(q_k)\mu(q_k)M^{-1}p_{k-1/2}. \tag{6.20b}$$

On the other hand we can expand the nonholonomic constraints around $q\,(0)$:

$$\mu\,(q\,(h))\,\dot{q}\,(h) = \mu\,(q\,(0))\,\dot{q}\,(0) + h\mu\,(q\,(0))\,\ddot{q}\,(0) + h\mu_q[\dot{q}\,(0)\,,\dot{q}\,(0)] + \mathcal{O}\left(h^2\right).$$

Since the constraints are satisfied at $t = 0$ and $t = h$, the previous expression becomes

$$h\mu\,(q\,(0))\,\ddot{q}\,(0) = -h\mu_q[\dot{q}\,(0)\,,\dot{q}\,(0)] + \mathcal{O}\left(h^2\right).$$

Now, taking standard approximations for first and second derivatives we deduce that

$$\begin{aligned} -2\mu(q_k)M^{-1}p_{k-1/2} &= -h\mu_q[M^{-1}p_{k+1/2}, M^{-1}p_{k+1/2}] \\ &+ h\mu(q_k)M^{-1}V_q(q_k) + \mathcal{O}(h^2). \end{aligned} \tag{6.21}$$

Therefore, substituting (6.21) in (6.20b) we recognize equation (6.19b) up to $\mathcal{O}(h^2)$ terms. Thus, we conclude that the nonholonomic Euler A method (6.14) is first-order convergent.

□

We recall now the definition of adjoint methods given in §3.4.4 since it will be useful in the following Theorem.

Definition 6.3.4. *For a one-step method $F : T^*Q \to T^*Q$, the adjoint method $F^* : T^*Q \to T^*Q$ is defined by*

$$(F^*)^h \circ F^{-h} = \mathrm{Id}_{T^*Q}.$$

Theorem 6.3.5. *The nonholonomic extensions of the Euler A and B methods are each other's adjoint.*

Proof. We will use a shorthand notation to define both integrators:

$$F_{A,B}(q_k, \tilde{p}_k, \tilde{\lambda}_k) = (q_{k+1}, \tilde{p}_{k+1}, \tilde{\lambda}_{k+1}).$$

Equations (6.14) and (6.15) can be rewritten to give a one step method instead of the leap-frog presented, namely:

$$q_{k+1}^A = q_k + hM^{-1}\tilde{p}_k - \frac{h^2}{2}M^{-1}V_q(q_k) - \frac{h^2}{2}M^{-1}\mu^T(q_k)\tilde{\lambda}_k, \tag{6.22a}$$

$$\tilde{p}_{k+1}^A = \tilde{p}_k - \frac{h}{2}V_q(q_k) - \frac{h}{2}\mu^T(q_k)\tilde{\lambda}_k - \frac{h}{2}V_q(q_{k+1}) - \frac{h}{2}\mu^T(q_{k+1})\tilde{\lambda}_{k+1}, \tag{6.22b}$$

$$\begin{aligned} 0 &= \mu(q_{k+1})M^{-1}\tilde{p}_k - \frac{h}{2}\mu(q_{k+1})M^{-1}V_q(q_k) - \frac{h}{2}\mu(q_{k+1})M^{-1}\mu^T(q_k)\tilde{\lambda}_k \\ &- \frac{h}{2}\mu(q_{k+1})M^{-1}\mu^T(q_{k+1})\tilde{\lambda}_{k+1}^A, \end{aligned} \tag{6.22c}$$

where $\tilde{p}_{k+1}^A$ and $\tilde{\lambda}_{k+1}^A$ are implicitly obtained from (6.22b) and (6.22c). The same occurs for F_B:

$$q_{k+1}^B = q_k + hM^{-1}\tilde{p}_k - \frac{h^2}{2}M^{-1}V_q(q_k) - \frac{h^2}{2}M^{-1}\mu^T(q_k)\tilde{\lambda}_k, \tag{6.23a}$$

$$\tilde{p}_{k+1}^B = \tilde{p}_k - \frac{h}{2}V_q(q_k) - \frac{h}{2}\mu^T(q_k)\tilde{\lambda}_k - \frac{h}{2}V_q(q_{k+1}) - \frac{h}{2}\mu^T(q_{k+1})\tilde{\lambda}_{k+1}, \tag{6.23b}$$

$$\begin{aligned} 0 &= -\mu(q_k)M^{-1}p_{k+1} - \frac{h}{2}\mu(q_k)M^{-1}\mu(q_k)\tilde{\lambda}_k - \frac{h}{2}\mu(q_k)M^{-1}V_q(q_{k+1}) \\ &- \frac{h}{2}\mu(q_k)M^{-1}\mu(q_{k+1})\tilde{\lambda}_{k+1}^B. \end{aligned} \tag{6.23c}$$

The point of the proof is to show that $F_A^h \circ F_B^{-h}(q_k, \tilde{p}_k, \tilde{\lambda}_k) = (q_k, \tilde{p}_k, \tilde{\lambda}_k)$. In order to that, we are going to use the following notation:

$$\begin{aligned} F_B^{-h}(q_k, \tilde{p}_k, \tilde{\lambda}_k) &= (q_{k+1}, \tilde{p}_{k+1}, \tilde{\lambda}_{k+1}) = (q_k', \tilde{p}_k', \tilde{\lambda}_k'), \\ F_B^{-h}(q_k', \tilde{p}_k', \tilde{\lambda}_k') &= (q_{k+1}', \tilde{p}_{k+1}', \tilde{\lambda}_{k+1}'), \end{aligned}$$

so that, the proof ends if $(q_{k+1}', \tilde{p}_{k+1}', \tilde{\lambda}_{k+1}') = (q_k, \tilde{p}_k, \tilde{\lambda}_k)$. After setting the time step as $-h$ and replacing (6.23a) and (6.23b) into (6.22a) is easy to check that $q_{k+1}' = q_k$. Furthermore, again fixing $-h$ as time step and taking into account equation (6.15e), from (6.23c) we arrive to

$$\begin{aligned} -\frac{h}{2}M^{-1}\mu(q_k')\tilde{\lambda}_k' &= \\ -M^{-1}\tilde{p}_k &- \frac{h}{2}V_q(q_k) - M^{-1}\tilde{p}_k' + \frac{h}{2}V_q(q_k') + \frac{h}{2}M^{-1}\mu^T(q_k)\tilde{\lambda}_k. \end{aligned}$$

Replacing this expression into (6.22c), considering that $q_{k+1}' = q_k$ and taking into account (6.14e) we find that

$$\frac{h}{2}\mu(q_k)M^{-1}\mu(q_k)^T\tilde{\lambda}_k - \frac{h}{2}\mu(q_k)M^{-1}\mu(q_k)^T\tilde{\lambda}_{k+1}',$$

which means

$$\tilde{\lambda}_{k+1}' = \tilde{\lambda}_k$$

since $\mathcal{C}(q_k)$ is regular. Finally, replacing (6.23b) into (6.22b) we find that $\tilde{p}_{k+1}' = \tilde{p}_k$. $\square$

Remark 6.3.6. *As shown in [124], the composition of Hamiltonian discrete flows, in the* **variational integrators** *sense, generated by the discrete Lagrangians (6.11) reproduces the RATTLE algorithm in the free case (that is, not constrained). More concretely, the composition*

$$F_{L_A}^{h/2} \circ F_{L_B}^{h/2}$$

provides the algorithm

$$\begin{aligned} p_{k+1/2} &= \tilde{p}_k - \frac{h}{2} V_q(q_k), \\ q_{k+1} &= q_k + h M^{-1} p_{k+1/2}, \\ \tilde{p}_{k+1} &= p_{k+1/2} - \frac{h}{2} V_q(q_{k+1}). \end{aligned}$$

Unfortunately, we have check that this is no longer true in the nonholonomic case, i.e., the composition (with time step $h/2$) of methods (6.14) and (6.15) does not reproduce the equations presented in remark 6.3.2. However, this composition still generates a second-order method since the intermediate steps are one-order methods which are each other's adjoint (as we have just proved).

6.4 Affine extension of the GNI

We consider in this section the case of affine noholonomic constraints (see §2.4.2 for the definition of affine constraints) determined by an affine bundle $\mathcal{A}$ of TQ modeled on a vector bundle $\mathcal{D}$. We will assume, in the sequel, that there exists a globally defined vector field $Y \in \mathfrak{X}(Q)$ such that $v_q \in \mathcal{A}_q$ if and only if $v_q - Y(q) \in \mathcal{D}_q$. Therefore, if $\mathcal{D}$ is determined by constraints $\mu_i^a(q)\dot{q}^i = 0$, then $\mathcal{A}$ is locally determined by the vanishing of the constraints

$$\phi^a\left(q^i, \dot{q}^i\right) = \mu_i^a(q)\left(\dot{q}^i - Y^i(q)\right) = 0, \;\; 1 \leq a \leq m. \tag{6.24}$$

where $Y = Y^i \frac{\partial}{\partial q^i}$.

In consequence, the initial data defining our **nonholonomic affine problem** is denoted by the 4-upla $(\mathcal{D}, \mathcal{G}, Y, V)$, where $\mathcal{D}$ is the distribution, $\mathcal{G}$ the Riemannian metric, Y the globally defined vector field and V is a potential function. By means of the metric, from Y, we can univocally define a 1-form, i.e. $\mathcal{G}(Y, \cdot) = \Pi \in \Lambda^1 Q$. Locally, $\Pi = \mathcal{G}_{ij} Y^j \, \mathrm{d}q^i$.

In terms of momenta the nonholonomic constraints (6.24) can be rewritten as

$$\mu_i^a(q)\, \mathcal{G}^{ij}\left(p_j - \Pi_j(q)\right) = 0. \tag{6.25}$$

where $p_i = \mathcal{G}_{ij}\dot{q}^j$.

Consider a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$. The proposed extension of the GNI method for affine nonholonomic equations is given by following equations:

$$\mathcal{P}_{q_k}^*\left(D_1 L_d\left(q_k, q_{k+1}\right)\right) + \mathcal{P}_{q_k}^*\left(D_2 L_d\left(q_{k-1}, q_k\right)\right) = 0, \tag{6.26a}$$
$$\mathcal{Q}_{q_k}^*\left(D_1 L_d\left(q_k, q_{k+1}\right)\right) - \mathcal{Q}_{q_k}^*\left(D_2 L_d\left(q_{k-1}, q_k\right)\right) + 2\mathcal{Q}_{q_k}^* \Pi = 0, \tag{6.26b}$$

where $\mathcal{Q}$ and $\mathcal{P}$ are the projectors defined in §6.1. Locally, the method (6.26) can be written as:

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = (\lambda_k)_b\, \mu^b(q_k), \tag{6.27a}$$

$$\mathcal{G}^{ij}(q_k)\mu_i^a(q_k)\left(\frac{\partial L_d}{\partial y^j}(q_k, q_{k+1}) - \frac{\partial L_d}{\partial x^j}(q_{k-1}, q_k) - 2\Pi_j(q_k)\right) = 0. \tag{6.27b}$$

Using the discrete Legendre transformations defined in §3.2, equation (6.27b) can be rewritten as

$$\mathcal{G}^{ij}(q_k)\mu_i^a(q_k)\left(\frac{\left(p^-_{k,k+1}\right)_j + \left(p^+_{k-1,k}\right)_j}{2} - \Pi_j(q_k)\right) = 0,$$

which corresponds to the discretization of the affine constraints (6.25) on the Hamiltonian side.

6.4.1 A theoretical example: nonholonomic SHAKE and RATTLE extensions for affine systems

Let consider again the mechanical Lagrangian $L(q,\dot q) = \frac{1}{2}\dot q^T\, M\, \dot q - V(q)$ and the discretization presented in (6.16). Applying the affine GNI equations (6.27) we obtain:

$$q_{k+1} - 2q_k + q_{k-1} = -h^2 M^{-1}\left(V_q(q_k) + \mu^T(q_k)\,\tilde\lambda_k\right), \tag{6.28a}$$

$$0 = \mu(q_k)\left(\frac{q_{k+1} - q_{k-1}}{2h} - Y(q_k)\right), \tag{6.28b}$$

which can be considered the nonholonomic extension of the SHAKE algorithm for affine systems. Denoting $\tilde p_k = M(q_{k+1} - q_{k-1})/2h$ and $p_{k+1/2} = M(q_{k+1} - q_k)/h$, from (6.28) we arrive to

$$\begin{aligned}
&p_{k+1/2} = \tilde p_k - \frac{h}{2}\left(V_q(q_k) + \mu^T(q_k)\tilde\lambda_k\right),\\
&q_{k+1} = q_k + hM^{-1}p_{k+1/2},\\
&\mu(q_k)M^{-1}\left(\tilde p_k - \Pi(q_k)\right) = 0,\\
&\tilde p_{k+1} = p_{k+1/2} - \frac{h}{2}\left(V_q(q_{k+1}) + \mu^T(q_{k+1})\tilde\lambda_{k+1}\right),\\
&\mu(q_{k+1})M^{-1}\left(\tilde p_{k+1} - \Pi(q_{k+1})\right) = 0,
\end{aligned}$$

which can be considered the nonholonomic extension of the RATTLE algorithm for affine nonholonomic systems.

6.5 Reduced systems

In this section we are going to consider configuration spaces of the form $Q = M \times G$, where M is a n-dimensional differentiable manifold and G is a m-finite-dimensional Lie group ($\mathfrak{g}$ will be its corresponding Lie algebra.) Therefore, there exists a global canonical splitting between variables describing the position and variables describing the orientation of the mechanical system. Then, we distinguish the pose coordinates $g \in G$, and the variables describing the internal shape of the system, that is $x \in M$ (in consequence $(x, \dot{x}) \in TM$). It is clear that $Q = M \times G$ is the total space of a trivial principal G-bundle over M, where the bundle projection $\phi : Q \to M$ is just the canonical projection on the first factor. We may consider the corresponding reduced tangent space $E = TQ/G$ over M. Identifying the tangent bundle to G with $G \times \mathfrak{g}$ by using left translation, thus, the reduced tangent space $E = TQ/G$ is isomorphic to the product manifold $TM \times \mathfrak{g}$ and the vector bundle projection is $\tau_M \circ pr_1$, where $pr_1 : TM \times \mathfrak{g} \to TM$ and $\tau_M : TM \to M$ are the canonical projections.

6.5.1 The case of linear constraints

Now suppose that $(\mathcal{G}, \mathcal{D}, V)$ is an standard mechanical nonholonomic system on TQ such that all the ingredients are G-invariant. In other words

$$
\begin{array}{rcl}
\mathcal{G}_{x,g}((X_x, g\xi), (Y_x, g\eta)) & = & \mathcal{G}_{x,e}((X_x, \xi), (Y_x, \eta)), \ \forall X_x, Y_x \in T_xM, \xi, \eta \in \mathfrak{g} \\
(X_x, \xi) \in \mathcal{D}_{(x,e)} & & \text{then } (X_x, g\xi) \in \mathcal{D}_{(x,g)} \\
V(x, g) & = & V(x, e) \equiv \tilde{V}(x) \ .
\end{array}
$$

Therefore, we obtain a new triple $(\tilde{\mathcal{G}}, \widetilde{\mathcal{D}}, \tilde{V})$ on $TM \times \mathfrak{g}$ where $\tilde{\mathcal{G}} : (TM \times \mathfrak{g}) \times (TM \times \mathfrak{g}) \longrightarrow \mathbb{R}$ is a bundle metric, $\widetilde{\mathcal{D}}$ is a vector subbundle of $TM \times \mathfrak{g} \to M$ and $\tilde{V} : M \to \mathbb{R}$ is the reduced potential. With all these ingredients it is possible to write the reduced nonholonomic equations or **nonholonomic Lagrange-Poincaré equations** (see [16, 37] for all the details, also for the non-trivial case).

Our objective is to find a discrete version of the GNI for the nonholonomic Lagrange-Poincaré equations. As in the previous sections, we can split the total space E as $E = \tilde{\mathcal{D}} \oplus \tilde{\mathcal{D}}^{\perp}$, now using the fibered metric $\tilde{\mathcal{G}}$, and regard the corresponding projectors $\mathcal{P} : E \to \tilde{\mathcal{D}}$, $\mathcal{Q} : E \to \tilde{\mathcal{D}}^{\perp}$. In order to propose the discrete nonholonomic equations, is necessary to set a discrete Lagrangian $L_d : M \times M \times G \to \mathbb{R}$, and the discrete Legendre transforms. Namely (see [118]):

$$
\begin{array}{rcl}
\mathbb{F}L_d^- : M \times M \times G & \to & T^*M \times \mathfrak{g}^* \\
(x_k, x_{k+1}, g_k) & \longmapsto & (x_k, -D_1L_d(x_k, x_{k+1}, g_k), R^*_{g_k} D_3L_d(x_k, x_{k+1}, g_k)), \\
\mathbb{F}L_d^+ : M \times M \times G & \to & T^*M \times \mathfrak{g}^* \\
(x_k, x_{k+1}, g_k) & \longmapsto & (x_{k+1}, D_2L_d(x_k, x_{k+1}, g_k), L^*_{g_k} D_3L_d(x_k, x_{k+1}, g_k)) \ .
\end{array}
$$

Thus, the proposed **reduced GNI equations** are

$$
\mathcal{P}^*_{x_k}\left(\mathbb{F}L_d^-(x_k, x_{k+1}, g_k)\right) - \mathcal{P}^*_{x_k}\left(\mathbb{F}L_d^+(x_k, x_{k+1}, g_k)\right) = 0, \tag{6.29a}
$$

$$
\mathcal{Q}^*_{x_k}\left(\mathbb{F}L_d^-(x_k, x_{k+1}, g_k)\right) + \mathcal{Q}^*_{x_k}\left(\mathbb{F}L_d^+(x_k, x_{k+1}, g_k)\right) = 0, \tag{6.29b}
$$

where the subscript x_k emphasizes the fact that the projections take place in the fiber over x_k.

To understand why (6.29b) represents a discretization of the nonholonomic constraints, we will work in local coordinates. Take now local coordinates (x^i) on M and a local basis of sections $\{\tilde{e}_\alpha, \tilde{e}_a\}$ of $\Gamma(TM \times \mathfrak{g})$ adapted to the decomposition $\tilde{\mathcal{D}} \oplus \tilde{\mathcal{D}}^\perp$, that is $\tilde{e}_\alpha(x) \in \widetilde{\mathcal{D}}_x$ and $\tilde{e}_a(x) \in \widetilde{\mathcal{D}}_x^\perp$, for all $x \in M$. We have that

$$\tilde{\mathcal{G}}(\tilde{e}_\alpha, \tilde{e}_\beta) = \tilde{\mathcal{G}}_{\alpha\beta}, \quad \tilde{\mathcal{G}}(\tilde{e}_a, \tilde{e}_\beta) = 0, \quad \tilde{\mathcal{G}}(\tilde{e}_a, \tilde{e}_b) = \tilde{\mathcal{G}}_{ab}\ .$$

Consider the induced adapted local coordinates (x^i, y^α, y^a) for $\Gamma(TM \times \mathfrak{g})$. The nonholonomic constraints are represented by $y^a = 0$ on E. Taking the dual basis $\{\tilde{e}^\alpha, \tilde{e}^a\}$ of $\Gamma(T^*M \times \mathfrak{g}^*)$, we have induced local coordinates (x^i, p_α, p_a) on the hamiltonian side, now the nonholonomic constraints are represented $p_a = 0$.

On the other hand, the projector $\mathcal{Q}$ has the following expression in this basis

$$\mathcal{Q} = \tilde{e}^a \otimes \tilde{e}_a, \tag{6.30}$$

Define the pre- and post-momenta by

$$\begin{aligned} p^-_{x_k} &= \mathbb{F}L^-_d(x_k, x_{k+1}, g_k) \in T^*_{x_k}M \times \mathfrak{g}^*, \\ p^+_{x_k} &= \mathbb{F}L^+_d(x_k, x_{k+1}, g_k) \in T^*_{x_k}M \times \mathfrak{g}^*. \end{aligned}$$

Finally, looking at equations (6.29b) and (6.30) we realize that

$$\mathcal{Q}^*_{x_k}\left(\frac{p^+_{x_k} + p^-_{x_k}}{2}\right) = 0\ . \tag{6.31}$$

If $p^+_{x_k} = p^+_\alpha e^\alpha(x_k) + p^+_a e^a(x_k)$ and $p^-_{x_k} = p^-_\alpha e^\alpha(x_k) + p^-_a e^a(x_k)$, then condition (6.31) is expressed as

$$\frac{p^+_a + p^-_a}{2} = 0,$$

which means that the average of post and pre-momenta satisfies the nonholonomic constraints written in the Hamiltonian side.

6.5.2 A theoretical example: RATTLE algorithm for reduced spaces

Let consider $M = \mathbb{R}^n$. Thus, $Q = \mathbb{R}^n \times G$ and $E = TQ/G \cong T\mathbb{R}^n \times \mathfrak{g}$. Take a basis $\{E_s\}$ of the Lie algebra $\mathfrak{g}$, and consider the following local basis of $\Gamma(T\mathbb{R}^n \times \mathfrak{g})$

$$\left\{\left(\frac{\partial}{\partial x^i}, 0\right), (0, E_s)\right\}.$$

Therefore, its dual basis is generated by

$$\left\{\left(\mathrm{d}x^i, 0\right), (0, E^s)\right\} \simeq \left\{\mathrm{d}x^i, E^s\right\}.$$

In this base of sections the metric $\tilde{\mathcal{G}}$ is written as

$$\tilde{\mathcal{G}} = \tilde{\mathcal{G}}_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j + \tilde{\mathcal{G}}_{it}\mathrm{d}x^i \otimes E^t + \tilde{\mathcal{G}}_{sj}E^s \otimes \mathrm{d}x^j + \tilde{\mathcal{G}}_{st}E^s \otimes E^t,$$

Assume that, in this local basis of section, the coefficients of the metric are constant, that is, they do not depend on the base coordinates x. For instance, a typical example would be

$$L(x, \dot{x}, \xi) = \frac{1}{2}\dot{x}^T M \dot{x} + \frac{1}{2}\langle \xi, \mathbb{I}\xi\rangle$$

where M is a constant regular symmetric matrix and the configuration manifold $Q = \mathbb{R}^n \times G$. In addition, $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$ is a symmetric positive definite inertia operator. More concretely, M has the following form:

$$M = \begin{pmatrix} \tilde{\mathcal{G}}_{ij} & \tilde{\mathcal{G}}_{it} \\ \tilde{\mathcal{G}}_{sj} & \tilde{\mathcal{G}}_{st} \end{pmatrix}$$

Consider the following discrete Lagrangian $L_d \colon \mathbb{R}^n \times \mathbb{R}^n \times G \to \mathbb{R}$ defined by

$$\begin{aligned} L_d(x_k, x_{k+1}, g_k) = \frac{h}{2}\tilde{\mathcal{G}}_{ij}\left(\frac{x^i_{k+1} - x^i_k}{h}\right)\left(\frac{x^j_{k+1} - x^j_k}{h}\right) + h\tilde{\mathcal{G}}_{it}\left(\frac{x^i_{k+1} - x^i_k}{h}\right)\frac{\left(\tau^{-1}(g_k)\right)^t}{h} \\ + \frac{h}{2}\tilde{\mathcal{G}}_{st}\frac{\left(\tau^{-1}(g_k)\right)^s}{h}\frac{\left(\tau^{-1}(g_k)\right)^t}{h} - \frac{h}{2}\left(V(x_k) + V(x_{k+1})\right), \end{aligned}$$

where $\tau : \mathfrak{g} \to G$ is a general retraction map (see §5.2.3 for its definition and properties). Observe that $\tau^{-1}(g_k) \in \mathfrak{g}$ and $\tau^{-1}(g_k) = \left(\tau^{-1}(g_k)\right)^s E_s$.

Additionally, we have the vector subbundle $\widetilde{\mathcal{D}}$ of $T\mathbb{R}^n \times \mathfrak{g}$ prescribing the nonholonomic constraints. Assume that $\widetilde{\mathcal{D}}^0 = \text{span}\,\{\mu^a_i\,(dx^i, 0) + \eta^a_s(0, E^s)\}$.

Under these statements, the equation of the GNI method (6.29a) splits into

$$\begin{aligned} \frac{1}{h}\tilde{\mathcal{G}}_{ij}(x^j_{k+1} - 2x^j_k + x^j_{k-1}) + \frac{1}{h}\tilde{\mathcal{G}}_{is}\left((\tau^{-1}(g_k))^s - (\tau^{-1}(g_{k-1})^t)\right) \\ + hV_{x^i}(x_k) = -\lambda_{a,k}\,\mu^a_i(x_k), \end{aligned} \tag{6.32a}$$

$$L^*_{g_{k-1}} D_3 L_d(x_{k-1}, x_k, g_{k-1}) - R^*_{g_k} D_3 L_d(x_k, x_{k+1}, g_k) = \lambda_{a,k}\,\eta^a_s(x_k)\,E^s, \tag{6.32b}$$

where V_{x^i} stands for $\partial V/\partial x^i$, and $\lambda_{a,k}$ are the Lagrange multipliers which might vary in each step. From (6.29a) is easy to see that

$$\mathbb{F}L^-_d(x_k, x_{k+1}, g_k) - \mathbb{F}L^+_d(x_{k-1}, x_k, g_{k-1}) \in \widetilde{\mathcal{D}}^0(x_k).$$

Taking into account the definition of the right-trivialized tangent retraction map (also given in §5.2.3) and lemmae presented in Appendix A, equation (6.32b) can be rewritten as

$$(\mathrm{d}\tau^{-1}_{-h\xi_k})^*\mathrm{d}l_d\big|_{(x_{k-1},x_k)}(\xi_k) - (\mathrm{d}\tau^{-1}_{h\xi_{k+1}})^*\mathrm{d}l_d\big|_{(x_k,x_{k+1})}(\xi_{k+1}) = \lambda_{a,k}\,\eta^a_t(x_k)\,E^t, \tag{6.33}$$

where $l_d\big|_{(x_k,x_{k+1})}(\xi_k) = L_d\big|_{(x_k,x_{k+1})}(\tau(h\xi_k))$, being $g_k = \tau(h\xi_k)$. However, thanks to the retraction map, in most applications the choice for l_d is just $l_d\big|_{(x,\dot{x})}(\xi_k) = hL\big|_{(x,\dot{x})}(\xi_k)$.

As we already know (6.29b) provides a discretization of the nonholonomic constraints on the hamiltonian side:

$$\begin{aligned} & A^{i,a}(x_k)\left(\tilde{\mathcal{G}}_{ij}\frac{(x^j_{k+1}-x^j_{k-1})}{2h}+\frac{1}{2h}\tilde{\mathcal{G}}_{is}((\tau^{-1}(g_k))^s+(\tau^{-1}(g_{k-1}))^s)\right) \\ & +\frac{1}{2}B^{a,s}(x_k)\left(L^*_{g_{k-1}}D_3L_d(x_{k-1},x_k,g_{k-1})+R^*_{g_k}D_3L_d(x_k,x_{k+1},g_k)\right)_s=0, \end{aligned} \tag{6.34}$$

or, equivalently,

$$\begin{aligned} & A^{i,a}(x_k)\left(\tilde{\mathcal{G}}_{ij}\frac{(x^j_{k+1}-x^j_{k-1})}{2h}+\frac{1}{2}\tilde{\mathcal{G}}_{it}(\xi^t_k+\xi^t_{k-1})\right) \\ & +\frac{1}{2}B^{a,t}(x_k)\left((\mathrm{d}\tau^{-1}_{-h\xi_{k-1}})^*\mathrm{d}l_d\big|_{(x_{k-1},x_k)}(\xi_{k-1})+(\mathrm{d}\tau^{-1}_{h\xi_k})^*\mathrm{d}l_d\big|_{(x_k,x_{k+1})}(\xi_k)\right)_t=0, \end{aligned}$$

where

$$\begin{aligned} A^{i,a}(x_k) &= (\tilde{\mathcal{G}}^{-1})^{ij}\mu^a_j(x_k)+(\tilde{\mathcal{G}}^{-1})^{it}\eta^a_t(x_k), \\ B^{t,a}(x_k) &= (\tilde{\mathcal{G}}^{-1})^{ti}\mu^a_i(x_k)+(\tilde{\mathcal{G}}^{-1})^{ts}\eta^a_s(x_k), \end{aligned}$$

being $(\tilde{\mathcal{G}}^{-1})$ the inverse matrix of $(\tilde{\mathcal{G}})=\begin{pmatrix}\tilde{\mathcal{G}}_{ij} & \tilde{\mathcal{G}}_{sj}\\ \tilde{\mathcal{G}}_{it} & \tilde{\mathcal{G}}_{st}\end{pmatrix}$.

Our aim in the following is to find an extension to the nonholonomic RATTLE algorithm presented in remark 6.3.2 for systems defined on $T\mathbb{R}^n\times\mathfrak{g}$. With that purpose we define

$$\begin{aligned} (\tilde{p}_k)_i &= \tilde{\mathcal{G}}_{ij}\frac{(x^j_{k+1}-x^j_{k-1})}{2h}+\frac{1}{2}\tilde{\mathcal{G}}_{is}(\xi^s_k+\xi^s_{k-1}), \\ (p_{k+1/2})_i &= \tilde{\mathcal{G}}_{ij}\frac{(x^j_{k+1}-x^j_k)}{h}+\frac{1}{h}\tilde{\mathcal{G}}_{is}\xi^s_{k+1}\in T^*_{x_k}M, \\ \tilde{\Phi}_k &= (\mathrm{d}\tau^{-1}_{h\xi_k})^*\mathrm{d}l_d\big|_{(x_k,x_{k+1})}(\xi_k), \\ \Phi_{k+1/2} &= \mathrm{Ad}^*_{\tau(h\xi_k)}\tilde{\Phi}_k-\frac{1}{2}\tilde{\lambda}_{a,k}\,\eta^a_s(x_k)\,E^s, \end{aligned}$$

where $\tilde{\lambda}_{a,k}=\lambda_{a,k}/h$. We also recall that $\xi_k=\tau^{-1}(g_k)/h$. After these redefinitions, equations (6.32a), (6.33) and (6.34) can be translated into the following algorithm

$$p_{k+1/2} = \tilde{p}_k - \frac{h}{2}\left(V_x(x_k) + \tilde{\lambda}_{a,k}\,\mu^a(x_k)\right), \tag{6.35a}$$

$$\Phi_{k+1/2} = \mathrm{Ad}^*_{\tau(h\xi_k)}\tilde{\Phi}_k - \frac{1}{2}\tilde{\lambda}_{a,k}\,\eta^a(x_k), \tag{6.35b}$$

$$x^i_{k+1} = x^i_k + h(\tilde{\mathfrak{G}}^{-1})^{ij}\left((p_{k+1/2})_j - \tilde{\mathfrak{G}}_{it}\,\xi^t_{k+1}\right), \tag{6.35c}$$

$$A^a(x_k)\,\tilde{p}_k + B^a(x_k)\left(\mathrm{Ad}^*_{\tau(h\xi_k)}\,\tilde{\Phi}_k + \tilde{\Phi}_{k+1}\right) = 0, \tag{6.35d}$$

$$\tilde{p}_{k+1} = p_{k+1/2} - \frac{h}{2}\left(V_x(x_{k+1}) + \tilde{\lambda}_{a,k+1}\,\mu^a(x_{k+1})\right), \tag{6.35e}$$

$$\tilde{\Phi}_{k+1} = \Phi_{k+1/2} - \frac{1}{2}\tilde{\lambda}_{a,k}\,\eta^a(x_k), \tag{6.35f}$$

$$A^a(x_{k+1})\,\tilde{p}_{k+1} + B^a(x_{k+1})\left(\mathrm{Ad}^*_{\tau(h\xi_{k+1})}\,\tilde{\Phi}_{k+1} + \tilde{\Phi}_{k+2}\right) = 0, \tag{6.35g}$$

where $\eta^a(x_k) = \eta^a_t(x_k)\,E^t$; moreover, most of the equations are written in a matrix form. Departing from initial data $(x_k, \tilde{p}_k, \xi_k, \tilde{\Phi}_k, \tilde{\lambda}_{a,k})$, equations (6.35a) and (6.35b) determine $p_{k+1/2}$ and $\Phi_{k+1/2}$. Furthermore, (6.35c) and (6.35d) determine x_{k+1} and ξ_{k+1} once we know $p_{k+1/2}$ and $\Phi_{k+1/2}$ (notice, from (6.35b) and (6.35f), that $\tilde{\Phi}_{k+1}$ is completely determined in terms of $\tilde{\Phi}_k$, $\tilde{\lambda}_{a,k}$ and x_k). The definition of $\tilde{p}_{k+1}$ and $\tilde{\Phi}_{k+1}$ needs a step forward in (6.32a), (6.33) and the constraint equation (6.34). That step is encoded in the last three equations of the algorithm. Once we have determined $(x_{k+1}, p_{k+1/2}, \xi_{k+1}, \Phi_{k+1/2}, \tilde{\lambda}_{a,k})$, we calculate $\tilde{\Phi}_{k+1}$ by (6.35f) and the remaining $\tilde{p}_{k+1}$ and $\tilde{\lambda}_{a,k+1}$ using (6.35e) and (6.35g) (notice that $\tilde{\Phi}_{k+2}$ is completely determined in terms of $\tilde{\Phi}_{k+1}$, $\tilde{\lambda}_{a,k+1}$). Summarizing, our algorithm follows the squeme:

$$(x_k, \tilde{p}_k, \xi_k, \tilde{\Phi}_k, \tilde{\lambda}_{a,k}) \to (x_{k+1}, p_{k+1/2}, \xi_{k+1}, \Phi_{k+1/2}, \tilde{\lambda}_{a,k}),$$
$$(x_{k+1}, p_{k+1/2}, \xi_{k+1}, \Phi_{k+1/2}, \tilde{\lambda}_{a,k}) \to (x_{k+1}, \tilde{p}_{k+1}, \xi_{k+1}, \tilde{\Phi}_{k+1}, \tilde{\lambda}_{a,k+1}),$$

and that's why we consider it the **nonholonomic RATTLE extension** for $T\mathbb{R}^n \times \mathfrak{g}$ reduced systems.

Remark 6.5.1. *After the redefinition of our variables presented above, equation* (6.34) *can be translated into*

$$\tilde{\Phi}_{k+1} = \mathrm{Ad}^*_{\tau(h\xi_k)}\tilde{\Phi}_k - \tilde{\lambda}_{a,k}\,\eta^a(x_k).$$

This equation clearly shows that $\tilde{\Phi}_{k+1}$ can be completely determined in terms of $\tilde{\Phi}_k, \tilde{\lambda}_{a,k}, x_k$; $\tilde{\Phi}_{k+2}$ in terms of $\tilde{\Phi}_{k+1}, \tilde{\lambda}_{a,k+1}, x_{k+1}$, etc. One could say that the intermediate $k+1/2$ step given in equations (6.35b) *and* (6.35f) *is arbitrary and completely unnecessary. Nevertheless, we fix that intermediate step in order to preserve the RATTLE structure.*

6.6 Extension to Lie groupoids

Let $G \rightrightarrows Q$ be a Lie groupoid and $\tau_{AG} : AG \to Q$ its associated Lie algebroid (recall both notions introduced in §1.5). Consider a mechanical system subjected to linear nonholonomic constraints, that is, a pair $(L, \mathcal{D})$ (see [118, 70] for more details), where *i)* $L : AG \to \mathbb{R}$ is a Lagrangian function of mechanical type

$$L(a) = \frac{1}{2}\,\mathcal{G}(a,a) - V(\tau_{AG}(a)), \;\; \text{where}\, a \in AG.$$

ii) $\mathcal{D}$ is the total space of a vector subbundle $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ of AG.

Here $\mathcal{G} : AG \times_Q AG \to \mathbb{R}$ is a bundle metric on AG. We also consider the orthogonal decomposition $AG = \mathcal{D} \oplus \mathcal{D}^{\perp}$ and the associated projectors $\mathcal{P} : AG \to \mathcal{D}$ and $\mathcal{Q} : AG \to \mathcal{D}^{\perp}$.

Consider a discretization $L_d : G \to \mathbb{R}$ of the Lagrangian L. It is possible to define two Legendre transformations $\mathbb{F}L_d^{\pm} \colon G \to A^*G$ by

$$\begin{aligned}\mathbb{F}L_d^{-}(h)(v_{\epsilon(\alpha(h))}) &= -v_{\epsilon(\alpha(h))}(L_d \circ r_h \circ i),\\ \mathbb{F}L_d^{+}(g)(v_{\epsilon(\beta(g))}) &= v_{\epsilon(\beta(g))}(L_d \circ L_g),\end{aligned}$$

(L_d denotes the discrete Lagrangian while here L_g denotes the left-translation in the groupoid by $g \in G$, see §1.5) where $v_{\epsilon(\alpha(h))} \in A_{\alpha(h)}G$ and $v_{\epsilon(\beta(g))} \in A_{\beta(g)}G$. Therefore $\mathbb{F}L_d^{-}(h) \in A^*_{\alpha(h)}G$ and $\mathbb{F}L_d^{+}(g) \in A^*_{\beta(g)}G$. Since the Euler-Lagrange equations are given by the matching of momenta, in the Lie groupoid setting they read

$$\mathbb{F}L_d^{-}(h) = \mathbb{F}L_d^{+}(g),$$

where (g,h) is in the set G_2.

The proposed **nonholonomic integrator** is

$$\mathcal{P}_q^*\left(\mathbb{F}L_d^{-}(h) - \mathbb{F}L_d^{+}(g)\right) = 0 \tag{6.36a}$$
$$\mathcal{Q}_q^*\left(\mathbb{F}L_d^{-}(h) + \mathbb{F}L_d^{+}(g)\right) = 0, \tag{6.36b}$$

where the subscript q emphasizes the fact that the projections take place in the fiber over q_k. Let $\{X_\alpha, X_a\}$ be a local basis adapted to $\mathcal{D} \oplus \mathcal{D}^{\perp}$, in the sense that locally $\mathcal{D} = \operatorname{span}\{X_\alpha\}$ and $\mathcal{D}^{\perp} = \operatorname{span}\{X_a\}$. We can rewrite equations (6.36) as

$$\mathbb{F}L_d^{-}(h)\,(X_\alpha(q)) - \mathbb{F}L_d^{+}(g)\,(X_\alpha(q)) = 0, \tag{6.37a}$$
$$\mathbb{F}L_d^{-}(h)\,(X_a(q)) + \mathbb{F}L_d^{+}(g)\,(X_a(q)) = 0, \tag{6.37b}$$

where $q \in Q$ and $(g,h) \in G_2$ (that is, are composable). Let us denote

$$\begin{aligned}p_g^{+} &= \mathbb{F}L_d^{+}(g) \in A_q^*G,\\ p_h^{-} &= \mathbb{F}L_d^{-}(h) \in A_q^*G,\end{aligned}$$

so equation (6.37b) becomes

$$\left(\frac{p_g^{+} + p_h^{-}}{2}\right)(X_a(q)) = 0.$$

If $\mu^a \in \Gamma(A^*G)$ are such that $\mathcal{D}^{\circ} = \operatorname{span}\{\mu^a\}$, then this last equation becomes

$$\mathcal{G}\left(\frac{p_g^{+} + p_h^{-}}{2}, \mu^a\right) = 0,$$

where, by a slight abuse of notation, we denote the bundle metric A^*G naturally induced by the bundle metric on AG using the same symbol $\mathcal{G}$. Note that the set of $\eta \in A^*G$ such that $\mathcal{G}(\eta, \mu^a) = 0$ for all a form the constraint submanifold $\bar{\mathcal{D}} = \mathrm{Leg}_{\mathcal{G}}(\mathcal{D})$. Therefore the average momentum $\tilde{p} = (p_g^+ + p_h^-)/2 \in \tilde{\mathcal{D}}$ satisfies in this sense the constraint equations.

Conclusions and future work

The closing chapter of this memory is devoted to summarize the contributions of the work. An outlook of the future research is also provided.

Chapter 4 has been devoted to the geometric study of the relationship between Hamiltonian dynamics and constrained variational calculus. Under regularity conditions we find that both are equivalent. This result, which is one of the main contributions of this thesis, is presented in Theorem 4.2.4. Furthermore, we have extended the interest in this connection to the discrete framework. As a result, we find the relationship between symplectic integrators and discrete variational calculus in presence of constraints, which is enclosed in Theorem 4.3.4. This relationship enlightens the geometric structure of symplectic integrators in terms of Lagrangian submanifolds. How to take advantage of the previous result in the context of geometrical integration of Hamiltonian systems is illustrated with several examples in §4.4. We have also analyzed in parallel the case of classical nonholonomic mechanics in the discrete (§4.3.3) and continuous cases (§4.2.3). Moreover, a natural algorithm for comparison of nonholonomic solutions and constrained variational solutions is given, generalizing the one provided in [35].

In chapter 5 we have built numerical methods for mechanical optimal control problems. More concretely, we extend the the theory of discrete mechanics to enable solutions for this kind of problems by means of the discretization of variational principles. Our method is based in the following key idea: to solve the optimal control problem as a variational integrator of a specially constructed higher-dimensional system. More concretely, the Lagrangian formalism where the mentioned variational integrators come from is introduced in §5.1.4 and §5.2.4, for systems defined on tangent bundles and Lie algebras, respectively. On the other hand, the specific relationships between the control space and the velocity phase space of the studied problems, relationships our approach is built upon, are established in the following definitions:

- TQ: def.5.1.5 and def.5.1.7, fully-actuated case and under-actuated case respectively.
- $\mathfrak{g}$: def.5.2.4 and def.5.2.5, fully-actuated case and under-actuated case respectively.

Simulations for the optimal control of an underwater vehicle are shown in fig.5.1 and fig.5.2. We would like to stress the adaptability of the proposed techniques to reduced problems and systems subject to nonholonomic constraints. These extensions are discussed in §5.3. Moreover, one of the outstanding features of the proposed method is the capacity to allow discontinuous controls. This capacity is presented in §5.2.10 and depicted in fig.5.3. As a consequence of this property, we expect that our techniques will be widely applied in future research concerning optimal control and its applications to engineering, since discontinuous input controls are very usual in practical problems.

Chapter 6 accounts for new developments regarding the Geometric Nonholonomic Integrator. Namely, we have extended the results in [44, 45, 93] to the case of affine constraints (§6.4) and reduced systems (§6.5). Both cases are illustrated with a theoretical example: the nonholonomic extension of the RATTLE algorithm within the GNI setting. Finally, we find a generalized setting for the GNI integrator in the framework of Lie groupoids (§6.6).

In [44, 45] was shown that the nonholonomic SHAKE method can be obtained by applying the GNI equations to a specific discretization of the mechanical Lagrangian. Moreover, it is proved that the associated RATTLE method (which is equivalent to the SHAKE) is globally second-order convergent. This result establishes a clear parallelism with the unconstrained case. Following this line, we have applied the GNI to the discretizations that further produce the symplectic-Euler methods as variational integrators. We have denoted these methods as Euler A and Euler B. In Theorem 6.3.3 we prove that Euler A and Euler B methods are globally first-order convergent. Furthermore, we prove in Theorem 6.3.5 that Euler A and Euler B methods are adjoint to each other. Since symplectic-Euler methods are first-globally convergent and adjoint to each other, these two results, apart from being some of the more important contributions of chapter 6, reinforce the parallelism with the unconstrained mechanics mentioned above.

Finally, in §6.5 we have studied how to extend the GNI method to other cases of interest, namely reduced systems and systems subjected to affine nonholonomic constraints.

The results presented in this thesis have been published in

- KOBILAROV M, JIMÉNEZ F AND MARTÍN DE DIEGO D, *Discrete Variational Optimal Control.* Preprint, arXiv:1203.0580, (2012), submitted for publication to *Journal of Nonlinear Science.* Some results in this paper have been further applied in
  - COLOMBO L, JIMÉNEZ F AND MARTÍN DE DIEGO D, *Discrete second-order Euler-Poincaré equations. Applications to optimal control.* Journal of Geometric methods of Modern Physics, **9**, (2012), arXiv:1109.4716.
- JIMÉNEZ F, DE LEÓN M, AND MARTÍN DE DIEGO D, *Hamiltonian dynamics and constrained variational calculus: continuous and discrete settings.* Accepted for publication in *Journal of Physics A.* arXiv:1108.5570, (2011).
- JIMÉNEZ F AND MARTÍN DE DIEGO D, *Continuous and discrete approaches to vakonomic mechanics.* Revista de la Real Academia de las Ciencias Exactas, Físicas y Naturales, Serie A. Springer, **106(1)**, pp. 75-87, (2011), DOI.
- JIMÉNEZ F AND MARTÍN DE DIEGO D, *A geometric approach to Discrete mechanics for optimal control theory.* Proceedings of the 49th IEEE Conference on Decision

and Control (CDC), pp. 5426-5431. Atlanta, Georgia, USA, (2010).

In addition, the results contained in this thesis have been presented in the following international meetings:

- Discrete variational optimal control (F Jiménez). *Friedrich-Alexander University of Nürnberg-Erlangen, Chair of Applied Dynamics*, invited talk, Erlangen, Germany, January 2012.

- On discrete mechanics for optimal control theory (F Jiménez). *IMAC Symposium on Dynamical Systems: Trends and Perspectives*, Castellón, Spain, September 2011.

- A new approach to discrete optimal control theory (F Jiménez). *California Institute of Technology*, invited talk, Pasadena, U.S.A., August 2011.

- New developments in discrete variational calculus and discrete optimal control theory (D Martín de Diego). *Mathematisches Forschungsinstitut Oberwolfach*, Germany, August 2011.

- Lagrangian submanifolds and constrained variational calculus (F Jiménez). *Workshop on Rough Paths and Combinatorics in Control Theory*, San Diego, U.S.A., July 2011.

- Lagrangian submanifolds and constrained variational calculus (F Jiménez). 5^{th} *International Summer School on Geometry, Mechanics and Control*, Madrid, Spain, June 2011.

- Lagrangian submanifolds and discrete constrained mechanics (D Martín de Diego). *XI Congreso Dr. Antonio Monteiro*, Bahía Blanca, Argentina, May 2011.

- A geometric approach to discrete mechanics for optimal control theory (D Martín de Diego). *49th IEEE Conference on Decision and Control*, Atlanta, U.S.A., December 2010.

- Optimal control of mechanical systems on Lie groups (D Martín de Diego). *VI International symposium HAMSYS-2010, Honoring EA Lacomba in his 65th Anniversary*, invited talk, Mexico city, Mexico, December 2010.

- Optimal control of mechanical systems on Lie groups: continuous and discrete (D Martín de Diego). *Workshop on Geometry of Constraints and Control - New Developments*, Banach Center, Warszawa, Poland, November 2010.

Future work

Once we have explored the insight that the geometry (both continuous and discrete) and numerical methods can provide on mechanical systems, there remains a lot to be done.

Higher-order methods. Constrained Lagrangian systems with symmetries

Many higher-order symplectic methods are obtained using composition of methods of lower-order. In the approach developed in chapter §4, this notion seems to be related with the notion of composability of canonical relations (which may fail even to be a manifold without appropriate transversality conditions [164]). This study is a promising line for numerical simulation of Hamiltonian dynamics and also constrained systems.

Other interesting case that is worth to explore is the extension of the theory developed in chapter 4 to reduced systems using the geometric framework given by the Lie algebroid and Lie groupoid formalisms [114]. For instance, in the discrete case, is interesting to derive the dynamics using also Lagrangian submanifolds of the so-called tangent groupoid [39]. In this context, the departing point would be a submanifold N of a Lie groupoid $G \rightrightarrows Q$, and a discrete Lagrangian $L_d : N \to \mathbb{R}$. Following Theorem 4.1.1, the Lagrangian submanifold Σ_{L_d} of the cotangent groupoid T^*G would be introduced. The Hamiltonian side is determined by a Poisson flow defined on A^*G, the dual bundle of the associated Lie algebroid AG to G. Therefore, we would have the following scheme:

$$
\begin{tikzcd}
\Sigma_{L_d} \arrow[rr, hook, "i_{\Sigma_{L_d}}"] \arrow[d, hook] & & T^*G \arrow[dl, "\pi_G"] \arrow[dr, shift left, "\tilde{\alpha}"] \arrow[dr, shift right, "\tilde{\beta}"'] & \\
N \arrow[r, hook, "i_N"] \arrow[d, "L_d"] & G \arrow[dr, shift left, "\alpha"] \arrow[dr, shift right, "\beta"'] & & A^*G \arrow[dl, "\tau_{A^*G}"] \\
\mathbb{R} & & Q &
\end{tikzcd}
$$

which, in principle, will allow us to extend the theory presented chapter 4 and give a unified framework to understand symmetries in constrained Lagrangian systems. Furthermore, due to the relationship between constrained Lagrangian systems and Hamiltonian systems established in Theorem 4.2.4, this is the suitable setup to study generalized Hamiltonian dynamics with symmetries and its relationship with morphisms of Lie algebroids (see [119]). Examples of systems this theory could be applied to are the following: subriemannian geometry and optimal control problems.

Advances on discrete variational optimal control

As shown in fig.5.1 and fig.5.2, simulations of an underactuated underwater vehicle illustrate an application of the method developed in chapter 5. Yet, further numerical studies and comparisons would be necessary to exactly quantify the advantages and the limitations of the proposed algorithm. An important future direction is thus to study the convergence properties of the optimal control system. Convergence for general nonlinear systems is a complex issue (not only in what optimal control regards but also concerning nonholonomic mechanics, as will be discussed in the next subsection). In this respect, it is interesting to note that the discrete mechanics and optimal control on Lie groups such as the example in using the Cayley map results in polynomial form without further approximation or Taylor series truncation. A useful future direction is then to study the regions of attraction of the numerical continuation using tools from algebraic geometry.

To be more precise, it has been shown in chapter 5 that many interesting optimal control problems defined on Lie algebras the momenta can be expressed as

$$\mu = (\mathrm{d}\tau_{h\xi}^{-1})^{*}\mathbb{I}\,\xi.$$

where $\mu \in \mathfrak{g}^{*}$, $\xi \in \mathfrak{g}$, h is the time step, $\frac{1}{2}\langle \mathbb{I}(\xi)\,,\,\xi\rangle$ is the kinetic energy of the system and $\mathbb{I}:\mathfrak{g}\to\mathfrak{g}^{*}$ is the inertia tensor. On one hand, the convergence of the optimal control problem is related to the invertibility of this expression when $h\to 0$ (which is the desired situation when looking for accuracy in the simulations of real systems). On the other hand, the previous expression turns into an algebraic system of equations when fixing the Lie group and τ as the Cayley map. Therefore, it is definitely interesting to understand algebraically the properties of the mentioned system of equations in the limit $h\to 0$.

More generally, the theoretical framework introduced in §5.3 can serve as a basis for deriving algorithms for control systems such as multi-body locomotion systems or robotic vehicles with nonholonomic constraints. Furthermore, the developed classes of systems can be unified through the recently developed groupoid framework [70, 165]. Each of the considered product spaces (e.g. $Q\times Q$) can be regarded as a single groupoid space with equations of motion resulting from a single generalized discrete variational principle. This will enable the automatic solution of optimal control problems for various complex systems and a convenient unified framework for implementing practical optimization schemes such as [15, 91, 101, 139]. More importantly, this viewpoint can be used to apply standard discrete Lagrangian regularity conditions (e.g. [124]) to optimal control problems evolving on the groupoid space. This would provide a deeper insight into the solvability of the resulting optimization schemes.

In §5.3.1 a geometric integrator for optimal control problems in trivial principal bundles has been built. This particular choice of the velocity phase space has been made because of its usual presence in practical applications (indeed, if the velocity phase space is eventually non-trivial it is treated locally). In any case, it is interesting from the geometric point of view to extend the developed setting to general principal bundles, which means to use tools associated to discrete connections (see [118]).

The discrete framework presented in §5.3.2 to integrate nonholonomic problems is, so far, completely theoretical. In the next future it will be applied to specific examples of nonholonomic problems. We expect that our techniques will be of great use when modelizing wheeled vehicles (which is an instance of nonholonomic system) with discontinuous (bang-bang like) input controls.

Regarding the regularity of solutions, in chapter 5 only regular solutions for the optimal control problem have been considered. It would be of great interest to extend our analysis to abnormal solutions. Some examples of systems with abnormal solutions and their study can be found in [6, 7, 8].

New perspectives in discrete nonholonomic mechanics

One of the main problems of discrete nonholonomic mechanics is the lack of preservation of geometric structures (non-preservation of the nonholonomic bracket, non-preservation of the nonholonomic momentum in general) mimicking the non-preservation of the continuous nonholonomic system. As has been already mentioned, [44, 45] introduced the GNI integrator, which is an integrator for nonholonomic problems with interesting geometric properties. In chapter 6 we have applied this method to different kinds of systems, such as affine of reduced systems. Due to the lack of preservation properties just mentioned, it is difficult to compare from the geometric perspective the nonholonomic integrators obtained from discrete nonholonomic mechanics with standard methods. One possibility is to use the comparison algorithms developed in §4.2.3 and §4.3.3 in order to detect if one particular nonholonomic integrator is preserving the common solutions of the continuous nonholonomic problem and its associated constrained variational problem.

GNI method, in some cases, gives rise to explicit integrators. As a consequence, these integrators show high efficiency at a computational level (see [93]). Thus, it would be interesting to connect the GNI method developed in chapter 6 to the techniques presented in chapter 5 in order to derive efficient numerical methods to integrate optimal control problems in the presence of nonholonomic constraints (type of problems which usually suppose high computational cost).

The lack of attention the numerical community has paid to nonholonomic problems so far, implies also a lack of knowledge about the convergence behavior of their numerical integrators. It was shown in [45] that the SHAKE extension to nonholonomic mechanics via GNI method is globally second-order convergent. In this work has been also proved (Theorem 6.3.3) that the Euler-symplectic extensions are globally first-order convergent. These proofs are obviously interesting but non-systematic, that is each case is treated independently. On the other hand, a clear parallelism between variational integrators and their extensions to nonholonomic mechanics via GNI has been established in [45] and Theorem 6.3.3 and Theorem 6.3.5 in §6.3. In regard to variational integrators, a systematic study of their convergence behavior has been accomplished by means of Backward Error Analysis (see [61]) in [124, 140]; it has been carefully introduced in §3.4.3. The key tool in this analysis is the exact Lagrangian

$$L_d^E(q_0, q_1, h) = \int_0^h L(q(t), \dot{q}(t))\,\mathrm{d}t,$$

presented in (3.3.2). Thus, it is extremely interesting to figure an exact Lagrangian for nonholonomic systems (in other words, when our system is defined on a discrete constraint submanifold) and find the way to, mimicking the process in the case of variational integrators, systematically understand their error behavior.

Related to this last topic is the systematic study of the error behaviour of the nonholonomic integrators. These integrators, either based on discrete variational calculus (see [38]) or on the GNI approach, present a bounded oscillation of the total energy of the system. A possible tool to approach the understanding of this behavior would be also Backward Error Analysis, as happens in the unconstrained case.

Conclusiones

El capítulo final de esta memoria está dedicado a resumir las contribuciones presentadas en este trabajo de tesis.

El capítulo 4 se ha dedicado al estudio geométrico de la relación entre la dinámica Hamililtoniana y el cálculo variacional con ligaduras (calcúlo vakónomo). Bajo ciertas condiciones de regularidad encontramos que ambos son equivalentes. Este resultado, que es una de las principales contribuciones de esta tesis, ha sido presentado en el Teorema 4.2.4. Además, hemos extendido esta relación al marco discreto. Como resultado encontramos la conexión entre los integradores simplécticos y el cálculo variacional en presencia de ligaduras, resultado contenido en el Teorema 4.3.4. Esta relación esclarece la estructura geométrica de los integradores simplécticos en términos de subvariedades Lagrangianas. El modo de aprovecharse del resultado anterior en el contexto de la integración geométrica de sistemas Hamiltonianos es presentado con varios ejemplos en §4.4. Asimismo, hemos analizado en paralelo el caso de mecánica noholónoma en los marcos discreto (§4.3.3) y contínuo (§4.2.3). Finalmente, hemos encontrado un algoritmo para la comparación de soluciones noholónomas y soluciones vakónomas, generalizando el ya presentado en [35].

En el capítulo 5 hemos construido métodos numéricos para problemas de control óptimo de tipo mecánico. Más concretamente, extendemos la teoría de de la mecánica discreta con el objetivo de habilitar soluciones para este tipo de problemas por medio de la discretización de los principios variacionales. Nuestro método se basa en la siguiente idea: resolver el problema de control óptimo como un integrador variacional construido especialmente para un sistema de dimensión mayor. El formalismo Lagrangiano del que los mencionados integradores variacionales proceden se introduce en §5.1.4 y §5.2.4 para sistemas definidos en fibrados tangentes y álgebras de Lie respectivamente. Por otro lado, la relación concreta entre el espacio de control y el espacio de fase (tipo velocidades) de los problemas bajo estudio, relación sobre la que se basa nuestro método, se establece en las siguientes definiciones:

- TQ: def.5.1.5 y def.5.1.7: sistemas completamente actuados e infraactuados respectivamente.
- $\mathfrak{g}$: def.5.2.4 y def.5.2.5: sistemas completamente actuados e infraactuados respectivamente.

Simulaciones del control óptimo de un vehículo subacuático son mostradas en fig.5.1 y fig.5.2.

Nos gustaría enfatizar la capacidad de adaptación de las técnicas propuestas a problemas reducidos o sujetos a ligaduras noholónomas. Estas extensiones se discuten en §5.3. Además, una de las propiedades más destacadas del método propuesto es su capacidad de admitir controles discontinuos. Esta capacidad se presenta en §5.2.10 y se ilustra en fig.5.3. Como consecuencia de esta propiedad, esperamos que nuestras técnicas sean ampliamente utilizadas en investigación futura relacionada con control óptimo y sus aplicaciones a ingeniería. Esta expectativa se basa en el hecho de que los controles discontinuous son muy habituales en problemas prácticos.

El capítulo 6 contiene nuevos desarrollos relacionados con el Integrador Noholónomo Geométrico (GNI). Hemos extendido los resultados presentados en [44, 45, 93] al caso de ligaduras afines (§6.4) y sistemas reducidos (§6.5). Ambos casos se ilustran con un ejemplo teórico: la extensión noholónoma del algoritmo RATTLE en el contexto del GNI. Finalmente, encontramos un marco generalizado para el integrador GNI en el contexto de los grupoides de Lie (§6.6).

En las referencias [44, 45] se demostró que el método SHAKE noholónomo puede ser obtenido mediante la aplicación de las ecuaciones GNI a una discretización concreta del Lagrangiano mecánico. Además, se prueba que el método RATTLE asociado (que es equivalente al SHAKE) es globalmente convergente de segundo orden. Este resultado establece un paralelismo claro con el caso libre. Siguiendo esta línea, hemos aplicado las ecuaciones GNI a la discretización que produce los métodos Euler-simpléctico como integradores variacionales. Hemos denotado estos métodos como Euler A y Euler B. En el Teorema 6.3.3 demostramos que los métodos Euler A y Euler B son globalmente convergentes de orden uno. Además, demostramos en el Teorema 6.3.5 que los métodos Euler A y Euler B son adjuntos el uno del otro. Puesto que los métodos Euler-simpléctico son también globalmente convergentes de primer orden y adjuntos el uno del otro, estos dos resultados, aparte de ser dos de las contribuciones más importantes del capítulo 6, refuerzan el paralelismo con el caso libre mencionado anteriormente.

Finalmente, en §6.5 hemos estudiado cómo extender el método GNI a otros casos de interés: sistemas reducidos y sistemas sujetos a ligaduras noholónomas de tipo afín.

Los resultados presentados en esta tesis han sido publicados en

- KOBILAROV M, JIMÉNEZ F AND MARTÍN DE DIEGO D, *Discrete Variational Optimal Control.* Preprint, arXiv:1203.0580, (2012), enviado para publicación a *Journal of Nonlinear Science.* Algunos resultados en este artícuo han sido utilizados en
  - COLOMBO L, JIMÉNEZ F AND MARTÍN DE DIEGO D, *Discrete second-order Euler-Poincaré equations. Applications to optimal control.* Journal of Geometric methods of Modern Physics, **9**, (2012), arXiv:1109.4716.
- JIMÉNEZ F, DE LEÓN M AND MARTÍN DE DIEGO D, *Hamiltonian dynamics and constrained variational calculus: continuous and discrete settings.* Aceptado para publicación en *Journal of Physics A*. arXiv:1108.5570, (2011).

- JIMÉNEZ F AND MARTÍN DE DIEGO D, *Continuous and discrete approaches to vakonomic mechanics.* Revista de la Real Academia de las Ciencias Exactas, Físicas y Naturales, Serie A. Springer, **106(1)**, pp. 75-87, (2011), DOI.

- JIMÉNEZ F AND MARTÍN DE DIEGO D, *A geometric approach to Discrete mechanics for optimal control theory.* Proceedings of the 49th IEEE Conference on Decision and Control (CDC), pp. 5426-5431. Atlanta, Georgia, USA, (2010).

Además, los resultados contenidos en esta tesis han sido presentados en los siguientes encuentros internacionales:

- Discrete variational optimal control (F Jiménez). *Friedrich-Alexander University of Nürnberg-Erlangen, Chair of Applied Dynamics*, charla invitada, Erlangen, Alemania, Enero 2012.

- On discrete mechanics for optimal control theory (F Jiménez). *IMAC Symposium on Dynamical Systems: Trends and Perspectives*, Castellón, España, Septiembre 2011.

- A new approach to discrete optimal control theory (F Jiménez). *California Institute of Technology*, charla invitada, Pasadena, EE.UU., Agosto 2011.

- New developments in discrete variational calculus and discrete optimal control theory (D Martín de Diego). *Mathematisches Forschungsinstitut Oberwolfach*, Alemania, Agosto 2011.

- Lagrangian submanifolds and constrained variational calculus (F Jiménez). *Workshop on Rough Paths and Combinatorics in Control Theory*, San Diego, EE.UU., Julio 2011.

- Lagrangian submanifolds and constrained variational calculus (F Jiménez). 5^{th} *International Summer School on Geometry, Mechanics and Control*, Madrid, España, Junio 2011.

- Lagrangian submanifolds and discrete constrained mechanics (D Martín de Diego). *XI Congreso Dr. Antonio Monteiro*, Bahía Blanca, Argentina, Mayo 2011.

- A geometric approach to discrete mechanics for optimal control theory (D Martín de Diego). *49th IEEE Conference on Decision and Control*, Atlanta, EE.UU., Diciembre 2010.

- Optimal control of mechanical systems on Lie groups (D Martín de Diego). *VI International symposium HAMSYS-2010, Honoring EA Lacomba in his 65th Anniversary*, charla invitada, Ciudad de México, México, Diciembre 2010.

- Optimal control of mechanical systems on Lie groups: continuous and discrete (D Martín de Diego). *Workshop on Geometry of Constraints and Control - New Developments*, Banach Center, Varsovia, Polonia, Noviembre 2010.

Agradecimientos

Agradecer algo a alguien supone la mayoría de las veces un acto de injusticia, ya que enfatiza la contribución de ese alguien a algún logro, olvidando la de otros muchos (incluido el azar) que de una forma u otra también puede ser importante. Por tanto declaro que voy a ser inevitablemente injusto, que para mí mencionar es igual a agradecer que es igual a recordar, y que (no siempre) trataré de restringirme a los tres últimos años (tiempo que he dedicado a esta tesis doctoral) a la hora de hacer memoria.

En primer lugar está David, que ha sido mi director de tesis y guía científico durante este periodo, pero que se ha convertido, sin dejar de ser lo anterior, en un compañero, amigo, consejero y cosufridor (y quien conozca esta enfermedad sabrá lo mucho que une a quienes la padecen) de la maldición roja y blanca. Nunca le estaré lo suficientemente agradecido por todo el trabajo que ha realizado para que este proyecto se lleve a cabo, por formarme como científico y por ser un ejemplo de cómo sacar las cosas adelante con $2 \cdot 10^6$ (cota inferior) asuntos en la cabeza. Le agradezco a Manolo los esfuerzos que hizo en su día, cuando un problema burocrático hizo peligrar el comienzo de mi andadura en las Matemáticas, por luchar contra aquellos molestos flecos y permitir que todo empezara. También ha sido mi coloborador científico y un inigualable ejemplo de tesón: sobre sus hombros ha edificado el actual Instituto de Ciencias Matemáticas, cuya existencia sería más que dudosa sin su aportación. Debo agradecer también a Alberto Casas, con quien di mis primeros y demasiado vacilantes pasos en la ciencia, por haberme introducido en este maravilloso mundo. Matteo, mi antiguo compañero de despacho en el departamento de Física Teórica de la Universidad Autónoma de Madrid, es a fin de cuentas el responsable de que yo esté aquí: un día de Marzo de 2008 me envió via e-mail una *posible oferta de trabajo* a la que, felizmente, decidí presentarme como candidato. Por eso lo recuerdo aquí.

Siento también profunda gratitud al departamento de Física Teórica de la Universidad de Zaragoza por haberme albergado como estudiante de doctorado y permitirme así defender la tesis que ahora presento. Dentro de la Universidad de Zaragoza debo mencionar especialmente a Pepín, que fue el primero en sugerir la posibilidad que ahora se lleva a cabo, a Eduardo, que ha sido mi tutor de tesis, y a Pedro Justes, secretario del departamento de Física Teórica, que se ha esforzado siempre por hacerme la vida más fácil cuando de enfrentarse a la burocracia (la gran enemiga de todos los científicos) se trataba. Lo mismo hace Rosa en el ICMAT (primero en Serrano y ahora en Cantoblanco), lidiando con viajes, facturas, tarjetas de embarque, vales de desplazamiento, liquidaciones, y un infinito etc, para ahorrarnos tantísimos dolores de cabeza.

De entre los miembros de mi *primera familia científica*, debo mencionar a mis dos *her-*

manos: Cédric, el mayor, por abrirme camino y Leo, el menor, a quien quiero creer que yo se lo abro, pero que, en realidad, se ha convertido en un voluntarioso e infatigable colaborador.

La Red de Geometría, Mecánica y Control supone una *segunda familia científica*, a la que debo viajes, encuentros, aprendizaje cientfico, camaradería e innumerables momentos de diversión. De entre todos ellos mencionaré a Edith, Juan Carlos, David Iglesias (añorado compañero de mi añoradísimo Serrano), María (que me ha sido de una ayuda inigualable a la hora de solicitar postdocs), Elisa (mi primita) y, más recientemente, Kurusch (primo consorte, ejemplo de vocación científica, capacidad de esfuerzo y, actualmente, compañero de piso).

De allende los mares no puedo olvidar a Vivi, que, intermitentemente, ha sido compañera de despacho, viaje, maleta y, durante dos meses, también de piso. Marin Kobilarov (en California) ha sido mi colaborador, generoso anfitrión en mi visita a Caltech, valedor en la caza de postdocs y una de las personas que también han contribuido en el plano matemático a que este trabajo llegue a su fin. Por todos esos motivos le estoy muy agradecido. *Marin Kobilarov (in California) has been my collaborator, generous host during my visit to Caltech, referee in my post-doc hunting and one of the people that have contributed at a mathematical level to the happy ending of this thesis. For all these reasons I am grateful to him.*

Mi andadura en el Instituto de Ciencias Matemáticas se inició en los antiguos cuarteles centrales de la calle Serrano. Como menciono antes, los echo mucho de menos, mayoritariamente por ser parte de mi más añorada aún Madrid. En Serrano me alojaba en el despacho 1001, lo que, inmediatamente y para siempre, me convirtió en un *milunista*. Siempre llevaré la etiqueta con orgullo, como si se tratara de un rojo emblema de valor, de la misma forma que lo hacen quienes han sido, antes y después del traslado a Cantoblanco, mis compañeros: Mario, Marti, Róber, Ana, Marina (mi vecina y profesora) y, especialmente, Emilio. Con él estoy compartiendo el proceso de redacción de nuestras tesis y la evolución de nuestros temores, dudas y alegrías a la hora de encarar el futuro en la ciencia. Por encima de todo es para mí un *agujero blanco*: de la misma forma que los agujeros negros no dejan escapar la luz, Emilio atrapa la oscuridad haciendo de lo que hay a su alrededor algo más limpio.

Pero ellos no son mis únicos compañeros. En nuestra nueva localización (*tan lejos de todo...*), comparto cafés y descansos con Ángel (el Doctor Castro), Jeza, Rafa, Miguel y Javi, lo que sin duda me hace la vida más amena. Mis otros compañeros, con los que hace casi ocho años comencé el proceso que de algún modo hoy termina, ya no están aquí (pues andan desperdigados por Europa), pero siempre estarán en mi corazón: Antón, Irene, Josemi, Jacobo, Alfonso y Adolfo (nunca olvidaré la conversación que mantuvimos por teléfono un día de septiembre de 2004 - él en Badajoz, yo en Madrid- en la que, recién acabada la licenciatura, dudábamos sobre nuestro futuro académico; casi simultáneamente nos dijimos *qué hacemos!!* al poco de descolgar).

He vivido en Madrid durante los primeros dos años y medio de los tres que he dedicado a este trabajo de tesis. En ese tiempo la ciudad se ha convertido en mi patio de recreo, mi banco de pruebas, mi librería, mi oscuro agujero, mi cine, mi café, mi cómoda cama, mi bar nocturno y, al mismo tiempo, mi Última Thule y mi Ítaca (a la que, no importa que pasen 20 años, deseo regresar). Algunas de las personas que me han acompañado en mi incansable peregrinaje por sus calles, haciéndolo más rico si cabe, son: Carmelo (mi Coronel), Amalio (Corazón de metal), Arantxa (Gatuna), Rocío (subiendo montes), Patri

(Acho1), David (Acho2) y Vital (Acho3).

Sin embargo, por mucho que quiera huir, no puedo evitar mi naturaleza *tricantina* y el atractor extraño que Tres Cantos supone para mí. En el pueblo crecí y a él, inevitablemente, vuelvo. Aquí es donde seré más injusto, pues son muchos los que han compartido conmigo la niñez, adolescencia y primera juventud. No obstante, atenerme a los tres últimos años me facilita las cosas y, del pasado que pervive en el presente, menciono a Carlos (con quien, viviendo siempre en *El filo de la navaja*, comencé hace años la difícil tarea de entender los imponderables de la vida, entre ellos *El innombrable* que cada vez más toma la forma de una voluntad primera, y quien me ha acompañado en su camino paralelo por el mundo de la investigación, como lo es el *Trópico de Cáncer* del *Trópico de Capricornio*, hasta que la diáspora postdoctoral nos condujo al *Jardín de los senderos que se bifurcan*), Manolo (con dos ases en la mano), Laura (tarde pero siempre a tiempo), Rodrigo (que siempre llama), Javier (Beitia) (desde la más tierna infancia) y, especialmente, a Pablo (siempre junto a mí paseando por la calle, o bien siempre junto a mí por teléfono, ya sea desde Nueva York, Londres, Barcelona, Paris o Helsinki). No puedo olvidarme de la pequeña Laura (hija de Javi), cuyo nacimiento ha sido sin duda la mejor noticia de este periodo.

Finalmente (y por tanto con importancia máxima) agradezco a toda mi familia haber acogido y alimentado durante las malas épocas los despojos en que, ahora con el fuego en mí[1], ya nunca volveré a convertirme.

[1] Ver cita al comienzo.

Appendix A: Lemmae

This Appendix deals with the set of Lemmae and proofs, involving the right-trivialized tangent and inverse right-trivialized tangent of a general retraction map (see definition 5.2.2), necessary for the derivation of the algorithms obtained in §5.

Lemma 6.6.1. *(See [123]) 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*

$$\delta(\mathrm{Ad}_g\,\lambda) = -\mathrm{Ad}_g\,[\lambda\,,\,g^{-1}\delta g],$$

where $[\cdot\,,\,\cdot] : \mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$ denotes the Lie bracket operating or equivalently $[\xi\,,\,\eta]\equiv \mathrm{ad}_\xi\eta$, for given η, $\xi\in\mathfrak{g}$.

Lemma 6.6.2. *The following identity holds*

$$\partial_\xi(\mathrm{Ad}_{\tau(\xi)}\lambda) = -[\mathrm{Ad}_{\tau(\xi)}\,,\,\mathrm{d}\tau_\xi].$$

Proof. By lemma 6.6.1

$$\begin{aligned}\partial_\xi(\mathrm{Ad}_{\tau(\xi)}\eta) &= -\mathrm{Ad}_{\tau(\xi)}[\lambda\,,\,\tau(-\xi)\,\delta\tau(\xi)]\\ &= -[\mathrm{Ad}_{\tau(\xi)}\lambda\,,\,\partial\tau(\xi)\,\tau(-\xi)]\\ &= -[\mathrm{Ad}_{\tau(\xi)}\eta\,,\,\mathrm{d}\tau_\xi],\end{aligned}$$

obtained from the tangent Definition 5.2.2 and using the fact that $\mathrm{Ad}_g[\lambda\,,\,\eta] = [\mathrm{Ad}_g\lambda\,,\,\mathrm{Ad}_g\eta]$. □

The following lemmae can be found in [22].

Lemma 6.6.3. *The following identity holds*

$$\mathrm{d}\tau_\xi\,\eta = \mathrm{Ad}_{\tau(\xi)}\,\mathrm{d}\tau_{-\xi}\,\eta,$$

for any $\xi,\eta\in\mathfrak{g}$.

Proof. Differentiation of $\tau(\xi)\,\tau(-\xi) = e$ implies that

$$\partial_\xi \tau(-\xi)\,\eta = -TL_{\tau(-\xi)}\,TR_{\tau(-\xi)}\,(\partial_\xi \tau(\xi)\,\eta)\,,$$

where TL , TR are the tangent of the left and right translations in the group respectively. On the other hand, the chain rule implies that

$$\partial_\xi \tau(-\xi)\,\eta = -TR_{\tau(-\xi)}\,\mathrm{d}\tau_{-\xi}\,\eta.$$

Combining both expressions we obtain

$$TL_{\tau(\xi)}\,\mathrm{d}\tau_{-\xi}\,\eta = TR_{\tau(\xi)}\,\mathrm{d}\tau_\xi\,\eta,$$

which proves the identity. □

Lemma 6.6.4. *The following identity holds*

$$\mathrm{d}\tau_\xi^{-1}\,\eta = \mathrm{d}\tau_{-\xi}^{-1}\left(\mathrm{Ad}_{\tau(-\xi)}\,\eta\right),$$

for any $\xi, \eta \in \mathfrak{g}$.

Proof. The proof follows directly from Lemma 6.6.3. Let $\eta \to \mathrm{d}\tau_\xi^{-1}\,\eta$ in that identity to obtain

$$\eta = \mathrm{Ad}_{\tau(\xi)}\,\mathrm{d}\tau_{-\xi}\,\mathrm{d}\tau_\xi^{-1}\,\eta.$$

Solving this last equation for $\mathrm{d}\tau_\xi^{-1}\,\eta$ we prove the identity

$$\mathrm{d}\tau_\xi^{-1}\,\eta = \mathrm{d}\tau_{-\xi}^{-1}\left(\mathrm{Ad}_{\tau(-\xi)}\,\eta\right).$$

□

Appendix B: Cayley transform for matrix groups

This Appendix is devoted to specify the operators required to implement equations (5.37), (5.50) and (5.51) for typical rigid body motion groups and general real matrix subgroups. While we have given more than one general choice for τ in §5.2.3, for computational efficiency we recommend the Cayley map since it is simple and does not involve trigonometric functions. 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.

A. $SO(3)$**:**

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. Define the map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ by

$$\hat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \tag{6.38}$$

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 some $\rho \in \mathbb{R}^3$. Using this identification we have

$$\mathrm{cay}(\hat{\omega}) = I_3 + \frac{4}{4 + \parallel \omega \parallel^2} \left(\hat{\omega} + \frac{\hat{\omega}^2}{2} \right), \tag{6.39}$$

where I_3 is the 3×3 identity. The linear maps $\mathrm{d}\tau_\xi$ and $\mathrm{d}\tau_\xi^{-1}$ are expressed as the 3×3 matrices

$$\mathrm{dcay}_\omega = \frac{2}{4 + \parallel \omega \parallel^2} (2I_2 + \hat{\omega}), \quad \mathrm{dcay}_\omega^{-1} = I_3 - \frac{\hat{\omega}}{2} + \frac{\omega \, \omega^T}{4}. \tag{6.40}$$

We point out that with the choice $\tau = \mathrm{cay}$ the optimization domain is not restricted, i.e. $\mathfrak{D}_{\mathrm{cay}} = \mathfrak{g}$ since the maps (6.40) are not singular for any $\xi \in \mathfrak{g}$. This is not the case for the exponential map for which $\mathfrak{D}_{\mathrm{exp}} = \{\xi \in \mathfrak{g}, \parallel \xi \parallel < 2\pi/h\}$ since the exponential map derivative is singular whenever the norm of its argument is a multiple of 2π [61], and the origin requires special handling.

B. $SE(2)$**:**

The coordinates of $SE(2)$ are (θ, x, y) with matrix representation $g \in SE(2)$ given by

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

Using the isomorphic map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{se}(2)$ given by:

$$\hat{v} = \begin{pmatrix} 0 & -v_1 & v_2 \\ v_1 & 0 & v_3 \\ 0 & 0 & 0 \end{pmatrix},$$

where $v = (v_1, v_2, v_2)^T \in \mathbb{R}^3$. Thus, $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ can be used as a basis for $\mathfrak{se}(2)$, where $\{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{R}^3$. The map $\mathrm{cay} : \mathfrak{se}(2) \to SE(2)$ is given by

$$\mathrm{cay}(\hat{v}) = \begin{pmatrix} \frac{1}{4+v_1^2} \begin{pmatrix} 4 - v_1^2 & -4v_1 & -2v_1v_3 + 4v_2 \\ 4v_1 & 4 - v_1^2 & 2v_1v_2 + 4v_3 \end{pmatrix} \\ \begin{matrix} 0 & \quad 0 & \quad 1 \end{matrix} \end{pmatrix},$$

while the map $\mathrm{d}\tau_\xi^{-1}$ becomes the 3×3 matrix

$$\mathrm{dcay}_{\hat{v}}^{-1} = I_3 - \frac{1}{2}\mathrm{ad}_v + \frac{1}{4}\left(v_1 v \;\; 0_{3\times 2}\right),$$

where

$$\mathrm{ad}_v = \begin{pmatrix} 0 & 0 & 0 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}.$$

C. $SE(3)$**:**

We make the identification $SE(3) \sim SO(3) \times \mathbb{R}^3$ using elements $R \in SO(3)$ and $x \in \mathbb{R}^3$ through

$$g = \begin{pmatrix} R & x \\ 0_{3\times 3} & 1 \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} R^T & -R^T x \\ 0_{3\times 3} & 1 \end{pmatrix}.$$

Elements of the Lie algebra $\xi \in \mathfrak{se}(3)$ are identified with **body-fixed** angular and linear velocities denoted $\omega \in \mathbb{R}^3$ and $v \in \mathbb{R}^3$, respectively, through

$$\xi = \begin{pmatrix} \hat{\omega} & v \\ 0_{3\times 3} & 0 \end{pmatrix},$$

where the map $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined in (6.38). Using this

$$\mathrm{cay}(\xi) = \begin{pmatrix} \mathrm{cay}(\hat{\omega}) & \mathrm{d}\tau_\omega\, v \\ 0 & 1 \end{pmatrix},$$

where cay : $\mathfrak{so}(3) \to SO(3)$ is given by (6.39) and dcay : $\mathbb{R}^3 \to \mathbb{R}^3$ by (6.40)[2].

D. General matrix subgroups:

The Lie algebra of a matrix Lie group coincides with the one-parameter subgroup generators of the group. Assume that we are given a k-dimensional Lie subalgebra denoted $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R})$. It is isomorphic to the space of generators of a unique connected k-dimensional matrix subgroup $G \subset GL(n, \mathbb{R})$. Therefore, a subalgebra $\mathfrak{g}$ determines the subgroup G in a one-to-one fashion

$$\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R}) \Longleftrightarrow G \subset GL(n, \mathbb{R}).$$

Assume that the Lie algebra basis elements are $\{E_\alpha\}_{\alpha=1}^{k}$, $E_\alpha \in \mathfrak{g}$, i.e., every element $\xi \in \mathfrak{g}$ can be written as $\xi = \xi^\alpha E_\alpha$. Define the following inner product for any $\xi, \eta \in \mathfrak{g}$

$$\ll \xi, \eta \gg = \mathrm{tr}\,(B\,\xi^T\,\eta),$$

where B is an $n \times n$ matrix such that $\ll E_\alpha\,,\,E_\beta \gg = \delta_{\alpha\,\beta}$. Correspondingly, a pairing between any $\mu \in \mathfrak{g}^*$ and $\xi \in \mathfrak{g}$ can be defined by

$$\langle \mu\,,\,\xi \rangle = \mathrm{tr}(B\,\mu\,\eta),$$

since the dual basis for $\mathfrak{g}^*$ is $\left\{[E_\alpha]^T\right\}_{\alpha=1}^{k}$.

As example, we can consider the **Kinetic Energy-Type Metric**: after having defined a metric pairing, a kinetic energy operator $\mathbb{I}$ can be expressed as

$$\langle \mathbb{I}(\xi)\,,\,\eta \rangle = \mathrm{tr}(B\,I_d\,\xi^T\,\eta),$$

for some symmetric matrix $I_d \in GL(n, \mathbb{R})$.

[2] note that cay denotes a map to either $SO(3)$ or $SE(3)$ which should be clear from its argument.

...At this time he had no messages for anyone. Nothing. Not a single word.

Saul Bellow, "Herzog".

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] Aebischer B, Borer M, Kälin M, Leuenberger Ch and Reimann HM, *Symplectic Geometry.* Prog. Math. vol. 124 (1994).

[4] Agrachev AA, Bonnard B, Chyba M and Kupka I, *Sub-Riemannian sphere in Martinet flat case.* ESAIM/COCV (Control, Optimizatin and Calculus of Variations), pp. 377-448, (1997).

[5] Arnold VI, *Dynamical Systems III (second edition).* Encyclopaedia of Mathematical Sciences, Springer-Verlag, (1993).

[6] Barbero-Liñán M, *A geometric study of abnormality in optimal control problems for control and mechanical control systems.* PhD thesis, Universidad Politécnica de Catalunya, (2008).

[7] Barbero-Liñán M and Muñoz-Lecanda MC, *Constraint algorithm for extremals in optimal control problems.* Int. J. Geom. Methods Mod. Phys. **6(7)**, pp. 1221-1233, (2009).

[8] Barbero-Liñán M and Muñoz-Lecanda MC, *Strict abnormal extremals in non-holonomic and kinematic control systems.* Discrete Contin. Dyn. Syst. Ser. **S3(1)**, pp. 1-17, (2010).

[9] Barth E and Leimkuhler B, *Symplectic methods for conservative multibody systems.* in Integration Algorithms and Classical Mechanics, American Mathematical Society, Providence, RI, pp. 25-43, (1996).

[10] Benettin G and Giorgilli A, *On the Hamiltonian interpolation of near-to- the-identity symplectic mappings with application to symplectic integration algorithms.* J. Statist. Phys. **74**, pp. 1117-1143, (1994).

[11] Benito R and Martín de Diego D, *Discrete vakonomic Mechanics.* Journal of Mathematical Physics, **46**, (2005).

[12] Berndt R, *An introduction to Symplectic Geometry.* Graduate studies in Mathematics, vol. 26, American Mathematical Society..

[13] Binz E, Sniatycki J and Fischer H, *Geometry of classical fields.* North-Holland Mathematics Studies **154**, North-Holland Publishing Co., Amsterdam, (1988).

[14] Bloch AM, *Nonholonomic Mechanics and Control.* Interdisciplinary Applied Mathematics Series **24**, Springer-Verlag New-York, (2003).

[15] Bloch AM, Hussein I, Leok M and Sanyal AK, *Geometric structure-preserving optimal control of a rigid body.* Journal of Dynamical and Control Systems, **15**, (2009).

[16] Bloch AM, Krishnaprasad PS, Marsden JE and Murray R, *Nonholonomic mechanical systems with symmetry.* Arch. Rational Mech. Anal. **136**, pp. 21-99, (1996).

[17] Bobenko AI and Suris YB, *Discrete Lagrangian reduction, discrete Euler-Poincaré equations and semidirect products.* Lett. Math. Phys. **49**, (1999).

[18] 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).

[19] Bonnard B, Chyba M and Kupka I, *Non Integrable Geodesics in SR-Martinet Geometry.* Proceedings of the 1997 Summer Research Institute on Differential Geometry and Control, pp. 119-130, (1999).

[20] 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).

[21] Bou-Rabee N, *Hamilton-Pontryagin integrators on Lie groups.* PhD thesis, California Institute for Technology, (2007).

[22] Bou-Rabee N and Marsden JE, *Hamilton-Pontryagin Integrators on Lie Groups: Introduction and Structure-Preserving Properties.* Foundations of Computational Mathematics, **9** (2), (2009).

[23] 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).

[24] Cadzow JA, *Discrete calculus of variations.* Internat. J. Control. **11**, pp. 393-407, (1970).

[25] Cadzow JA, *Discrete-Time Systems: An Introduction with Interdisciplinary Applications.* Prentice-Hall, (1973).

[26] Cano B and Sanz-Serna JM, *Error growth in the numerical integration of periodic orbits, with application to Hamiltonian and reversible systems.* SIAM J. Numer. Anal. **34**, pp. 1391-1417, (1997).

[27] Cantrijn F, Cortés J, de León M and Martín de Diego D, *On the geometry of generalized Chaplygin systems.* Math. Proc. Cambridge Philos. Soc., **132(2)**, pp. 323-351, (2002).

[28] Cariñena JF, Nunes da Costa JM and Santos P, *Reduction of Lagrangian mechanics on Lie algebroids.* J. Geom. Phys. **57(3)**, pp. 977-990, (2007).

[29] Cariñena JF and Rañada MF, *Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers.* J. Phys. A **26(6)**, pp. 1335-1351, (1993).

[30] do Carmo MP, *Riemannian geometry.* Birkhäuser, Boston-Basel- Berlin, (1992).

[31] Channell PJ and Scovel C, *Symplectic integration of Hamiltonian systems.* Nonlinearity, **3**(2): pp. 231-259, (1990).

[32] Chyba M, Hairer E and Vilmart G, *Symplectic integrators in sub-Riemannian geometry: the Martinet case* arXiv: math/0611380v1.

[33] Coddington EA and Levinson N, *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, Inc., New York-Toronto-London,(1955).

[34] Cortés J, *Geometric, control and numerical aspects of nonholonomic systems.* Volume **1793** of Lecture Notes in Mathematics. Springer-Verlag, Berlin, (2002).

[35] Cortés J,de León M, Martín de Diego D and Martínez S, *Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions.* SIAM J. Control Optim. **41(5)**, pp. 1389-1412, (2002).

[36] Cortés J, de León M, Marrero JC, Martín de Diego D and Martínez E, *A survey of Lagrangian mechanics and control on Lie algebroids and groupoids.* Int. J. Geom. Methods Mod. Phys. **3**, pp.509-558 (2006).

[37] Cortés J, de León M, Marrero JC and Martínez E, *Nonholonomic Lagrangian systems on Lie algebroids.* Discrete and Continuous Dynamical Systems: Series A, **24(2)**, pp. 213-271, (2009).

[38] Cortés J and Martínez S, *Nonholonomic integrators.* Nonlinearity, **14**, pp. 1365–1392, (2001).

[39] Coste A, Dazord P and Weinstein A, *Grupoïdes symplectiques.* Pub. Dép. Math. Lyon, **2/A**, pp. 1-62, (1987).

[40] Crainic M and Fernandez RL, *Integrability of Poisson brackets.* Annals of Mathematics, **157**, (2003) .

[41] Favretti M, *Equivalence of dynamics for nonholonomic systems with transverse constraints.* J. Dynam. Differential Equations, **10(4)**, pp. 511-536, (1998).

[42] Fedorov YN and Zenkov DV, *Discrete nonholonomic LL systems on Lie groups.* Nonlinearity, **18(5)**, pp. 2211-2241, (2005).

[43] Feng K, *On difference squemes nad symplectic geometry.* Proceedings of the 5-th International Symposium on differential geometri & differential equations, **3**, pp. 42-58, (1985).

[44] Ferraro S, Iglesias D and Martín de Diego D, *Momentum and energy preserving integrators for nonholonomic dynamics.* Nonlinearity (2008).

[45] Ferraro S, Iglesias D and Martín de Diego D, *Numerical and geometric aspects of the nonholonomic SHAKE and RATTLE methods.* Discrete and Continuous Dynamical Systems Supplement, (2009).

[46] Fetecau RC, Marsden JE, Ortiz M and West M, *Nonsmooth Lagrangian Mechanics and Variational Collision Integrators.* SIAM Journal of Dynamical Systems **2(3)**, pp. 381-416, (2003).

[47] Ge Z and Marsden JE, *Lie-Poisson integrators and Lie-Poisson Hamilton-Jacobi theory.* Phys. Lett. A, 133:134-139, (1988).

[48] Godbillon C, *Géometrie diférentielle et mécanique analytique.* Hermann, Paris (1969).

[49] Godbillon C, *Géometrie différentielle et mécanique analytique.* Hermann, Paris, (1969).

[50] Gonzalez O, Higham DJ and Stuart AM, *Qualitative properties of modified equations.* IMA J. Numer. Anal. **19**, pp. 169-190, (1999).

[51] Gotay MJ and Nester JM, *Presymplectic Lagrangian systems I: The constraint algorithm and the equivalence theorem.* Ann. Ins. H. Poincaré Sect. A (N.S.), **30**, pp. 129-142, (1979).

[52] Gotay MJ and Nester JM, *Presymplectic Lagrangian systems II: The second-order equation problem.* Ann. Ins. H. Poincaré Sect. A (N.S.), **32**, pp. 1-13 (1980).

[53] Gotay MJ, Nester JM and Hinds G, *Presymplectic manifolds and the Dirac-Bergmann theory of constraints.* J. Math. Phys., **32**, pp. 2744-2763, (1991).

[54] Grabowska K and Grabowski J, *Variational calculus with constraints on general algebroids.* J. Phys. A: Math. Theor. **41**, pp. 175-204, (2008).

[55] Grifone J, *Structure presque tangente et connexions I.* Annales de l'institut Fourier, **22**, pp. 287-334, (1972).

[56] Guillemin V and Sternberg S, *Symplectic Techniques in Physics.* Cambridge University Press (1984).

[57] Hairer E, *Backward analysis of numerical integrators and symplectic methods.* Ann. Numer. Math. **1**, pp. 107-132, (1994).

[58] Hairer E and Lubich C, *The life-span of backward error analysis for numerical integrators.* Numer. Math. **76**, pp. 441-462, (1997).

[59] Hairer E and Lubich C, *Invariant tori of dissipatively perturbed Hamiltonian systems under symplectic discretization.* Appl. Numer. Math. **29**, pp. 57-71, (1999).

[60] Hairer E and Lubich C, *Long-time energy conservation of numerical methods for oscillatory differential equations.* SIAM J. Numer. Anal. **38**, pp. 414-441, (2000).

[61] Hairer E, Lubich C, Wanner G, *Geometric Numerical Integration: Structure-Preserving algorithms for Ordinary Differential Equations* Springer, (2002).

[62] Hairer E, Norsett SP and Wanner G, *Solving Ordinary Differential Equations I : Nonstiff problems.* Volume 8 of Springer Series in Computational Mathematics. Springer-Verlag, second edition, (1993).

[63] Hardy DJ, Okunbor DI and Skeel R, *Symplectic variable step size integration for N-body problems.* in Proceedings of the NSF/CBMS Re- gional Conference on Numerical Analysis of Hamiltonian Differential Equations, Appl. Numer. Math. **29**, pp. 19-30, (1999).

[64] Higgins PJ and Mackenzie KCH, *Algebraic constructions in the category of Lie algebroids.* J. Algebra, **129**, pp. 194-230, (1990).

[65] Holm D, *Geometric Mechanics I, II.* Imperial College Press, (2009).

[66] Holm D, Marsden JE and Ratiu T, *The Euler-Poincaré equations and semidirect products with applications to continuum theories.* Adv. Math. 137, **1**, pp. 1-25, (1998).

[67] Husemoller D, *Fibre Bundles.* Springer series in Graduate Texts in Mathematics. Springer-Verlag, New York, (1994).

[68] Hwang CL and Fan LT, *A discrete version of Pontryagin's maximum principle.* Oper. Res. **15**, pp. 139-146, (1967).

[69] Ibort A, de León M, Lacomba E, Marrero JC and Martín de Diego D, *Geometric formulation of Carnot's theorem.* J. Phys. A, **34**(8):1691-1712, (2001).

[70] Iglesias D, Marrero JC, Martín de Diego and Martínez E, Discrete Nonholonomic Lagrangian Systems on Lie Groupoids. Journal of Nonlinear Science **18**, (2008).

[71] Iglesias D, Marrero JC, Martín de Diego D and Padrón E, *Discrete Dynamics in Implicit Form.* Preprint, arXiv:1011.3724v2 (2011).

[72] Iglesias D, Marrero JC, Martín de Diego D and Sosa D, *Lagrangian systems and variational constrained mechanics on Lie algebroids.* Dyn. Syst. **23(3)**, pp. 351-397, (2008).

[73] Iserles A, Munthe-Kaas H, Norsett S and Zanna A, *Lie-group methods.* Acta Numerica, (2005).

[74] Jaroszkiewicz G and Norton K, *Principles of discrete time mechanics, I: Particle systems.* J. Phys. A **30**, pp. 3115-3144, (1997).

[75] Jaroszkiewicz G and Norton K, *Principles of discrete time mechanics, II: Classical field theory.* J. Phys. A **30**, pp. 3145-3163, (1997b).

[76] Jiménez F, Kobilarov M and Martín de Diego D, *Discrete Variational Optimal Control.* Preprint, arXiv:1203.0580, (2012).

[77] Jiménez F, de León M and Martín de Diego D, *Hamiltonian dynamics and constrained variational calculus: continuous and discrete settings.* Accepted for publication in Journal of Physics A. arXiv:1108.5570, (2011).

[78] Jiménez F and Martín de Diego D, *A geometric approach to Discrete mechanics for optimal control theory.* Proceedings of the 49th IEEE Conference on Decision and Control (CDC), pp. 5426-5431. Atlanta, Georgia, USA, (2010).

[79] Jiménez F and Martín de Diego D, *Continuous and discrete approaches to vakonomic mechanics.* Revista de la Real Academia de las Ciencias Exactas, Físicas y Naturales, Serie A. Springer, **106(1)**, pp. 75-87, (2011).

[80] Jordan BW and Polak E, *Theory of a class of discrete optimal control systems.* J. Electron. Control **17**, pp. 697-711, (1964).

[81] Junge O, Marsden JE, and Ober-Blöbaum S, *Discrete mechanics and optimal control.* In 16th IFAC World Congress, Prague, Czech Republic, (2005).

[82] Junge O, Marsden JE and Ober-Blöbaum S, *Optimal reconfiguration of formation flying spacecraft - a decentralized approach.* In IEEE Conference on Decision and Control and European Control Conference ECC, San Diego, CA, USA, pp. 5210-5215, (2006).

[83] Junge O and Ober-Blöbaum S, *Optimal reconfiguration of formation flying satellites.* In IEEE Conference on Decision and Control and European Control Conference ECC, Seville, Spain, (2005).

[84] Kane C, Marsden JE and Ortiz M, *Symplectic energy-momentum integrators.* J. Math. Phys. 40, pp. 3353-3371, (1999).

[85] 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).

[86] Kane C, Repetto EA, Ortiz M and Marsden JE, *Finite element analysis of nonsmooth contact.* Comput. Meth. Appl. Mech. Eng. **180**, pp. 1-26, (1999).

[87] Kanso E and Marsden JE, *Optimal motion of an articulated body in a perfect fluid.* In IEEE Conference on Decision and Control and European Control Conference ECC, Seville, Spain, (2005).

[88] Klein J, *Espacies variationnels et Mécanique.* Annales de l'institut Fourier, **12**, pp. 1-124, (1962).

[89] Kobayashi S and Nomizu K, *Foundations of Differential Geometry. Vol I and II.* Wiley Classics Library. John Wiley and Sons Inc., New York, (1996).

[90] Kobilarov M, Desbrun M, Marsden JE and Sukhatme GS, *A discrete geometric optimal control framework for systems with symmetries.* In Robotics: Science and Systems (R:SS), Zurich, Switzerland, (2007).

[91] Kobilarov M and Marsden JE, *Discrete Geometric Optimal Control on Lie Groups.* IEEE Transactions on Robotics, **27(4)**, pp. 641-655, (2010).

[92] Kobilarov M and Sukhatme GS, *Optimal control using nonholonomic integrators.* In IEEE International Conference on Robotics and Automation (ICRA), Rome, Italy, pp 1832-1837, (2007).

[93] Kobilarov M, Martín de Diego D and Ferraro S, *Simulating Nonholonomic Dynamics.* Boletín de la Sociedad Espanola de Matemática Aplicada (S$\vec{e}$MA), **50**, pp. 61-81, (2010).

[94] Koiller J, *Reduction of some classical nonholonomic systems with symmetry.* Arch. Rational Mech. Anal., **118(2)**, pp. 113-148, (1992).

[95] Koon WS and Marsden JE, *Optimal Control for Holonomic and Nonholonomic Mechanical Systems with Symmetry and Lagrangian Reduction.* SIAM J. Control Optim. **35**, pp. 901-929, (1997).

[96] Koslov VV, *Realization of nonintegrable constraints in classical mechanics.* Dokl. Akad. Nauk. SSSL 273 **3**, 550-554, (1983).

[97] Lee TD, *Can time be a discrete dynamical variable?.* Phys. Lett. B **122**, pp. 217-220, (1983).

[98] Lee TD, *Difference equations and conservation laws.* J. Stat. Phys. **46**, pp. 843-860, (1987).

[99] Lee JM, *Riemannian Manifolds: An Introduction to Curvature.* Springer series in graduate Texts in Mathematics, vol 176. Springer-Verlag, New York, (1997).

[100] 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).

[101] 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).

[102] Leimkmuhler B and Skeel R, *Symplectic Numerical Integrators in Constrained Hamiltonian Systems.* J. Comput. Physics, **112**, (1994).

[103] de León M, *A historical review on nonholonomic mechanics.* Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A: Matematicas, **105**, Springer-Verlag, DOI: 10.1007/s13398-011-0046-2, (2011).

[104] de León M and Martín de Diego D, *On the geometry of nonholonomic Lagrangian systems.* J. Math. Phys., **37(7)**, pp. 3389-3414, (1996).

[105] de León M, Martín de Diego D and Martínez E, *Lagrangian submanifolds and dynamics on Lie algebroids.* J. Phys. A: Math. Gen. **38**, R241-R308, (2005).

[106] de León M, Martín de Diego D and Santamaría-Merino A, *Discrete variational integrators and optimal control theory.* Advances in Computational Mathematics (2006).

[107] de León M, Marrero JC and Martínez E, *Lagrangian submanifolds and dynamics on Lie algebroids.* J. Phys. A: Math. Gen. **38**, pp. 241-308, (2005).

[108] de León M and Rodrigues PR, *Methods of Differential Geometry in Analytical Mechanics.* North-Holland, Amsterdam (1989).

[109] Leyendecker S, Ober-Blöbaum S, Marsden JE and Ortiz M, *Discrete mechanics and optimal control for constrained multibody dynamics.* In 6th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, ASME International Design Engineering Technical Conferences, Las Vegas, Nevada, USA, (2007).

[110] Leyendecker S, Ober-Blöbaum S, Marsden JE and Ortiz M, *Discrete mechanics and optimal control for constrained systems.* Optimal Control Appl. Methods, **31(6)**, pp. 505-528, (2010).

[111] Libermann P and Marle CM, *Symplectic Geometry and Analytical Mechanics.* D. Reidel Publishing Company, Holland, (1987).

[112] Libermann P and Marle ChM, *Symplectic geometry and Analytical Mechanics.* Mathematics and its Applications, **35**. D. Reidel Publishing Co., Dordrecht, (1987).

[113] Logan JD, *First integrals in the discrete calculus of variations.* Aequa- tiones Mathematicae **9**, pp. 210-220, (1973).

[114] Mackenzie KCH, *Lie groupoids and Lie algebroids in differential geometry.* Cambridge University Press, (1987).

[115] Maeda S, *Canonical structure and symmetries for discrete systems.* Math. Japonica **25**, pp. 405-420, (1980).

[116] Maeda S, *Extension of discrete Noether theorem.* Math. Japonica **26**, pp. 85-90, (1981).

[117] Maeda S, *Lagrangian formulation of discrete systems and concept of difference space.* Math. Japonica **27**, pp. 345-356, (1981).

[118] Marrero JC, Martín de Diego D and Martínez E, *Discrete Lagrangian and Hamiltonian Mechanics on Lie groupoids.* Nonlinearity **19** (2006), no. 6, 1313-1348. Corrigendum: Nonlinearity **19**, no. 12, 3003-3004, (2006).

[119] Marrero JC, Martín de Diego D and Stern A, *Lagrangian submanifolds and discrete constrained mechanics on Lie groupoids.* Preprint, (2010).

[120] Marsden JE and Ostrowski J, *Symmetries in motion: Geometric foundations of motion control.* Nonlinear Sci. Today, (1998).

[121] Marsden JE, Pekarsky S and Shkoller S, *Discrete Euler-Poincaré and Lie-Poisson equations.* Nonlinearity, **12(6)**, pp. 1647-1662, (1999).

[122] Marsden JE, Pekarsky S and Shkoller S, *Symmetry reduction of discrete Lagrangian mechanics on Lie groups.* J. Geom. Phys. **36**, (1999).

[123] Marsden JE and Ratiu TS, *Introduction to mechanics and symmetry.* **17** of Texts in Appied Mathematics, Springer-Verlag, New York, (1999).

[124] Marsden JE and West M, *Discrete Mechanics and Variational Integrators.* Acta Numerica, Cambridge University Press, Cambridge, pp. 357-514, (2001).

[125] Marsden JE and Yoshimura H, *Dirac structures in Lagrangian mechanics, Part II: Variational structures.* Journal of Geometry and Physics, **57**, (2006).

[126] Martin J, *Discrete mechanics and optimal control.* Master's thesis, Department of Control and Dynamical Systems, California Institute of Technology, (2006).

[127] 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).

[128] Martínez E, *Lagrangian Mechanics on Lie algebroids.* Acta Appl. Math., **67**, pp. 295-320, (2001).

[129] Martínez E, *Reduction in optimal control theory.* Reports on Mathematical Physics **53**, pp. 79-90, (2004) .

[130] Martínez S, Cortés J and de León M, *The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems. The vakonomic bracket.* J. Math. Phys. **41(4)**, pp. 2090-2120, (2000).

[131] McLachlan RI and Marsland S, *Discrete mechanics and optimal control for image registration.* In Computational Techniques and Applications Conference (CTAC), (2006).

[132] McLachlan R and Perlmutter M, *Integrators for nonholonomic Mechanical Systems.* J. Nonlinear Sci., (2006).

[133] Mendella G, Marmo G and Tulczyjew WM, *Integrability of implicit differential equations.* Journal of physics A: Mathematical and general **28** 1, 149-164, (1995).

[134] Moser J and Veselov P, *Discrete versions of some classical integrable systems and factorization of matrix polynomials.* Comm. Math. Phys. **139(2)**, pp. 217-243, (1991).

[135] Nakahara M, *Geometry, Topology and Physics.* Adam Higer imprint by IOP Publishing Ltd., Bristol, New York, Philadelphia, (1990).

[136] Neimark J and Fufaev N, *Dynamics on Nonholonomic systems.* Translations of Mathematics Monographs, **33**, AMS, Providence, RI (1972).

[137] Norton K and Jaroszkiewicz G, *Principles of discrete time mechanics, III: Quantum field theory.* J. Phys. A **31**, pp. 977-1000, (1998).

[138] Ober-Blöbaum S, *Discrete Mechanics and Optimal Control* PhD thesis, University of Paderborn(2008).

[139] Ober-Blöbaum S, Junge O and Marsden JE, *Discrete Mechanics and Optimal Control: an Analysis.* ESAIM Control Optim. Calc. Var. **17**, (2011).

[140] Patrick GW and Cuell C, *Error analysis of variational integrators of unconstrained Lagrangian systems.* Numerische Mathematik **113**, pp. 243-264, (2009).

[141] Pekarek D, Ames AD and Marsden JE, *Discrete mechanics and optimal control applied to the compass gait biped.* In IEEE Conference on Decision and Control and European Control Conference ECC, New Orleans, USA, (2007).

[142] Poincaré H, *Sur une forme nouvelle des équations de la mécanique.* C.R. Acad. Sci. Paris, **132**, pp. 369-371, (1901).

[143] Pontryagin LS, Boltyanskii VG, Gamkrelidze RV and Mishchenko EF, *The Mathematical Theory of Optimal Processes.* Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt. Interscience Publishers John Wiley & Sons, Inc. New York-London, (1962).

[144] 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).

[145] 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).

[146] Pradines J, *Fibrés vectoriels doubles et calcul des jets non holonomes.* Amiens (1974).

[147] Reich S, *Backward error analysis for numerical integrators.* SIAM J. Numer. Anal. **36**, pp. 1549-1570, (1999).

[148] Reich S, *Preservation of adiabatic invariants under symplectic discretization.* Appl. Numer. Math. **29**, pp. 45-55, (1999).

[149] Ross SD, *Optimal flapping strokes for self-propulsion in a perfect fluid.* In American Control Conference, Minneapolis, Minnesota, USA, pp. 4118-4122, (2006).

[150] Ruth RD, *A canonical integration technique.* IEEE Trans. Nuclear Sci. **30**, pp. 2669-2671, (1983).

[151] Ryckaert JP, Ciccotti G and Berendsen H, *Numerical integration of the cartesian equations of motion of a system with constraint: molecular dynamics of n-alkanes.* J. Comput. Physics, **23**, (1977).

[152] Sanz-Serna JM, *Symplectic integrators for Hamiltonian problems: An overview.* Acta Numerica, **1**, Cambridge University Press, pp. 243- 286, (1992).

[153] Sanz-Serna JM and Calvo MP, *Numerical Hamiltonian problems.* Applied Mathematics and Mathematical Computation **7**, Chapman and Hall, London, (1994).

[154] Saunders D, *Prolongations of Lie groupoids and Lie algebroids.* Houston J. Math. 30 (3), pp. 637-655, (2004).

[155] Schlick T, Skeel R et al, *Algorithmic challenges in computational molecular biophysics.* J. Comput. Phys. **151**, pp. 9-48, (1999).

[156] Shimada M and Yoshida H, (1996) *Long-term conservation of adiabatic invariants by using symplectic integrators.* Publ. Astronomical Soc. Japan **48**, pp. 147-155, (1996).

[157] Tulczyjew WM, *Les sous-variétés lagrangiennes et la dynamique hamiltonienne.* C. R. Acad. Sc. Paris **283** Série A, 15-18, (1976).

[158] Tulczyjew WM, *Les sous-variétés lagrangiennes et la dynamique lagrangienne.* C. R. Acad. Sc. Paris **283** Série A, 675-678, (1976).

[159] Vershik AM and Faddeev LD, *Differential geometry and Lagrangian mechanics with constraints.* Sov. Phys. Dokl., **17**, pp. 34-36, (1972).

[160] Veselov AP, *Integrable discrete-time systems and dierence operators.* Funct. Anal. Appl. **22**, pp. 83-93, (1988).

[161] Veselov AP, *Integrable Lagrangian correspondences and the factorization of matrix polynomials.* Funct. Anal. Appl. **25**, pp. 112-122, (1991).

[162] de Vogelaére R, *Methods of integration which preserve the contact transformation property of the Hamiltonian equations.* University of Notre Dame preprint, (1956).

[163] Warner F, *Foundations of differentiable manifolds and Lie groups.* Scott, Foresman, Glenview, Ill, (1973).

[164] Weinstein A, *Symplectic categories.* Port. Math. **67(2)**, pp. 261–278, (2010).

[165] Weinstein A, *Lagrangian Mechanics and groupoids.* Fields Inst. Comm. **7**, (1996).

[166] Weinstein A, *Lectures on symplectic manifolds.* CBMS Regional Conference Series in Mathematics, **29**. American Mathematical Society, Providence, R.I., (1979).

[167] Wendlandt JM and Marsden JE, *Mechanical integrators derived from a discrete variational principle.* Physica D **106**, pp. 223-246, (1997).

[168] Wendlandt JM and Marsden JE, *Mechanical systems with symmetry, variational principles and integration algorithms.* Current and Future Di- rections in Applied Mathematics, Birkhäuser, pp. 219-261, (1997).